Skript zur Vorlesung: Lineare Algebra für Informatiker, SS2012

Hannes Diener

4. September 2013

Vorwort

Wie so oft, soll auch an dieser Stelle die Warnung stehen, daß das Lesen des Skriptes nicht den Besuch der Vorlesung ersetzt. Es erhebt keinen Anspruch auf Vollständigkeit. Der große weiße Rand auf der rechten Seite ist für Notizen und Anmerkungen gedacht. Für das Auffinden eventueller Fehler (Rechtschreib- oder schwerwiegendere) bin ich immer dankbar.

Inhaltsverzeichnis

Literaturverzeichnis

[1] Teschl, Gerald und Susanne. *Diskrete Mathematik für Informatiker. 3. Auflage*, 2008. Als ebook kostenlos verfügbar (an der Uni oder über VPN).
Sehr leserlich, manchmal etwas verspielt.

[2] Fischer, Gerd. *Lineare Algebra. 17. Auflage*,
als ebook kostenlos verfügbar (an der Uni oder über VPN).
Das Standardwerk im deutschen Sprachraum. Eher abstrakter Zugang zum Thema.

[3] Beutelspacher, Albrecht. *Lineare Algebra. 7. Auflage*, 2010. Mehrere Versionen und Auflagen in der Unibib erhältlich.
Wie alle Bücher von Beutelspacher äusserst empfehlenswert.

[4] Beutelspacher, Albrecht. *Das ist o.B.d.A. trivial*, 2011.
als ebook kostenlos verfügbar (an der Uni oder über VPN).
Eine hervorragende Einführung über das Formulieren und Beweisen mathematischer Aussagen.

[5] Beutelspacher, Albrecht. *Survival-Kit Mathematik*, 2009.
als ebook kostenlos verfügbar (an der Uni oder über VPN).
Der Schnellzugang zu den Basics

[6] David Austin *How Google Finds Your Needle in the Web's Haystack* AMS Feature Column,
http://www.ams.org/featurecolumn/archive/pagerank.html
Zum Ende der Vorlesung

0 Vorwissen, Notation

0.1 Logik

Die Einsicht, die der formalen Logik zu Grunde liegt, ist, daß einige Argumente nur auf Grund ihrer Struktur richtig sind. So ist

> „Wenn alle Kreter lügen und Epimenides Kreter ist, dann lügt Epimenides."

immer korrekt – und das unabhängig davon ob wirklich alle Kreter lügen. Der Wahrheitsgehalt ändert sich ebenfalls nicht, wenn man "Epimenides" durch "Angela" ersetzt; oder "Epimenides durch "5" die Aussage "Kreter" durch "Primzahl" und "lügen" durch "ungerade sein" ersetzt.

Wie sich im 20. Jhd. zeigte kann die gesamte Mathematik auf extrem wenige Grundannahmen, welche man als wahr akzeptieren muss, nur durch logische Schlüsse, die immer korrekt sind, aufgebaut werden. Um dies einwandfrei zu tun muss man zunächst die verwendete Sprache präzisieren.

Als erstes benötigt man Namen für Objekte – die sogenannten **Terme**.[1]

Definition 0.1.1. Ein Term ist induktiv definiert durch die folgenden Regeln:

- Gewisse Konstanten wie z.B. $0, 1, e, \pi$ sind Terme. Oft verwenden wir die Buchstaben $c, d, \ldots, c_1, c_2, c_3, \ldots$.
- Alle Variablen sind Terme. Normalerweise verwenden wir für Variablen die Buchstaben x, y, z oder auch $x_1, x_2, \ldots$.
- Sind $t_1, \ldots, t_n$ Terme und ist f ein n-stelliges Funktionssymbol, so ist auch $f(t_1, \ldots, t_n)$ ein Term.

[1] Welche Konstanten, Funktionssymbole usw. wir benötigen hängt natürlich vom Autor und Teilgebiet der Mathematik ab. Prinzipiell könnte man sich aber auf genau ein System einigen.

Beispiele für Terme sind also x oder c aber auch $f(x, y)$ oder $g(f(c, g(y)))$ – natürlich unter der Annahme, daß f und g die richtigen Stelligkeiten haben.

Als nächstes wollen wir Aussagen über Objekte machen. Hierzu wieder eine induktive Definition, welche Sätze wir überhaupt betrachten.

Definition 0.1.2. Eine Aussage ist induktiv definiert durch die folgenden Regeln:

- Grundaussagen sind die sogenannten Propositionale bzw. (n-stelligen) Prädikate. Für abstrakte Propositionale bzw. Prädikate verwenden wir oft die Großbuchstaben $P, Q, \ldots$. Ist P ein n-stelliges Prädikat und $t_1, \ldots, t_n$ Terme, so ist $P(t_1, \ldots, t_n)$ eine Aussage. Ebenso ist jedes Propositional eine Aussage.

- Sind α, β Aussagen, so auch

  1. $\alpha \wedge \beta$ – „α und β“,
  2. $\alpha \vee \beta$ – „α oder β“,
  3. $\neg\alpha$ – „nicht α“,
  4. $\alpha \implies \beta$ – „α impliziert β“, „wenn α dann β“, usw.
  5. $\exists x : \alpha$ – „Es existiert (mindestens) ein x so dass α gilt“,
  6. $\forall x : \alpha$ – „Für alle x gilt α “.

  Ausserdem benutzen wir auch $\alpha \iff \beta$ als Abkürzung für

$$(\alpha \implies \beta) \wedge (\beta \implies \alpha) \ .$$

Ein Satz, in der keine Variable vorkommt, oder jede Variable durch einen Quantor $\forall$ oder $\exists$ gebunden wird, heißt **Aussage**. Kommen Variablen ungebunden (frei) vor, so sprechen wir von einer **Aussageform**.

Bis jetzt haben wir nur die Syntax erklärt, d.h. nur formal die Zeichen identifiziert, mit denen wir arbeiten wollen. Was noch fehlt ist die Bedeutung dieser Zeichenketten; die sogenannte Semantik. Unser Ziel ist es hierbei jeder Aussage einen Wahrheitswert zuzuordnen. Hier müssen wir zunächst festlegen wie die Terme interpretiert werden. Danach können wir jeder Aussage α einen Wahrheitswert zuordnen. Ist α ein Propositional bzw. ein Prädikat, in welches Terme eingesetzt sind, können wir einen der Interpretation angemessenen Wahrheitswert wählen. Haben wir uns hier festgelegt folgt der Wahrheitswert für kompliziertere Terme zwangsweise:

- $\alpha \wedge \beta$ soll genau dann wahr sein wenn sowohl α als auch β wahr sind.

- $\alpha \vee \beta$ soll genau dann wahr sein wenn entweder α oder β (oder beide) wahr sind.
- $\neg\alpha$ soll genau dann wahr sein, wenn α falsch ist.
- $\alpha \implies \beta$ soll genau dann wahr sein, wenn entweder α falsch ist, oder sowohl α als auch β wahr sind.
- $\exists x : \alpha$ soll genau dann wahr sein, wenn es ein Objekt gibt, so daß für dieses Objekt die Aussage α wahr ist.
- $\forall x : \alpha$ soll genau dann wahr sein, wenn für alle Objekte, die wir betrachten die Aussage α wahr ist.

Mit diesen Regeln können wir den Wahrheitswert aller Aussagen "berechnen". Insbesondere interessieren uns logische Schlüsse, d.h Aussagen der Form $\alpha \implies \beta$, welche *immer* wahr sind, d.h. unabhängig davon, wie wir die Terme, Propositionale und Prädikate interpretieren.

Satz 0.1.3. *Die folgenden Aussagen sind logische Schlüsse:*

1. $(\alpha \wedge \beta) \implies \alpha$
2. $(\alpha \wedge (\alpha \implies \beta)) \implies \beta$
3. $(\alpha \implies \beta) \iff (\neg\beta \implies \neg\alpha)$
4. $((\alpha \implies \beta) \wedge (\neg\alpha \implies \beta)) \implies \beta$
5. *Weitere kommen im Laufe der Vorlesung hinzu.*

Als letztes interessiert uns das Konzept eines **Beweises**. Unter einem Beweis verstehen wir die Herleitung von Aussagen aus wenigen Grundannahmen (Axiomen) und bereits bewiesenen Aussagen durch logische Schlüsse.

0.2 Mengen

Mengen bilden die Grundlage der modernen Mathematik. Auch wenn man Mengen axiomatisch formal einführen und betrachten kann wollen wir uns auf ein naives Verständnis beschränken,[2] welches für diese Vorlesung vollkommen ausreicht.

Eine Menge ist eine Ansammlung von Objekten. Meist verwenden wir für generische (nicht näher spezifizierten) Mengen große, lateinische

[2] Lange Zeit dachte man, daß dieser naive, intuitive Zugang sich problemlos formalisieren lässt. Diese Hoffnung wurde durch Bertrand Russel zerstört, indem er vorschlug die Menge zu betrachten, die alle Mengen enthält, die sich nicht selbst enthalten. Eine solche Menge kann nicht existieren, da sowohl die Annahme, daß sie sich selbst enthält wie auch die Annahme, daß sie sich nicht enthält zu Widersprüchen führt. Die Implikation dieses sogenannten *Russell'schen Paradox* ist, daß wir nicht beliebige Mengen durch Beschreibung einer Eigenschaft formen können, sondern nur nach gewissen Regeln geformte Mengen.

Buchstaben, also $A, B, C, \ldots$. Ist ein Objekt x Element einer Menge A, so schreiben wir

$$x \in A \ .$$

Ist dies nicht der Fall schreiben wir $x \notin A$. Für Objekte verwenden wir normalerweise kleine lateinische Buchstaben $a, b, c, \ldots$. Allerdings kann es natürlich auch sein, daß Mengen wieder selber Elemente von Mengen sind, in welchen Falle diese Groß- und Kleinschreibungsregel nicht mehr funktioniert.

Mengen können wir entweder durch Aufzählung oder Angabe von einer definierenden Eigenschaft beschreiben. In beiden Fällen verwenden wir geschweifte Klammern { und }. Erstere Möglichkeit funktioniert nur bei endlichen Mengen. So wird die Menge, die die Zahlen $1, 2, 4, 5$ enthält als $\{1, 2, 4, 5\}$ bezeichnet. Die zweite Notation ist besser für große und komplizierte Mengen geeignet. Die Struktur ist die folgende:

$$\{\text{Objekttyp} \mid \text{Eigenschaft}\}$$

also z.B. $\{x$ natürliche Zahl $\mid x$ ist gerade$\}$. Eine Mischform ist die "unendliche Aufzählung". Z.B. ist klar, daß $\{1, 3, 5, 7, 9, \ldots\}$ die Menge aller ungeraden, natürlichen Zahlen ist. Es sollte allerdings immer klar sein, welche Menge gemeint ist. Es ist beispielsweise nicht klar, ob $\left\{2, 4, \ldots, 2^{200}\right\}$ die Menge aller geraden Zahlen oder die aller Zweierpotenzen zwischen 2 und 2^{200} sein soll. Diese Schreibweise ist also nur eine abkürzende Notation zur Vereinfachung der Kommunikation zwischen Mathematikern.

Die langweiligste Menge ist die, die gar kein Objekt enthält; die sog. **leere Menge**. In Zeichen verwenden wir $\{\}$ oder auch $\emptyset$. Man beachte den Unterschied zwischen der leeren Menge $\{\}$ und der Menge die die leere Menge enthält $\{\{\}\}$.

Eine Menge A ist **Teilmenge** einer Menge B, in Zeichen $A \subseteq B$, wenn alle Elemente $x \in A$ auch in B sind. Formal

$$A \subseteq B \iff \forall x : x \in A \implies x \in B \ .$$

Dies liefert uns auch eine gute Beweisstrategie um zu zeigen, daß eine Menge A Teilmenge einer Menge B ist: Wir nehmen uns ein beliebiges Element x in A und zeigen, daß es dann auch in B ist. Sind die Mengen A und B durch Eigenschaften bestimmt, sagen wir einmal $A = \{x \mid P(x)\}$ und $B = \{x \mid Q(x)\}$, so müssen wir zeigen, daß wenn ein Element Eigenschaft P besitzt, es auch Eigenschaft Q hat.

Eine Menge A ist eine **echte Teilmenge** einer Menge B, wenn $A \subseteq B$ ist, aber es zumindest ein Element in B gibt, das nicht in A ist. (Also wenn $\neg(B \subseteq A)$ gilt). In Zeichen schreiben wir $A \subsetneq B$.[3]

[3] Man beachte, daß oft auch das Zeichen $\subset$ verwendet wird. Allerdings verwenden es einige Autoren für eine echte Teilmenge, während andere Autoren es für beliebige Teilmengen benutzen. Um dieser Verwirrung zu entgehen benutzen wir es einfach

Zwei Mengen sind gleich, wenn sie die gleichen Elemente enthalten; also

$$A = B \iff A \subseteq B \wedge B \subseteq A$$

Auch in dieser Definition steckt eine gute Beweisstrategie um zu zeigen, daß zwei Mengen A und B gleich sind. Wir zeigen sowohl $A \subseteq B$, als auch $B \subseteq A$ (mit obiger Strategie).

Aus bereits vorhandenen Mengen A, B können wir mit folgenden Konstruktionen neue bilden:

- Die Vereinigung $A \cup B$ von zwei Mengen ist die Menge, die alle die Objekte enthält, die entweder in A oder B sind. Also

$$A \cup B = \{x \mid x \in A \vee x \in B\} .$$

- Der Durchschnitt $A \cap B$ von zwei Mengen ist die Menge, die alle die Objekte enthält, die sowohl in A als auch B sind. Also

$$A \cup B = \{x \mid x \in A \wedge x \in B\} .$$

- Die Mengentheoretische Differenz $A \setminus B$ von zwei Mengen ist die Menge, die alle die Objekte enthält, die in A aber nicht in B sind. Also

$$A \setminus B = \{x \mid x \in A \wedge x \notin B\} .$$

Für zwei beliebige Objekte a, b wird (a, b) als **geordnetes Paar** bezeichnet. Zwei geordnete Paare (a, b) und (a', b') sind genau dann gleich, wenn $a = a'$ und $b = b'$. Analoge Definitionen kann man natürlich auch für drei oder mehr Objekte machen (Tripel, Quadrupel, usw. bzw. n-Tupel). Man beachte den Unterschied zwischen (a, b) und $\{a, b\}$: so gilt $\{1, 2\} = \{2, 1\}$ aber $(1, 2) \neq (2, 1)$. Das **Kreuzprodukt** oder auch **kartesische Produkt** zweier Mengen A und B ist die Menge

$$A \times B = \{\, (a, b) \mid a \in A \wedge b \in B \,\}.$$

0.3 Funktionen

Definition 0.3.1.[4] Seien X, Y Mengen. Eine **Abbildung** oder **Funktion** von X nach Y (formal schreiben wir $f : X \to Y$) ist eine Vorschrift, welche jedem $x \in X$ genau ein $y \in Y$ zuordnet; da dieses eindeutig ist schreiben wir auch $f(x)$ für dieses Element.

Im einfachsten Falle können wir eine Funktion durch einen Term beschreiben. Z.B. ist die Funktion $f : \mathbb{N} \to \mathbb{N}$ definiert durch $x \mapsto x^2$

gar nicht.

[4] Im letzten Semester haben wir gesehen, wie wir auch Funktionen als Spezialfälle von Mengen (nämlich totale, rechtseindeutige Teilmengen von $X \times Y$) auffassen können.

einfach nur die Quadratfunktion. Man beachte jedoch den feinen Unterschied zwischen dem Term x^2 (Syntax) und der Funktion f (mathematisches Objekt). Unser Funktionsbegriff ist viel allgemeiner, als der, der sich nur auf "Kurven" beschränkt: so ist die Funktion

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

graphisch nur schwer vorstellbar.

Die einfachste Funktion ist die **Identität** bzw. **identische Abbildung** $\mathrm{id}_X : X \to X$, die jedes Element $x \in X$ auf sich selbst abbildet; für die also gilt $\mathrm{id}_X(x) = x$.

Wichtig sind die folgenden Eigenschaften einer Funktion $f : X \to Y$

1. f heißt **injektiv**, wenn $f(x) = f(y) \implies x = y$ für alle $x, y \in X$. Dies ist äquivalent zu $x \neq y \implies f(x) \neq f(y)$.

2. f heißt **surjektiv**, wenn für alle $y \in Y$ ein $x \in X$ mit $f(x) = y$ existiert.

3. f heißt **bijektiv**, wenn f sowohl injektiv, als auch surjektiv ist.

Ist $f : X \to Y$ eine Abbildung und $g : Y \to Z$, so heißt die Funktion $g \circ f : X \to Z$ definiert durch

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

die **Komposition** oder auch **Hintereinanderausführung** von f und g.

Ist $f : X \to Y$ eine Funktion und $A \subseteq X$ und $B \subseteq Y$ Teilmengen, so ist

$$f(A) = \{\, f(x) \mid x \in A \,\}$$

das **Bild** von A und

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

das **Urbild** von A.

Man beachte, daß man das Urbild immer bilden kann, auch wenn die Funktion keine **Umkehrabbildung** hat. Letztere ist die eindeutige Abbildung, ebenfalls mit $f^{-1} : Y \to X$ bezeichnet, für die gilt $f^{-1} \circ f = \mathrm{id}_X$ und $f \circ f^{-1} = \mathrm{id}_Y$. Man kann zeigen, daß die Umkehrabbildung genau dann existiert wenn f bijektiv ist.

0.4 Arbeiten mit Summen: Σ-Notation

Oft arbeiten wir mit Summen endlich vieler, aber möglicherweise sehr vielen, Summanden. Solche Summen können wir beschreiben indem

wir die meisten Terme auslassen und durch drei Punkte ersetzen, also z.B.

$$1 + 2 + 3 + \cdots + 100 \ .$$

Dies geht allerdings nur, wenn das Bildungsgesetz klar ist. Nicht klar ist beispielsweise, ob

$$1 + 2 + \cdots + 2^n$$

die Summe der ersten 2^n natürlichen Zahlen, oder der ersten $n + 1$ Potenzen von 2 beschreibt.
Besser ist die Σ-Notation, die auch für Körper (und später Ringe und Vektorräume) Sinn macht. Wir definieren:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n \ .$$

Damit kann obiges Problem nicht mehr auftreten, da wir uns entweder für

$$\sum_{i=1}^{n} 2^i \quad \text{oder für} \quad \sum_{i=1}^{2^n} i$$

entscheiden müssen.

Die Summation muß nicht unbedingt bei 1 anfangen und bei n aufhören.[5] Auch ein Ausdruck wie

$$\sum_{i=-3}^{n+3} b_i$$

macht Sinn. Alternativ können wir auch den Summationsindex i (oder eine andere Variable) mit Bedingungen unter das $\sum$ schreiben:

$$\sum_{0 \leq i \leqslant 5} a_i \ ;$$

dies ermöglicht auch Konstruktionen wie

$$\sum_{0 \leqslant k \leqslant n, k \text{ prim}} a_k \ .$$

Die folgenden Rechenregeln sind nicht tiefgreifend, aber äußerst hilfreich, wenn es um die Manipulation von Summen geht.

(a)

$$c \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} c a_i$$

(b)

$$\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i = \sum_{i=1}^{n} (a_i + b_i)$$

[5] In der Analysis gibt man auch Ausdrücken der Form $\sum_{i=1}^{\infty} a_i$ Bedeutung.

(c)

$$\sum_{i=0}^{n} a_i = \sum_{i=k}^{n+k} a_{i-k}$$

(d)

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{m} b_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{m}(a_i b_j)$$

(e) Man beachte, daß im allgemeinen

$$\left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} b_i\right) \neq \sum_{i=1}^{n}(a_i b_i)$$

(f) Der gleiche Fehler wäre auch Gleichheit hier anzunehmen:

$$\left(\sum_{i=1}^{n} a_i\right)^2 \neq \sum_{i=1}^{n} a_i^2$$

1 Grundlagen

1.1 Gleichungen

Eine Gleichung ist eine Aussageform der Form

$$t_1 = t_2 \ ,$$

wobei t_1 und t_2 Terme sind. Terme wiederum sind wohlgeformte Ausdrücke aus Variablen, Konstanten und Funktionen.

Eine **Lösung einer Gleichung** ist eine Belegung der Variablen, so daß die Aussage wahr ist.

1.2 Zahlenbereiche

Als Lösungen und Konstanten in oben erwähnten Gleichungen müssen wir uns noch auf einen „Zahlenbereich“ festlegen. Wie aus der Schule bekannt hat z.B. die Gleichung

$$x + 7 = 3$$

keine Lösung in den natürlichen Zahlen $\mathbb{N}$ (obwohl alle Konstanten natürlich Zahlen sind), sehr wohl aber in den ganzen Zahlen $\mathbb{Z}$.

Historisch und praktisch am wichtigsten sind die folgenden Zahlenbereiche:

- Die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$.
- Die natürlichen Zahlen mit der Null $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$.
- Die ganze Zahlen $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$.
- Die rationalen Zahlen $\mathbb{Q}$.
- Die reellen Zahlen $\mathbb{R}$.

1.3 Andere Zahlenbereiche

In der Vorlesung *Diskreter Mathematik*[1] haben Sie die modulare Arithmetik kennengelernt. Hier werden, vereinfacht gesagt, nur mit den Resten bei der Division mit einer festgewählten Zahl n gerechnet. Z.B. gilt

$$7 + 7 \equiv 3 \mod 11 \ ,$$

da $7 + 7 = 14 = 11 + 3$. Verwendet man dazu noch eine vereinfachte Schreibweise ($=$ an Stelle von $\equiv$ und Weglassen von $\mod n$), so gilt in diesem Zahlenbereich $\mathbb{Z}_{11}$, daß

$$7 + 7 = 3 \ .$$

1.4 Körper

Was haben all diese Zahlenbereiche gemeinsam; bzw. welche Strukturen wollen wir als Zahlenbereiche betrachten?

Definition 1.4.1. Eine Menge K zusammen mit zwei Abbildungen $+_K : K^2 \to K$ und $\cdot_K : K^2 \to K$ heißt **Körper**, wenn die folgenden neun Axiome für alle $x, y, z \in K$ erfüllt sind:
Addition

A1 (Assoziativität) $(x +_K y) +_K z = x +_K +(y +_K z)$

A2 (Existenz des neutralen Elements) es gibt ein Element in K, das wir 0_K(**Nullelement**) nennen, für das gilt $x +_K 0_K = x$

A3 (Existenz des inversen Elements) zu jedem $x \in K$ gibt es ein Element das wir $-x$ nennen, für das gilt: $x +_K -x = 0_K$

A4 (Kommutativität) $x +_K y = y +_K x$

Multiplikation

M1 (Assoziativität) $(x \cdot_K y) \cdot_K z = x \cdot_K \cdot(y \cdot_K z)$

M2 (Existenz des neutralen Elements) es gibt ein Element in K, das wir 1_K(**Einselement**) nennen, für das gilt $x \cdot_K 1_K = x$ und ausserdem $1_K \neq 0_K$

M3 (Existenz des inversen Elements) zu jedem $x \neq 0_K$ gibt es ein Element das wir x^{-1} nennen, für das gilt: $x \cdot_K x^{-1} = 1_K$

M4 (Kommutativität) $x \cdot_K y = y \cdot_K x$

Distributivität

[1] Wenn sie diese Vorlesung nicht gehört haben, sollten Sie diesen Paragraphen und modulare Arithmetik, ignorieren. Interessierte können sich natürlich auch die entsprechende Kapitel in [1] durchlesen.

D $x \cdot_K (y +_K z) = (x \cdot_K y) +_K (x \cdot_K z)$

Schreibweisen 1.4.2. Zur Vereinfachung wollen wir uns auf folgende Schreibweisen einigen:

- Die Verknüpfungen $+_K$, $\cdot_K$ schreiben wir in Infix-Notation. Also $a +_K b$ an Stelle von $+_K(a, b)$.
- Anstelle von 0_K, 1_K, $+_K$, $\cdot_K$ schreiben wir, wenn keine Verwirrung ensteht, oft einfach auch 0, 1, +, ·.
- An Stelle von $a \cdot_K b$ schreiben wir oft ab.
- Die Assoziativgesetze erlauben uns oft überflüssige Klammern wegzulassen. Also z.B. $a + b + c$.
- Um noch mehr Klammern zu sparen wollen wir wie üblich Punkt vor Strich gelten lassen.

Im folgenden wollen wir einige Beispiele von Körpern betrachten.

Beispiel 1.4.3. *Daß $\mathbb{Q}$ und $\mathbb{R}$ (mit der üblichen Addition und Multiplikation) Körper sind, ist klar.*

Beispiel 1.4.4. *Aus dem letzten Semester (DMI):*
Ist p eine Primzahl, so ist $\mathbb{Z}_p$ also Arithmetik modulo p zusammen mit der Addition und Multiplikation ein Körper. (Wenn Sie sich erinnern ist die Tatsache, daß p Primzahl ist, genau dazu notwendig, daß jede Restklasse außer 0 ein multiplikatives Inverses besitzt.) Dieser Körper wird mit $\mathbb{F}_p$ bezeichnet.

Beispiel 1.4.5. *Der kleinste mögliche Körper besteht aus der Menge $K = \{0, 1\}$ und den Verknüpfungen $+_K$ und $\cdot_K$ die durch folgende Verknüpfungstabellen definiert sind:*

$+_K$	0	1
0	0	1
1	1	0

und

$\cdot_K$	0	1
0	0	0
1	0	1

Die Axiome lassen sich leicht nachrechnen.
(Mit etwas Überlegung kann man zeigen, daß dies (bis auf Isomorphismen) der einzige Körper mit zwei Elementen ist. Beispielsweise ist es prinzipiell der gleiche Körper wie $\mathbb{F}_2$).

Wer mit Gruppen vertraut ist, dem hilft eventuell folgende Sichtweise:

Satz 1.4.6. *Ist $(K, +, \cdot)$ ein Körper, so ist $(K, +)$ und $(K \setminus \{0_K\}, \cdot)$ eine kommutative ("abelsche") Gruppe.*

Beweis. Das einzige Problem bei dieser Aussage ist zu zeigen, daß $K \setminus \{0_K\}$ unter der Multiplikation abgeschlossen ist. D.h., daß, wenn $a, b \in K \setminus \{0_K\}$ sind, muss auch $a \cdot b \neq 0_K$ sein. Dies folgt aber aus dem Lemma 1.4.9. □

Aus der Gruppensichtweise erhalten wir ausserdem:

Lemma 1.4.7. *Ist $(K, +, \cdot)$ ein Körper, so gibt es genau ein Nullelement, genau ein Einselement und zu jedem Element gibt es genau ein additives Inverses, und ein multiplikatives inverses (natürlich nicht für die 0).*

Beweis. Dies folgt aus der entsprechenden Aussage für Gruppen. □

Dies ermöglicht uns folgende Schreibweise zu verwenden:

Schreibweisen 1.4.8.

- Ist $a \in K$ ein Element in einem Körper, so wollen wir mit $-a$ das eindeutig bestimmte additive Inverse und mit a^{-1} das eindeutig bestimmte multiplikative Inverse (natürlich für $a \neq 0$) bezeichnen.
- Ausserdem wollen wir $a - b$ an Stelle von $a + (-b)$ schreiben.
- Manchmal ist es auch nett $\frac{a}{b}$ an Stelle von ab^{-1} zu schreiben.
- Für $\underbrace{a \cdot a \cdot \cdots \cdot a}_{n \text{ mal}}$ schreiben wir a^n und für $\underbrace{a + a + \cdots + a}_{n \text{ mal}}$ schreiben wir na. Bei dieser Schreibweise müssen wir aufpassen die natürlichen Zahlen und die Elemente des Körpers nicht durcheinanderzubringen!

Man beachte, das all diese Schreibweisen zwar nützlich, aber auch gefährlich sein können, da man nie vergessen darf, daß man in einem Körper und nicht unbedingt in den reellen Zahlen arbeitet. So folgt in unserem zwei-Elemente-Körper z.B. nicht aus $2a = 2b$, daß auch $a = b$ ist.

Lemma 1.4.9. *Ist $(K, +, \cdot)$ ein Körper, so gilt:*

1. *Es ist $a \cdot 0 = 0$ für alle $a \in K$.*
2. *Sind $a, b \in K$, $a \neq 0$ und $b \neq 0$, so ist auch $ab \neq 0$.*
3. *Ist $ab = 0$, dann ist entweder $a = 0$ oder $b = 0$.*
4. *$a(-b) = -(ab)$ und $(-a)(-b) = ab$*
5. *Ist $a \neq 0$ und $a \cdot b = a \cdot c$, so ist $b = c$.*

Beweis. 1. Es ist $a \cdot 1 = a \cdot (1 + 0) = a \cdot 1 + a \cdot 0$. Addieren wir $(-(a \cdot 1))$ zu beiden Seiten dieser Gleichung erhalten wir $0 = a \cdot 0$.

2. Nehmen wir an $ab = 0$. Da $a \neq 0$ existiert das multiplikative Inverse zu a. Multiplizieren wir $ab = 0$ mit diesem, so erhalten wir $b = 0$, ein Widerspruch zur Annahme. Also ist $ab \neq 0$.

3. Diese Aussage ist logisch äquivalent zu vorherigen Aussage.

4. Da das additive Inverse eindeutig ist, müssen wir zeigen, daß $ab + a(-b) = 0$. Dies folgt leicht mit dem Distributivgesetz. Für die nächste Aussage wenden wir diese gerade gezeigte Tatsache zweimal an:

$$(-a)(-b) = -(a(-b)) = -(-(ab)) \ .$$

   Ausserdem ist $-(-x) = x$ für alle Elemente eines Körpers (dies folgt aus dem entsprechenden Beweis für Gruppen).

5. Dies folgt leicht durch Multiplikation der Gleichung $ab = ac$ mit a^{-1}. Allerdings werden wir später in Lemma 2.0.5 sehen, daß diese Kürzungsregel auch in allgemeineren Fällen gilt.

□

Beispiel 1.4.10. *In dem Körper mit nur zwei Elementen hat die Gleichung*

$$x^2 = -1$$

eine Lösung: Da hier $1_K +_K 1_K = 0_K$ *ist, ist* $1_K = -1_K$*, d.h. die* 1 *ist ihr eigenes Inverses. Ausserdem ist* $1_K \cdot_K 1_K = 1_K$*. Also* $1^2 = 1 \cdot 1 = 1 = -1$*.*

1.5 Komplexe Zahlen

In diesem Kapitel *konstruieren* wir einen Körper, in dem die Gleichung

$$x^2 = -1$$

eine Lösung hat. Der (durchaus interessante) historische Hintergrund soll uns genauso wenige interessieren wie die Anwendungen beispielsweise in der Physik.

Um einen Körper zu konstruieren, müssen wir eine Menge und zwei Abbildungen angeben, und danach zeigen, daß alle neun Körperaxiome erfüllt sind.

Die Grundmenge ist $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, also die Menge aller Paare von reellen Zahlen. Ausserdem ist die Addition $+_{\mathbb{C}} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ punkteweise definiert durch $(a, b) +_{\mathbb{C}} (c, d) \mapsto (a + c, b + d)$ und die Multiplikation durch $\cdot_{\mathbb{C}} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ durch

$$(a, b) \cdot_{\mathbb{C}} (c, d) \mapsto (ac - bd, ad + bc) \ .$$

Als nächstes müssen wir noch die Axiome der Addition, die der Multiplikation und das Distributivgesetz überprüfen. Die meisten dieser Axiome sind einfach nachzuweisen, und der Beweis soll hier nicht geführt werden. Es sei nur erwähnt, daß $(0,0)$ das Nullelement und $(1,0)$ das Einselement ist. Die einzigen problematischen Axiome sind die Assoziativität der Multiplikation, die Distributivität und die Existenz des multiplikativen Inversen. Die ersten beiden sollen in den Übungen behandelt werden.
Zum Inversen der Multiplikation:
Sei $x = (a,b) \neq 0_{\mathbb{C}} = (0,0)$. Da zwei Paare ungleich sind, wenn die ersten oder die zweiten Komponenten nicht übereinstimmen, heißt das, daß dann entweder a oder b ungleich 0 sind. Auf alle Fälle ist $a^2 + b^2 \neq 0$. Wir behaupten, daß

$$z = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right)$$

das Inverse zu $x = (a,b)$ ist, wie man einfach nachrechnet:

$$\begin{aligned} x \cdot_{\mathbb{C}} z &= (a,b) \cdot_{\mathbb{C}} \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right) \\ &= \left(\frac{aa}{a^2+b^2} - \frac{-bb}{a^2+b^2}, \frac{ab}{a^2+b^2} - \frac{ba}{a^2+b^2} \right) \\ &= \left(\frac{a^2+b^2}{a^2+b^2}, \frac{ab-ab}{a^2+b^2} \right) \\ &= (1,0) = 1_{\mathbb{C}} \end{aligned}$$

Damit haben wir gezeigt, daß $\mathbb{C}$ ein Körper ist, in dem man also (meistens) wie gewohnt rechnen und Gleichungen lösen kann. Zusätzlich gilt aber auch:

$$(0,1)(0,1) = (-1,0) = -1_{\mathbb{C}} \ ;$$

es gibt in diesem Körper also auch ein Element, dessen Quadrat -1 ist.

Im folgenden wollen wir die Darstellung komplexer Zahlen vereinfachen. Dazu fassen wir zunächst jede reelle Zahl r als die komplexe Zahl $(r,0)$ auf. Wie man leicht nachrechnet gilt:

$$r(a,b) = (r,0)(a,b) = (ra, rb) \ ,$$

und

$$(a,b) = a(1,0) + b(0,1) \ . \tag{1.1}$$

Der entscheidende Trick ist nun, daß wir das Symbol i für die Zahl $(0,1)$ verwenden. Diese Zahl nennen wir die **imaginäre Einheit**. Dank der Gleichung 1.1 heißt das, das wir jede komplexe Zahl (a,b) auch als $a + bi$ schreiben können, wobei a und b reelle Zahlen sind. Mit

diesen Term können wir wie mit reellen Zahlen rechnen! Ausserdem gilt $i^2 = -1$. Um durch eine komplexe Zahl zu teilen – also mit dem multiplikativen Inversen zu multiplizieren – müssen wir etwas arbeiten. Z.B. ist $-i$ das Inverse zu i, da $(-i)i = -i^2 = -(-1) = 1$. Anders ausgedrückt gilt also

$$\frac{1}{i} = -i \ .$$

Für kompliziertere komplexe Zahlen hilft oben hergeleitete Formel, die in unser neuen Schreibweise lautet für $a, b \in \mathbb{R}$:

$$(a + bi)^{-1} = \frac{a - bi}{a^2 + b^2} \ ,$$

natürlich nur für $a + bi \neq 0$.

1.6 Mehr über komplexe Zahlen

Definition 1.6.1. Ist $x = a + bi$ eine komplexe Zahl mit $a, b \in \mathbb{R}$, so wollen wir mit $\Re(x)$ ihren **Realteil** a und mit $\Im(x)$ ihren **Imaginärteil** b bezeichnen.
Die Zahl $\overline{x} = a - bi$ heißt die zu $x = a + bi$ **konjugiert komplexe Zahl**

Satz 1.6.2. *Für die komplexe Konjugation gelten folgende Rechenregeln, für alle $x, y \in \mathbb{C}$*

1. $\overline{x + y} = \overline{x} + \overline{y}$
2. $\overline{x \cdot y} = \overline{x} \cdot \overline{y}$
3. $x \in \mathbb{R} \iff x = \overline{x}$

Wie man leicht nachrechnet ist $x\overline{x} = a^2 + b^2 \in \mathbb{R}_0^+$. Dies ermöglicht folgende Definition:

Definition 1.6.3. Ist $x = a + bi$ eine komplexe Zahl mit $a, b \in \mathbb{R}$, so ist ihr **absoluter Betrag** $|x|$ definiert durch $\sqrt{x\overline{x}} = \sqrt{a^2 + b^2}$.

Wie man leicht nachrechnet gilt:

Satz 1.6.4. *Für alle $x, y \in \mathbb{C}$ ist*

$$|x + y| \leqslant |x| + |y|$$

und

$$|xy| = |x||y| \ .$$

Man beachte allerdings, daß sich auf $\mathbb{C}$ keine vernünftige Ordnung definieren lässt. Die Ordnung in obigen Satz bezieht sich auf die Ordnung in $\mathbb{R}$!

1.7 geometrische Interpretation der komplexen Zahlen

Siehe Vorlesung.

1.8 Lineare Gleichungen in Körpern

Satz 1.8.1. *Ist K ein Körper, mit $a, b, c \in K$ und $a \neq 0$, so haben die Gleichungen*

$$ax = b$$

bzw.

$$ax + c = 0$$

genau eine Lösung.

Beweis. Es ist leicht zu überprüfen, daß $x = a^{-1}b$ eine Lösung für die erste Gleichung ist – die Gleichung hat also mindestens eine Lösung. Dies ist die einzige Lösung. Denn sind x, y Lösung, gilt also $ax = b$ und $ay = b$, dann gilt zusammen $ax = ay$. Da $a \neq 0$ gilt $x = y$.
Ebenso argumentiert man für $ax + c = 0$. □

Auch wenn dieser Beweis etwas kompliziert erscheint, sollte man ihn gut verstanden haben, denn dieser Typ Beweis und die Technik wird uns immer wieder begegnen: Um zu zeigen, daß es *genau* eine Lösung gibt zeigen wir:[2]

- Es gibt mindestens eine Lösung (z.B. indem wir, wie hier, die Lösung explizit angeben).
- Es gibt maximal eine Lösung (meistens, wie hier, indem wir zwei beliebige Lösungen hernehmen und zeigen, daß diese identisch sein müssen).

Komplizierter wird die Situation, wenn wir nach Lösungen suchen, die nicht nur eine Gleichung erfüllen, sondern mehrere. Aber dazu mehr im Teil 3.

[2] Die Reihenfolge ist hier egal. Oft ist es sogar besser zuerst die Eindeutigkeit zu zeigen, und dann die Existenz, da der Nachweis der Eindeutigkeit meist einfacher ist und man vielleicht die Idee zum Nachweis der Existenz bekommt.

2

Ringe und Polynome

Da es kein multiplikatives Inverses zu jeder Zahl in $\mathbb{Z}$ gibt, ist $\mathbb{Z}$ kein Körper. Trotzdem ist es ein Beispiel einer weiteren wichtigen algebraischen Struktur, die uns noch öfters begegnen wird.

Definition 2.0.2. Ein **Ring** besteht wie ein Körper aus einer Menge R und zwei Verknüpfungen "Multiplikation" und "Addition"; $+ : R^2 \to R$ und $\cdot : R^2 \to R$. Ausserdem müssen alle Körperaxiome erfüllt sein, bis auf die Kommutativität der Multiplikation und der Existenz der multiplikativen Inversen. Durch das Fehlen der Kommutativität benötigen wir allerdings zwei Distributivgesetze: Für alle $x, y, z \in R$

$$x(y+z) = xy + xz \text{ und } (y+z)x = yx + zx \ .$$

Ist die Multiplikation kommutativ, so sprechen wir von einem **kommutativen Ring**.

Natürlich ist jeder Körper ein Ring. Beispiel einer Struktur, die ein Ring, aber kein Körper ist, sind die ganzen Zahlen. Ein weiteres Beispiel bietet wiederum die Arithmetik modulo einer Zahl n, die jetzt aber keine Primzahl sein muß.

Beispiel 2.0.3. *Man beachte, daß man in Ringen nicht wie in Körpern rechnen kann: es gilt nicht unbedingt, daß $ab = ac$ für $a \neq 0$ folgt, daß $b = c$. Man kann also nicht beliebig kürzen. Betrachten wir z.B. den Restklasse-ring $\mathbb{Z}_6$. Hier gilt zwar*

$$3 \cdot 2 = 3 \cdot 4 \ ,$$

aber $2 \neq 4$.

Definition 2.0.4. Gilt in einem Ring, daß aus

$$ab = 0$$

folgt, daß entweder $a = 0$ oder $b = 0$, dann heißt der Ring **nullteilerfrei**.

In einem nullteilerfreien Ring können wir (wie in einem Körper) kürzen. Mehr noch Nullteilerfreiheit ist äquivalent zum Kürzen.

Lemma 2.0.5. *Sei R ein Ring. R ist genau dann nullteilerfrei, wenn für alle $a, b, c \in R$ mit $a \neq 0$ aus $ab = ac$ folgt, daß $b = c$.*

Beweis. Nehmen wir zunächst an, daß R nullteilerfrei ist, und a, b, c wie in der Voraussetzung. Addieren wir $a(-c)$ auf beiden Seiten der Gleichung $ab = ac$, so erhalten wir $ab + a(-c) = ac + a(-c)$.
Wenden wir nun das Distributivgesetz auf beiden Seiten an, so erhalten wir

$$a(b - c) = a(c - c) \ .$$

Wie im Beweis für Körper können wir auch in Ringen zeigen, daß $x \cdot 0 = 0$ ist. Also folgt, daß

$$a(b - c) = 0 \ .$$

Aus der Nullteilerfreiheit folgt, daß entweder $a = 0$ ist, was durch die Annahme ausgeschlossen ist, oder das $b - c = 0$ ist. Daraus folgt, daß $b = c$ gilt.

Umgekehrt, nehmen wir an, daß wir in einem Ring kürzen können, und sei $a, b \in R$ mit $ab = 0$. Wie bei Körpern kann man auch bei Ringen zeigen (gleicher Beweis) daß $x \cdot 0 = 0$ für alle $x \in R$. Also ist $ab = a \cdot 0$. Nun ist entweder $a = 0$ und wir sind fertig, oder $a \neq 0$ und wir können kürzen und $b = 0$. □

Einen wichtigen nullteilerfreien Ring bilden die Polynome über einem Körper. Sei hierfür X ein ansonsten unbenutztes Zeichen.

Definition 2.0.6. Sei $(R, +, \cdot)$ ein Ring. Ein **Polynom** mit **Koeffizienten** in R ist ein *formaler Ausdruck* der Gestalt

$$f(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n \ ,$$

wobei $a_0, \ldots, a_n \in R$.

Mit $R[X]$ bezeichnen wir die Menge all dieser Polynome.
Sind alle $a_i = 0$, so sprechen wir vom **Nullpolynom**.
Der **Grad** eines Polynoms f ist erklärt durch

$$\deg f = \begin{cases} -\infty & \text{falls } f = 0 \\ \max\{i \in \mathbb{N} \mid a_i \neq 0\} & \text{sonst.} \end{cases}$$

Es ist naheliegend in ein Polynom $f \in R[X]$ ein Element $\lambda \in R$ einzusetzen:

$$f(\lambda) := a_0 + a_1 \lambda + \ldots a_n \lambda^n \ .$$

Man beachte, daß $f(\lambda) \in R$ für $\lambda \in R$. Ein Polynom $f \in R[X]$ induziert also eine Abbildung $\tilde{f} : R \to R$ durch $\lambda \mapsto f(\lambda)$.

Der Unterschied zwischen dem formalen Polynom f und der induzierten Abbildung $\tilde{f}$ ist notwendig: sei beispielsweise in unserem Körper mit zwei Elementen das Polynom $f(X) = X + X^2$. Dann ist f ungleich dem Nullpolynom, da nicht alle Koeffizienten null sind. Allerdings gilt in K: $f(0) = 0 + 0^2 = 0$ und $f(1) = 1 + 1^2 = 1 + 1 = 0$. Die induzierte Abbildung ist also die Nullabbildung.

Man vergleiche diese Situation mit Programmen, wo es sehr leicht ist einzusehen, daß zwei nicht-identische Programme die gleiche Funktion berechnen können.

Um zu zeigen, daß $R[X]$ ein Ring ist, müssen wir noch die Addition und Multiplikation erklären. Für $f = a_0 + a_1X + \cdots + a_nX^n, g = b_0 + b_1X + \cdots + b_mX^m \in R[X]$ mit o.B.d.A. $n = m$ sei

$$f + g = (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_n + b_n)X^n$$

und

$$fg = c_0 + c_1X + \cdots + c_{n+m}X^{n+m} \ ,$$

wobei $c_k = \sum_{i+j=k} a_ib_j$.
Also ist z.B.

$$\begin{aligned} c_0 &= a_0b_0 \\ c_1 &= a_0b_1 + a_1b_0 \\ c_2 &= a_0b_2 + a_1b_1 + a_2b_0 \\ c_3 &= a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0 \\ &\vdots \end{aligned}$$

Wie man nachrechnen kann ist $R[X]$ wieder ein Ring. Mehr noch

Lemma 2.0.7. *Ist R nullteilerfrei, so gilt*

$$\deg(f \cdot g) = \deg f + \deg g \ .$$

Dabei soll formal gelten: $-\infty + n = m + (-\infty) = -\infty + (-\infty) = -\infty$. In jedem Ring gilt außerdem, daß

$$\deg(f + g) \leqslant \deg(f), \deg(g) \ .$$

Beweis. Vorlesung. □

Wie bei ganzen Zahlen, kann man auch in Polynomringen über Körpern(!) mit Rest dividieren.

Satz 2.0.8. *Ist K ein Körper und sind $f, g \in K[X]$ und $g \neq 0$, so gibt es eindeutig bestimmte Polynome $q, r \in K[X]$, so daß*

$$f = q \cdot g + r \text{ und } \deg r < \deg g \ .$$

Beweis. Der Beweis ist nicht schwer. An dieser Stelle soll er aber nur durch ein Beispiel illustriert werden. Das Verfahren ist im Allgemeinen unter dem Namen **Polynomdivision** bekannt.
Sei $f = 3X^5 + 2X^4 + 2$ und $g = X + 2$ zwei Polynome aus $\mathbb{Q}[X]$. Wir würden gerne f durch g teilen, also ein q finden, so daß $f = qg$. Konzentrieren wir uns auf den höchsten Koeffizienten so muss q ja auf alle Fälle mit $q_1 = 3X^4$ beginnen. Dies passt natürlich nicht genau, allerdings ist

$$f = q_1 g + \ldots,$$

wobei $\cdots = f - q_1 g$ ist, also: $f_1 = -4X^4 + 2$. Auch dieses können wir wieder fast durch g teilen: $-4X^4 + 2 = q_2 g + f_2$, wobei $q_2 = -4X^3$ und $f_2 = -4X^4 + 2 - 4X^4 - 8X^3$. Wir haben also schon

$$f = q_1 g + f_1 = q_1 g + q_2 g + f_2 = (q_1 + q_2)g + f_2 \ .$$

Führen wir dieses Verfahren weiter, bis irgendwann (nach höchstens Grad von f Schritten) $\deg f_i < \deg g$ ist. An der Stelle können wir aufhören, da $f_i = r$ und $q = q_1 + q_2 + \cdots + q_i$ die gesuchten Polynome sind. Man beachte, daß wir nur die Körperaxiome und keine sonstigen Eigenschaften von $\mathbb{Q}$ verwendet haben. □

Dieses Verfahren funktioniert übrigens auch in nullteilerfreien Ringen, falls

$$g = b_0 + b_1 X + \cdots + b_{m-1} X^{m-1} + X^m \ ,$$

also $b_m = 1$ ist. Solche Polynome heißen **normiert**.

Definition 2.0.9. Eine Nullstelle eines Polynoms $f \in R[X]$ ist ein Element $\lambda \in R$, so daß $f(\lambda) = 0$.

Lemma 2.0.10. *Ist R nullteilerfreier Ring und ist $\lambda \in R$ eine Nullstelle von $f \in R[X]$, so gibt es ein eindeutig bestimmtes $g \in R[X]$ mit*

1. $f = (X - \lambda)g$
2. $\deg g = \deg f - 1$

Beweis. Vorlesung. □

Korollar 2.0.11. *Ist R ein nullteilerfreier Ring und $f \in R[X]$ ein Polynom vom Grad k, so besitzt f höchstens k Nullstellen.*

Theorem 2.0.12 (Fundamentalsatz der Algebra)**.** *Jedes Polynom $f \in \mathbb{C}[X]$ mit $\deg(f) > 0$ besitzt eine Nullstelle.*

Beweis. Es gibt eine Vielzahl von Beweisen, alle benötigen jedoch Hilfsmittel aus der Analysis und sind recht anspruchsvoll. □

Korollar 2.0.13. *Jedes Polynom $f \in \mathbb{C}[X]$ zerfällt in Linearfaktoren. D.h. es gibt $a, \lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit $n = \deg f$ und*

$$f = a(X - \lambda_1) \cdot \cdots \cdot (X - \lambda_n)$$

Lemma 2.0.14. *Ist $\lambda \in \mathbb{C}$ eine Nullstelle von $f \in \mathbb{R}[X] \subseteq \mathbb{C}[X]$, so ist auch $\overline{\lambda}$ Nullstelle von f.*

Korollar 2.0.15. *Jedes Polynom $f \in \mathbb{C}[X]$ mit ungeradem Grad besitzt eine reelle Nullstelle.*

3 Lineare Gleichungssysteme

Im folgenden suchen wir nicht nur Lösungen, die eine Gleichung erfüllen, sondern solche, die gleich mehrere gleichzeitig erfüllen.
Sofern nicht anders vermerkt, ist K ein beliebiger Körper und R ein beliebiger Ring.

3.1 Einführung

Zunächst, aber wie üblich ein paar Definitionen.

Definition 3.1.1.

- Eine **lineare Gleichung** in n Unbekannten über einem Ring R ist eine Gleichung der Form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \ ,$$

  wobei $a_1, \ldots, a_n, b \in R$.

- Ist $b = 0$, so sprechen wir von einer **homogenen linearen Gleichung**.
- Ein (homogenes) lineares Gleichungssystem besteht aus endlich vielen (homogenen) linearen Gleichungen.
- Eine Lösung eines Gleichungssystems ist eine Belegung der Variablen, so daß alle Gleichungen des Gleichungssystems erfüllt sind.

Allgemein können wir ein lineares Gleichungssystem schreiben als

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

Ein solches Gleichungssystem ist vollständig durch die **Koeffizienten** $a_{ij} \in R$ und $b_i \in R$ bestimmt. Formal kurz können wir auch

$$\sum_{i=1}^{n} a_{ji} x_i = b_j$$

schreiben.

Beim Lösen von Gleichungssystemen können die folgenden Situationen auftreten:

- Das Gleichungssystem besitzt genau eine Lösung
- Das Gleichungssystem besitzt mehrere Lösungen
- Gleichungssystem besitzt keine Lösung

Beispiel 3.1.2. *Beispiele in der Vorlesung. Welche der Situationen der Fall ist hängt auch von dem zugrunde liegenden Ring ab.*

Wie schon aus der Schule bekannt können wir Lösungen suchen, indem wir eine Gleichung mit Konstanten multiplizieren und eine Gleichung zu einer weiteren addieren. Gehen wir einigermaßen systematisch vor erhalten wir so Gleichungen, von der Form $x_j = r$. Das gleiche Verfahren funktioniert auch in Körpern!

Beispiel 3.1.3. *Betrachten wir über $\mathbb{Z}_7$ das Gleichungssystem*

$$\begin{aligned} 3x + 4y &= 2 \\ 2x + 4y &= 3 \end{aligned}$$

so erhalten wir als einzige Lösung $x = 6, y = 3$

Bei wenigen Variablen ist es üblich x, y, z zu verwenden, bei mehreren ist $x_1, \ldots, x_n$ besser (da uns sonst die Buchstaben ausgehen würden).

Beispiel 3.1.4. *Betrachten wir über $\mathbb{R}$ das Gleichungssystem*

$$\begin{aligned} 2x + 4y + 6z &= 2 \\ x + 4y + z &= 1 \\ x + 2y + 3z &= 1 \end{aligned}$$

3.2 Matrizenschreibweise, das Verfahren von Gauss

Kommen wir nochmal zurück zu einem allgemeinen linearen Gleichungssystem

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

Hiermit zu arbeiten ist recht umständlich. Kompakter ist es die Koeffizienten $a_{i,j}$ als $m \times n$ **Matrix** A aufzuschreiben, also in rechteckiger Anordnung der Form

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

Kurz schreiben wir auch einfach

$$A = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$$

oder sogar einfach nur (a_{ij}). Die Mengen aller $m \times n$ Matrizen mit Koeffizienten in einem Ring R bezeichnen wir mit $M(m \times n, K)$. Ebenso schreiben wir die Koeffizienten b_i als **Spaltenvektor** der Form

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

oder alles gleich zusammen als **erweiterte Koeffizientenmatrix** des Gleichungssystem als

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right),$$

oder kurz $(A \mid b)$. Der Strich hat hierbei keinerlei mathematische Bedeutung ist nur eine Gedächtnisstütze. Man beachte, daß wir diese Matrix, wie das ursprüngliche Gleichungssystem manipulieren können d.h. wir können

- je zwei Zeilen/Gleichungen vertauschen
- eine Zeile/Gleichung mit einem Element multiplizieren
- eine Zeile/Gleichung zu einer anderen hinzuaddieren

Diese Manipulationen heißen **elementare Zeilenumformungen**. Die wichtigste Einsicht ist hierbei

Satz 3.2.1. *Sei R ein nullteilerfreier Ring. Sei $(A'|b')$ das Gleichungssystem, das durch Anwendung einer dieser obigen drei Manipulationen aus $(A|b)$ hervorgehende Gleichungssystem über R. Dann haben beide genau die gleichen Lösungen.*

Beweis. Vorlesung. □

All diese Operationen funktionieren natürlich nur auf den Zeilen und nicht mit Spalten! Man beachte auch, daß man die folgende Operation durch Hintereinanderausführung von elementaren Zeilenumformungen ausführen kann.

- Das Vielfache einer Zeile/Gleichung zu einer anderen hinzuaddieren

Als nächstes wollen wir festhalten, daß für einige Gleichungssystem sich in einem Körper sehr einfach alle Lösungen bestimmen lassen. Die erweiterten Koeffizientenmatrizen dieser einfach zu lösenden Gleichungssytem haben genau die folgende Form:

Siehe Vorlesung

Definition 3.2.2. Eine solche Matrix ist in **Zeilenstufenform** d.h. formal

1. Die letzten (möglicherweise keine) Zeilen bestehen nur aus dem Nullelement.
2. Ist j_k so daß links von dem j_k-ten Element in der k-ten Zeile nur Nullen stehen ($a_{ki} = 0$ für $i < j_k$), dieses Element aber nicht null ist ($a_{kj_k} \neq 0$), so stehen unterhalb von a_{kj_k} nur Nullen ($a_{ij_k} = 0$ für $i > k$).

Beispiel 3.2.3. *Die folgenden Matrizen sind in Zeilenstufenform*

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

Die folgenden Matrizen sind nicht *in Zeilenstufenform*

$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 4 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 2 & 2 & -5 \\ 1 & 2 & 3 & -1 \end{pmatrix}.$$

Ist ein Gleichungssystem über einem Körper K[1] durch eine erweiterte Matrix $(A \mid b)$ gegeben, die in Zeilenstufenform ist, so lassen sich die Lösungen durch folgendes Verfahren finden:

1. Gibt es Zeilen in denen nur ein Eintrag (und da die Matrix in ZSF ist der letzte) Eintrag nicht-null ist, so existiert keine Lösung:
Eine solche Zeile entspricht einer Gleichung der Form

$$0x_1 + 0x_2 + \cdots + 0x_n = b_k \neq 0 \ .$$

Offensichtlich gibt es keine Elemente, die diese Gleichung und damit das Gleichungssystem erfüllen, da die linke Seite der Gleichung immer null ist.

2. Nehmen wir nun zur Einfachheit halber an, daß die Matrix die folgende Form hat:

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ 0 & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ & 0 & \vdots & & & \vdots \\ & \vdots & a_{rr} & \dots & a_{rn} & b_r \\ & 0 & & \dots & 0 & 0 \\ & & & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \end{array}\right) ,$$

Das heißt, die letzten $n-r$ Zeilen bestehen nur aus Nullen, und $a_{ii} \neq 0$ für alle $1 \leqslant i \leqslant r$.
Wir nennen $x_1, \dots, x_r$ **gebundene Variablen**, und $x_{r+1}, \dots, x_n$ **freie Variablen**.[2] Sei $k = n-r$ die Anzahl der freien Variablen. Wir können nun $\lambda_1, \dots, \lambda_k$ frei wählen (daher der Name freie Variable) und setzen $x_{r+1} = \lambda_1, \dots, x_n = \lambda_r$.
Damit ist die Belegung der ersten r Variablen eindeutig bestimmt: aus der r-ten Gleichung

$$a_{rr}x_r + a_{r(r+1)}\lambda_{r+1} + \cdots + a_{rn}\lambda_n = b_r$$

folgt

$$x_r = a_{rr}^{-1}\left(b_r - a_{r(r+1)}\lambda_{r+1} - \cdots - a_{rn}\lambda_n\right) \ .$$

Man beachte, daß wir hierfür die Körperaxiome benötigen, um a_{rr}^{-1} zu bilden. Mit diesem Wert können wir nun x_{r-1} bestimmen indem

[1] Zur Umformung benötigen wir im Allgemeinen die Körpereigenschaften. Von hier an wollen wir also annehmen, daß wir über einem Körper K arbeiten, auch wenn in einigen Fällen es auch in Ringen klappen würde.

[2] Man beachte, daß man diese Rollen teilweise auch vertauschen kann; allerdings werden wir später sehen, daß die Anzahl der gebundenen und damit auch der freien Variablen für ein Gleichungssytem immer konstant sind.

wir in die $(r-1)$-te Gleichung einsetzen. Diesen Schritt können wir wiederholen, bis wir schließlich einen Wert für x_1 gefunden haben.

3. Der allgemeine Fall funktioniert analog, allerdings mit den freien unter die gebundenen Variablen "gemischt".

Beispiel 3.2.4. *Beispiel in der Vorlesung.*

Als letztes wollen wir noch Zeigen, daß wir eine Matrix über einem Körper *immer* mit elementaren Zeilenumformungen in Zeilenstufenform überführen können. Das erste Element einer Zeile, daß nicht null ist wollen wir als **Pivotelement** bezeichnen.

1. Wir wollen die Zeilen durch Vertauschen so anordnen, daß Zeilen die nur aus Nullen bestehen unten stehen. Ausserdem sollen die Zeilen, die ein Element enthalten, welches nicht null ist so geordnet sein, daß, je weiter das Pivotelement links ist, desto weiter oben soll diese Zeile stehen. Es gibt also schon Stufen, allerdings sind einige dieser Stufen noch zu steil.

2. Sagen wir, das Pivotelement der 1. Zeile steht an der Stelle k. D.h. $a_{1k} \neq 0$ ist dieses Pivotelement. Zu jeder Zeile ℓ, die ein Element besitzt, für das $a_{\ell k} \neq 0$ ist wollen wir das $a_{\ell k}a_{1k}^{-1}$-fache der ersten Zeile addieren. Dadurch entsteht an dieser k-ten Stelle in der ℓ-ten Zeile eine 0.
So können wir erreichen, daß unterhalb a_{1k} nur Nullen stehen.

3. Wir wiederholen Schritt 1 und 2, mit der Matrix, in der wir die jeweils neue 1. Zeile ignorieren, bis es keine oder nur eine Zeile gibt, welche nicht aus nur aus Nullen besteht.

Schreibweisen 3.2.5. Manipulieren wir eine Matrix A durch elementare Zeilenumformungen in eine Matrix A' so wollen wir schreiben

$$A \rightsquigarrow A' \ .$$

Welche Zeilenumformung verwendet wird können wir entweder aus dem Zusammenhang erkennen oder explizit angeben (z.B. indem wir Details über oder unter den Pfeil schreiben).
Ausser in trivialen Fällen ist $A = A'$ eine inkorrekte Schreibweise.[3]

[3] Um ganz ehrlich zu sein: Man kann zeigen, daß $\rightsquigarrow$ eine Äquivalenzrelation ist. Legt man nun noch fest, daß man Repräsentanten, wenn aus dem Zusammenhang klar, als Äquivalenzklasse auffasst, ist es sogar richtig $=$ zu verwenden. Sollten Sie diese Schreibweise bevorzugen, liefern Sie bitte in der Klausur bzw. in Übungsaufgaben den Beweis, daß $\rightsquigarrow$ Äquivalenzrelation ist.

Beispiel 3.2.6. *In $\mathbb{Q}$ können wir die folgende Matrix A in die Matrix A' umwandeln, welche in Zeilenstufenform ist.*

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 & 9 \\ 0 & 3 & 4 & 5 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & -1 & -2 & -9 \\ 0 & 0 & -3 & -6 & -18 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A'$$

Das Verfahren zum Lösen eines lineares Gleichungssystems, indem man zuerst die Matrix in Zeilenstufenform bringt und anschliessend die Lösungen von unten nach oben durch einsetzen erhält, heißt **Eliminationsverfahren von Gauss**.

Bemerkung 3.2.7. *Manche Bücher/Mathematiker empfehlen auch eine Matrix in Zeilenstufenform weiter in die sogenannte erweiterte Zeilenstufenform umzuwandeln. Hierbei wird sichergestellt, daß auch oberhalb der Pivotelemente Nullen stehen. Das Auffinden einer Lösung ist dann einfacher. Die Arbeit die man so spart hat man allerdings damit einfach schon durch weitere Zeilenumformungen geleistet.*

3.3 Rechnen mit Matrizen

Auch aus dem Zusammenhang der linearen Gleichungssystem gerissen sind Matrizen als mathematische Objekte nicht uninteressant.

Zunächst zwei einfache Verknüpfungen zwischen Matrizen, bzw. zwischen einem Körperelement und einer Matrix.

Definition 3.3.1. Sind $A = (a_{ij}), B = (b_{ij}) \in M(m \times n, K)$ Matrizen über K, so definieren wir die Summe $A + B$ als die $m \times n$-Matrix mit dem Eintrag $(a_{ij} + b_{ij})$ in der i-ten Zeile und j-ten Spalte.
Ist $k \in K$ ein Element so definieren wir des weiteren $k \cdot A$ bzw. kA als die $m \times n$-Matrix mit dem Eintrag ka_{ij} in der i-ten Zeile und j-ten Spalte. Wir sprechen in diesem Fall auch vom **Skalar** k.
Man beachte, daß bei dieser sogenannten **Skalarmultiplikation** der Skalar immer links und die Matrix immer rechts steht.

Definition 3.3.2. Für jeden Körper K sei $0_{m \times n}$ bzw. oft einfach nur 0 (wenn n und m aus dem Zusammenhang klar hervorgehen) die

Nullmatrix; die Matrix, welche nur Nullen als Einträge hat.

Wir stellen fest, daß für diese Verknüpfungen folgende algebraischen Rechenregeln gelten. Hierbei schreiben wir $-A$ für $(-1)A$ und $A - B$ für $A + (-B)$.

Satz 3.3.3. *Sind $A, B \in M(m \times n, K)$ und $a, b \in K$ so gilt:*

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $A - A = 0$
5. $0A = 0$
6. $a(A + B) = aA + aB$
7. $(a + b)A = aA + bA$
8. $(ab)A = a(bA)$
9. *Ist $cA = 0$, so ist entweder $c = 0$ oder $A = 0$.*

Beweis. Vorlesung. Einfach nachzurechnen. □

Bemerkung 3.3.4. *Aussagen 1-4 sagen, daß für alle n, m,*

$$(M(m \times n, K), +)$$

eine Gruppe[4] *ist.*

Als nächstes wollen wir noch erklären, wie wir zwei Matrizen multiplizieren. Die Definition ist auf den ersten Blick vielleicht etwas esoterisch, macht aber durchaus Sinn, wie wir später sehen werden.

Definition 3.3.5. Sei A eine $m \times n$ Matrix und B eine $n \times k$ Matrix über einem Körper K. So ist das Produkt AB die $m \times k$ Matrix, deren Eintrag in der i-ten Zeile und j-ten Spalte gleich

$$\sum_{\ell=1}^{n} a_{i\ell} b_{\ell j}$$

ist.

[4] Wie schon erwähnt wurden Gruppen in der Vorlesung "Diskrete Mathematiker für Informatiker" eingeführt. Nachdem Sie jetzt aber auch Ringe/Körper kennen sehen wir, daß Gruppen vereinfachte Ringe/Körper sind, wenn man die Multiplikation ignoriert und nur eine Verknüpfung hat.

Man beachte, daß sich nicht zwei beliebige Matrizen multiplizieren lassen, sondern nur zwei von denen die erste die gleiche Anzahl an Spalten wie die zweite an Zeilen hat.

Beispiel 3.3.6. *Vorlesung.*

Satz 3.3.7. *Sind A, B, C Matrizen über K und $a, b \in K$ so gilt, (falls definiert!):*

1. $(AB)C = A(BC)$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $A(aB) = a(AB)$[5]
5. $A0 = 0$ *und* $0A = 0$ *(für die Nullmatrix)*
6. $(aA)B = a(AB)$

Beweis. Vorlesung. Einfach nachzurechnen. □

Allerdings gilt:

Satz 3.3.8.

1. *Im Allgemeinen gilt für zwei Matrizen A, B, C über K* nicht*, daß*

$$AB = BA \ ,$$

   selbst wenn beide Produkte definiert sind.

2. *Es gilt auch nicht daß aus $AB = 0$ folgt, daß $A = 0$ oder $B = 0$ (natürlich auch hier selbst wenn die Produkte definiert sind).*

3. *Bei Matrizen darf man auch nicht kürzen. Genauer: es gibt Matrizen A, B, C, die alle nicht null (d.h. nicht die Nullmatrix) sind und für die gilt:*

$$AB = AC \quad \textit{aber} \quad B \neq C \ .$$

Beweis. Es reicht Beispiele anzugeben. Sei $K = \mathbb{Q}$. Im ersten Fall wählen wir

$$A = \begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}.$$

[5] Man beachte, daß auch hier bei der Skalarmultiplikation auf der linken Seite der Skalar links und die Matrix rechts steht. Ausserdem ist zu beachten, daß dieser Satz nur gilt wenn $a \in K$ und nicht für Matrizen.

Dann ist

$$AB = \begin{pmatrix} -1 & -2 \\ 11 & 4 \end{pmatrix} \quad \text{und} \quad BA = \begin{pmatrix} 3 & 6 \\ -3 & 0 \end{pmatrix} ,$$

also $AB \neq BA$.
Im zweiten Fall wählen wir

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 3 & 7 \\ 0 & 0 \end{pmatrix} .$$

Wie man leicht nachrechnet ist $AB = 0$.

Ein Beispiel für den dritten Fall ist:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} \quad \text{und} \quad C = \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix} .$$

□

Auch bei der Matrixmultiplikation gibt es ein Einselement:

Definition 3.3.9. Die $n \times n$ Matrix, die überall Nullen als Einträge hat, bis auf der "Diagonalen" wo Einsen stehen heißt die **Einheitsmatrix**. Formal ist die Einheitsmatrix also $I_n = (a_{ij}) \in M(n \times n, K)$ mit

$$a_{ij} = \begin{cases} 0 & \text{falls } i \neq j \\ 1 & \text{falls } i = j \end{cases} .$$

Matrizen die die gleiche Anzahl and Zeilen wie Spalten haben bezeichnen wir übrigens als **quadratisch**. Die ersten Einheitsmatrizen sind:

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Ist die Größe durch den Zusammenhang klar wollen wir vereinfacht auch I an Stelle von I_n schreiben.

Satz 3.3.10. *Ist A eine $m \times n$ Matrix, so gilt*

$$AI_n = A = I_m A .$$

Korollar 3.3.11. *$M(n \times n, K)$ ist mit obiger Addition und Multiplikation zweier Matrizen ein Ring, der aber nicht kommutativ und auch nicht nullteilerfrei ist.*

Diese Aussage übersieht zwar etwas Struktur, z.B. das wir auch mit Elementen des Körpers multiplizieren können und Matrizen verschiedener Dimensionen multiplizieren können.

3.4 Nochmals lineare Gleichungssyteme in Matrizenschreibweise

Mit oben eingeführter Multiplikation können wir lineare Gleichungssyteme noch besser aufschreiben. Sei hierzu

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

$n \times 1$ Matrizen, und A wie oben. Unser lineares Gleichungssystem ist dann genau

$$Ax = b \ .$$

Ein homogenes lineares Gleichungssystem können wir nun schreiben als

$$Ax = 0 \ .$$

So kompakt diese Schreibweise ist, sollten wir uns immer daran erinnern, daß A nicht eine einzelne Zahl und x nicht eine einzelne Variable ist. Eine Lösung zu finden, die dieses Gleichungssystem löst ist mit dem Gauss'schen Algorithmus ohne Probleme machbar, benötigt aber im allgemeinen viele Schritte.

Satz 3.4.1. *Sind x und y Lösungen des homogenen linearen Gleichungssytems und $k \in K$. So sind auch $x + y$ und kx Lösungen.*

Satz 3.4.2. *Ist z eine beliebige aber fest gewählte Lösung des linearen Gleichungssytems $Ax = b$ und sei y eine Lösung des homogenen linearen Gleichungssystem $Ax = 0$, so ist auch $z + y$ eine Lösung des (inhomogenen) linearen Gleichungssystems.*
Umgekehrt kann jede Lösung z' des (inhomogenen) linearen Gleichungssystems $Ax = b$ geschrieben werden als $z' = z + x$

Anders ausgedrückt sagt dieser Satz, daß wir, um ein inhomogenes Gleichungssystem zu lösen es ausreicht *eine einzige* beliebige Lösung zu $Ax = b$ zu finden, und dann alle Lösungen des homogenen linearen Gleichungssystems $Ax = 0$ zu finden, was im Allgemeinen einfacher ist.

4 Vektorräume

4.1 Definitionen

Sei, wie immer, K ein beliebiger Körper.
Wie wir im letzten Kapitel gesehen haben sind sowohl die Menge aller Lösungen als auch die Menge aller $m \times n$ Matrizen eine interessante Struktur. Was haben beide gemeinsam?

- Wir können Elemente addieren (d.h. wir haben eine Gruppenstruktur).
- Wir können Elemente mit Elementen aus dem Körper multiplizieren.
- Im Allgemeinen können wir nicht zwei Elemente multiplizieren. (Bei Matrizen geht es gut, aber was sollte das Produkt zweier Lösungen aus n-Elementen sein?)

Diese Art von Struktur ist Hauptinhalt der linearen Algebra. Es handelt sich um die sogenannten Vektorräume.

Definition 4.1.1. Ein K-**Vektorraum** besteht aus einer Menge V und zwei Verknüpfungen $+$ und $\cdot$. Hierbei ist $+ : V \times V \to V$ und $\cdot : K \times V \to V$. Die erste wollen wir Addition nennen, die zweite Skalarmultiplikation.
Außerdem sollen die folgenden Axiome gelten: $(V, +)$ ist eine kommutative Gruppe. D.h.

(V1) Für alle $u, v, w \in V$ gilt $(u + v) + w = u + (v + w)$.

(V2) Es gibt ein Element $0 \in V$, so daß für alle $u \in V$ gilt $u + 0 = 0 + u = u$. Dieses Element heißt der **Nullvektor**.

(V3) Für alle $u \in V$ gibt es ein Element $-u$, so daß gilt $u + (-u) = 0$.

(V4) Für alle $u, v \in V$ gilt $u + v = v + u$.

Außerdem verträgt sich die Skalarmultiplikation wie folgt:

(V5) Für alle $u \in V$ und $k, h \in K$ gilt $(k+h)u = ku + hu$.

(V6) Für alle $u \in V$ und $k, h \in K$ gilt $(kh)u = k(hu)$

(V7) Für alle $u \in V$ gilt $1 \cdot u = u$.

(V8) Für alle $u, v \in V$ und $k \in K$ gilt $k(u+v) = ku + kv$.

Elemente eines Vektorraumes bezeichnen wir als **Vektoren**.

Zwei wichtige Spezialfälle:

Definition 4.1.2. Vektorräume über $\mathbb{R}$ werden als **reelle Vektorräume** bezeichnet; solche über $\mathbb{C}$ als **komplexe Vektorräume**.

Satz 4.1.3. *Das Nullelement* 0 *und zu jedem* u *ist der negative Vektor* $-u$ *eindeutig bestimmt.*

Beweis. Auch hier könnte man auf den entsprechenden Satz aus der Gruppentheorie verweisen. Da aber ein Beweis recht kurz ist:
Nehmen wir an es gäbe noch ein weiteres Element $0'$, so daß für alle $u \in V$ gilt $0' + u = u + 0' = u$. Dann gilt ja insbesondere für $u = 0$:

$$0 = 0 + 0' = 0' \ .$$

Sei jetzt v ein Element, das sich wie $-u$ verhält; d.h.

$$u + v = v + u = 0 \ .$$

Dann gilt:

$$-u = 0 + (-u) = (v+u) - u = v + (u + (-u)) = v + 0 = v \ . \quad \square$$

Schreibweisen 4.1.4. Wie immer wollen wir $u - v$ an Stelle von $u + (-v)$ und av an Stelle von $a \cdot v$ schreiben.

4.2 Beispiele von Vektorräumen

1. Jeder Körper K ist ein K-Vektorraum mit der normalen Addition und der Körper-Multiplikation als Skalarmultiplikation.
   Etwas tiefgründiger: ist K' ein Unterkörper von K, so ist K ein K'-Vektorraum. Z.B. ist $\mathbb{R}$ ein $\mathbb{Q}$-Vektorraum, und $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum.

2. Wie wir gesehen haben ist $M(m \times n, K)$ ein K-Vektorraum. (Das man obendrein im Falle $m = n$ Matrizen multiplizieren kann ist für die Vektorraumstruktur unerheblich).

3. Das Standardbeispiel eines K Vektorraumes ist K^n die Menge aller Tupel der Länge n. Zwei solche Tupel $(k_1, \ldots, k_n)$ und

$(k'_1, \dots, k'_n)$ addieren wir einfach durch

$$(k_1, \dots, k_n) + (k'_1, \dots, k'_n) = (k_1 + k'_1, \dots, k_n + k'_n) \ .$$

Ein Tupel multiplizieren wir mit einem Skalar durch

$$k(k_1, \dots, k_n) = (kk_1, \dots, kk_n) \ .$$

4. Ähnlich wie oben können wir uns auch den K Vektorraum K^∞ aller unendlichen Tupel betrachten. Zwei solche Tupel $(k_1, k_2, \dots)$ und $(k'_1, k'_2, \dots)$ addieren wir einfach durch

$$(k_1, k_2, \dots) + (k'_1, k'_2, \dots) = (k_1 + k'_1, k_2 + k'_2, \dots) \ .$$

Ein Tupel multiplizieren wir mit einem Skalar durch

$$k(k_1, k_2, \dots) = (kk_1, kk_2, \dots) \ .$$

5. Genau genommen sind die letzten zwei Beispiele nur Spezialfälle des folgenden Vektorraumes: Sei X eine beliebige nicht-leere und K^X die Menge aller Funktionen $X \to K$. Zwei solche Funktionen $f, g \in K^X$ addieren wir durch

$$(f + g)(x) = f(x) + g(x) \ ,$$

und multiplizieren mit einem Skalar durch

$$(kf)(x) = k(f(x)) \ .$$

Man sollte begriffen haben, was die letzten beiden Gleichungen überhaupt aussagen! K^n ist dann nichts anderes als $K^{\{1,\dots,n\}}$ und K^∞ ist $K^{\mathbb{N}}$.

6. Ist $(A \mid 0)$ ein homogenes lineares Gleichungssystem über einem Körper K, so ist die Menge aller Lösungen ein Vektorraum.

Bemerkung 4.2.1. *Oft fassen wir K^n auch als den Vektorraum aller $1 \times n$ Matrizen (als Zeilenvektoren), oder aller $n \times 1$ Matrizen (als Spaltenvektoren) auf.*[1]

Bemerkung 4.2.2. *Die Vektorräume $\mathbb{R}^2$ und $\mathbb{R}^3$ eignen sich hervorragend zur mathematischen Modellierung des uns umgebenden Raumes.*

4.3 Linearkombinationen, Lineare Unabhängigkeit, Erzeugendensysteme

Leicht zu übersehen ist die nächste zentrale Definition:

[1] Einige Bücher treffen hier eine feste Wahl zwischen beiden Alternativen.

Definition 4.3.1. Sind $v_1, \ldots, v_n$ Vektoren eines K-Vektorraumes V und $k_1, \ldots, k_n \in K$ Skalare, so bezeichnen wir mit

$$\sum_{i=1}^{n} k_i v_i = k_1 v_1 + \cdots + k_n v_n$$

eine **Linearkombination** der Vektoren $v_1, \ldots, v_n$.
Ist $S \subseteq V$ eine Menge von Vektoren, so heißt v Linearkombination von S, wenn es *endlich* viele Vektoren in S gibt, so daß v eine Linearkombination dieser ist.

Man beachte, daß wir oft auch den Nullvektor als Linearkombination von Vektoren schreiben können, die nicht gleich dem Nullvektor sind und bei denen die Koeffizienten der Linearkombination ebenfalls nicht null sind. Z.B. ist in $M(2 \times 2, \mathbb{Q})$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Besonders interessiert sind wir an Mengen von Vektoren, bei denen wir den Nullvektor nicht als (nicht-triviale) Linearkombination schreiben können.

Definition 4.3.2. Vektoren $v_1, \ldots, v_n$ eines K-Vektorraumes V heißen **linear unabhängig** genau dann, wenn immer für Koeffizienten $k_1, \ldots, k_n \in K$ aus

$$k_1 v_1 + \cdots + k_n v_n = 0$$

folgt, daß

$$k_1 = 0, k_2 = 0, \ldots, k_n = 0 \ .$$

Anders ausgedrückt sind $v_1, \ldots, v_n$ linear unabhängig, wenn sich der Nullvektor nur auf triviale Weise als Linearkombination bilden lässt.

Definition 4.3.3. Vektoren $v_1, \ldots, v_n$ eines K-Vektorraumes V heißen **linear abhängig** genau dann, wenn sie nicht linear unabhängig sind. D.h., wenn es Skalar $k_1, \ldots, k_n \in K$ gibt, die nicht alle null sind[2] so daß

$$k_1 v_1 + \cdots + k_n v_n = 0 \ .$$

Definition 4.3.4. Eine Menge S von Vektoren heißt linear unabhängig, wenn jede endliche Auswahl an verschiedenen Vektoren linear unabhängig sind. Sie heißt linear abhängig, falls es endlich viele Vektoren aus S gibt, die linear abhängig sind.[3]

[2] Man beachte, daß einige null sein können. Bloss nicht alle gleichzeitig.
[3] Diese Definition ist vor Allem für unendliche Mengen S notwendig.

Lemma 4.3.5. *Eine Menge S mit zwei oder mehr Vektoren ist genau dann linear abhängig, falls sich einer*[4] *der Vektoren als Linearkombination der anderen darstellen lässt.*

Diese Charakterisierung ist zwar praktisch oft einfacher zu handhaben als unsere Definition, welche allerdings theoretisch dieser weit überlegen ist.

Einige einfache hinreichende Kriterien für lineare (Un)-abhängigkeit sind:

Bemerkung 4.3.6. *In jedem K-Vektorraum V gilt:*

1. *Ein einziger Vektor v ist genau dann linear unabhängig, wenn $v \neq 0$.*
2. *Enthält eine endliche Menge $\{v_1, \ldots, v_n\}$ von Vektoren den Nullvektor, so sind die Vektoren $v_1, \ldots, v_n$ linear abhängig.*
3. *Kommt der gleiche Vektor zweimal vor, so ist die Menge ebenfalls linear abhängig.*

Eine wichtig Eigenschaft von linear unabhängigen Vektoren ist der **Koeffizientenvergleich**:

Satz 4.3.7. *Seien $v_1, \ldots, v_n$ linear unabhängige Vektoren. Sind dann $k_1, \ldots, k_n$, $h_1, \ldots, h_n$ Skalare, so daß*

$$k_1 v_1 + \cdots + k_n v_n = h_1 v_1 + \cdots + h_n v_n \ ,$$

so ist $k_i = h_i$ für alle $1 \leqslant i \leqslant n$.

Definition 4.3.8. Ist S eine Menge von Vektoren, so sei $\langle S \rangle$ die Menge aller Linearkombinationen von S. Im Falle daß $S = \{v_1, \ldots, v_n\}$ endlich ist, schreiben wir $\langle v_1, \ldots, v_n \rangle$ an Stelle von $\langle \{v_1, \ldots, v_n\} \rangle$.

Lemma 4.3.9. *Sind B und B' Teilmengen eines Vektorraums V, so gilt*

$$\langle B' \rangle \subseteq \langle B \rangle \quad \Longleftrightarrow \quad B' \subseteq \langle B \rangle$$

Beweis. Für die Richtung $\Longrightarrow$ reicht es zu zeigen, daß $B' \subseteq \langle B' \rangle$. Dies ist aber klar, da jedes $v \in B'$ Linearkombination von Vektoren aus B ist (nämlich $v = 1 \cdot v$).
Für die Richtung $\Longleftarrow$ müssen wir etwas mehr arbeiten. Sei $v \in \langle B' \rangle$ beliebig. Wir müssen zeigen, daß v Linearkombination wen endlich vielen Vektoren aus B ist. Da $v \in \langle B' \rangle$ gibt es $v_1, \ldots v_n \in B'$ und $k_1, \ldots, k_n \in K$ so daß

$$v = \sum_{i=1}^{n} k_i v_i \ . \tag{4.1}$$

[4] Man beachte, daß dies nicht bedeutet, daß sich jeder der Vektoren als Linearkombination der anderen schreiben lässt.

Nach Vorraussetzung sind aber auch die v_i Linearkombinationvon Vektoren aus B. Es gibt also für jedes $1 \leqslant i \leqslant n$ $w_1^i, \ldots, w_{n_i}^i \in B$ und $\ell_1^i, \ldots, \ell_{n_i}^i$ so daß

$$v_i = \sum_{j=1}^{n_i} \ell_j^i w_j^i \ . \tag{4.2}$$

Setzen wir nun die Gleichungen 4.2 in Gleichung 4.1 ein, so erhalten wir

$$v = \sum_{i=1}^{n} k_i \left(\sum_{j=1}^{n_i} \ell_j^i w_j^i \right) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} k_i \ell_j^i w_j^i \ .$$

D.h. daß v Linearkombination von Vektoren aus B ist, was wir ja zeigen wollten. □

Allgemein gilt:

Satz 4.3.10. *Ist S eine Menge von Vektoren eines K-Vektorraumes V, so ist $\langle S \rangle$ ein K-Vektorraum.*

Beweis. Der Beweis hierzu folgt leicht aus einem allgemeinen Satz. □

Definition 4.3.11. Ist $V = \langle S \rangle$, so heißt S **Erzeugendensystem**.
Äusserst interessant sind minimale Erzeugendensysteme:
Eine Menge $B \subseteq V$ eines Vektorraumes heißt **Basis**, falls gilt

1. B ist Erzeugendensystem von V und
2. B ist linear unabhängig.

Eine Menge $B \subseteq V$ ist also Basis, wenn sich jeder Vektor aus V aus den Vektoren in B zusammensetzen lässt, und ausserdem wir keine Vektoren aus B weglassen können.

Beispiel 4.3.12. *Die **Standardbasis** von K^n sind die **Einheitsvektoren** $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0 \ldots, 0)$, $\ldots$, $e_n = (0, 0, \ldots, 1)$.*

Im Allgemeinen ist es schwer zu entscheiden, ob eine Menge von Vektoren Basis eines Vektorraums ist.

4.4 Lineare Unabhängigkeit und Erzeugendensysteme im K^n

Seien $v_1 = (a_{1,1}, \ldots, a_{m,1}), \ldots, v_n = (a_{1,n}, \ldots, a_{m,n})$ Vektoren im K^m. Fassen wir diese als Spaltenvektoren auf, so können wir sie in einer Matrix A zusammenfassen. Sind jetzt $k_1, \ldots, k_n \in K$ Skalare, so lässt sich die Linearkombination folgendermassen umschreiben:

$$
\begin{aligned}
&k_1 v_1, + \cdots + k_n v_n \\
&\qquad = k_1 \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + \cdots + k_n \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix} \\
&\qquad = \begin{pmatrix} k_1 a_{1,1} \\ \vdots \\ k_1 a_{m,1} \end{pmatrix} + \cdots + \begin{pmatrix} k_n a_{1,n} \\ \vdots \\ k_n a_{m,n} \end{pmatrix} \\
&\qquad = \begin{pmatrix} k_1 a_{1,1} + \cdots + k_n a_{1,n} \\ \vdots \\ k_1 a_{m,1} + \cdots + k_n a_{m,n} \end{pmatrix} \\
&= \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = A \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}
\end{aligned}
$$

D.h. wir können eine Linearkombination mit etwas Umformung als Matrixmultiplikation schreiben.

Satz 4.4.1. *Seien $v_1, \dots, v_n$ und A wie oben. Dann sind $v_1 \dots, v_n$ genau dann linear unabhängig, wenn das homogene lineare Gleichungssystem $Ax = 0$ nur die triviale Lösung hat.*

Satz 4.4.2. *Seien $v_1, \dots, v_n$ und A wie oben, und sei noch $b \in K^n$ ein weiterer Vektor. Dann ist $b \in \langle v_1 \dots, v_n \rangle$ genau dann, wenn das lineare Gleichungssystem $Ax = b$ (mindestens) eine Lösung hat.*

Das heißt um im K^n lineare Unabhängigkeit zu überprüfen oder um zu sehen ob ein Vektor als Linearkombination anderer Vektoren gebildet werden kann müssen wir ein lineares Gleichungssytem lösen, was wir aber mit dem Gauss'schen Verfahren *immer* können. Damit ist auch die Frage ob Vektoren im K^n eine Basis bilden immer mit dem Gauss'schen Verfahren entscheidbar.

4.5 Dimension

Als erstes wollen wir eine kurze Charakterisierung von Basen[5] geben.

Satz 4.5.1. *Sei $B \subseteq V$ eine Teilmenge eines Vektorraumes V. Die folgenden Aussagen sind äquivalent:*

1. *B ist Basis.*

[5] Plural von Basis!

2. *B ist maximale linear unabhängige Menge; d.h. B ist linear unabhängig, aber für alle $v \in V$ ist $B \cup \{v\}$ linear abhängig.*

3. *B ist minimales Erzeugendensystem; d.h. B ist Erzeugendensystem, aber für alle $v \in B$ ist $B \setminus \{v\}$ nicht mehr Erzeugendensystem.*

Beweis. Um zu sehen, daß (1) $\Rightarrow$ (2), sei B Basis und $v \in V$ beliebig. Da B Basis, lässt sich v als Linearkombination von Vektoren aus B darstellen. D.h. aber, daß $\{v\} \cup B$ linear abhängig ist.
(2) $\Rightarrow$ (1): Sei umgekehrt B maximale linear unabhängige Menge. Damit B Basis ist, müssen wir zeigen, daß B Erzeugendensystem ist (linear unabhängig ist die Menge ja bereits). Sei also $v \in V$ beliebig. Nach Voraussetzung ist $B \cup \{v\}$ linear abhängig, was impliziert daß es Koeffizienten $k_0, \ldots, k_n \in K$ und Vektoren $v_1, \ldots, v_n$ gibt mit

$$0 = k_0 v + \sum_{i=1}^{n} k_i v_i \ .$$

Nun kann k_0 nicht null sein, da sonst B linear abhängig wäre. Also ist $k_0 \neq 0$. Und wir können obige Gleichung umformen zu

$$v = \sum_{i=1}^{n} -k_0^{-1} k_i v_i \ ,$$

und also ist v Linearkombination aus Vektoren in B.
Die Implikationen (1) $\iff$ (3) sind als Übung gelassen. □

Ein Vektorraum V heißt **endlich erzeugt** falls es Vektoren $v_1, \ldots, v_n$ gibt, so daß

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

Korollar 4.5.2. *Jeder nicht-trivial endlich erzeugte Vektorraum besitzt eine endliche Basis.*

Bemerkung 4.5.3. *Die harmlos wirkende Aussage, daß jeder Vektorraum eine Basis besitzt ist äquivalent zum sogenannten Zorn'schen Lemma oder Auswahlaxiom. Dieses wird zwar mehrheitlich von Mathematikern akzeptiert, besitzt aber Folgerungen wie das Banach-Tarski Paradoxon, welche doch einen üblen Beigeschmack lassen.*

Lemma 4.5.4 (Austauschlemma)**.** *Sei $B = \{v_1, \ldots, v_n\}$ Basis eines Vektorraums V. Sei des Weiteren $w \in V$ und $k_1, \ldots, k_n \in K$ so daß*

$$w = \sum_{i=1}^{n} k_i v_i \ .$$

Für jedes j mit $k_j \neq 0$ ist

$$B' = \{v_1, \ldots, v_{j-1}, w, v_{j+1}, v_n\} \tag{4.3}$$

Basis von V. (Anderst ausgedrückt: wir können v_j in der Basis durch w ersetzen).

Beweis. Seien alle Elemente wie in der Aussage. Wir müssen zeigen, daß B' Basis ist.

- B' ist Erzeugendensystem: sei $v \in V$ beliebig. Dank des Lemmas 4.3.9 reicht es zu zeigen, daß $B \subseteq \langle B' \rangle$. Es ist klar, daß

$$v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in \langle B' \rangle$$

  sind. Wir müssen nur noch zeigen, daß $v_j \in \langle B' \rangle$ ist. Da $k_j \neq 0$ ist und also ein multiplikatives Inverses besitzt (wir arbeiten ja über einem Körper) können wir (4.3) umformen zu:

$$v_j = w - \sum_{\substack{i=1 \\ i \neq j}}^{n} k_j^{-1} k_i v_i \ .$$

  Dies sieht zwar kompliziert aus, wenn man sich die Gleichung jedoch ansieht sieht man, daß w eine Linearkombination der Vektoren in B' ist.

- B' ist linear unabhängig: Sei $\ell_1, \ldots, \ell_n \in K$ so daß

$$0 = \ell_1 v_1 + \cdots + \ell_{j-1} v_{j-1} + \ell_j w + \ell_{j+1} v_{j+1} + \cdots + \ell_n v_n \ . \tag{4.4}$$

  Wäre $l_j \neq 0$, so könnten wir diese Gleichung umformen zu

$$w = \sum_{\substack{i=1 \\ i \neq j}}^{n} \ell_j^{-1} \ell_i v_i \ ,$$

  also als Linearkombination der linear unabhängigen Vektoren $v_1, \ldots, v_j$, wobei der Koeffizient von v_j null ist. Koeffizientenvergleich mit der Gleichung (4.3) liefert aber für die j-ten Koeffizienten dann, daß auch $k_j = 0$; ein Widerspruch zu den Annahmen des Satzes. Also muss $\ell_j = 0$ gelten. Damit vereinfacht sich die Gleichung (4.4) zu

$$0 = \ell_1 v_1 + \cdots + \ell_{j-1} v_{j-1} + \ell_{j+1} v_{j+1} + \cdots + \ell_n v_n \ .$$

  Da B linear unabhängig ist und damit auch jede Teilmenge folgt daß $\ell_1, \ldots, \ell_{j-1}, \ell_{j+1}, \ldots, \ell_n = 0$ ist. Zusammen also $\ell_1, \ldots, \ell_n = 0$, was bedeutet, daß B' linear unabhängig ist.

□

Satz 4.5.5 (Austauschsatz von Steinitz)**.** *Sei $B = \{v_1, \ldots, v_n\}$ Basis eines Vektorraums V und $C = \{w_1, \ldots, w_m\}$ eine Menge linear un-*

abhängiger Vektoren in V. Dann gilt $m \leqslant n$ und es gibt m Vektoren in B, die wir durch C ersetzen können.

Korollar 4.5.6. *Sei B eine endliche Basis eines Vektorraums und C eine Menge linear unabhängiger Vektoren. Dann gilt $|C| \leqslant |B|$. Insbesondere ist C endlich.*

Korollar 4.5.7. *Je zwei Basen haben die gleiche Anzahl an Elementen.*

Definition 4.5.8. Ist V ein endlich erzeugter Vektorraum, so besitzt nach obigen Korollar jede Basis B die gleiche Anzahl n and Vektoren. Ausserdem existiert eine Basis nach Korollar 4.5.2. Diese Zahl n bezeichnen wir als die **Dimension** von V. In Zeichen schreiben wir

$$\dim_K V = n \ .$$

Für den Fall das $V = \{0\}$ ist setzen wir $\dim_K V = 0$.[6].
Ist klar, über welchem K wir arbeiten, so schreiben wir auch einfach $\dim V$.

Korollar 4.5.9.

1. *Sei V Vektorraum mit $\dim V = n$. Besitzt eine linear unabhängige Menge B genau n Elemente, so ist B Basis.*

2. *Sei V Vektorraum mit $\dim V = n$. Ist eine Menge B Erzeugendensystem mit genau n Elementen, so ist B Basis.*

4.6 Unterräume

Definition 4.6.1. Sei V ein K-Vektorraum und $U \subseteq V$ eine Teilmenge. Ist U mit den gleichen Verknüpfungen wie V ein Vektorraum, so heisst U Unter(vektor)raum von V.

Um zu zeigen, daß eine Teilmenge $U \subseteq V$ eines Vektorraums ein Vektorraum ist müssen wir nur zeigen, daß U unter den Verknüpfungen $+$ und $\cdot$ abgeschlossen ist. Die Axiome vererben sich dann.

Satz 4.6.2. *Eine Teilmenge U eines Vektorraums V ist genau dann Unterraum, falls*

1. *für alle $u, u' \in U$ folgt, daß $u + u' \in U$ und*

2. *für alle $u \in U$ und $k \in K$ gilt $ku \in U$.*

Wo ist hier das Problem? Sind wir einmal etwas penibler. Wir haben einen Vektorraum V, dessen Verknüpfungen wir hier mit $+_V : V^2 \to V$ und $\cdot_V : K \times V \to V$ bezeichnen wollen. Sei ausserdem $U \subseteq V$ eine

[6] Man beachte, daß $\{0\}$ keine Basis besitzt

Teilmenge. Damit U selbst wieder ein Vektorraum ist müssen wir die zwei Verknüpfungen $+_U : U^2 \to U$ und $\cdot_U : K \times U \to U$ finden, so daß U mit diesen ein Vektorraum ist. Wir wollen auch, daß diese Verknüpfungen auf U mit $+_V$ und $\cdot_V$ übereinstimmen. Natürlich können wir einfach sagen, daß $u +_U u'$ gleich $u +_V u'$ sein soll. Allerdings wissen wir ja nicht ob, selbst wenn $u, u' \in U$ sind, dann $u +_V u' \in U$ ist. Abstrakt ausgedrückt: schränken wir $+_V : V^2 \to V$ auf U ein, so erhalten wir eine Abbildung $U^2 \to V$; wir wollen aber eine Abbildung vom Typ $U^2 \to U$.

Allgemein ist es mit Hilfe des obigen Satzes leicht zu zeigen, daß eine Untermenge ein Vektorraum ist – leichter zumindest als zu zeigen, daß es für sich genommen ein Vektorraum ist (wir sparen uns alle 8 Axiome nachzuweisen). Warum ist das so? Wenn wir schon wissen, daß V Vektorraum ist haben wir die Arbeit diese Axiome nachzuweisen ja schon geleistet.

Beispiel 4.6.3. *Ein wichtiges Beispiel eines Unterraumes sind die folgenden beiden Mengen: Ist V ein Vektorraum, so sind V und $\{0\}$ Untervektorräume.*

Lemma 4.6.4. *Sind U_1, U_2 Unterräume eines Vektorraums V, so ist auch $U_1 \cap U_2$ Unterraum, und insbesondere selbst Vektorraum.*

Beweis. Vorlesung. □

Man beachte, daß die Vereinigung zweier Unterräume im Allgemeinen kein Unterraum ist.

Beispiel 4.6.5.

1. *Die Menge aller stetigen (differenzierbaren, k-mal differenzierbaren) Funktion $[0,1] \to \mathbb{R}$ ist ein Untervektorraum der Menge aller Funktionen $[0,1] \to \mathbb{R}$.*

2. *Sei $U \subset \mathbb{R}^2$ ein Unterraum der Euklidischen Ebene. Ist $\dim U = 2$ so ist $U = \mathbb{R}^2$ und ist $\dim U = 0$, so ist $U = \{0\}$. Die einzig interessanten Unterräume sind die mit $\dim U = 1$ (andere kann es nach dem folgenden Satz nicht geben). U ist also von der Form $\langle v \rangle$, mit $v \neq 0$. Die Menge aller Linearkombinationen, die wir aus einem Vektor bilden können hat aber eine einfache Struktur:*

$$\langle v \rangle = \{\, kv \mid k \in \mathbb{R} \,\} \ .$$

Geometrisch sind dies alle Punkte der Ebene, welche auf der Geraden liegen, die durch den Punkt v und den Ursprung geht.

Satz 4.6.6. *Sei V Vektorraum mit $\dim V = n$. Jede Basis eines Unterraums U hat höchstens n Elemente.*
Ausserdem kann eine Basis B_U von U zu einer Basis aus V ergänzt werden. D.h. es gibt eine Basis B_V von V mit $B_U \subseteq B_V$.

Beweis. Dies folgt wie so viele Sätze direkt aus dem Austauschsatz. □

Seien U_1, U_2 Unterräume von V. Dann heißt U_1 **Komplement** von U_2, falls

- der Durchschnitt von U_1 und U_2 minimal ist, also

$$U_1 \cap U_2 = \{0\}$$

und

- beide zusammen V erzeugen; also

$$\langle U_1 \cup U_2 \rangle = V \ .$$

In diesem Fall sagen wir auch, daß V die **direkte Summe** von U_1 und U_2 ist und schreiben $V = U_1 \oplus U_2$.

Korollar 4.6.7. *Jeder Unterraum besitzt einen komplementären Unterraum.*

Satz 4.6.8. *Seien U_1 und U_2 komplementäre Unterräume von V; also $V = U_1 \oplus U_2$. Dann lässt sich jeder Vektor $v \in V$ eindeutig als Summe eines Vektors $u_1 \in U_1$ und $u_2 \in U_2$ zerlegen.*

Beweis. Das sich jeder Vektor als Summe zweier Vektoren schreiben lässt, folgt aus der Definition von $\langle\ \rangle$. Die Eindeutigkeit hat einen recht hübschen Beweis: Nehmen wir an, wir können v auf zwei verschiedene Arten schreiben:

$$u_1 + u_2 = v = u_1' + u_2' \ ,$$

mit $u_1, u_1' \in U_1$ und $u_2, u_2' \in U_2$. Dann gilt nach Umformung:

$$u_1 - u_1' = u_2 - u_2' \ .$$

Da U_1 und U_2 Unterräume sind ist der Vektor $u_1 - u_1' \in U_1$ und $u_2 - u_2' \in U_2$. Beide Vektoren sind aber gleich, also sowohl in U_1 als auch in U_2. Der einzige Vektor mit dieser Eigenschaft ist aber der Nullvektor. Also gilt $u_1 - u_1' = 0$ und $u_2 - u_2' = 0$ oder äquivalent $u_1 = u_1'$ und $u_2 = u_2'$. □

Satz 4.6.9 (Dimensionssatz für Unterräume)**.** *Seien U_1 und U_2 Unterräume des Vektorraums V. Dann gilt*

$$\dim\langle U_1 \cup U_2 \rangle = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \ .$$

Man beachte, daß die eckigen Klammern um $U_1 \cup U_2$ notwendig sind, da, wie wir gesehen haben, die Vereinigung zweier Unterräume nicht unbedingt wieder ein Untervektorraum ist.

Beweis. Sei $v_1, \ldots, v_k$ eine Basis von $U_1 \cap U_2$. Wegen Satz 4.6.6 können wir $u_1, \ldots, u_m$ und $w_1, \ldots, w_\ell$ so daß

$$v_1, \ldots, v_k, u_1, \ldots, u_m$$

eine Basis von U_1 ist und

$$v_1, \ldots, v_k, w_1, \ldots, w_\ell$$

eine Basis von U_2 ist.

Wenn wir zeigen können, daß

$$v_1, \ldots, v_k, u_1, \ldots, u_m, w_1, \ldots, w_\ell$$

eine Basis von $\langle U_1 \cup U_2 \rangle$ ist sind wir fertig, da dann

$$\begin{aligned} \dim\langle U_1 \cup U_2 \rangle &= k + m + \ell \\ &= k + m + k + \ell - \ell \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \ . \end{aligned}$$

Es ist einfach zu sehen, daß die angesprochenen Vektoren $\langle U_1 \cup U_2 \rangle$ erzeugen. Der trickreiche Teil des Beweises ist die lineare Unabhängigkeit. Seien also $q_1, \ldots, q_k, r_1, \ldots, r_m, p_1, \ldots, p_\ell$ Skalare, so daß

$$(\#) \quad 0 = q_1 v_1 + \cdots + q_k v_k + r_1 u_1 + \cdots + r_m u_m + p_1 w_1 + \cdots + p_\ell w_\ell \ .$$

Formen wir diese Gleichung um erhalten wir

$$-q_1 v_1 - \cdots - q_k v_k - r_1 u_1 - \cdots - r_m u_m = p_1 w_1 + \cdots + p_\ell w_\ell \ .$$

Da die linke Seite Element von U_1 ist und die rechte Element von U_2 ist der Vektor $p_1 w_1 + \cdots + p_\ell w_\ell \in U_1 \cap U_2$. Also gibt es weiter Skalare $a_1, \ldots, a_\ell$, so daß

$$a_1 v_1 + \cdots + a_k v_k = p_1 w_1 + \cdots + p_\ell w_\ell \ .$$

Man beachte, daß $v_1, \ldots, v_k, w_1, \ldots, w_\ell$ eine Basis von U_2 und insbesondere linear unabhängig ist. Schreiben wir die Gleichung komplizierter, so erhalten wir

$$\begin{aligned} a_1 v_1 + \cdots + a_k v_k + 0 w_1 + \cdots + 0 w_\ell & \\ = 0 v_1 + \cdots + 0 v_k + p_1 w_1 + \cdots + p_\ell w_\ell \ . & \end{aligned}$$

Durch Koeffizientenvergleich folgt also, daß $p_1 = \cdots = p_\ell = 0$. Damit vereinfacht sich die Gleichung (#) zu

$$0 = q_1 v_1 + \cdots + q_k v_k + r_1 u_1 + \cdots + r_m u_m \ .$$

Da $v_1, \ldots, v_k, u_1, \ldots, u_m$ linear unabhängig sind, folgt, daß auch $q_1 = \cdots = q_k = r_1 = \cdots = r_m = 0$. Also sind alle Skalare in (#) null und wir haben lineare Unabhängigkeit. □

4.7 Koordinaten

Sei $B = \{v_1, \ldots, v_n\}$ eine festgewählte Basis eines n-dimensionalen Vektorraums V. Da B Erzeugendensystem ist gibt es für jedes $w \in V$ Skalare $k_1, \ldots, k_n$, so daß

$$w = \sum_{i=1}^{n} k_i v_i$$

ist. Da B auch linear unabhängig ist, sind wegen Satz 4.3.7 diese Skalare eindeutig.

Definition 4.7.1. Die oben beschriebenen Skalare bezeichnen wir als die **Koordinaten** von w zur Basis B. In Zeichen wollen wir hierfür

$$(w)_B = (k_1, \ldots, k_n)$$

verwenden.

Haben wir eine Basis festgewählt, so ist in einem n-dimensionalen K-Vektorraum jeder Vektor durch n Elemente aus K eindeutig bestimmt. Allerdings hängt diese Darstellung von der gewählten Basis und sogar der Reihenfolge der Basis-vektoren ab!

Beispiel 4.7.2.

Bemerkung 4.7.3. *Sei B die Basis der Einheitsvektoren in K^n. So gilt:*

$$(a_1, \ldots, a_n)_B = (a_1, \ldots, a_n) \ .$$

Man mache sich klar, was diese Gleichung überhaupt bedeutet.

In einem späteren Kapitel werden wir sehen, wie wir einfach zwischen zwei verschiedenen Koordinatensystemen wechseln können.

Lemma 4.7.4. *Sind $v, u \in V$ Vektoren, $k \in K$ ein Skalar und $B = \{v_1, \ldots, v_n\}$ Basis von V, so gilt*

$$(v + u)_B = (v)_B + (u)_B$$

und

$$(kv)_B = k(v)_B \ .$$

5 Lineare Abbildungen

5.1 Definitionen und Beispiele

Definition 5.1.1. Seien V und W beide K-Vektorräume. Eine Funktion $\varphi : V \to W$ heißt **linear** wenn für alle $v, u \in V$ und $k \in K$ gilt:

- $\varphi(u + v) = \varphi(u) + \varphi(v)$ und
- $\varphi(ku) = k\varphi(u)$.

Beispiel 5.1.2.

1. *Die identische Abbildung* $\mathrm{id}_V : V \to V$ *definiert durch* $\mathrm{id}_V(v) = v$ *ist immer linear.*
2. *Ist* $k \in K$ *beliebig, so ist die Abbildung* $\varphi : V \to V$ *definiert durch* $\varphi(v) = kv$ *linear.*
3. *Ist* $V = U_1 \oplus U_2$*, so ist die* ***Projektion*** P_{U_1} *von* V *auf* U_1 *definiert dadurch, daß jeder Vektor* $v = u_1 + u_2$ *mit* $u_1 \in U_1$ *und* $u_2 \in U_2$ *auf* u_1 *abgebildet wird. (Da wir gesehen haben, daß diese Zerlegung eindeutig ist, ist diese Funktion wohldefiniert).*
4. *Fassen wir* K^n *als Vektorraum der* $n \times 1$ *Matrizen auf, so ist für jede* $m \times n$ *Matrix die Abbildung* $\varphi_A : K^n \to K^m$ *definiert durch* $\varphi_A(v) \mapsto Av$ *linear.*

Satz 5.1.3. *Ist* $\varphi : V \to W$ *eine lineare Abbildung, so sind das* ***Bild*** $\mathrm{Bi}(\varphi)$ *und der Kern* $\mathrm{Ke}(\varphi)$ *Unterräume von* W *bzw.* V*.*
Hierbei ist

- $\mathrm{Bi}(\varphi) = \{w \in W \mid \exists v \in V : \varphi(v) = w\}$ *und*
- $\mathrm{Ke}(\varphi) = \{v \in V \mid \varphi(v) = 0\}$

Satz 5.1.4. *Eine lineare Abbildung φ ist genau dann injektiv*[1]*, wenn* $\mathrm{Ke}(\varphi) = \{0\}$ *ist.*

Beweis. Es sind zwei Richtungen zu beweisen. Nehmen wir zunächst an, daß φ injektiv ist. Wir wollen zeigen, daß $\mathrm{Ke}(\varphi) = \{0\}$ ist. Sei also $v \in \mathrm{Ke}(\varphi)$, d.h. $\varphi(v) = 0$. Wie man leicht sieht ist auch $\varphi(0) = 0$ (da $\varphi(0) = \varphi(0_K 0) = 0_K \varphi(0) = 0$). Da φ injektiv ist gilt also $v = 0$.
Sei umgekehrt $\mathrm{Ke}(\varphi) = \{0\}$. Wir wollen zeigen, daß φ injektiv ist. Sei also $\varphi(v) = \varphi(v')$. Das ist äquivalent zu $\varphi(v) - \varphi(v') = 0$. Wir schließen:

$$0 = \varphi(v) - \varphi(v') = \varphi(v - v') \ .$$

Da das $\mathrm{Ke}(\varphi) = \{0\}$ ist das einzige Element, das auf die null abgebildet wird, die null selber. Also ist $v - v' = 0$, oder äquivalent $v = v'$. □

Satz 5.1.5. *Sind $\varphi : V \to W$ und $\psi : W \to U$ lineare Abbildungen zwischen Vektorräumen, so ist auch $\psi \circ \varphi : V \to U$ linear.*

Beweis. Einfach. Vorlesung. □

Definition 5.1.6. Eine bijektive lineare Abbildung $\varphi : V \to W$ heißt (Vektorraum)-**Isomorphismus**. Existiert ein Isomoprhismus zwischen V und W, so heißen diese **isomorph**.

Isomorphe Vektorräume sind, als Vektorräume quasi identisch und unterscheiden sich nur durch ihre Beschreibung.

Satz 5.1.7. *Ist $v_1, \ldots, v_n$ eine Basis eines Vektorraums V und $w_1, \ldots, w_n$ beliebige Vektoren in einem Vektorraum W, dann gibt es genau eine lineare Abbildung $\varphi : V \to W$ mit $\varphi(v_i) = w_i$ für $1 \leqslant i \leqslant n$.*

Beweis. Vorlesung. □

Dies heißt, daß eine Lineare Abbildung schon durch ihr Verhalten auf einer Basis eindeutig festgelegt ist.

Lemma 5.1.8. *Ist $\varphi : V \to W$ eine lineare Abbildung und $v_1, \ldots, v_n$ eine Basis von V. Sei ausserdem $w_i = \varphi(v_i)$. Dann gilt:*

1. *$w_1, \ldots, w_n$ sind genau dann linear unabhängig, wenn φ injektiv ist.*

2. *$w_1, \ldots, w_n$ sind genau dann Erzeugendensystem von W, wenn φ surjektiv ist.*

3. *$w_1, \ldots, w_n$ sind genau dann Basis, wenn φ bijektiv also Isomorphismus ist.*

[1] Zur Erinnerung: Eine Funktion heißt injektiv, wenn verschiedene Elemente auf verschiedene Element abgebildet werden. Logisch ausgedrückt: für alle $u, v \in V$ gilt $\varphi(u) = \varphi(v) \implies u = v$

Korollar 5.1.9. *Je zwei K-Vektorräume derselben Dimension sind isomorph.*

Wie wir gesehen haben ist für jede Matrix A die Funktion $\varphi_A : K^n \to K^m$ definiert durch Multiplikation mit eben der Matrix A linear. Die Umkehrung stimmt in folgendem Sinne:

Definition 5.1.10. Sei $\varphi : V \to W$ eine lineare Abbildung, $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $C = \{w_1, \ldots, w_m\}$ eine Basis von W. Für jedes Element v_j können wir das Bild $\varphi(v_j)$ eindeutig ausdrücken als Linearkombinationen der Basiselemente $w_1, \ldots, w_m$:

$$\varphi(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_j \ .$$

Die $m \times n$ Matrix mit eben diesen Einträgen a_{ij} heißt **Darstellungsmatrix** der linearen Abbildung φ. D.h. die Koeffizienten von $\varphi(v_j)$ bilden die j-te *Spalte* der Darstellungsmatrix. Diese wollen wir mit $M(\varphi, B, C)$ bezeichnen. Sind die Basen fest gewählt, so wollen wir auch kurz $M(\varphi)$ benutzen.

Man beachte, daß die Darstellungsmatrix *eindeutig* ist, aber von den gewählten Basen abhängt.

Satz 5.1.11. *Ist $\varphi : V \to W$ linear Abbildung und $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $C = \{w_1, \ldots, w_m\}$ eine von W.*
Es gilt:

$$(\varphi(v))_C = M(\varphi, B, C)(v)_B$$

Beweis. Wir wollen zunächst zeigen, daß $(\varphi(v_j))_C = A(v_j)_B$, der allgemeine Fall folgt dann wegen Lemma 4.7.4.
Wir berechnen

$$\begin{aligned}(\varphi(v_j))_C &= (a_{1j}, a_{2j}, \ldots, a_{mj}) \\ &= M(\varphi, B, C)e_j \\ &= M(\varphi, B, C)(v_j)_B \ .\end{aligned}$$

□

Satz 5.1.12. *Sind $\varphi : V \to W$ und $\psi : W \to U$ lineare Abbildungen zwischen endlichen Vektorräumen V, W, U mit Basen B, C, D. So gilt für die Darstellungsmatrizen:*

$$M(\psi \circ \varphi, B, D) = M(\varphi, B, C)M(\psi, C, D) \ .$$

D.h. Komposition von Abbildungen entspricht der Multiplikation ihrer darstellenden Matrizen.

Beweis. Sei $B = v_1, \ldots, v_n$, $C = u_1, \ldots, u_k$ und $D = w_1, \ldots, w_m$. Des Weiteren sei $(a_{ij}) = M(\varphi, B, C)$, $(b_{ij}) = M(\psi, C, D)$ und $(c_{ij}) =$

$M(\psi \circ \varphi, B, D)$. D.h.

$$\varphi(v_j) = \sum_{i=1}^{k} a_{ij} u_i$$
$$\psi(u_i) = \sum_{\ell=1}^{m} b_{\ell i} u_i$$
$$(\psi \circ \varphi)(v_j) = \sum_{\ell=1}^{m} c_{\ell j} w_\ell$$

Dann müssen wir zeigen, daß

$$c_{i,j} = \sum_{\ell=1}^{m} a_{i\ell} b_{\ell j} \ .$$

Zum Berechnen der c_{ij} müssen wir uns $\psi \circ \varphi(v_j)$ betrachten:

$$\begin{aligned}(\psi \circ \varphi)(v_j) &= \psi\left(\sum_{i=1}^{k} a_{ij} u_i\right) \\ &= \sum_{i=1}^{k} a_{ij} \psi\left(u_i\right) \\ &= \sum_{i=1}^{k} a_{ij} \sum_{\ell=1}^{m} b_{\ell i} w_l \\ &= \sum_{\ell=1}^{m} \sum_{i=1}^{k} a_{ij} b_{\ell i} w_l \\ &= \sum_{\ell=1}^{m} \left(\sum_{i=1}^{k} a_{ij} b_{\ell i}\right) w_l\end{aligned}$$

Dies ist aber genau was wir zeigen wollten. □

Bemerkung 5.1.13. *Sei A eine $m \times n$ Matrix, so ist A Darstellungsmatrix der Abbildung $\varphi_A : K^n \to K^m$ bezüglich der Standardbasen.*

Beispiel 5.1.14. *Die Identische Abbildung id_{K^n} hat die Einheitsmatrix als Darstellungsmatrix bezüglich der Standardbasis.*

Beispiel 5.1.15. *Sei V der Vektorraum der Polynome über $\mathbb{R}$ vom Grad kleiner gleich 4. (V ist ein Unterraum des Vektorraums aller Funktionen $\mathbb{R} \to \mathbb{R}$). Man kann leicht sehen, daß $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$, $p_3(x) = x^3$ und $p_4(x) = x^4$ eine Basis B von V bilden. Wie man leicht sehen kann ist die Ableitung $\varphi(p) = p'$ eine lineare*

Abbildung auf diesem Vektorraum. Die darstellende Matrix ist:

$$M(\varphi, B, B) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

5.2 Umkehrabbildung und Matrixinverse

Zur Erinnerung: ist $f : A \to B$ eine bijektive Abbildung, so können wir die Umkehrabbildung $f^{-1} : B \to A$ bilden, die jedem Element $b \in B$ das eindeutig bestimmte Element a zuordnet, für das $f(a) = b$ gilt.[2]
Die Umkehrabbildung ist dadurch charakterisiert, daß $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$ gilt.

Lemma 5.2.1. *Ist $\varphi : V \to W$ eine Isomorphismus, so ist die Umkehrabbildung φ^{-1} ebenfalls lineare Abbildung.*

Definition 5.2.2. Eine $n \times n$ Matrix A heißt **invertierbar**, falls es eine Matrix A^{-1} gibt, so daß $AA^{-1} = I_n = A^{-1}A$ gilt.

Fügen wir die bisherigen Ergebnisse zusammen, so erhalten wir folgendes

Korollar 5.2.3. *Ist A eine $n \times n$ Matrix, so daß $\varphi_A : K^n \to K^n$ bijektiv ist, so ist A invertierbar.*

Es gilt

Lemma 5.2.4. *Sind A, B invertierbar, so gilt*

1. *Es gibt genau eine Matrixinverse.*
2. $(A^{-1})^{-1} = A$
3. $(AB)^{-1} = B^{-1}A^{-1}$

Im einem folgenden Kapitel wollen wir uns überlegen, wann eine Matrix invertierbar ist, und wie man das Inverse konkret finden kann.

Definition 5.2.5. Sind B und C Basen eines endlich-dimensionalen Vektorraums, so heißt $M(\mathrm{id}, B, C)$ **Basistransformationsmatrix** zwischen B und C. Da id bijektiv ist, ist diese Matrix invertierbar.

Wie der Name sagt erlaubt die Basistransformationsmatrix zwischen zwei Koordinatendarstellungen hin- und herzuspringen. Dies folgt aus der Allgemeinen Aussage (Satz 5.1.11), da $\mathrm{id}(v) = v$:

$$(v)_C = M(\mathrm{id}, B, C)(v)_B$$

[2] Hierbei wird die Surjektivität dazu verwendet, daß ein solches a existiert, und die Injektivität, daß es höchstens ein solches a gibt.

Satz 5.2.6. *Ist $\varphi : V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen und B, B' und C, C' Basen von V und W, und $A = M(\varphi, B, C)$ und $A' = M(\varphi, B', C')$, so gilt*

$$A = SA'T$$

wobei S und T Basistransformationsmatrizen, also insbesondere invertierbar sind.

5.3 Algorithmus zum Invertieren von Matrizen bzw. Zerlegen in Elementarmatrizen

Zwar haben wir im letzten Kapitel gesehen, unter welchen Umständen die Existenz einer inversen Matrix gesichert ist, jedoch wissen wir nicht, wie wir diese praktisch berechnen können. Zunächst ein paar technische Hilfsmittel um elementare Zeilenumformungen auf Matrixmultiplikationen zurückführen zu können:
Sei $E_{i,j}$ die $(n \times n)$-Matrix, die wie die Einheitsmatrix aufgebaut ist, aber an (i,i)-ter und (j,j)-ter Stelle eine 0 hat und dafür an (i,j)-ter und (j,i)-ter Stelle eine 1. Dann gilt:

> Multiplizieren wir eine $n \times m$ Matrix A von links mit $E_{i,j}$, so erhalten wir die Matrix, die durch Vertauschen der i-ten mit der j-ten Zeile aus A hervorgeht.

Sei $F_{i,k}$ die $(n \times n)$-Matrix, die wie die Einheitsmatrix aufgebaut ist, aber an (i,i)-te Stelle den Eintrag $k \neq 0$ hat. Dann gilt:

> Multiplizieren wir eine $n \times m$ Matrix A von links mit $F_{i,k}$, so erhalten wir die Matrix, die durch Multiplikation der i-ten Zeile mit k aus A hervorgeht.

Sei $G_{k,i,j}$ die $(n \times n)$-Matrix, die wie die Einheitsmatrix aufgebaut ist, aber an (j,i)-ten Stelle den Eintrag k hat. Dann gilt:

> Multiplizieren wir eine $n \times m$ Matrix A von links mit $G_{k,i,j}$, so erhalten wir die Matrix, die durch Addition des k-fachen der i-ten Zeile zur j-ten Zeile aus A hervorgeht.

Definition 5.3.1. Diese drei Typen von Matrizen wollen wir als **Elementarmatrizen** bezeichnen.

Lemma 5.3.2. *Elementarmatrizen sind invertierbar.*

Dies führt uns zu folgendem Algorithmus um eine Matrix $A \in M(n \times n, K)$ zu invertieren.

Schritt 1: Wir schreiben A neben I_n. Also:

$$(A \; I_n)$$

Schritt 2: Durch elementare Zeilenumformungen wandeln wir die linke Seite A in die Einheitsmatrix um. Alle Änderungen haben auch einen Effekt auf die rechte Seite. Wir enden also mit

$$(I_n \; A)$$

Schritt 3: Auf der rechten Seite steht nun das Inverse zu A.

Dieser überraschend einfache Algorithmus funktioniert aus folgendem Grund:
Alle in Schritt 2 ausgeführten Umformungen lassen sich durch Multiplikation mit Elementarmatrizen beschreiben. Sagen wir diese sind der Reihe nach $D_1, \ldots, D_n$. Dann haben wir ja den Zusammenhang

$$I_n = D_n D_{n-1} \ldots D_1 A \; ,$$

was aber heißt das A invertierbar ist. Mehr noch: am Ende des zweiten Schritts steht eben $D_n D_{n-1} \ldots D_1$ auf der rechten Seite; wir können dies also einfach ablesen. (Dies ist der eigentliche Sinn der rechten Seite).
Mehr noch: Notieren wir uns welche D_i wir verwendet haben, so wissen wir, daß

$$A^{-1} = D_n D_{n-1} \ldots D_1$$

und damit

$$A = (D_n D_{n-1} \ldots D_1)^{-1} = D_1^{-1} \ldots D_{n-1}^{-1} D_n^{-1} \; .$$

Das heißt dieser Algorithmus ermöglicht uns es auch eine invertierbare Matrix als Produkt (sehr einfacher) Elementarmatrizen zu schreiben.

Bemerkung 5.3.3. *Eine Matrix ist genau dann invertierbar, wenn sie sich als Produkt von Elementarmatrizen schreiben lässt.*

Wie wir gleich sehen werden kann der Algorithmus sogar verwendet werden um zu testen, ob eine Matrix überhaupt invertierbar ist (Satz 5.5.8).

5.4 Mal wieder lineare Gleichungssysteme

Kehren wir nochmals zu linearen Gleichungssystemen zurück. Ist $Ax = b$ ein lineares GS mit invertierbaren A (also insbesondere quadratischen A), so können wir die Lösungen leicht ausrechnen:

Satz 5.4.1. *Ist $A \in M(n \times n, K)$ invertierbar, so hat*

$$Ax = b$$

genau ein Lösung, nämlich $x = A^{-1}b$. Insbesondere hat

$$Ax = 0$$

nur die triviale Lösung $x = 0$.

Man beachte jedoch, daß dies nur für invertierbare (quadratische) Matrizen gilt. Im allgemeinen Fall kann ein System $Ax = b$ immer noch keine, genau eine und unendlich viele Lösungen haben.

5.5 Der Rang

Wie wir gesehen haben sind $\mathrm{Bi}(\varphi)$ und $\mathrm{Ke}(\varphi)$ für eine lineare Abbildung $\varphi : V \to W$ Unterräume von W bzw. V.

Satz 5.5.1 (Dimensionssatz für lineare Abbildungen)**.** *Ist $\varphi : V \to W$ eine lineare Abbildung und V endlich-dimensional, so gilt:*

$$\dim V = \dim \mathrm{Bi}(\varphi) + \dim \mathrm{Ke}(\varphi)$$

Beweis. Sei $v_1, \ldots, v_r$ eine Basis von $\mathrm{Ke}(\varphi)$ und $w_1, \ldots, w_s$ Basis von $\mathrm{Bi}(\varphi)$ (man überlege sich, warum $\mathrm{Bi}(\varphi)$ endlich-dimensional ist). Nach Definition des Bildes gibt es $u_1, \ldots, u_s$ so daß $\varphi(u_i) = w_i$ für $1 \leqslant i \leqslant s$. Wir wollen zeigen, daß $B = \{u_1, \ldots, u_s, v_1, \ldots, v_r\}$ Basis von V ist. Sei zunächst $v \in V$ beliebig. Dann gibt es $k_1, \ldots, k_s \in K$ mit $\varphi(v) = \sum k_i w_i$. Der Trick des Beweises ist nun, daß wir uns den Vektor $v' = \sum k_i u_i$ betrachten. Es gilt $\varphi(v') = \varphi(v)$, und damit $v - v' \in \mathrm{Ke}\,\varphi$. Also gibt es $\ell_1, \ldots, \ell_r \in K$ mit $v - v' = \sum \ell_i v_i$. Zusammen gilt also

$$v = v - v' + v' = \sum \ell_i v_i + \sum k_i u_i \ .$$

Also ist B Erzeugendensystem. Um zu sehen, daß B auch linear unabhängig ist, betrachten wir $k_1, \ldots, k_s, \ell_1, \ldots, \ell_r$ mit

$$0 = k_1 u_1 + \cdots + k_s u_s + \ell_1 v_1 + \cdots + \ell_r v_r \ .$$

Wenden wir φ auf diese Gleichung an erhalten wir

$$0 = k_1 \varphi(u_1) + \cdots + k_s \varphi(u_s) = \sum k_i w_i \ .$$

Da $w_1, \ldots, w_s$ Basis ist folgt $k_1 = \cdots = k_s = 0$. Also gilt

$$0 = \ell_1 v_1 + \cdots + \ell_s v_s \ ,$$

aber da auch $v_1, \ldots, v_r$ Basis sind folgt auch, daß $\ell_1 = \cdots = \ell_r = 0$. □

Korollar 5.5.2. *Ist $\varphi : V \to W$ eine lineare Abbildung und* $\dim V = \dim W < \infty$ *so sind die folgenden Aussagen äquivalent:*

1. *φ ist bijektiv*

2. *φ ist injektiv*

3. *φ ist surjektiv.*

Für die speziellen linearen Abbildungen φ_A, d.h. solche vom Typ $K^n \to K^m$ die durch Multiplikation mit einer Matrix A gegeben sind, ist das Bild und der Kern leicht zu bestimmen.

Bemerkung 5.5.3. *Ist $A \in M(m \times n, K)$ so ist*

1. $\mathrm{Ke}(\varphi_A) = \{x \in K^n \mid Ax = 0\}$.
   D.h. um den Kern zu finden müssen wir ein homogenes lineares Gleichungssystem lösen.

2. $\mathrm{Bi}(\varphi_A) = \langle v_1, \ldots, v_n \rangle$, *wobei $v_1, \ldots, v_n$ die Spalten der Matrix A sind.*

Definition 5.5.4. Ist A eine Matrix, so ist der **Rang** von A definiert durch:

$$\mathrm{rang}(A) = \dim \mathrm{Bi}(\varphi_A) \ .$$

Bemerkung 5.5.5. *Sind A, B Matrizen, so daß $A = TBS$, wobei T, S Basistransformationsmatrizen sind, so ist* $\mathrm{rang}(A) = \mathrm{rang}(B)$

Aus dieser Darstellung folgt:

Satz 5.5.6. *Ist $\varphi : V \to W$ eine lineare Abbildung und V und W endlich-dimensional, so ist der Rang aller darstellenden Matrizen dieser Abbildung gleich.*

Beispiel 5.5.7. *Der Rang der Abbildung aus Beispiel 5.1.15 ist* 4.

Satz 5.5.8. *Die folgenden Aussagen sind für eine Matrix $A \in M(n \times n, K)$ äquivalent:*

1. *φ_A ist bijektiv.*

2. *A ist invertierbar,*

3. *Es existiert B so daß $BA = I_n$.*

4. $\mathrm{rang}(A) = n$.

5. *Die Spalten von A sind linear unabhängig.*

6. *$Ax = 0$ hat nur die triviale Lösung. D.h.* $\mathrm{Ke}(\varphi_A) = \{0\}$.

7. *A lässt sich durch elementare Zeilenumformungen in die Einheitsmatrix überführen.*

5.6 Die Geometrie von Lineare Abbildungen des $\mathbb{R}^2$

In diesem Kapitel wollen wir uns überlegen, wie invertierbare lineare Abbildungen $\mathbb{R}^2 \to \mathbb{R}^2$ überhaupt aussehen. Gehen wir zunächst einige Beispiel durch. Zur Erinnerung, da wir hier die Standardbasis verwenden ist die erste Spalte der darstellenden Matrix das Bild von e_1 und die zweite Spalte das Bild von e_2.

Matrix	Beschreibung	Bild
$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	Spiegelung an der y-Achse.	
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	Spiegelung an der x-Achse.	
$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	Spiegelung an der Winkelhalbierenden des 1./3. Quadranten.	
$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$	Rotation um den Winkel α	
$\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$	Kompression/Streckung entlang der y-Richtung um den Faktor k	
$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$	Kompression/Streckung entlang der x-Richtung um den Faktor k	
$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$	Scherung in x-Richtung mit Faktor k	

Satz 5.6.1. *Elementarmatrizen in $M(2\times 2, \mathbb{R})$ lassen sich als Produkt obiger Matrizen schreiben.*

Korollar 5.6.2. *Jede invertierbare lineare Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ lässt sich als Hintereinanderausführung von obigen Matrizen beschreiben.*

5.7 Die Geometrie von Lineare Abbildungen des $\mathbb{R}^3$

Wie im letzten Teil, sind auch viele natürliche Transformationen des $\mathbb{R}^3$ linear.

Matrix	Beschreibung	Bild
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$	Rotation um den Winkel α um die x-Achse	
$\begin{pmatrix} \cos(\alpha) & 0 & \sin(\alpha) \\ 0 & 1 & 0 \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{pmatrix}$	Rotation um den Winkel α um die y-Achse	
$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Rotation um den Winkel α um die z-Achse	
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$	Spiegelung an der xy-Ebene	
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Spiegelung an der xz-Ebene	
$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Spiegelung an der yz-Ebene	

5.8 Anwendung: CSS3-Transformationen

HTML-Seiten bestehen größtenteils aus geschachtelten Elemente, die den hierarchischen Aufbau einer Seite beschreiben. Ein Browser versucht all diese Elemente als Rechtecke ineinander und nebeneinander anzuordnen und gemäß der Vorgaben zu rendern. Zusätzlich können gewisse Eigenschaften all dieser Elemente bestimmt werden. Dies geschah anfänglich direkt im HTML-Code, was allerdings sehr schnell unübersichtlich und schwer zu verwalten war (Z.B. wenn man die Schriftart jedes Paragraphen von Comic-Sans auf Helvetica umstellen will musste man dies in jeder Paragraph-Box tun). Besser ist es das Design auf einer vom Inhalt getrennten Ebene zu verwalten, am Besten in einer separaten Datei. Die formale Sprache, die beschreibt welche

Elemente welche Eigenschaften erhalten heißt CSS. Der aktuellste Standard ist CSS3, der allerdings noch nicht vollkommen unterstützt wird: zur Zeit ist kein einziger Browser in der Lage mit allen Erweiterungen umzugehen[3]

Eine der in CSS3 neu hinzugefügten Eigenschaften[4] ist:

```
transform-matrix(a,b,c,d,0,0)
```

Auf das Element für das diese Eigenschaft angegeben wird, wird die Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

angewendet.[5] D.h., daß wir alle 2D-Transformationen des letzten Kapitels auf Elemente anwenden können! Ist z.B.

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} ,$$

dann wird der Punkt p durch $\varphi_A(p)$ ja um 90 Grad gedreht. D.h., daß beispielsweise durch

```
<h2 style="transform-matrix(0,1,-1,0,0,0)">Hallo</h2>
```

Die Überschrift “Hallo” gedreht dargestellt wird.
Für Rotationen, Spiegelungen und Skalierungen gibt es zwar auch explizite Eigenschaften (transform-rotate, transform-), die man verwenden kann, allerdings ist dies wohl eher, für Anwender gedacht, die nichts von Matrixtransformationen verstehen.

[3] Mal ganz abgesehen vom Internet-Explorer, der traditionell nicht nur sehr langsam ist die Standards umzusetzen, sondern dies auch noch fehlerhaft oder auf eine Art und Weise, die mit dem Standard inkompatibel ist macht. Siehe z.B. http://www.findmebyip.com/litmus/

[4] Zur Zeit muß man noch für jeden Browser einen Präfix verwenden. Für Firefox lautet die entsprechende Eigenschaft -moz-transform-matrix(a,b,c,d,x,y), für Safari/Chrome -webkit-transform-matrix(a,b,c,d,x,y) usw.

[5] die letzten beiden Argumente sind dafür da, das Element in der Ebene zu verschieben.

6 Determinanten

Kommen wir nun zu einem weiteren wichtigen Kapitel. Determinanten geben einem in gewisser Weise die Essenz einer Matrix. Zur Erinnerung aus DMI:

Definition 6.0.1. Eine Permutation σ der Länge n ist eine bijektive Funktion $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$.

Die Permutationen der Länge n bilden eine nicht-kommutative Gruppe S_n mit der Komposition von Funktionen als Verknüpfung. Jede Permutation kann als Produkt von Transpositionen geschrieben werden, das sind Permutationen, welche genau zwei Elemente vertauschen. Diese Zerlegung ist nicht eindeutig, allerdings werden immer gerade oder ungerade viele Transpositionen benötigt.

Satz 6.0.2. *Es gibt eine Funktion* $\operatorname{sgn} : S_n \to \{-1, 1\}$ *so daß*

$$\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \ ,$$

und $\operatorname{sgn}(\tau) = -1$ *für alle Transpositionen.*

Diese Signumsfunktion erlaubt uns folgendes Monster zu definieren:

Definition 6.0.3. Ist $A \in M(n \times n, K)$ eine quadratische Matrix mit Einträgen a_{ij}, so ist die Determinante von A definiert durch

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdot \cdots \cdot a_{n,\sigma(n)} \ .$$

Diese Darstellung ist etwas unhandlich. Deshalb wollen wir zunächst den 2×2 und 3×3 Fall behandeln.

Satz 6.0.4.

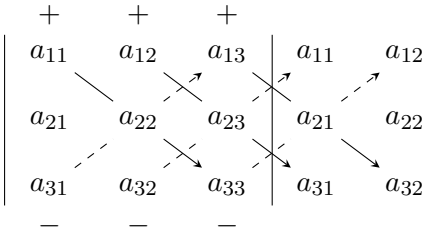
$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \ .$$

Satz 6.0.5 (Regel von Sarrus)**.** *Die Determinante einer* 3×3 *Matrix lässt sich durch folgenden Formalismus berechnen. Mit Hilfe der*

Leibnizformel:

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11} \tag{6.1}$$

Dies kann man sich wie folgt merken



Bei größeren Matrizen ist es nahezu unmöglich die Determinante explizit auszurechnen. Da bei Größe $n \times n$ die Summe über alle Permutationen in S_n geht und es von letzteren $n!$ gibt, nimmt die Komplexität schnell zu. Für besondere Matrizen geht es allerdings einfacher:

Satz 6.0.6.

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & & a_{nn} \end{pmatrix} = a_{11}a_{22}\dots a_{nn} \ .$$

Insbesondere ist $\det(I_n) = 1$

Beweis. Wie man leicht sieht gibt es für $\sigma \in S_n$ mit $\sigma \neq \mathrm{id}$ einen Index i so daß $\sigma(i) < i$ ist. Deshalb ist $a_{i\sigma(i)} = 0$ für dieses σ und damit

$$a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)} = 0 \ .$$

D.h.

$$\begin{aligned} \det(A) &= \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)} \\ &= \mathrm{sgn}(\mathrm{id}) a_{1\mathrm{id}(1)}a_{2\mathrm{id}(2)}\dots a_{n\mathrm{id}(n)} \\ &= a_{11}a_{22}\dots a_{nn} \end{aligned}$$

□

Satz 6.0.7.

1. *Geht A durch Vertauschen von zwei Zeilen aus B hervor, so ist*

$$\det A = -\det B \ .$$

2. *Besitzt eine Matrix zwei identische Zeilen, so ist ihre Determinante gleich null.*

3. *Geht A durch Multiplikation einer Zeile mit einem Faktor k aus B hervor, so ist*
$$\det A = k \det B \ .$$

4. *Die Determinante ist additiv in jeder Zeile:*
$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_i + z_i' \\ \vdots \\ z_n \end{pmatrix} = \det \begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix} + \det \begin{pmatrix} z_1 \\ \vdots \\ z_i' \\ \vdots \\ z_n \end{pmatrix}$$

5. *Geht A durch Addition einer Zeile zu einer anderen aus B hervor, so ist*
$$\det A = \det B \ .$$

Beweis. Vorlesung □

Hat eine Abbildung die Eigenschaften 3 und 4, so sagen wir auch, daß die Abbildung **eine linear Abbildung in jeder Zeile** ist.
Obiger Satz ermöglicht uns allgemein Determinanten zu berechnen. Wir werden noch weitere Möglichkeiten kennenlernen.

Beispiel 6.0.8.

Lemma 6.0.9. *Es gibt genau eine Abbildung $F : M(n \times n, K) \to K$, so daß*

1. *$F(A) = 0$, falls in A eine Zeile doppelt vorkommt.*

2. *$F(I_n) = 1$.*

3. *F ist eine lineare Abbildung in jeder Zeile.*

Mehr: hat eine Abbildung obige Eigenschaften, so ist $F(A) = 0$, falls die Zeilen von A linear abhängig sind.
Da die Determinante obige Eigenschaften hat, heißt das, daß sie eindeutig durch diese Eigenschaften bestimmt ist.

Beweis. Vorlesung. □

Satz 6.0.10 (Entwicklung nach einer Zeile)**.** *Sei $A = (a_{ij}) \in M(n \times n, K)$ und A_{ij} die Matrix, die durch Löschen der i-ten Zeile und j-ten Spalte entsteht, so gilt*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \ . \tag{6.2}$$

Da die es enorm viel Rechenaufwand benötigt, um die Determinante einer großen Matrix zu berechnen ist obiger Satz nützlich, denn er erlaubt es die Berechnung der Determinante auf die kleinerer Matrizen zurückzuführen. Besonders effizient ist das Ganze, wenn eine Zeile viele Nullen enthält.

Beweis. Wir zeigen, daß die durch die rechte Seite von (6.2) definierte Abbildung die Eigenschaften in Lemma 6.0.9 hat. □

Definition 6.0.11. Die **transponierte Matrix** A^T zu einer $m \times n$ Matrix A ist die $n \times m$ Matrix mit Einträgen a_{ji}.

Beispiel 6.0.12.

Satz 6.0.13. *Ist A quadratisch, so gilt:*

$$\det(A) = \det(A^T) \ .$$

Beweis. Setzt man A^T in die Leibnizformel ein, erhalten wir:

$$\det(A^T) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(n),n} \ .$$

Ordnen wir die Reihenfolge der Terme in $a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(n),n}$ um, so erhalten wir $a_{1,\sigma^{-1}(1)} \cdot \dots \cdot a_{n,\sigma^{-1}(n)}$. Also ist

$$\begin{aligned} \det(A^T) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot \dots \cdot a_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma^{-1}(1)} \cdot \dots \cdot a_{n,\sigma^{-1}(n)} \end{aligned}$$

Da jetzt noch $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ ist und

$$\{\sigma \in S_n\} = \{\sigma^{-1} \mid \sigma \in S_n\}$$

gilt also

$$\det(A^T) = \det(A) \ .$$ □

Korollar 6.0.14. *Alle Sätze dieses Kapitels gelten auch für Spalten an Stelle von Zeilen.*

Als letztes wollen wir uns noch an eine zentrale Aussage vorarbeiten:

Lemma 6.0.15.

1. *Für Elementarmatrizen D ist* $\det(D) \neq 0$.
2. *Für eine beliebige quadratische Matrix A und eine Elementarmatrix D gilt*
$$\det(DA) = \det(D)\det(A) \ .$$
3. *A ist genau dann invertierbar, falls* $\det(A) \neq 0$.

Satz 6.0.16. *Sind $A, B \in M(n \times n, K)$, so gilt*

$$\det(AB) = \det(A)\det(B) \ .$$

Korollar 6.0.17. *Ist $A \in M(n \times n, K)$ invertierbar, so ist*

$$\det(A^{-1}) = \det(A)^{-1} \ .$$

Zusammenfassend haben wir folgende Möglichkeiten Determinanten zu bestimmen:

- Explizit bei 2×2 und 3×3 Matrizen.
- Entwicklung nach einer Zeile/Spalte (insbesondere nützlich falls diese viele Nullen enthält)
- Durch Elementare Zeilenumformungen und Zurückführen auf Dreiecksmatrizen.

Satz 6.0.18 (Die Cramersche Regel)**.** *Ist $Ax = b$ ein lineares Gleichungssystem mit A invertierbar, so ist die eindeutige Lösung $x = (x_1, \ldots, x_n)$ gegeben durch*

$$x_i = \frac{\det A_i}{\det(A)}$$

wobei A_i die Matrix ist, die entsteht wenn man die i-te Spalte von A durch b ersetzt.

Beweis. Wir verwenden folgenden hübschen Trick:
Sei X_i die Matrix, die wie die Einheitsmatrix aussieht, aber an Stelle der i-ten Spalte x (als Spaltenvektor) hat. Wie man nachrechnen kann gilt dann $AX_i = A_i$. Ausserdem ist $\det(X_i) = x_i$. Mit Hilfe der Multiplikationsregel gilt also

$$\det(A_i) = \det(AX_i) = \det(A)\det(X_i) = \det(A)x_i \ .$$

Da $\det(A) \neq 0$ können wir umformen zu

$$x_i = \frac{\det(A_i)}{\det(A)} \ . \qquad \square$$

Korollar 6.0.19. *Ist $\det(A) \neq 0$, so kann die Inverse Matrix zu A berechnet werden durch*

$$A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)} \ ,$$

*wobei $\mathrm{adj}(A)$ die sogenannte **adjungierte** Matrix ist. Das ist die $n \times n$-Matrix, deren i,j-ter Eintrag gleich $(-1)^{i+j}\det(A_{ji})$ ist. (Diese Einträge heißen übrigen Cofaktoren).*

Bemerkung 6.0.20. *Die letzten beiden Resultate sind weniger von praktischen also von theoretischen Nutzen. So folgt z.B. leicht, daß die Lösungen zu* $Ax = b$ *und die Inverse* A^{-1} stetig *von* A *und* b *abhängen. Ausserdem ermöglichen diese Darstellungen Abschätzungen bei numerischen Verfahren.*

7 Eigenwerte und Eigenvektoren

7.1 Der "Google"-Algorithmus

Wie misst man Relevanz einer Webseite?[1] Der Kern des äussert erfolgreichen Google Algorithmus[2] ist folgende Idee: Eine Webseite ist umso wichtiger je mehr andere Seiten auf sie verlinken. Ausserdem vererbt sich die Wichtigkeit, d.h. eine Link von einer wichtigen Webseite verleiht einer Seite mehr Relevanz, als einer einer unwichtigen Seite. Betrachten wir das Ganze etwas formaler. Nehmen wir an wir haben die Seiten $P_1, \ldots, P_n$ (auf denen z.B. ein Suchwort vorkommt). Wir suchen ein Maß der Wichtigkeit $I : \{P_1, \ldots, P_n\} \to \mathbb{R}$ für jede Seite (je höher die Zahl $I(P_j)$ desto wichtiger die Seite P_j). Wie schon angedeutet soll gelten:

$$I(P_j) = \sum \frac{I(P_i)}{\ell_i} \ .$$

Wobei die Summe über alle Seiten P_i läuft, die auf P_j verweisen und ℓ_i die Anzahl der ausgehenden Links von P_j sein soll (verweist eine Seite auf 3 andere Seiten, so soll jede dieser Seiten $1/3$ der Wichtigkeit der ersten Seite vererbt bekommen).
Definieren wir nun eine Matrix H über

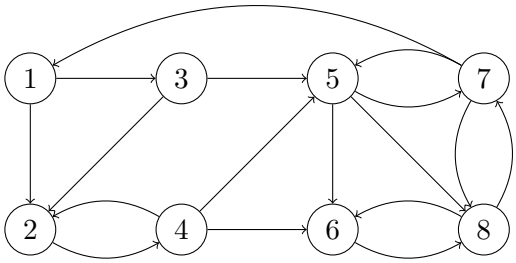
$$H_{i,j} = \begin{cases} \frac{1}{\ell_j} & \text{falls } P_j \text{ auf } P_i \text{ verlinkt,} \\ 0 & \text{sonst.} \end{cases}$$

Fassen wir jetzt noch I als Spaltenvektor auf, so können wir unser Problem umformen in eine Matrixgleichung

$$HI = I \ .$$

[1] Mehr Details kann man in [6] finden

[2] Offizieller Name ist "PageRank Algorithmus". Das Wort "Page" kommt allerdings nicht vom englischen Wort für Seite, sondern vom Ko-Gründer von Google Larry Page.



Das Problem hier ist, daß I auf beiden Seiten der Gleichung vorkommt. Solche Vektoren und wie wir diese finden können sind Inhalt dieses Kapitels.

Ein konkretes Beispiel: nehmen wir an das Internet hat nur 8 Seiten, welche folgendermaßen aufeinander verweisen.

Das entsprechende H ist dann

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 1 & \frac{1}{3} & 0 \end{pmatrix} .$$

Wie man nachrechnet ist dann ein mögliches I

$$I = \begin{pmatrix} 0.0600 \\ 0.0675 \\ 0.0300 \\ 0.0675 \\ 0.0975 \\ 0.2025 \\ 0.1800 \\ 0.2950 \end{pmatrix} .$$

D.h. Seite 8 ist die anscheinend relevanteste und 3 die am am wenigsten relevante Seite.

7.2 Eigenwerte und Eigenvektoren

Definition 7.2.1. Ein Element $k \in K$ heißt **Eigenwert** von $\varphi : V \to W$, falls es einen Vektor $v \neq 0$ gibt, so daß

$$\varphi(v) = kv \ .$$

Ein Vektor $v \in V \setminus \{0\}$ heißt **Eigenvektor** von φ zum Eigenwert k, falls

$$\varphi(v) = kv \ .$$

Sprechen wir von Eigenvektoren und -werten einer Matrix A, so bezieht sich das natürlich auf die Funktion φ_A.

Beispiel 7.2.2. *1. Die Einheitsmatrix hat den Eigenwert* 1*. Und jeder nicht-triviale Vektor ist Eigenvektor.*

2. Die Nullmatrix hat den Eigenwert 0*. Und jeder nicht-triviale Vektor ist Eigenvektor.*

3. Ist $\mathrm{Ke}(\varphi_A) \neq \{0\}$*, so ist* 0 *Eigenwert.*

*4. Ist A eine **Diagonalmatrix**, also*

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & \vdots \\ & \vdots & \ddots & & \\ 0 & \dots & 0 & a_{(n-1)(n-1)} & 0 \\ 0 & \dots & & 0 & a_{nn} \end{pmatrix} ,$$

so hat A die Eigenvektoren $e_1, \dots, e_n$ jeweils zu den Eigenwerten $a_{11}, \dots, a_{nn}$

5. Eine Abbildung hat nicht unbedingt Eigenvektoren bzw. -werte: beispielsweise ist die Multiplikation mit der Matrix

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

gleich der Rotation um den Winkel α gegen den Uhrzeigersinn des $\mathbb{R}^2$ um den Ursprung. Somit ist leicht einzusehen, daß kein Vektor auf ein Vielfaches seiner selbst abgebildet wird, es sei denn α ist ein Vielfaches von π.[3]

[3]Formal: nehmen wir an es gibt $a, b, k \in \mathbb{R}$ mit

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = k \begin{pmatrix} a \\ b \end{pmatrix} .$$

D.h. $a\cos(\alpha) - b\sin(\alpha) = ka$ und $a\sin(\alpha) + b\cos(\alpha) = kb$. Äquivalent lassen sich diese Gleichungen umformen zu $-b\sin(\alpha) = a(k - \cos(\alpha))$ und $a\sin(\alpha) = b(k - \cos(\alpha))$. Damit (a, b) Eigenvektor ist darf es nicht gleich dem Nullvektor sein; d.h. entweder $a \neq 0$ oder $b \neq 0$. Betrachten wir den Fall $a \neq 0$, der Fall $b \neq 0$ ist analog. Dann können wir die erste Gleichung durch a teilen und in die zweite einsetzen und erhalten:

$$\frac{-b^2}{a}\sin(\alpha) = a\sin(\alpha) \ .$$

Wenn α kein Vielfaches von π ist, dann ist $\sin(\alpha) \neq 0$ und wir erhalten $-b^2 = a^2$. Da es sich um reelle Zahlen handelt ist dies aber ein Widerspruch.

6. *Die Matrix*

$$\begin{pmatrix} -1 & 6 \\ 0 & 5 \end{pmatrix}$$

hat Eigenwerte -1 *und* 5 *jeweils mit Eigenvektor* $(1,0)$ *und* $(1,1)$.

Satz 7.2.3. *Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.*

Beweis. Sei $\varphi : V \to W$ eine lineare Abbildung mit $v_1, \ldots, v_m$ Eigenvektoren zu den Eigenwerten $k_1, \ldots, k_n$, wobei die k_i paarweise unterschiedlich sind. Wir beweisen die Aussage mit Induktion über m. Der Induktionsanfang ist klar, da ein Eigenvektor nicht trivial sein darf. Für den Induktionsschritt nehmen wir an, wir haben die Aussage schon für $m-1$ Vektoren gezeigt.

Sei nun $h_1, \ldots, h_m$ so daß

$$0 = h_1 v_1 + \cdots + h_m v_m \ .$$

Wenden wir φ auf diese Gleichung an und verwenden die Linearität, so erhalten wir

$$0 = h_1 \varphi(v_1) + \cdots + h_m \varphi(v_m) \ .$$

Da $v_1, \ldots, v_m$ Eigenvektoren sind, ist dann aber

$$0 = h_1 k_1 v_1 + \cdots + h_m k_m v_m \ .$$

Nun kommt ein wichtiger Trick: wir ziehen diese Gleichung von der ersten multipliziert mit k_m ab und erhalten nach Ausklammern:

$$0 = h_1 (k_1 - k_m) v_1 + \cdots + h_m (k_m - k_m) v_m \ .$$

Da der letzte Term wegfällt und die Vektoren nach Induktionsvoraussetzung linear unabhängig sind gilt $h_i(k_i - k_m) = 0$ für alle $1 \leqslant m-1$. Da aber $k_i \neq k_m$ ist muss $h_i = 0$ sein für alle $1 \leqslant m-1$. Setzen wir dies wiederum in die erste Gleichung ein, so bleibt nur noch $h_m v_m = 0$. Wie im Induktionsanfang also auch $h_m = 0$.

□

Definition 7.2.4. Eine lineare Abbildung $\varphi : V \to V$ heißt **diagonalisierbar** falls V eine Basis aus Eigenvektoren von φ besitzt.

Woher kommt der Name diagonalisierbar? Betrachten wir einen endlich-dimensionalen Vektorraum und sei B eine Basis aus Eigenvektoren. Wie man sieht ist dann die darstellende Matrix zu dieser Basis, also $M(\varphi, B, B)$ eine Diagonalmatrix, mit den Eigenwerten auf der Diagonale.

Satz 7.2.5. *Eine Matrix A ist genau dann diagonalisierbar, wenn es eine invertierbare Matrix S gibt, so daß $S^{-1}AS$ eine Diagonalmatrix ist.*[4]

Satz 7.2.6. *Besitzt eine $n \times n$ Matrix n verschieden Eigenwerte, so ist sie diagonalisierbar.*

Beweis. Besitzt $A \in M(n \times n, K)$ n verschieden Eigenwerte, so gibt es nach Definition zu den Eigenwerten auch Eigenvektoren $v_1, \ldots, v_n \in K^n$. Nach Satz 7.2.3 sind Eigenvektoren zu verschiedenen Eigenwerten aber linear unabhängig, bilden also nach Korollar 4.5.9 eine Basis. □

Alles wunderbar, aber wie können wir Eigenwerte und Eigenvektoren finden?

Satz 7.2.7. *Ist k ein Eigenwert einer Matrix A, so ist v genau dann Eigenvektor zu k, wenn v Lösung von $(A - \lambda I)x = 0$ ist.*
D.h. um Eigenvektoren zu einem Eigenwert zu finden müssen wir nur ein homogenes lineares GS lösen.

Satz 7.2.8. *Ist A eine quadratische Matrix, so ist k genau dann Eigenwert, falls*

$$\det(A - kI) = 0$$

gilt.

Beweis. Eigenwerte/Eigenvektoren erfüllen die Gleichung $Av = kv$, oder äquivalent $Av - kI_n v = 0$. Dies wiederum ist äquivalent zu $(A - kI_n)v = 0$. D.h. ist v Eigenvektor zum Eigenwert k, so ist $A - kI_n$ nicht invertierbar (denn es existiert ja eine nicht-triviale Lösung, nämlich v). Also ist $\det(A - kI_n) = 0$.
Die umgekehrte Argumentation ist ähnlich: ist $\det(A - kI_n) = 0$, so ist $A - kI_n$ nicht invertierbar, und darum hat $(A - kI_n)x = 0$ eine nicht-triviale Lösung, welche ein Eigenvektor ist. □

Um Eigenwerte zu finden, suchen wir die Nullstellen der Funktion $k \mapsto \det(A - kI)$, die, wie man leicht sieht, ein Polynom ist.

Definition 7.2.9. Der Ausdruck $\chi_A = \det(A - XI) \in K[X]$ heißt das **charakteristische Polynom** von A.

Korollar 7.2.10. *Die Eigenwerte von A sind genau die Nullstellen des charakteristischen Polynoms von A.*

Lemma 7.2.11. *Ist A eine $n \times n$-Matrix, so ist χ_A ein Polynom und hat Grad n.*

Korollar 7.2.12. *Ist A eine $n \times n$-Matrix, so hat A höchstens n verschiedene Eigenwerte.*

[4] Man beachte, daß es für jede Matrix A Matrizen invertierbare S, T gibt, so daß SAT Diagonalmatrix ist.

Bemerkung 7.2.13. *Ist D Diagonalmatrix, also*

$$D = \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ & \vdots & \ddots & \\ 0 & 0 & \dots & d_{nn} \end{pmatrix} \quad \text{so ist} \quad D^m = \begin{pmatrix} d_{11}^m & 0 & \dots & 0 \\ 0 & d_{22}^m & \dots & 0 \\ & \vdots & \ddots & \\ 0 & 0 & \dots & d_{nn}^m \end{pmatrix} .$$

Ist A diagonalisierbar, so gibt es, wie wir gesehen haben, eine invertierbare Matrix S und eine Diagonalmatrix D mit $S^{-1}AS = D$, oder äquivalent $A = SDS^{-1}$. Damit lassen sich aber die Potenzen von A leicht berechnen:

$$A^n = SDS^{-1}SDS^{-1}\dots SDS^{-1} = SD^nS^{-1} .$$

Beispiel 7.2.14. *Aus DMI erinnern Sie sich vielleicht ja noch an die Fibonacci Zahlen. Hierbei ist $F_0 = F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5$ usw. D.h. $F_{n+2} = F_{n+1} + F_n$. Wie kann man diese effizient berechnen? Wie man leicht sieht ist*

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{n+1} \\ F_n + F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{n+1} \\ F_{n+2} \end{pmatrix}$$

Iterieren wir dies, so erhalten wir

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Ist $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ diagonalisierbar, so ist lässt sich mit obiger Bemerkung eine einfache Formel für F_n finden.
Das charakteristische Polynom dieser Matrix ist: $-X(1-X)-1 = X^2 - X - 1$. Dies hat die Nullstellen $\lambda_1, \lambda_2 = \frac{1 \pm \sqrt{5}}{2}$. Man beachte, daß

$$\lambda_1\lambda_2 = -1 \quad \text{und} \quad \lambda_1 + \lambda_2 = 1 .$$

Löst man die entsprechenden linearen Gleichungssysteme, so erhält man die zugehörigen Eigenvektoren:

$$v_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} .$$

Sei jetzt $T = (v_1 v_2)$. Die Inverse Matrix lässt sich (ausnahmsweise am einfachsten) mit Korollar 6.0.19 berechnen. Wir erhalten

$$T^{-1} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix} .$$

Insgesamt gilt also

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Also

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1^n\lambda_2 & -\lambda_1^n \\ -\lambda_1\lambda_2^n & \lambda_2^n \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

und

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_1^n\lambda_2 - \lambda_1\lambda_2^n & \lambda_1^n - \lambda_2^n \\ * & * \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Die Einträge $*$ sind für uns uninteressant. Insgesamt:

$$\begin{aligned} F_n &= \frac{1}{\lambda_2 - \lambda_1}(\lambda_1^n\lambda_2 - \lambda_1\lambda_2^n + \lambda_1^n - \lambda_2^n) \\ &= \frac{1}{\lambda_2 - \lambda_1}(\lambda_1^n(\lambda_2 - 1) + \lambda_2^n(\lambda_1 - 1)) \\ &= \frac{1}{\lambda_2 - \lambda_1}(-\lambda_1^{n+1} + \lambda_2^{n+1}) \\ &= \frac{\lambda_2^{n+1} - \lambda_1^{n+1}}{\lambda_2 - \lambda_1} = \frac{\lambda_2^{n+1} - \lambda_1^{n+1}}{\sqrt{5}} \end{aligned}$$

7.3 Nochmal der Google Algorithmus

Wie wir gesehen haben, läuft die Suche nach einem objektiven Ranking von Webseiten darauf hinaus eine lineare Gleichung der Form

$$Hx = x$$

zu lösen. Mit der Terminologie des letzten Kapitels heißt das wir suchen zur "Suchmatrix" einen Eigenvektor mit Eigenwert 1. Ohne näher auf Details einzugehen, hat jede Suchmatrix (nur positive Werte, Summe der Spalten gleich 1) immer den größten Eigenwert 1.

Allerdings löst man, um x zu finden, in der Praxis nicht das Gleichungssystem $(A - I_n)x = 0$, da das zu rechenaufwendig wäre. Besser ist schon folgende Einsicht, die wir aber nicht beweisen wollen. Fangen wir mit einem zufälligen Vektor v and, und berechnen nach und nach $A^n v$ für immer größer werdende n, so konvergiert die Folge (meistens) gegen den Eigenvektor mit dem grössten Eigenwert. So wie die Suchmatrix aber definiert ist, ist 1 der grösste Eigenwert! Da A sehr viele (fast nur) Nullen enthält ist die Berechnung von A^n relativ einfach. Spekulationen von Insidern lassen vermuten, daß bereits für n zwischen 50 und 100 sich bereits gute Annäherungen an x finden lassen. Mehr Spekulationen lauten, daß Google ca. jeden Monat den Vektor x

neuberechnet, allerdings auch kleinere Änderungen für Nachrichten usw. zulässt. Die Details sind natürlich Betriebsgeheimnis.

8 Euklidische Vektorräume

8.1 Skalarprodukte

Aspekte, die sich in unserem abstrakten Raumbegriff bis jetzt nicht wiederfinden sind Längen und Winkel. Hierzu benötigen wir ein Konstrukt, welches zunächst wenig mit diesen Begriffen zu tun hat.

Definition 8.1.1. Sei V ein reeller Vektorraum. Ein **Skalarprodukt** in V ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ mit den folgenden Eigenschaften

(i) Bilinearität: Für jedes $v \in V$ sind die Abbildungen $\langle v, \cdot \rangle$ und $\langle \cdot, v \rangle$ linear

(ii) Symmetrie: $\langle v, w \rangle = \langle w, v \rangle$ für alle $v, w \in V$.

(iii) Positive Definitheit: $\langle v, v \rangle > 0$ für alle $v \neq 0$.

Das Skalarprodukt ($V \times V \to \mathbb{R}$) ist nicht mit der Skalarmultiplikation ($V \times \mathbb{R} \to V$) zu verwechseln!

Definition 8.1.2. Ein **Euklidischer Vektorraum** ist ein reeller Vektorraum V zusammen mit einem Skalarprodukt auf V.

Beispiel 8.1.3. *Die Abbildung* $\langle x, y \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *definiert durch*

$$\langle (x_1, \ldots, x_n), (v_1, \ldots, v_n) \rangle = x_1 y_1 + \cdots + x_n y_n$$

ist ein Skalarprodukt. In diesem Fall sprechen wir von dem üblichen bzw. dem ***Standard-Skalarprodukt****.*
Dies ist allerdings nicht das einzige Skalarprodukt für $\mathbb{R}^n$*: ist* $A \in M(n \times n, \mathbb{R})$ *invertierbar, so ist auch* $(x, y) \mapsto \langle Ax, Ay \rangle$ *ein Skalarprodukt.*

Ein weiteres wichtiges Beispiel ist das folgende:

Beispiel 8.1.4. *Sei V der Vektorraum der stetigen Funktionen von $[-1,1]$ nach $\mathbb{R}$. Die Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ definiert durch*

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$$

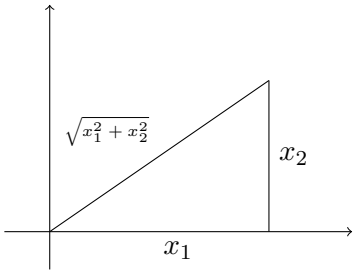
ist ein Skalarprodukt.

Was hat das Ganze jetzt mit Längen zu tun?

Definition 8.1.5. Ist $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und $x \in V$, so ist die **Norm** von x definiert durch

$$\|x\| = \sqrt{\langle x, x\rangle}\ .$$

Man beachte, daß im $\mathbb{R}^2$ mit dem Standard-Skalarprodukt die Norm $\|x\|$ eines Punktes $x = (x_1, x_2)$ genau dem Abstand zum Ursprung entspricht: $\sqrt{x_1^2 + x_2^2}$. Das Gleiche gilt für den $\mathbb{R}^3$.



Satz 8.1.6 (**Cauchy-Schwarz'sche Ungleichung**)**.** *In jedem euklidischen Vektorraum $(V, \langle\cdot,\cdot\rangle)$ gilt*

$$\langle x, y\rangle \leqslant \|x\|\|y\|$$

für $x, y \in V$.

Satz 8.1.7. *Ist V ein euklidischer Vektorraum, so hat die Norm folgende Eigenschaften:*

(i) $\|x\| \geqslant 0$ *für alle x.*

(ii) $\|x\| = 0 \iff x = 0$.

(iii) $\|kx\| = |k|\|x\|$ *für $k \in \mathbb{R}$ und $x \in V$.*

(iv) ***(Dreiecksungleichung)*** $\|x + y\| \leqslant \|x\| + \|y\|$ *für alle $x, y \in V$.*

Man beachte, daß, wegen der Cauchy-Schwarz'schen Ungleichung gilt, daß

$$-1 \leqslant \frac{\langle x, y\rangle}{\|x\|\|y\|} \leqslant 1\ .$$

Dies erlaubt uns die folgende Definition:

Definition 8.1.8. Sind $x, y \in V$ nicht null und V euklidischer Vektorraum, so ist der **Öffnungswinkel** $\alpha(x, y)$ zwischen x und y definiert durch

$$\cos(\alpha(x, y)) = \frac{\langle x, y\rangle}{\|x\|\|y\|}$$

Wie man sich leicht überlegt entspricht dies im $\mathbb{R}^2$ und $\mathbb{R}^3$ genau dem normalen Winkelbegriffs.

8.2 Orthogonalität

Definition 8.2.1. Zwei Elemente v, w eines euklidischen Vektorraumes heißen **orthogonal** oder **senkrecht zueinander**, wenn

$$\langle v, w\rangle = 0 \ .$$

In diesem Fall schreiben wir $v \perp w$.

Man beachte, daß für nicht-trivale v, w gilt, daß $v \perp w \iff \cos(\alpha(v, q)) = 0$.

Definition 8.2.2. Ist M eine Teilmenge des euklidischen Vektorraumes V, so heißt

$$M^\perp = \{\, v \in V \mid \forall u \in M : \langle u, v\rangle = 0 \,\}$$

das **orthogonale Komplement**.

Satz 8.2.3. *$M^\perp$ ist ein Untervektorraum.*

Definition 8.2.4. Eine Menge von Vektoren X heißen **Orthonormalsystem**, falls

- $\|v\| = 1$ für alle $v \in X$ und
- $\langle v, w\rangle = 0$ für $v, w \in X$ mit $v \neq w$.

Lemma 8.2.5. *Ein Orthonormalsystem ist linear unabhängig.*

Lemma 8.2.6. *Ist $v_1, \ldots, v_n$ ein Orthonormalsystem und U der von diesen Vektoren erzeugte Unterraum, so gibt es eindeutig bestimmte u, w mit $u \in U$ und $w \in U^\perp$ so daß*

$$u = \sum_{i=1}^{n} \langle v, v_i\rangle v_i \ .$$

Insbesondere lässt sich jeder Vektor $v \in V$ schreiben als

$$v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i \ ,$$

wenn $v_1, \ldots, v_n$ Basis von V sind. D.h. um die Koordinaten von v zur Basis $v_1, \ldots, v_n$ zu bestimmen müssen wir kein lineares Gleichungssystem lösen sondern nur Skalarprodukte ausrechnen!

8.3 Anwendungen

Definition 8.3.1. Sei V ein euklidischer, vollständiger, Vektorraum. $v_1, v_2, \ldots$ sind eine **Hilbert-Basis**, falls

1. $v_1, v_2, \ldots$ sind ein Orthonormalsystem
2. $\langle v_1, v_2, \ldots \rangle$ ist eine dichte[1] Teilmenge von V.

Satz 8.3.2. *Sei V und $v_1, v_2, \ldots$ wie oben. Für beliebige $v \in V$ ist dann*

$$v = \sum_{n=0}^{\infty} \langle v, v_n \rangle v_n \ .$$

Man beachte, daß die unendliche Summe im Sinne der Norm gemeint ist.

Beispiel 8.3.3. *Für $L^2 = \left\{ f : [0, 2\pi] \to \mathbb{R} \mid \int_0^{2\pi} f(x)f(x)dx < \infty \right\}$ mit dem Skalarprodukt $\langle f, g \rangle = \int_0^{2\pi} fgdx$ sind*

$$c_0 = \frac{1}{2\pi}, \quad s_n = \frac{1}{\pi}\sin(nx), \quad c_n = \frac{1}{\pi}\cos(nx)$$

eine Hilbertbasis.

D.h. die meisten Funktionen $f : [0, 2\pi] \to \mathbb{R}$ lassen sich in Sinus/Kosinus-Wellen zerlegen. Mehr noch die Koeffizienten lassen sich durch Integration bestimmen! Diese Technik hat eine Vielzahl von Anwendungen. Hier nur ein paar Beispiele aus der verlustbehafteten Datenkompression:

Beispiel 8.3.4 (mp3)**.** *Die Zerlegung ist der erste Schritt in den meisten verlust Kompressionsverfahren von Audio-dateien. (Alles über $20khz$ ist für das menschliche Ohr nicht hörbar, ähnliche Frequenzen lassen sich zusammenfassen, usw.)*

Beispiel 8.3.5 (jpeg)**.** *Auch Bilddaten können auf diese Art und Weise komprimiert werden. Dazu fassen wir eine Bitmap als Welle auf und zerlegen diese wieder in Sinus/Kosinus-Wellen. Höhere Frequenzen können wir nun weglassen.* [2]

Beispiel 8.3.6 (jpeg2000)**.** *Für Bilddateien ist die Verwendung der sin/cos-Hilbertbasis nicht so effizient wie bei Audiodateien, die ja schon natürliche Wellen sind. Anders als die meisten Audiostücke haben die meisten (Photo-) Bilder scharfe Kanten und plötzliche Farbübergänge. Verwendung finden hier sogenannte Wavelets, daß sind Basisfunktionen, die eher wie Zacken aussehen (es gibt je nach Anwendungen eine Vielzahl dieser Wavelets) und damit besser "kantige" Funktionen approximieren können.*

[1] Dies ist eine Definition aus der Analysis, auf die wir hier nicht weiter eingehen wollen.

[2] Jpeg verwendet eine diskrete Variante, die "Diskrete Kosinus Transformation".

Index