Lineare Algebra I

H. Lenzing: WS 2003/04

Inhaltsverzeichnis

Kapitel 1

Anschauliche Vektorrechnung

Montag, 13. Oktober 03

Einordnung

Dieses erste Kapitel hat motivierenden Charakter. Es führt — an die geometrische Anschauung anknüpfend — die zentralen Begriffe der Linearen Algebra unmittelbar und ohne pedantische Vorreden ein, was uns erlaubt, ohne Umwege mit diesen Konzepten zu arbeiten. Ein gewisser Nachteil dieses Vorgehens ist die — nach heutigem Standard — mangelnde begriffliche Präzision. Es sei daher — dieses erste Kapitel betreffend — an Sie appelliert, sich durch diese Unschärfe nicht irritieren zu lassen und den vorgestellten Betrachtungen mit Begeisterung zu folgen.

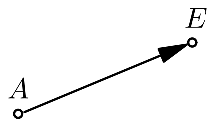
Eine saubere begriffliche Fundierung, die heute in den meisten Darstellungen der Linearen Algebra am Anfang steht, fußt auf einer axiomatischen Einführung des Vektorraumbegriffs. Trotz der begrifflichen Schärfe, die mit diesem Vorgehen verbunden ist, haben Anfänger/innen — bedingt durch den hohen Abstraktionsgrad — erfahrungsgemäß Schwierigkeiten, sich bei einer solchen Art des Vorgehens ein inhaltliches Verständnis zu erwerben. Ein weiteres Problem liegt darin, dass bei einem axiomatischen Vorgehen der Zusammenhang zur übrigen mathematischen Erfahrung nur schwer herzustellen ist und daher die Vorgabe eines solchen Axiomensystems (zunächst jedenfalls) recht künstlich erscheint.

Es ist daher eine Aufgabe dieses Vorspanns, für genügend Anschauungsmaterial zu sorgen und die für die Lineare Algebra grundlegenden Begriffsbildungen des **Vektorraum**s und der **linearen Abbildung** herzuleiten. Die zunächst vorhandene begriffliche Unschärfe nehmen wir dabei in Kauf; sie wird durch die am Wege mitgenommenen Einsichten mehr als kompensiert.

1.1 Vektoren und ihre Länge

Wir arbeiten im **3-dimensionalen Anschauungsraum**, den wir nicht weiter erklären wollen (und können). Die erwähnte begriffliche Unschärfe hat ihre Wurzel im hier notwendigen Rückgriff auf die Anschauung.

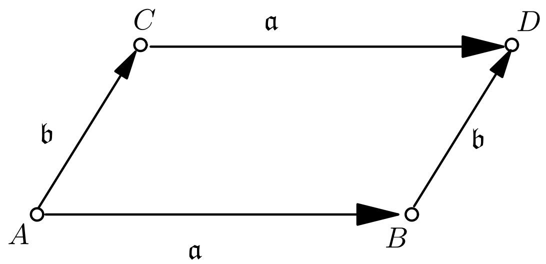
Ein **geordnetes Punktepaar** (A, E) definiert einen **Vektor** $\overrightarrow{AE}$



Vektor, dargestellt als gerichtete Strecke

A nennen wir den **Anfangspunkt**, B den **Endpunkt** der gerichteten Strecke von A nach B.

Hinsichtlich der **Gleichheit von Vektoren** treffen wir die Vereinbarung, dass zwei Vektoren $\mathfrak{a}$ und $\mathfrak{b}$ übereinstimmen, wenn sie durch eine Parallelverschiebung auseinander hervorgehen[1]. Ist somit



Gleichheit von Vektoren

ein Parallelogramm, so ist $\overrightarrow{AC} = \overrightarrow{BD}$ und natürlich auch $\overrightarrow{AB} = \overrightarrow{CD}$.

Es ist üblich, Vektoren mit kleinen deutschen Buchstaben zu bezeichnen

$$\mathfrak{a} = \overrightarrow{AB} = \overrightarrow{CD}$$

$$\mathfrak{b} = \overrightarrow{AC} = \overrightarrow{BD}.$$

Wir können — wie in der Physik üblich — auch die Schreibweise

$$\overrightarrow{a} = \overrightarrow{AB}, \quad \overrightarrow{b} = \overrightarrow{AC}$$

usf. verwenden.

Fassen wir zusammen: *Ein Vektor ist die Gesamtheit (Menge) aller Strecken von gleicher Richtung und gleicher Länge.*

[1]Diese Vereinbarung zieht nach sich, dass wir — streng genommen — nicht mehr vom Anfangspunkt bzw. Endpunkt eines Vektors sprechen können. Gleichwohl werden wir diese Redeweise im folgenden verwenden. Gemeint sind dann Anfangs- und Endpunkt einer gerichteten Strecke, die wir zur Repräsentation des Vektors gewählt haben.

Die **Länge eines Vektors** $\mathfrak{a} = \overrightarrow{AB}$, d.h. den Abstand der beiden Punkte A, B bezeichnen wir mit

$$|\mathfrak{a}|$$

und nennen diese Zahl ≥ 0 den **Betrag von $\mathfrak{a}$.**

Falls $A = B$ nennen wir $\overrightarrow{AB}$ den **Nullvektor**. Schreibweise

$$\mathfrak{o} = \overrightarrow{AA}.$$

Natürlich ist $\mathfrak{a} = \mathfrak{o}$ genau dann, wenn $|\mathfrak{a}| = 0$.

Wir kürzen diesen Sachverhalt wie folgt ab:

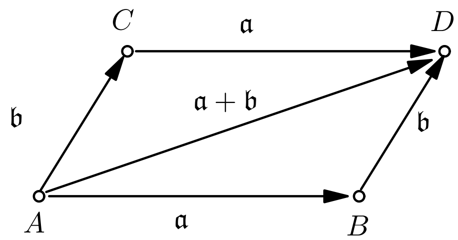
$$\mathfrak{a} = \mathfrak{o} \quad \Leftrightarrow \quad |\mathfrak{a}| = 0.$$

Wir beachten, dass hier zwischen der Zahl 0 und dem Vektor $\mathfrak{o}$ deutlich zu unterscheiden[2] ist.

[2] Später werden wir diese bezeichnungsmäßige Unterscheidung zwischen Nullvektor $\mathfrak{o}$ und Zahl 0 als zu pedantisch aufgeben, aber natürlich an der begrifflichen Verschiedenheit festhalten.

1.2 Die Addition von Vektoren

Der Vektor $\mathfrak{a} + \mathfrak{b}$ wird gebildet, indem der Anfangspunkt von $\mathfrak{b}$ an den Endpunkt von $\mathfrak{a}$ gefügt wird.



Addition von zwei Vektoren

Ist also $\mathfrak{a} = \overrightarrow{AB}$, $\mathfrak{b} = \overrightarrow{BC}$, dann ist $\mathfrak{a} + \mathfrak{b} = \overrightarrow{AC}$.

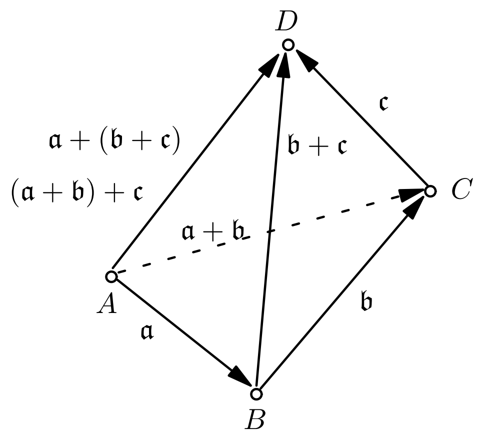
Es ist aber auch $\mathfrak{b} = \overrightarrow{AD}$, $\mathfrak{a} = \overrightarrow{DC}$ und folglich $\mathfrak{b} + \mathfrak{a} = \overrightarrow{AD} + \overrightarrow{DC} = \overrightarrow{AC}$.

Damit haben wir die **Kommutativität**

(*A* 1) $$\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a} \quad \text{für alle } \mathfrak{a}, \mathfrak{b}$$

der Vektoraddition nachgewiesen.

Um die Summe von drei Vektoren $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ zu bilden, betrachten wir die folgende Figur:



Assoziativität der Vektoraddition

Hieraus erhalten wir sofort die **Assoziativität** der Vektoraddition

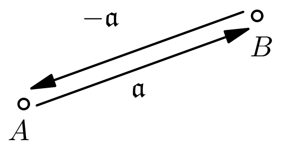
(*A* 2) $$(\mathfrak{a} + \mathfrak{b}) + \mathfrak{c} = \mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) \quad \text{für alle } \mathfrak{a}, \mathfrak{b}, \mathfrak{c}.$$

Ist $\mathfrak{a} = \overrightarrow{AB}$, so ist $\mathfrak{o} = \overrightarrow{AA} = \overrightarrow{BB}$. Es folgt $\mathfrak{a} + \mathfrak{o} = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} = \mathfrak{a}$, entsprechend (A 1) auch $\mathfrak{o} + \mathfrak{a} = \mathfrak{a}$. Folglich gilt

$$(A\ 3) \qquad \mathfrak{a} + \mathfrak{o} = \mathfrak{o} + \mathfrak{a} = \mathfrak{a} \quad \text{für alle } \mathfrak{a}.$$

Donnerstag, 16. Oktober 03

Erklären wir für $\mathfrak{a} = \overrightarrow{AB}$ den Vektor $-\mathfrak{a}$ als $\overrightarrow{BA}$



Negativer Vektor

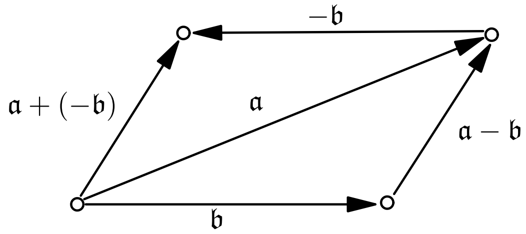
so erhalten wir $\mathfrak{a} + (-\mathfrak{a}) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \mathfrak{o}$. Somit unter erneuter Beachtung von (A 1) als charakterisierende **Eigenschaft des additiven Inversen**

$$(A\ 4) \qquad \mathfrak{a} + (-\mathfrak{a}) = \mathfrak{o} = (-\mathfrak{a}) + \mathfrak{a} \quad \text{für alle } \mathfrak{a}.$$

Schließlich erklären wir die **Differenz** $\mathfrak{a} - \mathfrak{b}$ von Vektoren $\mathfrak{a}$, $\mathfrak{b}$ durch

$$\mathfrak{a} - \mathfrak{b} = \mathfrak{a} + (-\mathfrak{b}).$$

Im Bild:



Differenz zweier Vektoren

Merkregel. $\mathfrak{a} - \mathfrak{b}$ *ist der Vektor, der vom Endpunkt von* $\mathfrak{b}$ *zum Endpunkt von* $\mathfrak{a}$ *weist, falls* $\mathfrak{a}$ *und* $\mathfrak{b}$ *so gelegt sind, dass ihre Anfangspunkte zusammenfallen.*

Es gilt natürlich

$$(\mathfrak{a} - \mathfrak{b}) + \mathfrak{b} = \mathfrak{a}.$$

Sie können dies direkt am obigen Bild ablesen, aber auch aus den bisher schon gewonnenen Rechenregeln herleiten:

$$(\mathfrak{a} - \mathfrak{b}) + \mathfrak{b} \overset{\text{nach Definition}}{=} (\mathfrak{a} + (-\mathfrak{b})) + \mathfrak{b} \overset{(A\ 2)}{=} \mathfrak{a} + ((-\mathfrak{b}) + \mathfrak{b}) \overset{\text{nach Definition}}{=} \mathfrak{a} + \mathfrak{o} \overset{(A\ 3)}{=} \mathfrak{a}.$$

Es handelt sich bei der Subtraktion von Vektoren somit tatsächlich um die Umkehrung der Addition.

1.3 Multiplikation eines Vektors mit einem Skalar

Ist a eine reelle Zahl (**Skalar**) und ist $a > 0$, so verstehen wir unter $a\mathfrak{a}$ denjenigen Vektor, der die gleiche Richtung hat wie $\mathfrak{a}$, aber a-mal so lang ist: $|a\mathfrak{a}| = a|\mathfrak{a}|$.

Wir setzen $0\mathfrak{a} = \mathfrak{o}$ und erklären für $a > 0$ $(-a)\mathfrak{a}$ als denjenigen Vektor, der die entgegengesetzte Richtung von $\mathfrak{a}$, aber seine a-fache Länge hat: $|(-a)\mathfrak{a}| = a|\mathfrak{a}|$.

Allgemein haben wir daher

$$|a\mathfrak{a}| = |a||\mathfrak{a}| \tag{3.1}$$

falls a ein Skalar und $\mathfrak{a}$ ein Vektor ist.

Hierbei ist

$$|a| = \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases} \tag{3.2}$$

der Abstand der reellen Zahl a vom Nullpunkt auf der reellen Zahlengeraden, der sog. **Absolutbetrag von** a.

Man verifiziert leicht die folgenden Rechenregeln für die Multiplikation von Vektoren mit Skalaren:

$$\begin{array}{lrcl} (M\ 1) & a(\mathfrak{a}+\mathfrak{b}) & = & a\mathfrak{a}+a\mathfrak{b} \\ (M\ 2) & (a+b)\mathfrak{a} & = & a\mathfrak{a}+b\mathfrak{a} \\ (M\ 3) & a(b\mathfrak{a}) & = & (ab)\mathfrak{a} \\ (M\ 4) & 1\mathfrak{a} & = & \mathfrak{a}. \end{array}$$

Verabredung: Anstelle von $\frac{1}{a}\mathfrak{a}$ (für einen Skalar $a \neq 0$) schreiben wir häufig $\frac{\mathfrak{a}}{a}$. Insbesondere erhalten wir mit $a = |\mathfrak{a}|$, falls $\mathfrak{a} \neq \mathfrak{o}$, mittels (3.1), dass $\frac{a}{\mathfrak{a}}$ die Länge Eins hat. Für jeden Vektor $\mathfrak{a} \neq \mathfrak{o}$ ist daher

$$\frac{\mathfrak{a}}{|\mathfrak{a}|}$$

stets ein **Einheitsvektor**, ein Vektor der Länge Eins, der zudem die selbe Richtung hat wie $\mathfrak{a}$. Den Übergang von $\mathfrak{a} \neq \mathfrak{o}$ zum Einheitsvektor $\frac{\mathfrak{a}}{|\mathfrak{a}|}$ nennt man **normieren**.

1.4 Die Axiome eines reellen Vektorraums

Aus den bisherigen Betrachtungen lässt sich leicht ein Axiomensystem für den grundlegenden Begriff des **reellen Vektorraums** gewinnen. $\mathbb{R}$ bezeichnet hier und im folgenden die **Menge der reellen Zahlen**[3].

Definition 4.1 *Eine Menge V zusammen mit zwei Zuordnungsvorschriften*

$$\begin{aligned} + &: V \times V \longrightarrow V, \quad (v,w) \mapsto v+w \\ \cdot &: \mathbb{R} \times V \longrightarrow V, \quad (a,v) \mapsto a\,v \end{aligned}$$

heißt ein reeller Vektorraum*, wenn gilt:*

(A 1) $v + w = w + v$ für alle $v, w \in V$.

(A 2) $u + (v + w) = (u + w) + v$ für alle $u, v, w \in V$.

(A 3) Es gibt $0 \in V$ mit $v + 0 = v = 0 + v$ für alle $v \in V$.

(A 4) Zu jedem $v \in V$ gibt es $w \in V$ mit $v + w = 0 = w + v$.

(M 1) $a(v + w) = av + aw$ für alle $a \in \mathbb{R}$, $v, w \in V$.

(M 2) $(a + b)v = av + bv$ für alle $a, b \in \mathbb{R}$, $v \in V$.

(M 3) $(a \cdot b)v = a(bv)$ für alle $a, b \in \mathbb{R}$, $v \in V$.

(M 4) $1v = v$ für alle $v \in V$.

Die Elemente v eines Vektorraums V nennen wir dann **Vektoren**, die Elemente von $\mathbb{R}$ **Skalare**. Dabei — und dies macht dieses Vorgehen so abstrakt — spielt es keine Rolle mehr, welche Natur die Elemente v des Vektorraums V haben. Ein Vektor — für sich genommen — hat auf diesem Definitionsniveau keine Existenz[4]: ein Element v wird — anders als in der anschaulichen Vektorrechnung — dann erst durch seine Mitgliedschaft in einem Vektorraum zu einem Vektor.

Wer unsere vorangehenden Betrachtungen mit Aufmerksamkeit verfolgt hat, wird in der obigen Definition das Auftreten der Länge von Vektoren vermissen. Dieses Manko werden wir später beseitigen. Lassen wir dieses Phänomen einmal beiseite, können wir sagen, dass eine Aussage über Vektoren genau dann gilt, wenn sie sich mittels gültiger Schlussregeln der Logik aus den Axiomen (A 1) – (A 4) und (M 1) – (M 4) herleiten lässt.

---

[3]Es hat sich in der mathematischen Grundausbildung eingebürgert, die Einführung und genauere Behandlung der reellen Zahlen der Analysis zu überlassen

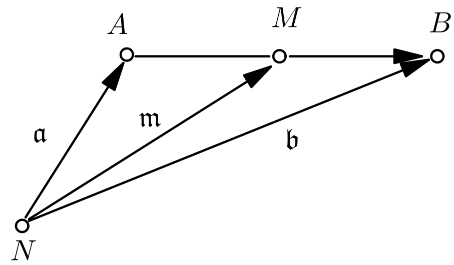
[4]Dies ist ein wesentlicher Unterschied zur anschaulichen Vektorrechnung.

1.5 Beispiele

Für die Betrachtungen dieses Abschnitts erweist es sich als bequem, einen Punkt N des Raumes auszuzeichnen (als "Bezugspunkt", "Nullpunkt" oder "Ursprung des Koordinatensystems"). Jeder Punkt P des Raumes ist dann relativ zu N durch seinen **Ortsvektor** $\mathfrak{a} = \overrightarrow{NP}$ eindeutig bestimmt. In den folgenden Untersuchungen werden wir ohne weiteren Kommentar die Punkte des Raumes durch ihre Ortsvektoren beschreiben.

1.5.1 Der Mittelpunkt M der Strecke $\overline{AB}$

.



Mittelpunkt einer Strecke

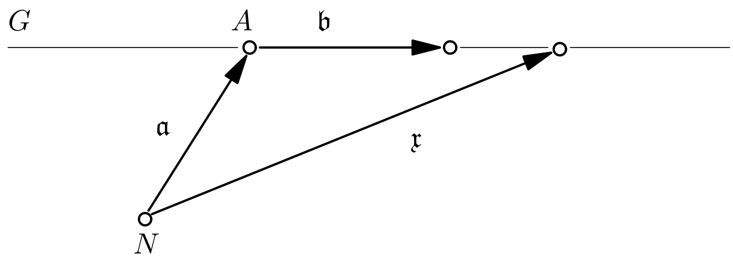
Es ist $\overrightarrow{AB} = \mathfrak{b} - \mathfrak{a}$, folglich $\overrightarrow{AM} = 1/2(\mathfrak{b} - \mathfrak{a})$, somit

$$\overrightarrow{NM} = \mathfrak{a} + \frac{1}{2}(\mathfrak{b} - \mathfrak{a}) = \frac{1}{2}(\mathfrak{a} + \mathfrak{b}).$$

Damit

$$\mathfrak{m} = \frac{1}{2}(\mathfrak{a} + \mathfrak{b}).$$

1.5.2 Gerade, gegeben durch Punkt und Richtung



Gerade, gegeben durch Punkt und RichtungDie Gerade G durch den Punkt $\mathfrak{a}$ mit

Richtungsvektor $\mathfrak{b}$ besteht aus allen 'Punkten $\mathfrak{x}$, welche die Form

$$\mathfrak{x} = \mathfrak{a} + t\mathfrak{b} \quad \text{mit einem Skalar } t$$

haben. In Mengenschreibweise ist

$$G = \{\mathfrak{x} \mid \mathfrak{x} = \mathfrak{a} + t\mathfrak{b} \text{ für ein } t \in \mathbb{R}\}\,.$$

Wir werden für diese einführenden Betrachtungen — im Interesse einer knappen Redeweise — jedoch häufig von der ‘Geraden’ $\mathfrak{x} = \mathfrak{a} + t\mathfrak{b}$ reden oder $\mathfrak{x} = \mathfrak{a} + t\mathfrak{b}$, $t \in \mathbb{R}$, als Parameterdarstellung der Geraden bezeichnen.

Fassen wir zusammen: *Ein Punkt $\mathfrak{x}$ liegt genau dann auf der Geraden G durch $\mathfrak{a}$ mit Richtung $\mathfrak{b}$ (also $\mathfrak{x} \in G$) wenn es ein $t \in \mathbb{R}$ gibt mit $\mathfrak{x} = \mathfrak{a} + t\mathfrak{b}$.*
Der Richtungsvektor $\mathfrak{b}$ muss natürlich verschieden von 0 sein, damit wirklich eine Gerade vorliegt.

Montag, 20. Oktober 03

Es ist leicht zu sehen, dass verschiedene Daten $(\mathfrak{a}, \mathfrak{b})$ dieselbe Gerade bestimmen können. Es ist sehr lehrreich, sich in diesem Zusammenhang folgendes klar zu machen.
Gleichheit von Geraden: *Die durch*

$$\begin{aligned} \mathfrak{x} &= \mathfrak{a} + t\mathfrak{b} \\ \mathfrak{x}_0 &= \mathfrak{a}_0 + t_0\mathfrak{b}_0 \end{aligned}$$

gegebenen Geraden G und G_0 stimmen genau dann überein, wenn

(1) $$\mathfrak{b} = a\mathfrak{b}_0 \quad \text{mit } 0 \neq a \in \mathbb{R}$$

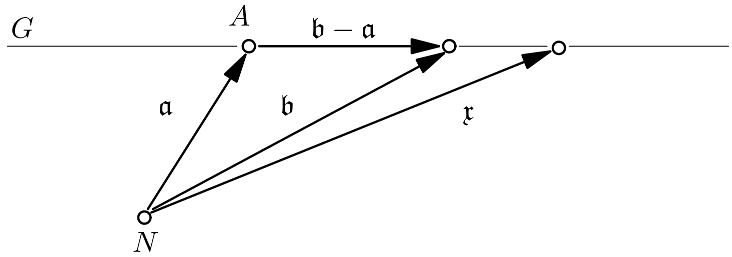
und

(2) $$\mathfrak{a} - \mathfrak{a}_0 = b\mathfrak{b}_0 \quad \text{mit } b \in \mathbb{R}$$

gilt.

Übung 5.1 *Wann sind die beiden Geraden (im Raum) parallel?*

1.5.3 Gerade, gegeben durch zwei Punkte



Gerade, gegeben durch zwei Punkte $\mathfrak{a} \neq \mathfrak{b}$Die gesuchte Parameterdarstellung ist

$\mathfrak{x} = \mathfrak{a} + t(\mathfrak{b} - \mathfrak{a})$.

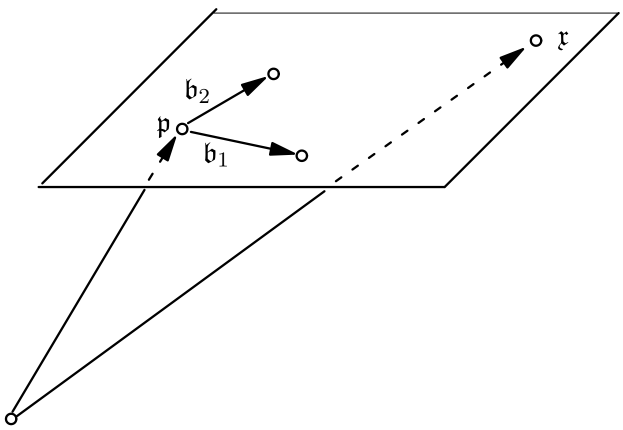
1.5.4 Ebene, gegeben durch einen Punkt und zwei Richtungen

Wir nennen zwei Vektoren $\mathfrak{a}$ und $\mathfrak{b}$ **parallel** (vgl. auch Übung 5.1), wenn $\mathfrak{a}$ ein skalares Vielfaches von $\mathfrak{b}$ oder $\mathfrak{b}$ ein skalares Vielfaches von $\mathfrak{a}$ ist[5]. Falls $\mathfrak{a} \neq 0$ ist, folgt aus $\mathfrak{a} = a\mathfrak{b}$, dass notwendig $a \neq 0$. In diesem Fall ist dann $\mathfrak{b} = \frac{1}{a}\mathfrak{a}$ seinerseits ein skalares Vielfaches von $\mathfrak{a}$.

Seien jetzt ein Vektor $\mathfrak{a}$ und zwei nicht parallele Vektoren $\mathfrak{b}_1 \neq \mathfrak{o}$, $\mathfrak{b}_2 \neq \mathfrak{o}$ gegeben. Die Menge aller ‘Punkte

$$\mathfrak{x} = \mathfrak{a} + s_1\mathfrak{b}_1 + s_2\mathfrak{b}_2 \tag{1}$$

mit Skalaren s_1, s_2 ist die **Ebene durch $\mathfrak{a}$ mit den "Richtungen" $\mathfrak{b}_1$ und $\mathfrak{b}_2$**.



Ebene, gegeben durch Punkt und zwei Richtungen

Die Formel (1) nennen wir auch eine **Parameterdarstellung der Ebene**.

1.5.5 Ebene durch 3 nicht auf einer Geraden liegende Punkte

Seien $\mathfrak{a}_0, \mathfrak{a}_1, \mathfrak{a}_2$ derartige Punkte. Dann sind $\mathfrak{a}_1 - \mathfrak{a}_0$ und $\mathfrak{a}_2 - \mathfrak{a}_0$ nicht parallel und

$$\mathfrak{x} = \mathfrak{a}_0 + s_1(\mathfrak{a}_1 - \mathfrak{a}_0) + s_2(\mathfrak{a}_2 - \mathfrak{a}_0)$$

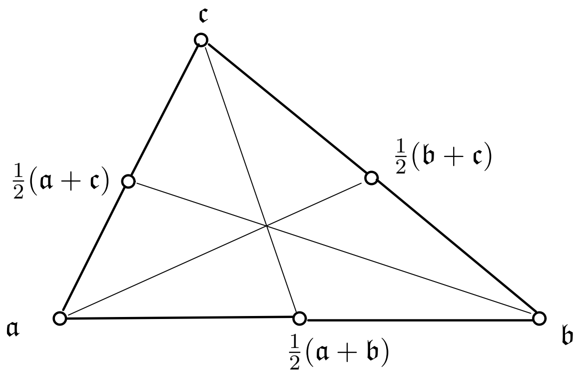
ist eine Parameterdarstellung der gesuchten Ebene.

Eine Ebene E hat ersichtlich verschiedene Parameterdarstellungen. Bei zwei in Parameterform gegebenen Ebenen ist es zudem schwieriger als im Geradenfall zu entscheiden, ob sie zusammenfallen oder zueinander parallel sind.

[5] Nach dieser Erklärung ist der Nullvektor zu jedem Vektor parallel.

1.5.6 Schnitt der Schwerlinien eines Dreiecks im Schwerpunkt

Die Schwerlinien eines Dreiecks verbinden definitionsgemäß die Ecken mit der Mitte der jeweils gegenüber liegenden Seite. Wir wollen zeigen, dass sich die Schwerlinien in einem Punkt, dem **Schwerpunkt des Dreiecks**, schneiden und dieser — von den Ecken aus gesehen — die Schwerlinien im Verhältnis 2 : 1 teilt.



Schwerlinien und Schwerpunkt eines Dreiecks

Es ist nämlich

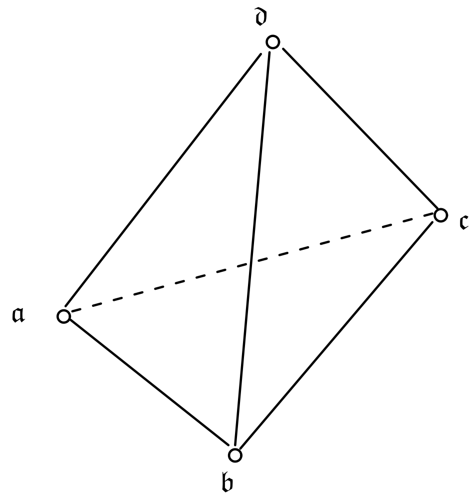
$$\begin{aligned} & \mathfrak{a}+\frac{2}{3}\left[\frac{1}{2}(\mathfrak{b}+\mathfrak{c})-\mathfrak{a}\right] \\ = \quad & \mathfrak{b}+\frac{2}{3}\left[\frac{1}{2}(\mathfrak{a}+\mathfrak{c})-\mathfrak{b}\right] \\ = \quad & \mathfrak{c}+\frac{2}{3}\left[\frac{1}{2}(\mathfrak{a}+\mathfrak{b})-\mathfrak{c}\right] \\ = \quad & \frac{1}{3}\,(\mathfrak{a}+\mathfrak{b}+\mathfrak{c})) \end{aligned}$$

Der Schwerpunkt $\mathfrak{s}$ des Dreiecks mit den Eckpunkten $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ ist somit durch die Formel

$$\mathfrak{s}=\frac{1}{3}\,(\mathfrak{a}+\mathfrak{b}+\mathfrak{c}))$$

gegeben, die in offensichtlicher Weise die Formel für den Mittelpunkt einer Strecke verallgemeinert.

Es ist jetzt leicht, die Formel für den **Schwerpunkt einer Pyramide** (Tetraeder) mit den Eckpunkten $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, $\mathfrak{d}$



Schwerpunkt eines Tetraeders

zu erraten und zu beweisen:

Übung 5.2 *Die Schwerlinien eines Tetraeders, die jeweils eine Ecke mit dem Schwerpunkt der gegenüberliegenden Dreiecksseite verbinden, schneiden sich in einem Punkt* $\mathfrak{s}$*, dem* Schwerpunkt des Tetraeders*. Ferner schneidet* $\mathfrak{s}$ *die Schwerlinien — jeweils vom Eckpunkt aus gesehen — im Verhältnis* $3:1$.

Donnerstag, 23. Oktober 2003

1.6 Lineare Unabhängigkeit, Dimensionsaxiom

6.1 *Ein System von Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ heißt* ***linear abhängig****, wenn es Skalare $a_1, a_2, \ldots, a_n$ gibt, die nicht sämtlich 0 sind, so dass die* ***Linearkombination***

$$a_1\mathfrak{a}_1 + a_2\mathfrak{a}_2 + \cdots + a_n\mathfrak{a}_n = \mathfrak{o} \tag{1}$$

ist. Gibt es solche Skalare $a_1, a_2, \ldots, a_n$ nicht, so heißen die Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ ***linear unabhängig****.*

Die lineare Unabhängigkeit von Vektoren können wir — im Einklang mit obiger Erklärung — wie folgt kennzeichnen:

6.2 Die Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ bilden genau dann ein **linear unabhängig**es System, wenn Gleichung (1) nur für $a_1 = a_2 = \cdots = a_n = 0$ möglich ist. Anders formuliert, wenn aus dem Bestehen der Gleichung (1) folgt, dass $a_1 = a_2 = \cdots = a_n = 0$ gelten muss.

Wir erläutern diese Begriffe an einer Reihe von Beispielen.

6.3 $\boxed{n = 1}$. *Ein Vektor $\mathfrak{a}$ ist genau dann linear abhängig, wenn $\mathfrak{a} = \mathfrak{o}$. Somit ist $\mathfrak{a}$ genau dann linear unabhängig, wenn $\mathfrak{a} \neq \mathfrak{o}$ ist.*

Beweis. 1. Falls $\mathfrak{a} = \mathfrak{o}$ ist, so folgt $1\mathfrak{a} = \mathfrak{a} = \mathfrak{o}$, also ist $\mathfrak{a}$ linear abhängig.
2. Falls $\mathfrak{a}$ linear abhängig ist, gibt es nach (6.1) eine Zahl $a \neq 0$ mit

$$a\mathfrak{a} = \mathfrak{o}.$$

Weil $a \neq 0$ können wir $1/a$ bilden und erhalten

$$\mathfrak{a} = 1\mathfrak{a} = (\frac{1}{a}a)\mathfrak{a} = \frac{1}{a}\,(a\mathfrak{a}) = \frac{1}{a}\mathfrak{o} = \mathfrak{o}.$$

$\square$

6.4 $\boxed{n = 2}$. Zwei Vektoren $\mathfrak{a}_1$, $\mathfrak{a}_2$ sind nach Definition genau dann linear abhängig, wenn es Skalare a_1, a_2 gibt, die nicht beide 0 sind und für die

$$a_1\mathfrak{a}_1 + a_2\mathfrak{a}_2 = \mathfrak{o}$$

gilt. Ohne Beschränkung der Allgemeinheit sei $a_1 \neq 0$. Durch Auflösen obiger Beziehung nach $\mathfrak{a}_1$ folgt dann

$$\mathfrak{a}_1 = -\frac{a_2}{a_1}\mathfrak{a}_2.$$

Zwei linear abhängige Ortsvektoren $\mathfrak{a}_1$, $\mathfrak{a}_2$ liegen daher auf einer Geraden durch den Nullpunkt N.

Mit etwas mehr Sorgfalt zeigt man, dass folgende Aussagen gleichbedeutend (äquivalent) sind:

(1) $\mathfrak{a}_1$, $\mathfrak{a}_2$ *sind linear abhängig.*

(2) *Es gibt ein* $a \in \mathbb{R}$, *so dass* $\mathfrak{a}_1 = a\mathfrak{a}_2$ **oder** $\mathfrak{a}_2 = a\mathfrak{a}_1$, *d.h.* $\mathfrak{a}_1$ *und* $\mathfrak{a}_2$ *sind parallel.*

(3) *Die Ortsvektoren* $\mathfrak{a}_1$ *und* $\mathfrak{a}_2$ *liegen in einer Geraden durch den Nullpunkt* N.

6.5 $\boxed{n = 3}$. $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ sind genau dann linear abhängig, wenn es Skalare a_1, a_2, a_3 gibt, von denen mindestens einer nicht 0 ist und für die

$$a_1\mathfrak{a}_1 + a_2\mathfrak{a}_2 + a_3\mathfrak{a}_3 = \mathfrak{o}$$

gilt. Ohne Einschränkung können wir durch Umnummerieren erreichen, dass $a_3 \neq 0$ gilt. Wir können die obige Gleichung nach $\mathfrak{a}_3$ auflösen und erhalten

$$\mathfrak{a}_3 = -\frac{a_1}{a_3}\mathfrak{a}_1 - \frac{a_2}{a_3}\mathfrak{a}_2.$$

Somit lässt sich $\mathfrak{a}_3$ aus $\mathfrak{a}_1$, $\mathfrak{a}_2$ linear kombinieren, liegt somit in der durch $\mathfrak{a}_1$, $\mathfrak{a}_2$ aufgespannten Ebene durch $\mathfrak{o}$. Dieser Schluss lässt sich umkehren und führt zur Äquivalenz der folgenden Aussagen

(1) $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ *sind linear abhängig.*

(2) *Es gibt Zahlen* c_1, c_2, *so dass*

$$\begin{array}{lcll} \mathfrak{a}_1 & = & c_1\mathfrak{a}_2 + c_2\mathfrak{a}_3 & \textbf{oder} \\ \mathfrak{a}_2 & = & c_1\mathfrak{a}_1 + c_2\mathfrak{a}_3 & \textbf{oder} \\ \mathfrak{a}_3 & = & c_1\mathfrak{a}_1 + c_2\mathfrak{a}_2 & \text{gilt.} \end{array}$$

(3) *Die Ortsvektoren* $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ *liegen sämtlich in einer Ebene durch den Nullpunkt* N.

6.6 Wir können nun das Dimensionsaxiom der anschaulichen Vektorrechnung formulieren:

Dimensionsaxiom. *Es gibt 3 linear unabhängige Vektoren* $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$. *Aber je 4 Vektoren sind linear abhängig.*

Wir wählen jetzt 3 linear unabhängige Vektoren $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3$ aus, eine sogenannte **Basis**. Dann gilt:

Satz 6.7 *Jeder Vektor $\mathfrak{a}$ lässt sich auf genau eine Weise als Linearkombination*

$$\mathfrak{a} = a_1\mathfrak{b}_1 + a_2\mathfrak{b}_2 + a_3\mathfrak{b}_3$$

der Basisvektoren $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3$ darstellen.

Beweis. Existenz. Wegen des Dimensionsaxioms sind die Vektoren $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3, \mathfrak{a}$ linear abhängig. Es gibt daher eine Linearkombination

$$(*) \qquad b_1\mathfrak{b}_1 + b_2\mathfrak{b}_2 + b_3\mathfrak{b}_3 + a\mathfrak{a} = \mathfrak{o},$$

wobei mindestens einer der Skalare b_1, b_2, b_3, a von 0 verschieden ist. Nehmen wir an, dass $a = 0$ ist, so folgt aus (*), dass die Vektoren $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3$ ein linear abhängiges System bilden, was der vorausgesetzten linearen Unabhängigkeit widerspricht. Folglich muss $a \neq 0$ sein, und es folgt, dass

$$\mathfrak{a} = (-\frac{b_1}{a})\mathfrak{b}_1 + (-\frac{b_2}{a})\mathfrak{b}_2 + (-\frac{b_3}{a})\mathfrak{b}_3$$

eine Linearkombination der drei Basisvektoren ist.

Eindeutigkeit. Wir nehmen nun an, dass

$$\begin{aligned} \mathfrak{a} &= a_1\mathfrak{b}_1 + a_2\mathfrak{b}_2 + a_3\mathfrak{b}_3 \\ \mathfrak{a} &= b_1\mathfrak{b}_1 + b_2\mathfrak{b}_2 + b_3\mathfrak{b}_3 \end{aligned}$$

zwei Darstellungen von $\mathfrak{a}$ als Linearkombination von $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3$ sind. Bilden der Differenz führt zur Gleichung

$$\mathfrak{o} = (a_1 - b_1)\mathfrak{b}_1 + (a_2 - b_2)\mathfrak{b}_2 + (a_3 - b_3)\mathfrak{b}_3,$$

woraus — die lineare Unabhängigkeit von $\mathfrak{b}_1, \mathfrak{b}_2, \mathfrak{b}_3$ berücksichtigend — das Verschwinden der Koeffizienten, also $a_1 - b_1 = 0$, $a_2 - b_2 = 0$, $a_3 - b_3 = 0$ folgt. Wir haben damit $a_1 = b_1$, $a_2 = b_2$, $a_3 = b_3$ und folglich die Eindeutigkeit der Darstellung gezeigt. □

6.8 Im Umgang mit linearer Abhängigkeit bzw. Unabhängigkeit erweisen sich die nachfolgend aufgelisteten — und leicht zu beweisenden — Eigenschaften als sehr nützlich:

(1) Ist das System $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_k$ linear abhängig, so auch jedes größere System $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_k, \mathfrak{a}_{k+1}, \ldots, \mathfrak{a}_n$.

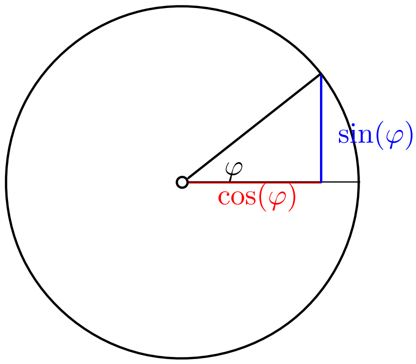
(2) Jedes System $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$, welches den Nullvektor enthält, ist linear abhängig.

(3) Ist das System $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ linear unabhängig, so ist auch jedes Teilsystem $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_k$ linear unabhängig.

1.7 Inneres Produkt (Skalarprodukt)

Montag, 27. Okt. 2003

7.1 Wir erinnern zunächst an die **Winkelfunktionen** sin und cos, deren Wirkung wir am Einheitskreis veranschaulichen:



Sinus und CosinusWinkel messen wir hier im Bogenmaß, in mathematisch positiver Richtung mit positivem, in entgegengesetzter Richtung mit negativem Vorzeichen.

7.2 *Das **innere Produkt** $\langle \mathfrak{a}, \mathfrak{b} \rangle$ (oder auch $\mathfrak{a} \cdot \mathfrak{b}$) der Vektoren $\mathfrak{a}$, $\mathfrak{b}$ ist die durch*

$$\langle \mathfrak{a}, \mathfrak{b} \rangle = |\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\sphericalangle(\mathfrak{a}, \mathfrak{b}))$$

erklärte reelle Zahl.

> Dabei wählen wir hier den Winkel $\varphi = \sphericalangle(\mathfrak{a}, \mathfrak{b})$ zwischen $\mathfrak{a}$ und $\mathfrak{b}$ stets im Intervallbereich $0 \leq \varphi \leq \pi$. Falls $\mathfrak{a}$ und $\mathfrak{b}$ nicht Null sind, ist φ in diesem Intervallbereich eindeutig bestimmt. Falls $\mathfrak{a}$ oder $\mathfrak{b}$ der Nullvektor ist, lassen wir für φ jeden der Werte $0 \leq \varphi \leq \pi$ zu.

Für das innere Produkt gelten die folgenden Gesetzmäßigkeiten:

(S 1) $\langle \mathfrak{a}, \mathfrak{b} \rangle = \langle \mathfrak{b}, \mathfrak{a} \rangle$.

(S 2) $\langle a\, \mathfrak{a}, \mathfrak{b} \rangle = a \langle \mathfrak{a}, \mathfrak{b} \rangle$.

(S 3) $\langle \mathfrak{a} + \mathfrak{b}, \mathfrak{c} \rangle = \langle \mathfrak{a}, \mathfrak{c} \rangle + \langle \mathfrak{b}, \mathfrak{c} \rangle$.

(S 4) $\langle \mathfrak{a}, \mathfrak{a} \rangle = |\mathfrak{a}|^2$.

Beweis. Zu (S 1):

$$\begin{aligned}\langle \mathfrak{a}, \mathfrak{b} \rangle &= |\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\sphericalangle(\mathfrak{a}, \mathfrak{b})) \\ &= |\mathfrak{b}| \cdot |\mathfrak{a}| \cdot \cos(\sphericalangle(\mathfrak{b}, \mathfrak{a})) \\ &= \langle \mathfrak{b}, \mathfrak{a} \rangle.\end{aligned}$$

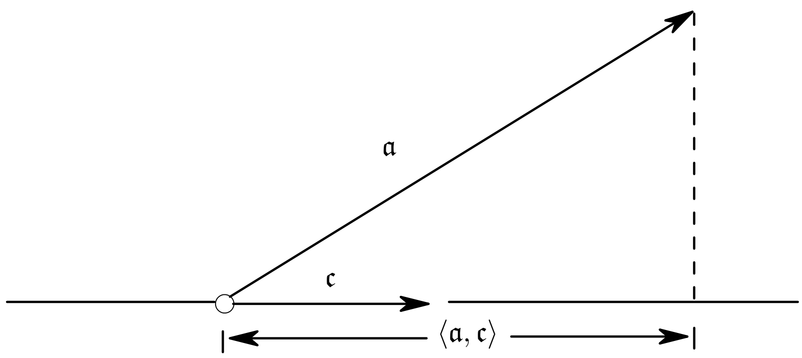
Zu (S 2): Für $a = 0$ ist alles klar. Sei nunmehr $a > 0$, so folgt

$$\begin{aligned}\langle a\mathfrak{a}, \mathfrak{b} \rangle &= |a\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\sphericalangle(a\mathfrak{a}, \mathfrak{b})) \\ &= a|\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\sphericalangle(\mathfrak{a}, \mathfrak{b})) \\ &= a\langle \mathfrak{a}, \mathfrak{b} \rangle.\end{aligned}$$

Es folgt ferner (immer noch $a > 0$ voraussetzend)

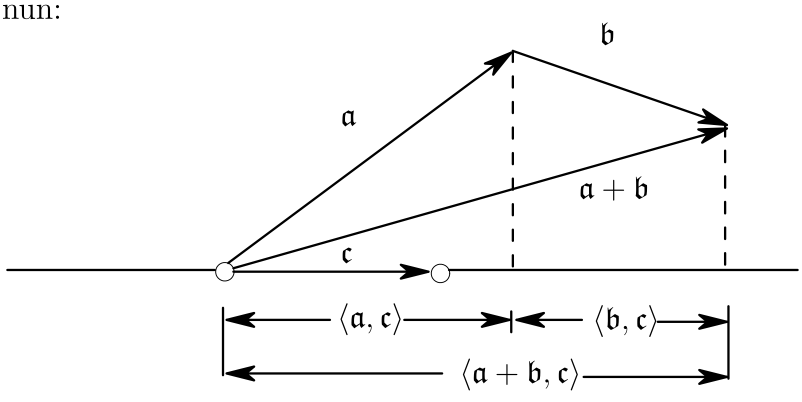
$$\begin{aligned}\langle -a\mathfrak{a}, \mathfrak{b} \rangle &= |-a\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\sphericalangle(-\mathfrak{a}, \mathfrak{b})) \\ &= a|\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\pi - \sphericalangle(\mathfrak{a}, \mathfrak{b})) \\ &= (-a)|\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \cos(\sphericalangle(\mathfrak{a}, \mathfrak{b})) \\ &= (-a)\langle \mathfrak{a}, \mathfrak{b} \rangle.\end{aligned}$$

Zu (S 3): Mit Hilfe von (S 2) können wir uns auf den Fall $|\mathfrak{c}| = 1$ beschränken. In diesem Fall ist $\langle \mathfrak{a}, \mathfrak{c} \rangle$ die Länge der Projektion von $\mathfrak{a}$ auf die Gerade mit Richtung $\mathfrak{c}$:



Berechnung von $\langle \mathfrak{a}, \mathfrak{c} \rangle$ für $|\mathfrak{c}| = 1$.

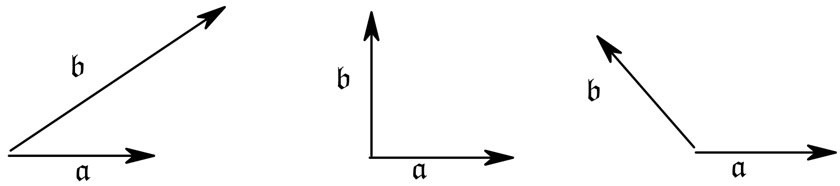
Betrachte nun:



Die Formel $\langle \mathfrak{a}+\mathfrak{b},\mathfrak{c}\rangle = \langle \mathfrak{a},\mathfrak{c}\rangle + \langle \mathfrak{b},\mathfrak{c}\rangle$.

Zu (S 4): Wir beachten $\sphericalangle(\mathfrak{a},\mathfrak{a}) = 0$ und $\cos 0 = 1$. □

7.3 Mit Hilfe des Skalarprodukts können wir erkennen, ob zwei Vektoren aufeinander senkrecht stehen oder einen spitzen bzw. stumpfen Winkel bilden:



$\langle \mathfrak{a},\mathfrak{b}\rangle > 0$ $\langle \mathfrak{a},\mathfrak{b}\rangle = 0$ $\langle \mathfrak{a},\mathfrak{b}\rangle < 0$

Spitze, rechte bzw. stumpfe Winkel.

*Falls $\langle \mathfrak{a},\mathfrak{b}\rangle = 0$ sagen wir, dass $\mathfrak{a}$ und $\mathfrak{b}$ zueinander **orthogonal** sind.*Der Nullvektor ist nach dieser Erklärung zu jedem Vektor orthogonal.

7.4 *Ein System $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ von paarweise orthogonalen (d.h. $\langle \mathfrak{a}_i, \mathfrak{a}_j\rangle = 0$ für $i \neq j$) Vektoren ungleich $\mathfrak{o}$ nennen wir ein **Orthogonalsystem**. Falls die $\mathfrak{a}_i$ zusätzlich Einheitsvektoren sind, sprechen wir von einem **Orthonormalsystem**.*

Satz 7.4 *Jedes Orthogonalsystem ist linear unabhängig.*

Beweis. Sei $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ ein Orthogonalsystem und $a_1\mathfrak{a}_1 + \cdots + a_i\mathfrak{a}_i + \cdots + a_n\mathfrak{a}_n = \mathfrak{o}$. Durch 'Multiplikation' mit $\mathfrak{a}_i$ folgt

$$a_1\langle \mathfrak{a}_1, \mathfrak{a}_i\rangle + \cdots + a_i\langle \mathfrak{a}_i, \mathfrak{a}_i\rangle + \cdots + a_n\langle \mathfrak{a}_n, \mathfrak{a}_i\rangle = 0,$$

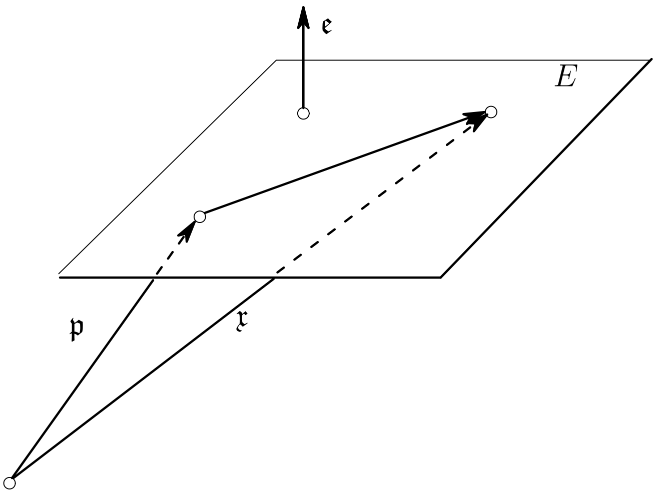
somit unter Beachtung der Orthogonalität

$$a_i|\mathfrak{a}_i|^2 = a_i\langle \mathfrak{a}_i, \mathfrak{a}_i\rangle = 0.$$

Da $\mathfrak{a}_i \neq \mathfrak{o}$, folgt hieraus $a_i = 0$. □

$\boxed{\textbf{Donnerstag 30. Okt. 2003}}$

7.5 Hessesche Normalform einer Ebene Gegeben seien ein Punkt $\mathfrak{p}$ der Ebene E und ein auf E senkrecht stehender **Einheitsvektor** $\mathfrak{e}$. (Es gibt deren zwei, nämlich $\mathfrak{e}$ und $-\mathfrak{e}$. Die Auszeichnung einer der beiden Möglichkeiten ermöglicht die Ebene zu orientieren. Wir wollen diesen Gedankengang hier jedoch nicht weiterverfolgen.)



Ebene, gegeben durch Punkt und Stellungsvektor.

Die "in der Ebene liegenden" Vektoren wie $\mathfrak{x} - \mathfrak{p}$ sind dadurch gekennzeichnet, dass sie zum Stellungsvektor $\mathfrak{e}$ orthogonal sind. Also ist $|\mathfrak{e}| = 1$ und

(1) $$\langle \mathfrak{x} - \mathfrak{p}, \mathfrak{e} \rangle = 0$$

eine Gleichung, der genau die Punkte $\mathfrak{x}$ von E genügen. D.h.

$$E = \{\mathfrak{x} \mid \langle \mathfrak{x} - \mathfrak{p}, \mathfrak{e} \rangle = 0\}.$$

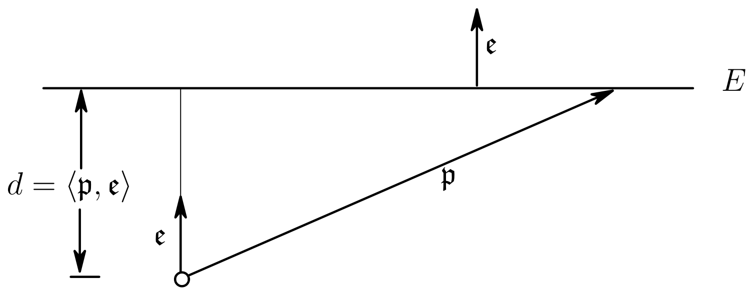
Wir können Gleichung (1) umformen zu

(2) $$\langle \mathfrak{x}, \mathfrak{e} \rangle = \langle \mathfrak{p}, \mathfrak{e} \rangle$$

oder indem wir die Größe $d = \langle \mathfrak{p}, \mathfrak{e} \rangle$ einführen zu

(3) $$\langle \mathfrak{x}, \mathfrak{e} \rangle = d.$$

Für $|\mathfrak{e}| = 1$ sind (1) bzw. (3) Darstellungen der Ebene E in **Hessescher Normalform**. Wir bemerken, dass in (3) der Punkt $\mathfrak{p}$ nicht mehr erscheint. Die Bedeutung der Zahl d ergibt sich aus der folgenden Skizze:

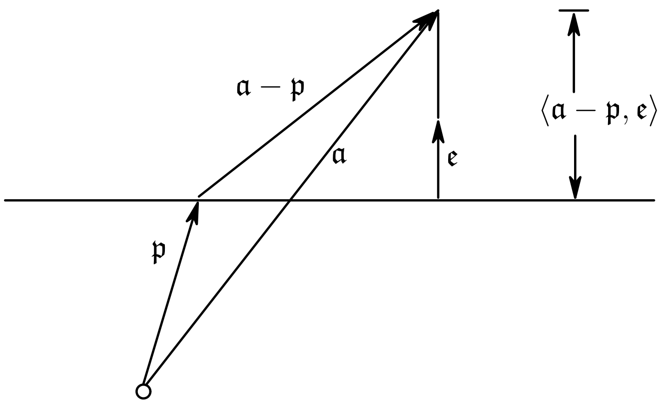


Der Abstand des Nullpunkts von der Ebene.

Die Größe $d = \langle \mathfrak{p}, \mathfrak{e} \rangle$ ist somit der **orientierte Abstand** des Nullpunkts N von der Ebene E. (Im skizzierten Fall ist $d = \langle \mathfrak{p}, \mathfrak{e} \rangle > 0$. Wir beachten hier, dass Ersetzen von $\mathfrak{e}$ durch $-\mathfrak{e}$ die Ebene E nicht, wohl aber das Vorzeichen des Abstands ändert.

Achtung: Für diese Interpretation ist wichtig, dass $\mathfrak{e}$ ein Einheitsvektor ist.

7.6 Abstand Punkt-Ebene.
Die Ebene E sei in Hessescher Normalform $\langle \mathfrak{x} - \mathfrak{p}, \mathfrak{e} \rangle = 0$ mit $|\mathfrak{e}| = 1$ gegeben.



Abstand eines Punktes von einer Ebene.

Somit ist $|\langle \mathfrak{a} - \mathfrak{p}, \mathfrak{e} \rangle|$ der Abstand und entsprechend $\langle \mathfrak{a} - \mathfrak{p}, \mathfrak{e} \rangle$ der **orientierte Abstand** des Punktes $\mathfrak{a}$ von der Ebene E. Die Punkte des durch E bestimmten Halbraumes, in welchen der Stellungsvektor $\mathfrak{e}$ weist, haben positiven Abstand, die des anderen Halbraumes haben negativen Abstand von E.

Halten wir fest: Den orientierten Abstand eines Punktes $\mathfrak{a}$ zu einer in Hessescher Normalform $\langle \mathfrak{x} - \mathfrak{p}, \mathfrak{e} \rangle = 0$, beziehungsweise $\langle \mathfrak{x}, \mathfrak{e} \rangle - d = 0$, gegebenen Ebene E erhalten wir durch Einsetzen von $\mathfrak{a}$ in die linke Seite der Ebenengleichung. Warum ist hier auf die Normierung des Stellungsvektors $\mathfrak{e}$ zu achten?

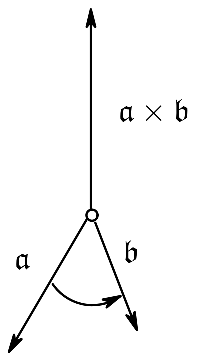
1.8 Äußeres Produkt (Vektorprodukt)

8.1 *Das* ***äußere Produkt*** $\mathfrak{a} \times \mathfrak{b}$ *(oder auch* $[\mathfrak{a}, \mathfrak{b}]$*) von zwei Vektoren* $\mathfrak{a}$*,* $\mathfrak{b}$ *ist wieder ein Vektor und wird bestimmt durch die folgenden Eigenschaften:*

(1) $\mathfrak{a} \times \mathfrak{b}$ *steht senkrecht auf* $\mathfrak{a}$ *und auf* $\mathfrak{b}$.

(1) $|\mathfrak{a} \times \mathfrak{b}| = |\mathfrak{a}| \cdot |\mathfrak{b}| \cdot \sin(\sphericalangle(\mathfrak{a}, \mathfrak{b}))$[6].

(1) $\mathfrak{a}$*,* $\mathfrak{b}$ *und* $\mathfrak{a} \times \mathfrak{b}$ *bilden in dieser Reihenfolge ein Rechtssystem.*



3-Fingerregel oder Schraubenregel

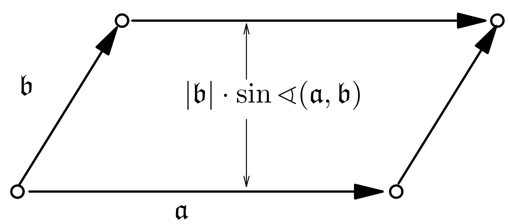
Man prüft nach, dass die Regeln

(V1) $\mathfrak{a} \times \mathfrak{b} = -\mathfrak{b} \times \mathfrak{a}$,

(V2) $(a\mathfrak{a}) \times \mathfrak{b} = a(\mathfrak{a} \times \mathfrak{b})$,

(V3) $(\mathfrak{a} + \mathfrak{b}) \times \mathfrak{c} = \mathfrak{a} \times \mathfrak{c} + \mathfrak{b} \times \mathfrak{c}$

gelten.



Hinweis: $|\mathfrak{a} \times \mathfrak{b}|$ ist die Fläche des von $\mathfrak{a}$, $\mathfrak{b}$ aufgespannten Parallelogramms. Insbesondere ist $\mathfrak{a} \times \mathfrak{b} = \mathfrak{o}$ genau dann, wenn $\mathfrak{a}$ und $\mathfrak{b}$ parallel sind. Somit: *Die folgenden drei Bedingungen sind äquivalent:*

(1) $\mathfrak{a}$*,* $\mathfrak{b}$ *sind linear abhängig.*

(2) $\mathfrak{a}$*,* $\mathfrak{b}$ *sind parallel.*

(3) $\mathfrak{a} \times \mathfrak{b} = \mathfrak{o}$.

[6]Wir erinnern an die Vereinbarung, den Winkel zwischen $\mathfrak{a}$ und $\mathfrak{b}$ im Intervall von 0 bis π zu wählen. Daher ist $\sin(\sphericalangle(\mathfrak{a}, \mathfrak{b})) \geq 0$.

Wie das Skalarprodukt ist daher auch das Vektorprodukt in jedem Argument linear[7].

8.2 Von Parameter- zur Hesseschen Normalform.
Die Ebene E sei in Parameterform

$$\mathfrak{x} = \mathfrak{a} + s\mathfrak{b}_1 + t\mathfrak{b}_2 \quad \text{mit } \mathfrak{b}_1 \times \mathfrak{b}_2 \neq \mathfrak{o}$$

gegeben. Offensichtlich steht der Vektor

$$\mathfrak{e} = \frac{\mathfrak{b}_1 \times \mathfrak{b}_2}{|\mathfrak{b}_1 \times \mathfrak{b}_2|}$$

sowohl auf $\mathfrak{b}_1$ wie auf $\mathfrak{b}_2$ senkrecht und hat die Länge 1. Damit ist

$$\langle \mathfrak{x}, \mathfrak{e} \rangle = \langle \mathfrak{a}, \mathfrak{e} \rangle$$

die zugehörige Ebenengleichung in Hessescher Normalform.

8.3 Schnitt zweier Ebenen.
Seien

$$\begin{aligned} \langle \mathfrak{x}, \mathfrak{e}_1 \rangle &= d_1, \quad |\mathfrak{e}_1| = 1 \\ \langle \mathfrak{x}, \mathfrak{e}_2 \rangle &= d_2, \quad |\mathfrak{e}_2| = 1 \end{aligned}$$

die Gleichungen für zwei Ebenen E_1, E_2. Wir nehmen an, dass E_1 und E_2 **nicht parallel** sind und somit $\mathfrak{e}_1 \times \mathfrak{e}_2 \neq \mathfrak{o}$ gilt. Ferner nehmen wir an, dass wir **einen** Punkt $\mathfrak{a}$ kennen, der beiden Ebenen gemeinsam ist, für den also

$$\begin{aligned} \langle \mathfrak{a}, \mathfrak{e}_1 \rangle &= d_1 \\ \langle \mathfrak{a}, \mathfrak{e}_2 \rangle &= d_2 \end{aligned}$$

gilt[8]. Der Vektor $\mathfrak{e}_1 \times \mathfrak{e}_2$ steht senkrecht auf $\mathfrak{e}_1$ und $\mathfrak{e}_2$. Alle Punkte der Form $\mathfrak{x} = \mathfrak{a} + t\,\mathfrak{e}_1 \times \mathfrak{e}_2$ liegen daher sowohl in E_1 als auch in E_2. Die resultierende **Schnittgerade**

$$G = E_1 \cap E_2 = \{\mathfrak{x} \mid \mathfrak{x} \in E_1 \text{ und } \mathfrak{x} \in E_2\}$$

erhalten wir somit als die Menge aller

$$\mathfrak{x} = \mathfrak{a} + t\,\mathfrak{e}_1 \times \mathfrak{e}_2, \quad \text{mit } t \in \mathbb{R}.$$

---

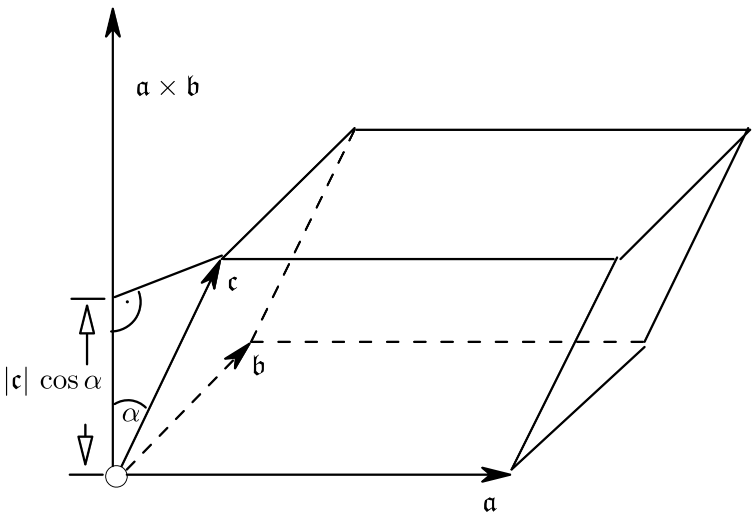
[7] Summen und Skalare lassen sich somit aus jedem Faktor herausziehen.
[8] Wie man einen solchen Schnittpunkt bestimmt, werden wir später sehen.

1.9 Das Spatprodukt

9.1 *Unter dem* ***Spatprodukt*** *der Vektoren* $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ *verstehen wir die Zahl*

$$(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) = \langle \mathfrak{a} \times \mathfrak{b}, \mathfrak{c} \rangle = \langle \mathfrak{c}, \mathfrak{a} \times \mathfrak{b} \rangle.$$

Somit ist $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) = |\mathfrak{a} \times \mathfrak{b}| \cdot |\mathfrak{c}| \cdot \cos\alpha$, wobei $\alpha = \sphericalangle(\mathfrak{a} \times \mathfrak{b}, \mathfrak{c})$ ist.



Volumen eines Spats

Im skizzierten Fall ist $|\mathfrak{a} \times \mathfrak{b}|$ die Grundfläche F des von $\mathfrak{a}$, $\mathfrak{b}$ aufgespannten Spats. Weiter ist $h = |\mathfrak{c}| \cdot \cos\alpha$ (für $0 \le \alpha \le \pi/2$) die Höhe des Spats, somit $V = F \cdot h = |\mathfrak{a} \times \mathfrak{b}||\mathfrak{c}| \cdot \cos(\alpha)$ sein **Volumen**. Sollte $\pi/2 \le \alpha \le \pi$ sein, so ist $V = -(\mathfrak{a}, \mathfrak{b}, \mathfrak{c})$.

Das Spatprodukt $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c})$ interpretieren wir geometrisch als das **orientierte Volumen des von $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ aufgespannten Spats**. Dabei ist $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) \ge 0$, falls $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ ein Rechtssystem bilden. Das Spatprodukt ändert sich folglich nicht bei zyklischer Vertauschung der Faktoren

(*Sp* 1) $$(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) = (\mathfrak{b}, \mathfrak{c}, \mathfrak{a}) = (\mathfrak{c}, \mathfrak{a}, \mathfrak{b}).$$

Aber

(*Sp* 2) $$(\mathfrak{b}, \mathfrak{a}, \mathfrak{c}) = (\mathfrak{a}, \mathfrak{c}, \mathfrak{b}) = (\mathfrak{c}, \mathfrak{b}, \mathfrak{a}) = -(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}).$$

Ferner folgt aus den Rechenregeln für inneres und äußeres Produkt:

(*Sp* 3) $$(\mathfrak{a} + \mathfrak{a}', \mathfrak{b}, \mathfrak{c}) = (\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) + (\mathfrak{a}', \mathfrak{b}, \mathfrak{c}).$$

(*Sp* 4) $$(a\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) = a(\mathfrak{a}, \mathfrak{b}, \mathfrak{c}).$$

Insbesondere ist das Spatprodukt in jedem Faktor linear.

Montag,3. November 2003

Satz 9.2 *Drei Vektoren $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ sind genau dann linear abhängig, wenn ihr Spatprodukt $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3) = 0$ ist.*

Beweis. 1. Falls $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3) = 0$, ist das Volumen des von $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ aufgespannten Spats gleich 0 und $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ liegen in einer Ebene durch 0.

2. Falls $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ linear abhängig sind, ist ohne Einschränkung

$$\mathfrak{a}_1 = a\mathfrak{a}_2 + b\mathfrak{a}_3$$

mit geeigneten Zahlen a, b. Es folgt

$$\begin{aligned} (\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3) &= (a\mathfrak{a}_2 + b\mathfrak{a}_3, \mathfrak{a}_2, \mathfrak{a}_3) \\ &= a(\mathfrak{a}_2, \mathfrak{a}_2, \mathfrak{a}_3) + b(\mathfrak{a}_3, \mathfrak{a}_2, \mathfrak{a}_3) \\ &= 0 + 0 = 0. \end{aligned}$$ □

9.3 Sei $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ ein System von 3 linear unabhängigen Vektoren, also eine Basis. Wir haben gesehen (6.7), dass man jeden Vektor $\mathfrak{a}$ eindeutig als Linearkombination

$$\mathfrak{a} = a_1\mathfrak{a}_1 + a_2\mathfrak{a}_2 + a_3\mathfrak{a}_3$$

darstellen kann. Die Eindeutigkeit der "Koordinaten" (a_1, a_2, a_3) legt den Versuch nahe, sie formelmäßig aus $\mathfrak{a}, \mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ zu ermitteln:

Satz 9.4 (Cramersche Regel) *Falls $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3) \neq 0$ und*

$$\mathfrak{a} = a_1\mathfrak{a}_1 + a_2\mathfrak{a}_2 + a_3\mathfrak{a}_3 \tag{1.1}$$

gilt, berechnen sich die Koeffizienten a_1, a_2, a_3 mittels der Formeln

$$a_1 = \frac{(\mathfrak{a}, \mathfrak{a}_2, \mathfrak{a}_3)}{(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)}, \quad a_2 = \frac{(\mathfrak{a}_1, \mathfrak{a}, \mathfrak{a}_3)}{(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)}, \quad a_3 = \frac{(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a})}{(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)}. \tag{1.2}$$

Beweis. Wir multiplizieren Formel (1.1) von rechts im Sinne des skalaren Produkts mit $\mathfrak{a}_2 \times \mathfrak{a}_3$:

$$\begin{aligned} (\mathfrak{a}, \mathfrak{a}_2, \mathfrak{a}_3) &= a_1(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3) + a_2(\mathfrak{a}_2, \mathfrak{a}_2, \mathfrak{a}_3) + a_3(\mathfrak{a}_3, \mathfrak{a}_2, \mathfrak{a}_3) \\ &= a_1(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3). \end{aligned}$$

Es folgt der behauptete Ausdruck für a_1. Entsprechend ermitteln wir a_2 und a_3. □

1.10 Mehrfache Produkte

10.1 Was lässt sich über $\mathfrak{a} \times (\mathfrak{b} \times \mathfrak{c})$ sagen? Ohne Einschränkung können wir $\mathfrak{b} \times \mathfrak{c} \neq \mathfrak{o}$ annehmen. Da $\mathfrak{a} \times (\mathfrak{b} \times \mathfrak{c})$ auf $\mathfrak{b} \times \mathfrak{c}$ senkrecht steht, muss $\mathfrak{a} \times (\mathfrak{b} \times \mathfrak{c})$ in der durch $\mathfrak{a}$ und $\mathfrak{b}$ aufgespannten Ebene durch den Nullpunkt liegen, daher die Form

$$\mathfrak{a} \times (\mathfrak{b} \times \mathfrak{c}) = b\,\mathfrak{b} + c\,\mathfrak{c}$$

haben.

Zur Bestimmung der Koeffizienten b und c sind zusätzliche Überlegungen erforderlich, die wir hier nicht im Einzelnen schildern. Wir beschränken uns aus Zeitgründen vielmehr auf die nachträgliche Verifizierung des Ergebnisses.

Satz 10.2 (Graßmannscher Entwicklungssatz) *Für je drei Vektoren* $\mathfrak{a}$, $\mathfrak{b}$ *und* $\mathfrak{c}$ *gilt:*

$$\mathfrak{a} \times (\mathfrak{b} \times \mathfrak{c}) = \langle \mathfrak{a}, \mathfrak{c} \rangle \mathfrak{b} - \langle \mathfrak{a}, \mathfrak{b} \rangle \mathfrak{c}. \tag{1.3}$$

Beweis. Wir beachten zunächst, dass beide Seiten der Gleichung (1.3) in jedem der drei Argumente $\mathfrak{a}$, $\mathfrak{b}$ und $\mathfrak{c}$ linear sind. Es reicht daher die Formel für den Fall nachzuweisen, dass $\mathfrak{a}$, $\mathfrak{b}$ und $\mathfrak{c}$ unabhängig voneinander die Vektoren einer Basis $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$ durchlaufen. Von dieser Basis können wir zusätzlich annehmen, dass sie orthonormal ist. Die verbleibende Überprüfung von Formeln wie

$$\mathfrak{e}_1 \times (\mathfrak{e}_3 \times \mathfrak{e}_1) = \langle \mathfrak{e}_1, \mathfrak{e}_1 \rangle \mathfrak{e}_3 - \langle \mathfrak{e}_1, \mathfrak{e}_3 \rangle \mathfrak{e}_1$$

ist dann offensichtlich. □

1.11 Koordinatendarstellungen

Donnerstag,6. November 2003

11.1 Wir fixieren für die folgenden Betrachtungen eine Orthonormalbasis

$$\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3 \tag{1.4}$$

des Anschauungsraumes. Damit verlangen wir

$$\langle \mathfrak{e}_i, \mathfrak{e}_k \rangle = \delta_{ik} := \begin{cases} 1 & \text{falls } i = k \\ 0 & \text{falls } i \neq k. \end{cases}$$

[Das hier eingeführte **Kronecker-Symbol** δ_{ik} wird Ihnen noch häufig begegnen.] Ferner setzen wir voraus, dass $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$ **in dieser Reihenfolge ein Rechtssystem** bilden, also

$$\mathfrak{e}_3 = \mathfrak{e}_1 \times \mathfrak{e}_2$$

gilt.

> *Eine solche Orthonormalbasis $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$, welche zugleich ein Rechtssystem ist, können wir leicht herstellen:* Wir starten mit zwei linear unabhängigen Vektoren $\mathfrak{a}$ und $\mathfrak{b}$, die wir ohne Einschränkung **normiert** annehmen und bilden
>
> $$\mathfrak{b}' = \mathfrak{b} - \langle \mathfrak{a}, \mathfrak{b} \rangle \mathfrak{a}.$$
>
> Es ist $\mathfrak{b}' \neq \mathfrak{o}$ (warum?) und ferner $\langle \mathfrak{a}, \mathfrak{b}' \rangle = 0$. Wir setzen nun
>
> $$\mathfrak{e}_1 = \mathfrak{a}, \quad \mathfrak{e}_2 = \frac{\mathfrak{b}'}{|\mathfrak{b}'|}, \quad \mathfrak{e}_3 = \mathfrak{e}_1 \times \mathfrak{e}_2$$
>
> und haben damit eine Orthonormalbasis gefunden.

11.2 Wir halten weiterhin die Orthonormalbasis $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$ fest. Nach Satz 6.7 — oder auch unter Verwendung der Cramerschen Regel Satz 9.4 — können wir jeden Vektor $\mathfrak{a}$ **eindeutig** als Linearkombination

$$\mathfrak{a} = a_1\mathfrak{e}_1 + a_2\mathfrak{e}_2 + a_3\mathfrak{e}_3 \tag{1.5}$$

schreiben. Nebenbei ist leicht zu sehen, dass im behandelten Kontext $a_i = \langle \mathfrak{a}, \mathfrak{e}_i \rangle$ gilt. Wir nennen das Tripel a_1, a_2, a_3 die **Koordinaten** von $\mathfrak{a}$ und schreiben anstelle von (1.5) auch

$$\mathfrak{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Mit dieser Übereinkunft gilt insbesondere

$$\mathfrak{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathfrak{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathfrak{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

11.3 Rechenregeln.

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix} \tag{1.6}$$

$$a\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} aa_1 \\ aa_2 \\ aa_3 \end{pmatrix} \tag{1.7}$$

$$\langle\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\rangle = a_1b_1 + a_2b_2 + a_3b_3 \tag{1.8}$$

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{pmatrix} \tag{1.9}$$

$$\begin{aligned} (\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}) &= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \\ &:= \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} c_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} c_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} c_3, \end{aligned} \tag{1.10}$$

wobei wir für die letzten beiden Eigenschaften von der Determinantenschreibweise Gebrauch machen. Die 2×2-**Determinante** wird durch die Festsetzung

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} := ad - bc$$

erklärt und die 3×3-**Determinante** wie in Formel (1.10) durch — wie man sagt — **Entwicklung nach der letzten Spalte** auf die Berechnung von 2×2-Determinanten zurückgeführt. Die auch im folgenden verwendete Bezeichnung := weist darauf hin, dass das neu eingeführte Symbol auf der linken Seite durch den Ausdruck auf der rechten Seite erklärt wird.

Beweis. Zu (1.6): Es ist

$$\begin{aligned} \mathfrak{a} &= a_1\mathfrak{e}_1 + a_2\mathfrak{e}_2 + a_3\mathfrak{e}_3 \\ \mathfrak{b} &= b_1\mathfrak{e}_1 + b_2\mathfrak{e}_2 + b_3\mathfrak{e}_3, \end{aligned}$$

woraus

$$\mathfrak{a} + \mathfrak{b} = (a_1+b_1)\mathfrak{e}_1 + (a_2+b_2)\mathfrak{e}_2 + (a_3+b_3)\mathfrak{e}_3$$

folgt.

Zu (1.7): Es ist

$$a(a_1\mathfrak{e}_1 + a_2\mathfrak{e}_2 + a_3\mathfrak{e}_3) = (aa_1)\mathfrak{e}_1 + (aa_2)\mathfrak{e}_2 + (aa_3)\mathfrak{e}_3.$$

Zu (1.8):

$$\begin{array}{rcl} \mathfrak{a} & = & a_1\mathfrak{e}_1 + a_2\mathfrak{e}_2 + a_3\mathfrak{e}_3 \\ \mathfrak{b} & = & b_1\mathfrak{e}_1 + b_2\mathfrak{e}_2 + b_3\mathfrak{e}_3 \end{array}$$

impliziert wegen $\langle \mathfrak{e}_i, \mathfrak{e}_k \rangle = \delta_{ik}$

$$\begin{array}{rcl} \langle \mathfrak{a}, \mathfrak{b} \rangle & = & \langle \sum_{i=1}^{3} a_i\mathfrak{e}_i, \sum_{k=1}^{3} b_k\mathfrak{e}_k \rangle \\ & = & \sum_{i,k=1}^{3} a_i b_k \langle \mathfrak{e}_i, \mathfrak{e}_k \rangle \\ & = & \sum_{i}^{3} a_i b_i. \end{array}$$

Zu (1.9): Mit obigen Bezeichnungen ist

$$\begin{array}{rcl} \mathfrak{a} \times \mathfrak{b} & = & \left(\sum_{i=1}^{3} a_i\mathfrak{e}_i \right) \times \left(\sum_{k=1}^{3} b_k\mathfrak{e}_k \right) \\ & = & \sum_{i,k=1}^{n} a_i b_k \, \mathfrak{e}_i \times \mathfrak{e}_k \\ & = & a_1b_2 \, \mathfrak{e}_1 \times \mathfrak{e}_2 + a_1b_3 \, \mathfrak{e}_1 \times \mathfrak{e}_3 + a_2b_1 \, \mathfrak{e}_2 \times \mathfrak{e}_1 \\ & + & a_2b_3 \, \mathfrak{e}_2 \times \mathfrak{e}_3 + a_3b_1 \, \mathfrak{e}_3 \times \mathfrak{e}_1 + a_3b_2 \, \mathfrak{e}_3 \times \mathfrak{e}_2 \\ & = & a_1b_2\mathfrak{e}_3 - a_1b_3 \, \mathfrak{e}_2 - a_2b_1\mathfrak{e}_3 \\ & + & a_2b_3 \, \mathfrak{e}_1 + a_3b_1 \, \mathfrak{e}_2 + a_3b_2 \, \mathfrak{e}_1 \\ & = & (a_2b_3 - a_3b_2) \, \mathfrak{e}_1 - (a_1b_3 - a_3b_1) \, \mathfrak{e}_2 + (a_1b_2 - a_2b_1) \, \mathfrak{e}_3 \\ & = & \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \mathfrak{e}_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \mathfrak{e}_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \mathfrak{e}_3. \end{array}$$

Zu (1.10): Es ist

$$\begin{array}{rcl} (\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) & = & \langle \mathfrak{a} \times \mathfrak{b}, \mathfrak{c} \rangle \\ & = & \left\langle \begin{pmatrix} \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right\rangle \end{array}$$

$$= \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} c_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} c_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} c_3.$$

□

Bemerkung 11.4 Dass $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3$ ein Orthonormalsystem ist, fand bereits in (1.8) Verwendung. Dass es sich zudem um ein Rechtssystem handelt

$$\mathfrak{e}_1 \times \mathfrak{e}_2 = \mathfrak{e}_3 \quad \mathfrak{e}_2 \times \mathfrak{e}_1 = -\mathfrak{e}_3$$

$$\mathfrak{e}_2 \times \mathfrak{e}_3 = \mathfrak{e}_1 \quad \mathfrak{e}_3 \times \mathfrak{e}_2 = -\mathfrak{e}_1$$

$$\mathfrak{e}_3 \times \mathfrak{e}_1 = \mathfrak{e}_2 \quad \mathfrak{e}_1 \times \mathfrak{e}_3 = -\mathfrak{e}_2$$

ist wichtig für (1.9) und folglich auch für (1.10).

Rückblick und Ausblick 11.5 Wir haben mit dem Nachweis der Rechenregeln (1.6) – (1.10) einen wichtigen Punkt erreicht:

1. Wir können von nun an mit Vektoren **ohne Rückgriff auf die räumliche Anschauung umgehen**, und sowohl Vektorrechnung wie geometrische Eigenschaften direkt auf das Rechnen mit reellen Zahlen zurückführen.

2. Durch die Darstellung als Zahlentripel liegt eine für Rechnungen geeignete zugleich kompakte wie elegante Bezeichnung von Vektoren vor.

3. Das Rechnen mit Spalten ist offensichtlich nicht auf 3 Koordinaten beschränkt[9]. Durch Bildung von n-Tupeln reeller Zahlen werden wir automatisch auf das Konzept eines n-dimensionalen Vektorraums aufmerksam gemacht.

4. Aus diesem erweiterten Blickwinkel erscheint die der anschaulichen Vektorrechnung zugrunde liegende Dreidimensionalität eher zufällig.

Wir beachten jedoch, dass die Auswahl von Koordinaten ein Element der Willkür enthält: die Koordinatendarstellung von Vektoren erfordert die künstliche Auswahl einer Basis (hier sogar Orthonormalbasis). Die früheren Überlegungen zeigen überdies, dass das Wechselspiel zwischen Geometrie einerseits und Algebra andererseits auch basisfrei und damit koordinatenunabhängig funktioniert.

[9]Vektor- und Spatprodukt nehmen wir von dieser Feststellung vorläufig aus, erst Determinantentheorie und äußere Produktbildung erlauben die korrekte Generalisierung von Vektor- und Spatprodukt auf höherdimensionale Fälle.

Kapitel 2

Systeme Linearer Gleichungen

2.1 Vorbetrachtungen

An späterer Stelle werden wir lineare Gleichungssysteme vom höheren Standpunkt behandeln. Hier wollen wir einsehen, dass schon mit geringem begrifflichen Aufwand die praktische (d.h. algorithmische) Lösung solcher Gleichungssysteme gelingt. Wir machen zugleich den Sprung von den 3-er-Spalten der anschaulichen Vektorrechnung zu n-Vektoren des $\mathbb{R}^n$.

1.1 Ist ein Gleichungssystem

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots \;\; \vdots \;\; \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \tag{2.1}$$

aus m Gleichungen in den n Unbekannten $x_1, x_2, \ldots, x_n$ vorgelegt, so nennen wir das $m \times n$-Zahlenschema

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

die **Koeffizientenmatrix** und die Spalte

$$\mathfrak{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \qquad \text{die } \textbf{rechte Seite}$$

des Gleichungssystems. Durch Zusammenfassen der $m \times n$-Matrix A mit der m-Spalte $\mathfrak{b}$ erhalten wir die sog. **erweiterte Matrix**

$$[A, \mathfrak{b}] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{mn} & \cdots & a_{1n} & b_m \end{array}\right],$$

welche das lineare Gleichungssystem (2.1) vollständig beschreibt. Lösungen des Gleichungssystems schreiben wir als Spaltenvektoren mit n Einträgen (n-**Vektoren**)

$$\mathfrak{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Beispiel 1.2 Im Fall $m = n$ nennen wir A eine **quadratische Matrix**. Die $n \times n$-Matrix

$$E = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

ist ein Beispiel einer quadratischen Matrix und heißt **Einheitsmatrix**. Zur Hervorhebung des Formats $n \times n$ werden wir sie manchmal auch mit E_n bezeichnen. Für $A = E_n$ hat das zugehörige Gleichungssystem die Form

$$\begin{array}{rcl} 1x_1 + 0x_2 + \cdots + 0x_n & = & b_1 \\ 0x_1 + 1x_2 + \cdots + 0x_n & = & b_2 \\ \vdots & & \vdots \\ 0x_1 + 0x_2 + \cdots + 1x_n & = & b_m. \end{array}$$

Dieses Gleichungssystem ist offensichtlich **eindeutig lösbar** mit der Lösung

$$\mathfrak{x} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

$\boxed{\textbf{Montag, 10. November 2003}}$

Satz 1.3 *Die* ***Lösungsmenge*** *eines linearen Gleichungssystems* **ändert sich nicht** *bei einer der folgenden Operationen*

(1) *Vertauschen zweier Gleichungen.*

(2) *Multiplikation beider Seiten einer Gleichung mit einem Faktor* $\neq 0$.

(3) *Addition eines Vielfachen einer Gleichung zu einer anderen Gleichung.*

Ein gegebenes System werden wir später mit solchen Umformungen solange vereinfachen, bis wir die Lösungsmenge des vereinfachten — und damit auch des ursprünglichen — Systems ablesen können.

Beweis. (1), (2) sind klar, nur (3) bedarf der Begründung. Seien also $i \neq k$ aus dem Bereich $1, \ldots, m$. Wir betrachten die i-te und die k-te Gleichung

$$\begin{array}{rcl} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n & = & b_i \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n & = & b_k. \end{array} \tag{2.2}$$

Jede Lösung der beiden Gleichungen ist auch eine Lösung von

$$(a_{i1} + aa_{k1})x_1 + (a_{i2} + aa_{k2})x_2 + \cdots + (a_{in} + aa_{kn})x_n = b_i + ab_k \tag{2.3}$$

und damit der Gleichungen

$$\begin{aligned}(a_{i1} + aa_{k1})x_1 + (a_{i2} + aa_{k2})x_2 + \cdots + (a_{in} + aa_{kn})x_n &= b_i + ab_k \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &= b_k.\end{aligned} \tag{2.4}$$

Halten wir fest: Jede Lösung von (2.2) ist auch eine von (2.4). Aus (2.4) erhalten wir aber (2.2) zurück, indem wir das $(-a)$-fache der zweiten Gleichung von (2.4) zur ersten Gleichung von (2.4) addieren. Mit obigem Schluss ist jede Lösung von (2.4) dann auch eine von (2.2).

Somit haben (2.2) und (2.4) dieselbe Lösungsmenge. □

1.4 Für die zugehörige erweiterte Matrix $[A|\mathfrak{b}]$ eines linearen Gleichungssystems liest sich das so:

Jede Operation einer der folgenden drei Typen

- Vertauschen zweier Zeilen,
- Multiplikation einer Zeile mit einem Faktor $\neq 0$,
- Addition eines Vielfachen einer Zeile zu einer **anderen** Zeile

macht aus $[A, \mathfrak{b}]$ eine neue "erweiterte" Matrix $[A'|\mathfrak{b}']$, für welche das zugeordnete Gleichungssystem dieselbe Lösungsmenge hat wie das Ausgangssystem.

1.5 Die Schritte 1.2 und 1.4, die beide für sich genommen ganz naheliegend[1] sind, führen zu folgendem **Lösungsversuch** für ein lineares Gleichungssystem mit quadratischer Matrix A:

Idee 1.5 *Forme $[A|\mathfrak{b}]$ solange durch* ***elementare Zeilenumformungen*** *um bis die Form*

$$[E|\mathfrak{c}]$$

erreicht ist, wobei E die $n \times n$-Einheitsmatrix ist. In diesem Fall ist

$$\mathfrak{x} = \mathfrak{c}$$

die eindeutig bestimmte Lösung von $(*)$.

Häufig klappt's:

[1] Mathematiker haben es sich angewöhnt stattdessen "trivial" zu sagen.

Beispiel 1.6 Das lineare Gleichungssystem

$$\begin{aligned} 1x_1 + 2x_2 + 0x_3 &= 1 \\ 2x_1 + 3x_2 + 0x_3 &= 1 \\ 3x_1 + 4x_2 + 1x_3 &= 3 \end{aligned}$$

führt zur erweiterten Matrix

$$\begin{array}{ccc}
\left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 2 & 3 & 0 & 1 \\ 3 & 4 & 1 & 3 \end{array}\right] & \mapsto & \left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & -1 & 0 & -1 \\ 3 & 4 & 1 & 3 \end{array}\right] \\
\left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & -2 & 1 & 0 \end{array}\right] & \mapsto & \left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & -2 & 1 & 0 \end{array}\right] \\
\left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & -2 & 1 & 0 \end{array}\right] & \mapsto & \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]
\end{array}$$

und den markierten elementaren Zeilenumformungen. Folglich ist das Gleichungssystem eindeutig lösbar mit Lösung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$$

Mitunter klappt's nicht:

Beispiel 1.7 Das lineare Gleichungssystem

$$\begin{aligned} 3x_1 + 1x_2 + 6x_3 &= 18 \\ 2x_1 + 1x_2 + 4x_3 &= 13 \\ 1x_1 + 1x_2 + 2x_3 &= 10 \end{aligned}$$

führt zur erweiterten Matrix

$$\begin{array}{ccc}
\left[\begin{array}{ccc|c} 3 & 1 & 6 & 18 \\ 2 & 1 & 4 & 13 \\ 1 & 1 & 2 & 10 \end{array}\right] & \mapsto & \left[\begin{array}{ccc|c} 1 & 1 & 2 & 10 \\ 2 & 1 & 4 & 13 \\ 3 & 1 & 6 & 18 \end{array}\right] \\
\left[\begin{array}{ccc|c} 1 & 1 & 2 & 10 \\ 0 & -1 & 0 & -7 \\ 0 & -2 & 0 & -12 \end{array}\right] & \mapsto & \left[\begin{array}{ccc|c} 1 & 1 & 2 & 10 \\ 0 & 1 & 0 & 7 \\ 0 & 2 & 0 & 12 \end{array}\right] \\
\left[\begin{array}{ccc|c} 1 & 0 & 2 & 3 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & -2 \end{array}\right] & &
\end{array}$$

und den markierten elementaren Zeilenumformungen. Das zur letzten erweiterten Matrix gehörige Gleichungssystem ist ersichtlich **nicht lösbar**, da seine letzte Gleichung $0x_1 + 0x_2 + 0x_3 = -2$ keine Lösung besitzt.

Donnerstag, 20. November 2003

Bisher haben wir Beispiele für eindeutig lösbare und für unlösbare lineare Gleichungssysteme gesehen. Es kommt auch vor, dass ein Gleichungssystem mehrfach lösbar ist:

Beispiel 1.8 Das lineare Gleichungssystem

$$\begin{array}{rcl} 3x_1 + 1x_2 + 6x_3 & = & 18 \\ 2x_1 + 1x_2 + 4x_3 & = & 13 \\ 1x_1 + 0x_2 + 2x_3 & = & 5 \end{array}$$

führt zur erweiterten Matrix

$$\left[\begin{array}{ccc|c} 3 & 1 & 6 & 18 \\ 2 & 1 & 4 & 13 \\ 1 & 0 & 2 & 5 \end{array}\right] \mapsto \left[\begin{array}{ccc|c} 1 & 0 & 2 & 5 \\ 2 & 1 & 4 & 13 \\ 3 & 1 & 6 & 18 \end{array}\right] \mapsto \left[\begin{array}{ccc|c} 1 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 1 & 0 & 3 \end{array}\right]$$
$$\mapsto \left[\begin{array}{ccc|c} 1 & 0 & 2 & 5 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Die Lösungen des Ausgangssystems stimmen daher mit den Lösungen des umgeformten Gleichungssystems $x_1 + 2x_3 = 5$, $x_2 = 3$ überein. Setzen wir abkürzend $x := -x_3$, so sind die Lösungen durch die Menge aller

$$\begin{pmatrix} 5 - 2x \\ 3 \\ x \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix} + x \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$$

gegeben, wobei x ein beliebiger Skalar ist. Wir haben hier somit unendlich viele Lösungen.

2.2 Die Struktur der Lösungsmenge

2.1 Motiviert durch die obigen Beispiele führen wir einige Überlegungen zur Lösbarkeit und zur Lösungsstruktur eines linearen Gleichungssystems durch. Sei — wie im letzten Abschnitt — ein lineares Gleichungssystem (*)

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

mit der Koeffizientenmatrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

und der rechten Seite

$$\mathfrak{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

gegeben. Wir nennen

$$\mathfrak{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathfrak{a}_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \mathfrak{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

die **Spalten** der Matrix A. (*) lässt sich dann umschreiben zur Linearkombinationsaufgabe

$$(*') \qquad x_1\mathfrak{a}_1 + x_2\mathfrak{a}_2 + \cdots + x_n\mathfrak{a}_n = \mathfrak{b}.$$

Wir nennen das zur erweiterten Matrix $[A|\mathfrak{o}]$ gehörige Gleichungssystem (**), äquivalent die Linearkombinationsaufgabe

$$(**') \qquad x_1\mathfrak{a}_1 + x_2\mathfrak{a}_2 + \cdots + x_n\mathfrak{a}_n = \mathfrak{o}$$

das zugehörige **homogene Gleichungssystem** und nennen (*) oder auch das gleichbedeutende (*') ein inhomogenes Gleichungssystem[2].

[2]Strikt gesprochen ist für die Inhomogenität $\mathfrak{b} \neq \mathfrak{o}$ vorauszusetzen. Für uns ist es nachfolgend jedoch bequemer, **inhomogenes Gleichungssystem** als Oberbegriff zu interpretieren, der den Begriff **hogenes Gleichungssystem** mitumfasst.

Die Menge $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ aller n-Spalten mit reellen Einträgen bildet einen Vektorraum, den wir mit $\mathbb{R}^n$ bezeichnen. Für $n = 3$ haben wir uns davon im Detail überzeugt, für allgemeines n gelten die für $n = 3$ durchgeführten Beweisführungen ganz analog. Die Lösungen von (*) oder (***) sind daher gewisse Elemente des $\mathbb{R}^n$.

Satz 2.1 (Struktursatz) *(a) Das System* $(*)$ *ist genau dann lösbar, wenn sich* $\mathfrak{b}$ *als Linearkombination der Spalten* $\mathfrak{a}_2, \mathfrak{a}_3, \ldots, \mathfrak{a}_n$ *von* A *schreiben lässt.*

(b) Sind $\mathfrak{x}$ *und* $\mathfrak{y}$ *Lösungen des homogenen Gleichungssystems, so auch jede Linearkombination* $a\mathfrak{x} + b\mathfrak{y}$*.*

(c) Ist $\mathfrak{y}$ *eine Lösung von* $(*)$*, so gilt:*

$\mathfrak{z}$ *ist genau dann ebenfalls eine Lösung von (*), wenn*

$$\mathfrak{z} = \mathfrak{y} + \mathfrak{x}$$

für eine Lösung $\mathfrak{x}$ *des homogenen Gleichungssystem* $(**)$*.*

Beweis. (a) folgt sofort aus (*').

(b) rechnet man leicht nach; verwende hierzu Form (*').

(c) Nach Voraussetzung ist

$$y_1\mathfrak{a}_1 + y_2\mathfrak{a}_2 + \cdots + y_n\mathfrak{a}_n = \mathfrak{b}.$$

Es gilt somit

$$z_1\mathfrak{a}_1 + z_2\mathfrak{a}_2 + \cdots + z_n\mathfrak{a}_n = \mathfrak{b}$$

genau dann, wenn

$$(z_1 - y_1)\mathfrak{a}_1 + (z_2 - y_2)\mathfrak{a}_2 + \cdots + (z_n - y_n)\mathfrak{a}_n = \mathfrak{o},$$

also

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =_{\text{def}} \begin{pmatrix} z_1 - y_1 \\ \vdots \\ z_n - y_n \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} - \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

eine Lösung des homogenen Systems (**) ist. □

Den vorstehenden Satz können wir wie folgt formulieren:

(1) Die Lösungen des **homogenen Gleichungssystems** (*) bilden einen Unterraum H des $\mathbb{R}^n$.

(2) Falls das **inhomogene Gleichungssystem** (**) eine Lösung, sagen wir $\mathfrak{x}_0$, hat[3], so ist die Lösungsmenge L des homogenen Gleichungssystems gerade

$$L = \mathfrak{x}_0 + H := \{\mathfrak{x}_0 + \mathfrak{h} \mid \mathfrak{h} \in H\}.$$

[3]Wir sprechen dann auch von einer speziellen Lösung.

In Worten: *Sei* $\mathfrak{x}_0$ *eine spezielle Lösung des inhomogenen linearen Gleichungssystems. Jede Lösung* $\mathfrak{x}$ *des inhomogenen Gleichungssystems hat dann die Form* $\mathfrak{x} = \mathfrak{x}_0 + \mathfrak{h}$*, wobei* $\mathfrak{h}$ *eine Lösung des homogenen Gleichungssystems ist. Umgekehrt ist jeder Vektor der Form* $\mathfrak{x} = \mathfrak{x}_0 + \mathfrak{h}$*, für den* $\mathfrak{h}$ *eine Lösung des homogenen Gleichungssystems ist, eine Lösung des inhomogenen Gleichungssystems.*

2.3 Umformung auf Zeilenstufenform

Montag, 17. November 2003

Wir sehen hier ein typisches Beispiel einer $m \times n$-Matrix in Zeilenstufenform.

$$A = \begin{array}{c} \begin{pmatrix} 0 & 0 & 1 & * & 0 & * & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{matrix} & & i_1 & & i_2 & & & i_3 & & i_4 & \end{matrix} \end{array}$$

Eine Treppenlinie trennt einen unteren Bereich ab, der nur aus Nulleinträgen besteht.

In unserem Fall haben wir $r = 4$ Stufen an den Positionen

$$1 \leq i_1 < i_2 < \cdots < i_r \leq n.$$

Eine Matrix hat per Definition genau dann **Zeilenstufenform**, wenn die folgenden Eigenschaften erfüllt sind:

(S1) In den r Stufen, d.h. an den markierten Stellen $i_1, i_2, \ldots, i_r$ stehen der Reihe nach die **Einheitsvektoren** $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_r$.

(S2) Die übrigen Einträge des oberen Bereichs können beliebig gewählt werden. Im Bereich unterhalb der Treppenlinie gibt es nur Einträge gleich Null.

(S3) Jeder Spaltenvektor ist Linearkombination vorangehender Stufenvektoren.

(S4) Kein Stufenvektor ist Linearkombination vorangehender Spaltenvektoren.

Es ist leicht zu sehen, dass (S2) aus den übrigen Anforderungen folgt. Warum?

Die Anzahl der Stufen nennen wir den **Rang**[4] von A.

Dabei verstehen wir hier unter den **Einheitsvektoren** die Vektoren der Form

$$\mathfrak{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathfrak{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathfrak{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

[4] Momentan macht dieser Begriff nur Sinn für Matrizen in Zeilenstufenform. Erst später werden wir den Rang beliebiger Matrizen erklären.

Satz 3.1 (Zeilenstufenform) *Jede $m \times n$-Matrix lässt sich durch elementare Zeilenumformungen auf Zeilenstufenform bringen.*

Wir behandeln zunächst ein Beispiel, bevor wir die Behauptung beweisen.

Vorschau: Es wird sich herausstellen, dass die entsprechende Umformung in Zeilenstufenform genau die Technik ist, die wir zur vollständigen Lösung eines linearen Gleichungssystems benötigen.

Beispiel 3.2 (Umformung in Zeilenstufenform)

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

$$\mapsto \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} |\underline{1} & 0 & 1 & 0 & 1 \\ 0 & |\underline{1} & \underline{0} & 0 & 0 \\ 0 & 0 & 0 & |\underline{1} & \underline{0} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Die letzte Matrix ist in Zeilenstufenform mit drei Stufen in den Spalten $1, 2$ *und* 4. *Sie hat daher den Rang drei.*

Beweis von Satz 3.1. Wir starten mit einer $m \times n$-Matrix $A = (a_{ik})$ mit dem Spaltenaufbau $[\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n]$. Falls alle Spalten Null sind, sind wir fertig. Andernfalls sei i_1 der kleinste Index, für den die zugehörige Spalte ungleich Null ist. Durch Vertauschen von Zeilen können wir erreichen, dass der erste Eintrag von $\mathfrak{a}_{i_1}$ ungleich Null und nach Multiplikation der ersten Zeile mit einem geeigneten Faktor sogar $a_{1,i_1} = 1$ gilt.

Die übrigen Einträge dieser Spalte können wir durch elementare Zeilenumformungen zu Null machen. Es ergibt sich dann eine Matrix der Form

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & & \vdots & & B & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}$$

Hierbei ist B eine $(m-1) \times (n-i_1)$-Matrix. Durch Induktion[5] nach der Zeilenzahl m können wir daher annehmen, dass sich B durch elementare Zeilenumformungen auf Zeilenstufenform B' bringen lässt. Diese Zeilenumformungen wenden wir auf die ganze Matrix an.

Für jede Stufe von B' können wir wegen des dort vorhandenen Eintrags 1 erreichen, dass durch elementare Zeilenumformung auch der Eintrag der ersten Zeile der großen Matrix Null wird. Das Ergebnis ist eine Matrix in Zeilenstufenform. □

[5] Die Induktionsverankerung für $m = 1$ bereitet kein Problem. Wir konzentrieren uns hier auf den Induktionsschritt.

2.4 Der Gauß-Algorithmus

Wir lassen uns in der folgenden Darstellung vom früher behandelten Struktursatz 2.1 leiten und behandeln der Reihe nach die Fragen

- Lösbarkeit
- Finden einer speziellen Lösung
- Bestimmung des Lösungsraum des homogenen Systems
- Zusammensetzen zur 'allgemeinen Lösung'.

Lösbarkeit

Durch Umformung auf Zeilenstufenform lässt sich die Lösbarkeit schnell entscheiden.

Satz 3.1 (Lösbarkeitskriterium) *Das lineare Gleichungssystem zur erweiterten Matrix $[A|\mathfrak{b}]$ ist genau dann lösbar, wenn nach der Umformung von $[A|\mathfrak{b}]$ in Zeilenstufenform $[A'|\mathfrak{b}']$ die letzte Spalte $\mathfrak{b}'$ kein Stufenvektor ist.*

Beweis. Da sich die Lösungsmenge durch elementare Zeilenumformungen nicht verändert, ist das lineare Gleichungssystem zu $[A|\mathfrak{b}]$ genau dann lösbar, wenn dasjenige zu $[A'|\mathfrak{b}']$ lösbar ist. Dies ist äquivalent dazu, dass sich $\mathfrak{b}'$ aus den Spaltenvektoren von A' linear kombinieren lässt.

Nach **(S3)** und **(S4)** ist dies genau dann der Fall, wenn $\mathfrak{b}'$ kein Stufenvektor ist. □

Finden einer speziellen Lösung

Wir behalten die obigen Bezeichnungen bei. Die Matrizen A und A' seien $m \times n$-Matrizen und $\mathfrak{b}$ sowie $\mathfrak{b}'$ seien entsprechend m-Spalten. *Wir nehmen an, dass das System lösbar ist*, also $\mathfrak{b}'$ kein Stufenvektor ist. Weglassen aller Nullzeilen[6] führt zu einer $r \times (n+1)$-Matrix, mit $r \leq n$ und gleicher Lösungsmenge, wobei r der Rang von A ist.

Wir füllen nun die resultierende Matrix solange mit Nullzeilen auf bzw. löschen "überflüssige" Nullzeilen, bis in der i-ten Stufe stets der Einheitsvektor $\mathfrak{e}_i$ steht (für alle $i = 1, \ldots, r$) und wir insgesamt n Zeilen haben. Dies liefert eine erweiterte Matrix $[A''|\mathfrak{b}'']$, wobei A'' das Format $n \times n$ hat.

[6] Weglassen bzw. Hinzufügen von Nullzeilen einer erweiterten Matrix verändert die Lösungsmenge nicht.

Satz 3.2 (Spezielle Lösung) *Mit obigen Voraussetzungen und Bezeichnungen ist $\mathfrak{b}''$ eine Lösung des Gleichungssystems zu $[A|\mathfrak{b}]$.*

Beweis. Wenn der i-te Eintrag von $\mathfrak{b}''$ von Null verschieden ist, so ist nach Konstruktion die i-te Spalte von A'' ein Stufenvektor, also gleich dem Einheitsvektor $\mathfrak{e}_i$. Interpretation des Gleichungssystems zu $[A''|\mathfrak{b}'']$ als Linearkombinationsaufgabe zeigt, dass $\mathfrak{b}''$ eine Lösung ist. □

Wir führen die Überlegungen an einem Beispiel durch.

Beispiel 3.3 (Spezielle Lösung) *Die erweiterte Matrix*

$$[A|\mathfrak{b}] = \left[\begin{array}{ccccc|c} |\underline{1} & 0 & 1 & 0 & 1 & 3 \\ 0 & |\underline{1} & \underline{2} & 0 & 2 & -1 \\ 0 & 0 & 0 & |\underline{1} & \underline{3} & \underline{-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

befindet sich schon in Zeilenstufenform. Die letzte Spalte ist keine Stufe, daher ist das System lösbar. Auffüllen mit Nullzeilen liefert das äquivalente System:

$$[A''|\mathfrak{b}''] = \left[\begin{array}{ccccc|c} |\underline{1} & 0 & 1 & 0 & 1 & 3 \\ 0 & |\underline{1} & \underline{2} & 0 & 2 & -1 \\ 0 & 0 & 0 & |0 & 2 & 0 \\ 0 & 0 & 0 & |\underline{1} & \underline{3} & \underline{-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Wir beachten, dass Nullen in der Hauptdiagonalen von A'' in derselben Zeile einen Nulleintrag in $\mathfrak{b}''$ hervorrufen. Grund?
Es folgt: $\mathfrak{b}''$ ist eine spezielle Lösung des Gleichungssystems zu $[A|\mathfrak{b}]$.

Hinweis: Dieses einfache Rezept zum Finden einer speziellen Lösung darf nicht zum Schluss verleiten, $\mathfrak{b}''$ sei die einzige Lösung dieses Gleichungssystems.

Homogene Systeme in Zeilenstufenform

Wir nehmen jetzt an, dass A eine Matrix in Zeilenstufenform ist. Durch Streichen und Neueinfügen von Nullzeilen sei schon erreicht, dass A quadratisch, vom Format $n \times n$ ist und die Stufenvektoren ihre Eins in der Hauptdiagonalen haben. Wir sprechen dann von **erweiterter Stufenform**. Die Anzahl der Stufen(vektoren), also der Rang von A, sei gleich r. Es seien $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ die Spalten der Matrix A.

Jeder Nichtstufenvektor von A hat als Hauptdiagonaleintrag eine Null. Abänderung dieser (diagonalen) Nullen zu -1 führt zu Vektoren

$$\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_{n-r}.$$

Satz 3.4 (Lösungsraum des homogenen Systems) *Die vorstehend konstruierten Vektoren* $\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_{n-r}$ *sind linear unabhängige Lösungen des homogenen Gleichungssystems. Ferner ist jede Lösung* $\mathfrak{h}$ *des homogenen Gleichungssystems eine Linearkombination von* $\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_{n-r}$ *und umgekehrt.*

Beweis. Wir erinnern daran, dass die Lösungen des homogenen Systems einen **Unterraum** H von $\mathbb{R}^n$ bilden, der aus allen n-Spalten

$$\mathfrak{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

besteht, für die

$$x_1\mathfrak{a}_1 + x_2\mathfrak{a}_2 + \cdots + x_n\mathfrak{a}_n = \mathfrak{o}$$

gilt. Es liegt daher der n-Vektor $\mathfrak{o}$ in H. Ferner ist H gegen Summenbildung und gegen Multiplikation mit Skalaren abgeschlossen. Folglich ist H auch gegen Bildung von Linearkombinationen abgeschlossen.

Wir haben noch die folgenden Eigenschaften zu zeigen:

(a) Jedes $\mathfrak{h}_j$ gehört zu H.

(b) Die Vektoren $\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_{n-r}$ sind linear unabhängig.

(c) Jedes $\mathfrak{h}$ aus H ist darstellbar als Linearkombination $\sum_{j=1}^{n-r} \alpha_j \, \mathfrak{h}_j$.

Durch Umnummerieren der Variablen $x_1, x_2, \ldots, x_n$ (und synchrones Umnummerieren der Spalten $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ von A) können wir annehmen, dass die Stufenvektoren der Reihe nach in den Positionen 1 bis r stehen und danach die Nichtstufenvektoren kommen. Also gilt mit diesen Vereinbarungen:

(1) $\mathfrak{a}_i = \mathfrak{e}_i$ für $i = 1, \ldots, r$.

(2) Jedes $\mathfrak{a}_{r+j}$, $j = 1, \ldots, n - r$, ist eine Linearkombination von $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_r$.

(3) Es ist $\mathfrak{h}_j = \mathfrak{a}_{r+j} - \mathfrak{e}_{r+j}$.

Zu (a): Wegen **(2)** haben $\mathfrak{a}_{r+j}$ und $\mathfrak{h}_j$ die Form

$$\mathfrak{a}_{r+j} = \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und } \mathfrak{h}_j = \begin{pmatrix} a_1 \\ \vdots \\ a_r \\ 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{pmatrix},$$

wobei der Eintrag -1 von $\mathfrak{h}_j$ in der Position $r+j$ steht. Es folgt wegen **(1)**

$$(a_1\mathfrak{a}_1 + \cdots + a_r\mathfrak{a}_r) + (-1)\mathfrak{a}_{r+j} = \mathfrak{a}_{r+j} - \mathfrak{a}_{r+j} = \mathfrak{o}.$$

Interpretation des homogenen Gleichungssystems als Linearkombinationsaufgabe zeigt uns, dass $\mathfrak{h}_j$ zu H gehört.

Zu (b): Es gilt

$$\mathfrak{h}_1 = \begin{pmatrix} * \\ \vdots \\ * \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathfrak{h}_2 = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \; \mathfrak{h}_{n-r} = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix}.$$

Nehmen wir nunmehr $\alpha_1\mathfrak{h}_1 + \alpha_2\mathfrak{h}_2 + \cdots + a_{n-r}\mathfrak{h}_{n-r} = \mathfrak{o}$ an, so folgt

$$\begin{pmatrix} * \\ \vdots \\ * \\ -\alpha_1 \\ -\alpha_2 \\ \vdots \\ -\alpha_{n-r} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

und somit $\alpha_1 = \alpha_2 = \cdots = \alpha_{n-r} = 0$. Damit ist die lineare Unabhängigkeit der $\mathfrak{h}_j$ gezeigt.

Zu (c): Es sei der Spaltenvektor $\mathfrak{x}$ mit den Einträgen $x_1, x_2, \ldots, x_n$ in H gelegen. Es gelte also $\sum_{j=1}^{n} x_i\mathfrak{a}_i = \mathfrak{o}$. Wir setzen $\alpha_j := -x_{r+j}$ für $j = 1, \ldots, n-r$ und zeigen, dass $\mathfrak{x} = \sum_{j=1}^{n-r} \alpha_j\mathfrak{h}_j$ gilt.

Es ist nämlich unter Berücksichtigung von **(1)** und **(3)**

$$\begin{aligned} \mathfrak{o} &= \sum_{j=1}^{r} x_i\mathfrak{a}_i + \sum_{j=1}^{n-r} x_{r+j}\mathfrak{a}_{r+j} \\ &= \left(\sum_{j=1}^{r} x_i\mathfrak{e}_i + \sum_{j=1}^{n-r} x_{r+j}\mathfrak{e}_{r+j}\right) - \sum_{j=1}^{n-r} \alpha_j\mathfrak{h}_j \\ &= \mathfrak{x} - \sum_{j=1}^{n-r} \alpha_j\mathfrak{h}_j. \end{aligned}$$

Damit ist der Beweis von Satz 3.4 abgeschlossen. □

Die drei Schritte des Gauß-Algorithmus

Wir haben jetzt alle Bausteine für den **Gauß-Algorithmus** zur Lösung eines linearen Gleichungssystems (*) beisammen, den wir nachfolgend zusammenfassen.

Schritt 1 (Lösbarkeit): Die erweiterte Matrix $[A|\mathfrak{b}]$ des linearen Gleichungssystems (*) wird durch elementare Zeilenumformungen auf Zeilenstufenform $[A'|\mathfrak{b}']$ gebracht. Falls die letzte Spalte $\mathfrak{b}'$ ein Stufenvektor ist, ist das System nicht lösbar, andernfalls ist es lösbar.

Schritt 2 (Spezielle Lösung): Durch Streichen bzw. Einfügen von Nullzeilen wird $[A'|\mathfrak{b}']$ so zu $[A''|\mathfrak{b}'']$ umgeformt, dass A'' eine quadratische Matrix ist und die Einsen der Stufenvektoren in der Hauptdiagonale von A' stehen (erweiterte Stufenform). *Im lösbaren Fall* ist dann $\mathfrak{b}''$ eine spezielle Lösung des Gleichungssystems zu $[A|\mathfrak{b}]$.

Schritt 3 (Lösungsraum des homogenen Systems): Für jeden der Nichtstufenvektoren von A'' ersetzen wir den Hauptdiagonaleintrag Null durch -1 und erhalten ein System von $n-r$ (r=Anzahl der Stufenvektoren) Spaltenvektoren $\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_{n-r}$, welches eine **Basis**[7] des Lösungsraums H des homogenen Gleichungssystems zu $[A|\mathfrak{o}]$ bildet.

Die 'allgemeine Lösung' des Gleichungssystems zu $[A|\mathfrak{b}]$ hat dann die Form

$$\mathfrak{x} = \mathfrak{b}'' + \sum_{i=1}^{n-r} \alpha_i . \mathfrak{h}_i, \qquad \text{mit beliebigen Skalaren } \alpha_i.$$

Die letzte Aussage erhalten wir aus dem Struktursatz 2.1.

Anwendungsbeispiel

Schritt 1: Lösbarkeit

Gegeben sei die erweiterte Matrix $[A|b] = \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 1 & -1 \end{array}\right]$

Umformung in Zeilenstufenform: [8]

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & 2 & 1 & 2 & 1 & -1 \end{array}\right] \mapsto \left[\begin{array}{ccccc|c} |\underline{1} & 1 & 1 & 0 & 1 & 2 \\ 0 & |\underline{1} & \underline{0} & 1 & 0 & -2 \\ 0 & 0 & 0 & |\underline{1} & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

[7] Unter einer Basis von H verstehen wir dabei ein linear unabhängiges System von Vektoren von H, aus denen sich jeder Vektor von H linear kombinieren lässt.

[8] Beachten Sie, dass die Summe der ersten drei Zeilen gleich der vierten ist.

$$\mapsto \left[\begin{array}{ccccc|c} \lvert\underline{1} & 0 & 1 & -1 & 1 & 4 \\ 0 & \lvert\underline{1} & \underline{0} & 1 & 0 & -2 \\ 0 & 0 & 0 & \lvert\underline{1} & \underline{0} & \underline{-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \quad \mapsto \left[\begin{array}{ccccc|c} \lvert\underline{1} & 0 & 1 & 0 & 1 & 3 \\ 0 & \lvert\underline{1} & \underline{0} & 0 & 0 & -1 \\ 0 & 0 & 0 & \lvert\underline{1} & \underline{0} & \underline{-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Wir lesen ab: Das System ist lösbar!

Schritt 2: Streichen/Einfügen von Nullzeilen

Aus

$$\left[\begin{array}{ccccc|c} \lvert\underline{1} & 0 & 1 & 0 & 1 & 3 \\ 0 & \lvert\underline{1} & \underline{0} & 0 & 0 & -1 \\ 0 & 0 & 0 & \lvert\underline{1} & \underline{0} & \underline{-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

wird die Matrix

$$\left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Wir lesen ab: Die rechte Seite[9] $\mathfrak{x}_0 = (3, -1, 0, -1, 0)$ ist eine spezielle Lösung.

Schritt 3: Lösungsraum des homogenen Systems

Die im Diagonaleintrag modifizierten Nichtstufen-Vektoren

$$\mathfrak{h}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathfrak{h}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

bilden eine Basis des Lösungsraums des homogenen Systems.

Das lineare Gleichungssystem (*) ist durch Angabe von $[\mathfrak{x}_0; \mathfrak{h}_1, \mathfrak{h}_2]$ vollständig gelöst! Die ‘allgemeine Lösung’ des Gleichungssystems zu $[A|\mathfrak{b}]$ hat nämlich die Form

$$\mathfrak{x} = \mathfrak{x}_0 + \alpha_1\mathfrak{h}_1 + \alpha_2\mathfrak{h}_2 \text{ mit beliebigen Skalaren } \alpha_1, \alpha_2.$$

[9] Aus Platzgründen schreiben wir hier $\mathfrak{x}_0$ als Zeile, nicht als Spalte.

Carl Friedrich Gauß (1777-1855)

./Bilder/gauss5.jpg

Als 19-jähriger löste Gauß ein seit der Antike offenes Problem. Er zeigte während eines Ferienaufenthalts "durch angestrengtes Nachdenken über den Zusammenhang der Wurzeln" (der Gleichung $x^{17} = 1$), wie er später schrieb "noch ehe ich aus dem Bette aufgestanden war", dass sich das reguläre 17-Eck allein mittels Zirkel und Lineal konstruieren lässt, ein Ereignis, das ihn bestimmte, Mathematik zu seinem Beruf zu machen.

Gauß entwickelte sich zum größten Mathematiker seiner Zeit. Nicht nur der Gaußsche Algorithmus, auch die komplexe Zahlenebene (Gaußsche Zahlenebene) und die Gaußsche Normalverteilung, dargestellt auf dem 10-DM-Schein, sind nach ihm benannt. Gearbeitet hat er auf allen Gebieten der Mathematik und ihrer Anwendungen.

Kapitel 3

Vektorräume und Lineare Abbildungen

Die Thematik "Vektorräume und Lineare Abbildungen" bildet den Kern dieser Veranstaltung. **Lineare Techniken** sind zentral für weite Bereiche mathematischen Argumentierens. Die in der Analysis thematisierte **Lineare Approximation** ermöglicht sehr häufig, komplizierte mathematische Fragestellungen auf **Lineare Probleme** zu reduzieren.

Lineare Probleme, ihrerseits, führen meist auf das Lösen **Lineare**r **Gleichungssysteme**. Dieselben genießen in der Mathematik entsprechend einen hohen Stellenwert. Lösungstechniken für lineare Gleichungssysteme (Struktursatz und Gauß-Algorithmus) haben wir schon im vergangenen Kapitel untersucht. Als weitere mächtige Technik wird die **Matrizenrechnung** hinzutreten.

Gemeinsame Basis aller linearen Techniken bildet der Begriff des **Vektorraum**s und der im engen Zusammenhang stehende Begriff der **Linearen Abbildung**. Wir starten demgemäß mit dem Begriff eines abstrakten Vektorraums, untersuchen seine elementaren Eigenschaften und unterlegen den Begriff mit einer Fülle von Beispielen. Im abstrakten Kontext dieses Kapitels ist es dabei hilfreich, die früher behandelte anschauliche Vektorrechnung wo immer möglich zum Vergleich heranzuziehen.

3.1 Der Vektorraumbegriff

Vektorräume und Beispiele

Wir greifen die im Rahmen der anschaulichen Vektorrechnung isolierten "Gesetzmäßigkeiten" **(A1)–(A4)**, **(M1)–(M4)** erneut auf und definieren den Begriff des (abstrakten) Vektorraums durch axiomatische Forderung dieser Bedingungen:

Definition 1.1 (Reller Vektorraum) *Eine Menge V versehen mit zwei Operationen*

$$\begin{aligned} + \ &: \ V \times V \longrightarrow V, \quad (v, w) \mapsto v + w \\ \cdot \ &: \ \mathbb{R} \times V \longrightarrow V, \quad (a, v) \mapsto a\, v \end{aligned}$$

heißt ***Vektorraum****, genauer reeller Vektorraum oder $\mathbb{R}$-Vektorraum, wenn gilt:*

(A 1) $v + w = w + v$ *für alle* $v, w \in V$.

(A 2) $u + (v + w) = (u + w) + v$ *für alle* $u, v, w \in V$.

(A 3) *Es gibt* $0 \in V$ *mit* $v + 0 = v = 0 + v$ *für alle* $v \in V$.

(A 4) *Zu jedem* $v \in V$ *gibt es* $w \in V$ *mit* $v + w = 0 = w + v$.

(M 1) $a(v + w) = av + aw$ *für alle* $a \in \mathbb{R}$, $v, w \in V$.

(M 2) $(a + b)v = av + bv$ *für alle* $a, b \in \mathbb{R}$, $v \in V$.

(M 3) $(a \cdot b)v = a(bv)$ *für alle* $a, b \in \mathbb{R}$, $v \in V$.

(M 4) $1v = v$ *für alle* $v \in V$.

Elemente aus V *heißen* ***Vektoren****, die aus* $\mathbb{R}$ ***Skalare***

Wir beachten, dass andere in der Diskussion des ersten Kapitels aufgetretene Strukturen des Anschauungsraumes (Skalarprodukt, Vektorprodukt, Spatprodukt) nicht Bestandteil dieser Definition sind. Hinsichtlich des Skalarprodukts werden wir in Kapitel **??** sogenannte Euklidische Vektorräume einführen, die mit einem Skalarprodukt ausgestattet sind.

Ein anderer Punkt verdient Hervorhebung. In der anschaulichen Vektorrechnung haben wir Vektoren individuell definiert (als Klassen zueinander paralleler gerichteter Strecken) und den Anschauungsraum als Menge aller dieser Vektoren aufgefasst. Für die anstehende abstrakte Behandlung von Vektorräumen müssen wir umdenken. Ein Vektor ist jetzt nichts anderes als ein Mitglied eines Vektorraums. Ohne diesen macht es keinen Sinn mehr, von Vektoren zu sprechen.

Noch eine Anmerkung zur Schreibweise: Anders als in der anschaulichen Vektorrechnung werden wir Vektoren in der Regel durch kleine lateinische Buchstaben wie v, w, usw. bezeichnen; entsprechend werden wir nunmehr weder die Schreibweise $\mathfrak{v}$ noch $\overrightarrow{v}$ verwenden. Ebenfalls entfällt bezeichnungsmäßig — nicht jedoch begrifflich — die Unterscheidung zwischen Nullvektor und Skalar Null.

Beispiel 1.2 (Anschauungsraum und $\mathbb{R}^3$) (i) *Die Vektoren des Anschauungsraumes bilden mit den Operationen* Vektoraddition *und* Multiplikation mit Skalaren *einen* $\mathbb{R}$*-Vektorraum.*

(ii) *Die Menge*

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R} \right\}$$

bildet mit den Operationen

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}, \quad a \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ ax_3 \end{pmatrix}$$

ebenfalls einen $\mathbb{R}$*-Vektorraum.*

Beispiel 1.3 (Lösungsmenge einer linearen Gleichung) *Wir fixieren vier reelle Zahlen* a_1, a_2, a_3, b. *Dann ist die Menge*

$$L = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R},\ a_1x_1 + a_2x_2 + a_3x_3 = b \right\}$$

bzgl. der Operationen von 1.2 (ii) *genau dann ein* $\mathbb{R}$*-Vektorraum, wenn* $b = 0$.

Satz 1.4 (Unser Standardbeispiel) *Für jedes $n \geq 0$ bildet die Menge*

$$K^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \dots, x_n \in K \right\}$$

mit den Operationen

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad a \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$$

einen K-Vektorraum.

Beweis. Koordinatenweises Rechnen in K zeigt die Gültigkeit von **(A1)** – **(A4)** sowie **(M1)** – **(M4)**. □

Das folgende Beispiel zeigt die vereinheitlichende Kraft und Denkökonomie des abstrakten Vektorraumbegriffs, der — wie sich an dem Beispiel ablesen lässt — auch für die Analysis sehr nützlich ist.

Beispiel 1.5 (Vektorräume von Funktionen) *Sei*

$$V = \{f \mid f : [0,1] \to \mathbb{R}\}$$

die Menge aller reellwertigen Funktionen, welche auf dem reellen Einheitsintervall $[0,1] = \{x \mid x \in \mathbb{R}, 0 \leq x \leq 1\}$ definiert sind. Wir erklären $f+g$ und $a.f$ ($a \in \mathbb{R}$, f, $g \in V$) durch

$$(f+g)(x) = f(x) + g(x), \quad (a.f)(x) = a \cdot f(x).$$

Es ist leicht zu sehen, dass V bzgl. dieser Operationen ein reeller Vektorraum ist.

Das Beispiel gestattet offensichtliche Variationen: "Polynomfunktionen auf $[0,1]$", "differenzierbare Funktionen auf $[0,1]$".

Montag, 1. Dezember 2003

Der Körperbegriff

Bemerkung 1.6 In die Definition eines Vektorraums — und die nachfolgende Behandlung — gehen nur die grundlegenden Eigenschaften der Addition und Multiplikation der reellen Zahlen ein, die man zu den gleich zu besprechenden **Körperbegriff** zusammenfasst. Wir werden daher später allgemeiner Vektorräume über einem beliebigen Körper K betrachten.

Definition 1.7 (Körper) *Eine Menge K versehen mit zwei Operationen*

$$\begin{aligned} + &: K \times K \longrightarrow K, \quad (x,y) \mapsto x+y \\ \cdot &: K \times K \longrightarrow K, \quad (x,y) \mapsto x \cdot y \end{aligned}$$

heißt ***Körper****, wenn gilt:*
Axiome der Addition

(A1) *Kommutativität:* $x + y = y + x$

(A2) *Assoziativität:* $(x + y) + z = x + (y + z)$

(A3) *Existenz der Null: Es gibt* $0 \in K$ *mit* $x + 0 = x$ *für alle* $x \in K$

(A4) *Existenz eines additiv Inversen: Zu jedem* $x \in K$ *gibt es* $y \in K$ *mit* $x + y = 0$.

Axiome der Multiplikation

(M1) *Kommutativität:* $x \cdot y = y \cdot x$

(M2) *Assoziativität:* $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

(M3) *Existenz einer Eins: Es gibt* $1 \in K$ *mit* $x \cdot 1 = x$ *für alle* $x \in K$.

(M4) *Existenz eines multiplikativ Inversen: Zu jedem* $x \in K$, $x \neq 0$, *gibt es ein* $y \in K$ *mit* $x \cdot y = 1$.

Distributivgesetz
Addition und Multiplikation sind verkoppelt durch die Distributivität
(D) $(x + y) \cdot z = x \cdot z + y \cdot z$ *für alle* $x, y, z \in K$.

Beispiele 1.8 1. Bezüglich der Addition und Multiplikation reeller Zahlen bilden die Mengen $\mathbb{R}$ aller reellen Zahlen einen Körper, ebenfalls die Menge $\mathbb{Q}$ aller rationalen Zahlen $\frac{m}{n}$ mit ganzzahligen m, n und $n \neq 0$.

2. Die Menge $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ bildet — bezüglich der Addition und Multiplikation reeller Zahlen — e ebenfalls einen Körper[1]

3. Die Menge $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$ aller komplexen Zahlen bildet bzgl. $(a+bi)+(a'+b'i) = (a+a')+(b+b')i$ und $(a+bi)\cdot(a'+b'i) = (aa'-bb')+(ab'+a'b)i$ — wie wir später zeigen werden — gleichfalls einen Körper.

4. Es gibt auch endliche Körper. So gibt es für jede Primzahl p einen endlichen Körper $\mathbb{F}_p$ mit genau p Elementen.

5. $\mathbb{N}$ und $\mathbb{Z}$ sind bzgl. der gewöhnlichen Addition und Multiplikation keine Körper. Grund?

---

[1] Der Nachweis von **(M 4)** ist in diesem Beispiel nicht offensichtlich. Er erfordert Rückgriff auf die Irrationalität von $\sqrt{2}$.

Bemerkung 1.9 Wie vorher angemerkt, gehen nur die Körpereigenschaften von $\mathbb{R}$ in die Definition eines Vektorraums über $\mathbb{R}$ ein. Es macht daher Sinn, allgemeiner einen Körper K zugrundezulegen und für Operationen

$$\begin{array}{rcl} + & : & V \times V \longrightarrow V, \quad (v,w) \mapsto v+w \\ \cdot & : & K \times V \longrightarrow V, \quad (a,v) \mapsto a\,v, \end{array}$$

genannt Addition und Multiplikation (mit Skalaren), die Eigenschaften **(A 1)** – **(A 4)** und **(M 1)** – **(M 4)** zu verlangen. Wir sprechen dann von einem Vektorraum über dem Körper K, kurz von einem K-Vektorraum.

Wir können daher insbesondere von Vektorräumen über $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sprechen. Beispielsweise ist $\mathbb{Q}^n$ ein Vektorraum über $\mathbb{Q}$.

Wir werden diese Erweiterung des Vektorraumbegriffs jetzt noch nicht benötigen und bleiben momentan bei Vektorräumen über $\mathbb{R}$. Es ist jedoch nützlich zu verfolgen, dass die folgenden Argumente nicht von speziellen Eigenschaften der reellen Zahlen Gebrauch machen[2].

Das Rechnen in Vektorräumen

(1) *Das Nullelement* $0 \in V$ *ist eindeutig bestimmt.*

Beweis. Seien 0 und $0'$ Nullelmente von V, dann gilt

$$0 = 0 + 0' = 0'. \qquad \square$$

(2) *Das additive Inverse zu* x *ist eindeutig bestimmt.*

Beweis. Sei $x + y = 0$ und $x + y' = 0$. Dann folgt

$$y = y + 0 = y + (x + y') = (y + x) + y' = (x + y) + y' = 0 + y' = y' \qquad \square$$

Verabredung: Für das additive Inverse y zu x schreiben wir $\boxed{y := -x}$.

(3) $a.x = 0 \quad \Leftrightarrow \quad (a = 0 \text{ oder } x = 0)$.

Beweis. "$\Leftarrow$": Es ist $a.0 = a.(0 + 0) = a.0 + a.0$ und Addition mit $-(a.0)$ liefert $0 = a.0$. Entsprechend ist $0.x = (0 + 0).x = 0.x + 0.x$, hier liefert Addition mit $-(0.x)$ das gewünschte Resultat $0 = 0.x$.

"$\Rightarrow$" $a.x = 0$ und $a \neq 0$ impliziert $x = 1.x = (\frac{1}{a} \cdot a).x = \frac{1}{a}.(a.x) = \frac{1}{a}.0 = 0$. $\square$

[2]Davon ausgenommen ist die spätere Behandlung von Vektorräumen mit Skalarprodukt, wo es wesentlich ist, reelle Vektorräume zugrunde zu legen und auf die Ordnung der reellen Zahlen zurückzugreifen.

Subtraktion in Vektorräumen

Wir setzen $\boxed{x - y := x + (-y)}$. Dann gilt

$$(4) \qquad -(-x) = x$$

Beweis. Nach Definition von $-x$ gilt $x + (-x) = 0 = (-x) + x$. Damit ist x zu $-x$ invers, folglich $x = -(-x)$. □

$$(5) \qquad -(x + y) = (-x) + (-y)$$

Beweis. Unter Verwendung von Kommutativität und Assoziativität folgt

$$(x + y) + ((-x) + (-y)) = (x + (-x)) + (y + (-y)) = 0 + 0 = 0.$$

Somit ist $(-x) + (-y)$ das additiv Inverse zu $x + y$. □

$$(6) \qquad -(x - y) = y - x$$

Beweis. Man kombiniere (5) und (6). □*(7)* Die Gleichung $x + b = c$ (b, $c \in V$ gegeben) ist eindeutig lösbar mit Lösung $\boxed{x = c - b}$.
Beweis. *Existenz*: Es ist $(c - b) + b = c + (-b + b) = c + 0 = c$, somit ist $x = c - b$ eine Lösung.

Eindeutigkeit: Falls x eine Lösung von $x + b = c$ ist, so folgt durch Addition von $-b$ auf beiden Seiten, dass $x = c - b$ ist. □

$$(8) \qquad (-a).x = -(a.x) = a.(-x).$$

Beweis. Es ist $a + (-a) = 0$, somit $0 = 0.x = (a + (-a)).x = a.x + (-a).x$, und die Behauptung folgt. □

Erzeugendensysteme und Basen

Definition 1.10 (Linearkombination) *V sei ein Vektorraum.*

(a) Ein Element der Form

$$v = a_1.v_1 + \cdots + a_n.v_n,$$

*mit Skalaren $a_1, a_2, \ldots, a_n$ heißt **Linearkombination** der Elemente $v_1, v_2, \ldots, v_n \in V$ mit den Koeffizienten $a_1, a_2, \ldots, a_n$.*

*(b) Ein System von Vektoren $(v_1, v_2, \ldots, v_n)$ von V heißt **Erzeugendensystem** von V, wenn sich jedes $v \in V$ als Linearkombination*

$$v = a_1.v_1 + \cdots + a_n.v_n,$$

mit geeigneten Skalaren $a_1, a_2, \ldots, a_n$ schreiben lässt.

Definition 1.11 (Lineare Abhängigkeit) *V sei ein Vektorraum. Ein System $(v_1, v_2, \dots, v_n)$ von Vektoren aus V heißt* ***linear abhängig****, wenn es Skalare $(a_1, a_2, \dots, a_n) \neq (0, \dots, 0)$*[3] *gibt mit*

$$a_1.v_1 + a_2.v_2 + \cdots + a_n.v_n = 0.$$

Andernfalls heißt das System $(v_1, v_2, \dots, v_n)$ ***linear unabhängig***[4].

Definition 1.12 (Basis) *Ein System $(b_1, b_2, \dots, b_n)$ von n Vektoren aus V heißt eine* ***Basis*** *von V, falls das System $b_1, b_2, \dots, b_n$ zugleich linear unabhängig und ein Erzeugendensystem von V ist.*

Beispiel 1.13 (Die Standardbasis des K^n) *Im K^n ist das System*

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

ein Erzeugendensystem, denn

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} &= \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_n \end{pmatrix} \\ &= a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \end{aligned}$$

Aus der obigen Formel folgt zugleich die lineare Unabhängigkeit von $e_1, e_2, \dots, e_n$. (Weshalb?) □

Folglich ist $(e_1, e_2, \dots, e_n)$ eine Basis von K^n, welche wir die Standardbasis[5] des K^n nennen. Insbesondere liegt für $n = 1$ mit $e_1 = 1$ eine Basis des Vektorraums $K = K^1$ vor. Aber auch jedes andere Element $0 \neq b \in K$ ist eine Basis von K^1.

---

[3] Dies bedeutet: **mindestens** ein a_i ist ungleich Null.

[4] Wir werden — üblichem Sprachgebrauch folgend, davon sprechen, dass die Vektoren $a_1, a_2, \dots, a_n$ linear abhängig (bzw. linear unabhängig) sind, obwohl dies missverständlich ist: Sind sämtliche $v_i \neq 0$, so ist $v_1, v_2, \dots, v_n$ ein System von Vektoren, die einzeln jeweils linear unabhängig sind, aber als Folge $v_1, v_2, \dots, v_n$ durchaus linear abhängig sein können.

[5] **Die** Basis eines Vektorraums $V \neq 0$ oder auch des Vektorraums K^n, $n \geq 1$ gibt es nicht, da es in jedem dieser Fälle mehrere (sogar unendlich viele) Basen gibt. Der bestimmte Artikel für die Bezeichnung der Standardbasis des K^n macht dagegen Sinn!

Satz 1.14 *Ist* $(b_1, b_2, \ldots, b_n)$ *eine Basis von* V*, so besitzt jeder Vektor* $v \in V$ *eine eindeutige Darstellung*

$$v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

als Linearkombination von $b_1, b_2, \ldots, b_n$*.*

Beweis. Vgl. auch (6.7). Die **Existenz** einer Linearkombination

$$v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

folgt, weil $b_1, b_2, \ldots, b_n$ ein Erzeugendensystem von V ist.

Zum Nachweis der **Eindeutigkeit** dieser Darstellung nehmen wir an, dass

$$\begin{aligned} v &= a_1.b_1 + a_2.b_2 + \cdots + a_n.b_n \\ v &= a'_1.b_1 + a'_2.b_2 + \cdots + a'_n.b_n \end{aligned}$$

zwei Darstellungen von v als Linearkombination von $b_1, b_2, \ldots, b_n$ sind. Bilden der Differenz führt zur Gleichung

$$0 = (a_1 - a'_1).b_1 + (a_2 - a'_2).b_2 + \cdots + (a_n - a'_n).b_n,$$

woraus — die lineare Unabhängigkeit von $v_1, v_2, \ldots, v_n$ berücksichtigend — das Verschwinden der Koeffizienten, also $a_1 - a'_1 = 0$, $a_2 - a'_2 = 0, \ldots, a_n - a'_n = 0$ folgt. Wir haben damit $a_1 = a'_1$, $a_2 = a'_2, \ldots, a_n = a'_n$ und folglich die Eindeutigkeit der Darstellung gezeigt. □

Wenn wir eine Basis von V haben, kennen wir daher den Vektorraum V vollständig, da wir seine Elemente als eindeutige Linearkombinationen seiner Elemente darstellen können. Im allgemeinen wird ein Vektorraum allerdings **mehrere**, häufig — siehe Anschauungsraum — sogar unendlich viele Basen haben.

Folgerung 1.15 *Ist* $b_1, b_2, \ldots, b_n$ *eine Basis von* V*, so ist die Abbildung*

$$h : K^n \longrightarrow V, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=0}^{n} x_i.b_i$$

bijektiv*, d.h. zu jedem* $v \in V$ *gibt es genau ein* $x \in K^n$ *mit* $h(x) = v$*.* □

Bemerkung 1.16 *(a) Die Vektoren einer Basis* $(b_1, b_2, \ldots, b_n)$ *sind immer paarweise verschieden. Ferner kommt es für die Basiseigenschaft nicht auf die Reihenfolge an, in der die Vektoren einer Basis aufgezählt werden. Immer, wenn es daher bequem ist, werden wir anstelle von* $(b_1, b_2, \ldots, b_n)$ *von der Basis* $B = \{b_1, b_2, \ldots, b_n\}$ *sprechen.*

(b) In $K = K^1$ liefert jeder Vektor ungleich Null, also jede reelle Zahl $b \neq 0$, eine Basis.

(c) Im K^2 bilden — wie wir in Kapitel I gesehen haben — zwei Vektoren

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

genau dann eine Basis, wenn $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 \neq 0$ *gilt.*

(d) Im K^3 bilden — wie wir in Kapitel I gesehen haben — drei Vektoren

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

genau dann eine Basis, wenn die Determinante

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_2 & c_3 \end{vmatrix}$$

nicht Null ist.

Die vorstehenden Kriterien liefern eine **formelmäßige** Ermittlung der linearen Abhängigkeit bzw. Unabhängigkeit. Später wird uns generell die Determinantentheorie für $n \times n$-Matrizen (Kapitel 5) und ein entsprechendes Kriterium für die lineare Unabhängigkeit von n Vektoren des K^n zur Verfügung stehen. Die Berechnung von Determinanten ist für große Formate jedenfalls zeitaufwendig. Andere Verfahren zur Bestimmung der linearen Unabhängigkeit — etwa Varianten des **Gauß-Algorithmus** — führen für Spaltenvektoren in der Regel sehr viel schneller zum Ziel und sind entsprechend vorzuziehen.

3.2 Unterräume und Lineare Hülle

Definition 2.1 *Eine Teilmenge U eines R-Vektorraums V heißt* ***Unterraum*** *von V, wenn gilt:*

(U 1) $0 \in U$.

(U 2) $U + U \subseteq U$, *d.h.* $x,\, y \in U \Rightarrow x + y \in U$.

(U 2) $K.U \subseteq U$, *d.h.* $x \in U$ *und* $a \in K$ *impliziert* $a.x \in U$.

Dabei verwenden wir die folgenden abkürzenden Bezeichnungen:

$$\begin{aligned} U + U &= \{u_1 + u_2 \mid u_1 \in U_1 \text{ und } u_2 \in U_2\}, \\ K.U &= \{r.u \mid r \in K \text{ und } u \in U\}, \\ -U &= \{-u \mid u \in U\}. \end{aligned}$$

Donnerstag, 11. Dezember 03

Satz 2.2 *Der Name Unter***raum** *ist gerechtfertigt, denn jeder Unterraum U von V ist bzgl. der Verknüpfungen*

$$\begin{aligned} + &: U \times U \longrightarrow U, \quad (u,v) \mapsto u + v \\ . &: K \times U \longrightarrow U, \quad (a,u) \mapsto a.u \end{aligned}$$

wieder ein Vektorraum.

Beweis. Wegen **(U 1)** und **(U 2)** machen beide Verknüpfungen Sinn, sie sind also — wie man sagt — **wohldefiniert**. Die Bedingungen **(A 1)**, **(A 2)** eines Vektorraums sind offensichtlich und **(A 3)** ist wegen **(U 1)** erfüllt. Wegen **(U 3)** ist insbesondere $-U \subseteq U$, d.h. $x \in U$ impliziert $-x \in U$. Somit gilt auch **(A 4)**. Schließlich gelten **(M 1)** – **(M 4)** in U, da sie im umfassenden Vektorraum V gelten. □

Beispiele 2.3 (a) Für jeden Vektorraum V sind stets V selbst und $\{0\}$ Unterräume von V.

(b) Ist ferner v ein beliebiger Vektor aus V, so ist $K.v := \{a.v \mid a \in K\}$ ein Unterraum von V. Wir nennen ihn den von v **erzeugten Unterraum**.

(b) V sei der Anschauungsraum. Die nur aus 0 bestehende Teilmenge $\{0\}$, eine Gerade G durch 0 oder eine Ebene E durch 0 sind Beispiele für Unterräume von V. Im Zusammenhang mit der Diskussion des Dimensionsbegriffs werden wir später sehen, dass mit dieser Aufzählung alle Unterräume U des Anschauungsraumes (gleichbedeutend des $\mathbb{R}^3$) erfasst sind.

Weitere Unterräume eines K-Vektorraums V können wir uns nach folgendem Muster verschaffen:

Satz 2.4 *Sind $v_1, v_2, \ldots, v_t \in V$, so ist*

$$\langle v_1, v_2, \ldots, v_t\rangle := \{a_1.v_1 + \cdots + a_n.v_n \mid a_1, a_2, \ldots, a_n \in K\}$$

ein Unterraum von V, welchen wir den von $v_1, v_2, \ldots, v_t$ ***aufgespannten***[6] ***Unterraum*** *oder auch die* ***lineare Hülle*** *von $v_1, v_2, \ldots, v_t$ oder $\{v_1, v_2, \ldots, v_t\}$ nennen.*

Beweis. Es ist klar, dass $0 = 0.v_1 + 0.v_2 + \cdots + 0.v_t$ in $H = \langle v_1, v_2, \ldots, v_t\rangle$ liegt und dass H ferner wegen

$$\begin{aligned}(a_1.v_1 + \cdots + a_t.v_t) + (b_1.v_1 + \cdots + b_t.v_t) &= (a_1 + b_1).v_1 + \cdots + (a_t + b_t).v_t \\ a.\,(a_1.v_1 + \cdots + a_t.v_t) &= (aa_1).v_1 + \cdots + aa_t).v_t\end{aligned}$$

gegen Bildung von Summen und von Produkten mit Skalaren abgeschlossen ist. H ist somit ein Unterraum von V. $\square$

Satz 2.5 *Es ist $\langle v_1, v_2, \ldots, v_t\rangle$ der bezüglich Inklusion "$\subseteq$" kleinste Unterraum von V, welcher $v_1, v_2, \ldots, v_t$ enthält.*

Beweis. Wir wissen schon, dass $H = \langle v_1, v_2, \ldots, v_t\rangle$ ein Unterraum ist. Da $v_i = 0.v_1 + \cdots + 1.v_i + \cdots + 0.v_t$ gilt, liegen die Elemente $v_1, v_2, \ldots, v_t$ sämtlich in H. Ist nun U irgendein Unterraum von V, welcher $v_1, v_2, \ldots, v_t$ enthält, so liegen alle Linearkombinationen $a_1.v_1 + \cdots + a_t.v_t$ ebenfalls in U, woraus $H \subseteq U$ folgt. $\square$

Für spätere Verwendung notieren wir noch:

Satz 2.6 *Sind U_1 und U_2 Unterräume von V, so auch ihr Durchschnitt $U_1 \cap U_2 = \{u \mid u \in U_1$ und $u \in U_2\}$ und ihre Summe $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\ u_2 \in U_2\}$.*

Beweis. Wir zeigen, dass $U = U_1 \cap U_2$ ein Unterraum von V ist. Da 0_V sowohl in U_1 als auch in U_2 liegt, so folgt $0_V \in U$ und (U 1) ist erfüllt. Seien nun x und y Elemente aus U. Insbesondere sind dann x und y Elemente von U_1 und folglich liegt auch ihre Summe $x + y$ in U_1. Entsprechend liegt $x + y$ in U_2. Es folgt, dass $x + y$ in U liegt und somit (U 2) erfüllt ist. Seien schließlich r ein Skalar und u ein Mitglied von U. Dann liegt mit u auch $r.u$ in U_1. Entsprechend folgt $r.u \in U_2$, folglich liegt $r.u$ in U, womit auch (U 3) gezeigt ist.

Mit analogen Argumenten behandelt man den Fall $U_1 + U_2$. $\square$

Bemerkung 2.7 *Es seien U_1 und U_2 Unterräume von V. Im allgemeinen ist die Vereinigung $U_1 \cup U_2$ nicht wieder ein Unterraum*[7] *von V. Beispielsweise sind im $\mathbb{R}^2$ die Koordinatenachsen $U_1 = \{(x, 0) \mid x \in \mathbb{R}\}$ und $U_2 = \{(0, y) \mid y \in \mathbb{R}\}$ Unterräume, aber ihre Vereinigung $U_1 \cup U_2$ ist kein Unterraum des $\mathbb{R}^2$.*

---

[6] Auch die Bezeichnung 'von $v_1, v_2, \ldots, v_t$ erzeugter Unterraum' ist gebräuchlich.

[7] Genauer gilt hier: Die Vereinigung $U_1 \cup U_2$ ist genau dann ein Unterraum, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ gilt.

Einschub: Vollständige Induktion

Wir werden im folgenden häufig Beweise durch vollständige Induktion führen und schieben daher einen Exkurs über das Prinzip der vollständigen Induktion ein. Wir werden dabei sehen, dass es sich bei demselben um eine besondere Eigenschaft der natürlichen Zahlen handelt, die wir nachfolgend herausarbeiten.

Die natürlichen Zahlen

Für uns reicht die Vorstellung, dass die **natürlichen Zahlen** diejenigen sind, die man zum Zählen und daher zur Anzahlbestimmung endlicher Mengen verwendet. Es sind dies also die Zahlen

$$0, 1, 2, 3, 4, 5, \ldots$$

Wie schon gelegentlich zuvor, bezeichnen wir mit $\mathbb{N}$ die Menge aller natürlichen Zahlen[8]. Wir setzen als bekannt voraus, wie man hinsichtlich Addition und Multiplikation mit natürlichen Zahlen umgeht und wie man dieselben der Größe nach vergleicht ($m \leq n$, $m < n$).

Das Prinzip der kleinsten natürlichen Zahl

Die bei weitem wichtigste Eigenschaft der natürlichen Zahlen ist das sehr anschauliche, und daher hier nicht weiter begründete

Prinzip der kleinsten natürlichen Zahl: Gegeben sei irgendeine *nichtleere* Menge M von natürlichen Zahlen (M darf dabei endlich oder unendlich sein, muss aber, wie verlangt, mindestens ein Mitglied enthalten). Dann gibt es unter allen Zahlen von M eine **kleinste.**

Formelmäßig ausgedrückt: Falls $\emptyset \neq M \subseteq \mathbb{N}$, so existiert eine Element $m_0 \in M$ mit $m_0 \leq m$ für jedes $m \in M$.

Die Problemstellung

Wir wollen auf einen Streich unendlich viele Behauptungen (Aussagen) zu beweisen. Wir stellen uns dazu vor, dass wir **jeder** natürlichen Zahl $n \geq 1$ eine Aussage $A(n)$ zugeordnet haben, etwa die Behauptung

$$A(n): \quad 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

Den Beweis von $A(1), A(2), \ldots$ wollen wir **mit einem einzigen Beweis** erledigen! Dies gelingt mit dem Prinzip der vollständigen Induktion, das wir gleich aus dem Prinzip der kleinsten natürlichen Zahl gewinnen werden.

[8] Eine andere Tradition reserviert das Symbol $\mathbb{N}$ für die Menge der natürlichen Zahlen ≥ 1. In diesem Fall wird das Symbol $\mathbb{N}_0$ verwendet, um die 0 einzubeziehen.

Prinzip der vollständigen Induktion

Satz (Vollständige Induktion) *Für jede*[9] *natürliche Zahl n sei eine Behauptung $A(n)$ vorgelegt. Es gelte*

(I 1) *$A(0)$ ist wahr.*

(I 2) *Immer, wenn $A(n)$ wahr ist, ist auch $A(n+1)$ wahr.*

Dann ist $A(n)$ wahr für ***jede*** *natürliche Zahl n.*

Beweis. Wir führen einen sogenannten **Widerspruchsbeweis**. Dazu **nehmen wir an**, dass es eine natürliche Zahl m gibt, für die $A(m)$ **falsch** ist und zeigen anschließend, dass diese Annahme zu einem **Widerspruch** führt!

Nach Annahme ist die Menge V aller "Verbrecher" v, für die $A(v)$ falsch ist, nicht leer. Sie enthält nämlich mindestens das Element m. Nach dem Prinzip der kleinsten natürlichen Zahl hat V folglich ein kleinstes Mitglied v_0; es gibt somit einen *kleinsten Verbrecher.*

Wegen (I 1) ist $v_0 \neq 0$, daher $v_0 > 0$. Somit ist $n = v_0 - 1 \geq 0$ eine natürliche Zahl, die kleiner ist als v_0 und daher **nicht** in V liegt. Folglich ist $A(n)$ wahr. Wegen der Richtigkeit von $A(n)$ ist dann wegen (I 2) auch die Aussage $A(n+1) = A(v_0)$ wahr, **Widerspruch!**

Unsere **ursprüngliche Annahme ist daher falsch**, das behauptete Induktionsprinzip damit bewiesen. □

Anwendungsbeispiel

Wir zeigen durch vollständige Induktion, dass für jedes $n \geq 1$ die Behauptung

$$A(n): \quad 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

richtig ist.

Induktionsverankerung: Es gilt $A(0)$[10].

Induktionsschritt von n auf $n+1$: Wir nehmen an, dass $A(n)$ richtig ist, d.h. wir nehmen an, dass (für dieses n) die Formel

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \quad \text{gilt.}$$

[9] Es ist auch möglich, die Behauptungen $A(n)$ nur für alle natürlichen $n \geq n_0$ zu betrachten. In diesem Fall ist (I 1) zu "$A(n_0)$ ist wahr" zu modifizieren.

[10] Hier verwenden wir die Vereinbarung, dass eine Summe von null Summanden gleich 0 ist.

Wir zeigen, dass unter dieser Voraussetzung auch $A(n+1)$ gilt:

$$\begin{aligned}(1+2+3+\cdots+n)+(n+1) &\overset{I.V.}{=} \frac{n(n+1)}{2}+(n+1)\\ &= \frac{n(n+1)+2(n+1)}{2}\\ &= \frac{(n+1)(n+2)}{2}\end{aligned}$$

Dies ist gerade die Aussage $A(n+1)$. □

Variante des Induktionsprinzips

Für praktische Zwecke ist häufig die folgende Fassung des Induktionsprinzips nützlich:

Satz 3.1 (Induktion, Variante) *Für jede natürliche Zahl $n \geq n_0$ sei eine Aussage $A(n)$ gegeben. Es gelte*

(I 1)* *$A(n_0)$ ist wahr.*

(I 2)* *Es sei $m > n_0$. Falls $A(k)$ für alle $n_0 \leq k < m$ wahr ist, so ist auch $A(m)$ wahr.*

Dann ist $A(n)$ für jedes $n \geq n_0$ wahr.

Beweis. Wir nehmen an, unsere Behauptung sei falsch. Die Menge V aller natürlichen Zahlen (Verbrecher) $n \geq n_0$, für die $A(n)$ falsch ist, *ist dann nicht leer*. Nach dem Prinzip der kleinsten natürlichen Zahl hat V ein kleinstes Mitglied v (einen *kleinsten Verbrecher*). Wegen (I 1)* ist $v > n_0$. Für alle k mit $n_0 \leq k < v$ ist daher $A(k)$ wahr, wegen (I 2)* ist dann auch $A(v)$ wahr, **Widerspruch!**

Damit ist unsere ursprüngliche Annahme falsch und die obige Behauptung bewiesen. □

Rückblick: Widerspruchsbeweis

Sie haben bemerkt, dass das generelle Beweisschema beider Induktionsprinzipien ziemlich ähnlich ist: in beiden Fällen haben wir zudem einen **Beweis durch Widerspruch** geführt.

Dasselbe beruht darauf, dass *jede Aussage* entweder *wahr* oder *falsch* ist und ferner sich aus einer wahren Aussage durch zulässige logische Schlüsse stets wieder eine wahre Aussage ergibt.

Nehmen wir hypothetisch an, eine zu beweisende Aussage A *sei falsch* und ferner, dass sich aus dieser Annahme durch zulässige Schlüsse eine Aussage B ergibt, die falsch ist. Dann kann unsere Annahme (A *sei falsch*) nicht wahr sein; *A ist somit wahr.*

3.3 Austauschsatz, Basisergänzungssatz und Dimension

Montag, 15. Dezember 2003

Es sei V ein Vektorraum. Jedes Teilsystem eines linear unabhängigen Systems von V ist dann wieder linear unabhängig in V. Ferner führt jede Erweiterung eines Erzeugendensystems von V wieder zu einem Erzeugendensystem von V. Wir nennen ein Erzeugendensystem $v_1, v_2, \ldots, v_n$ von V **unverkürzbar**, falls kein echtes Teilsystem von $v_1, v_2, \ldots, v_n$ ein Erzeugendensystem von V ist. Entsprechend nennen wir eine linear unabhängiges System $v_1, v_2, \ldots, v_n$ **maximal linear unabhängig**, falls jede echte Erweiterung linear abhängig ist.

Um die Unverkürzbarkeit eines Erzeugendensystems $v_1, v_2, \ldots, v_n$ einzusehen, reicht dabei der Nachweis, dass jedes durch Weglassen eines einzigen Mitglieds v_i entstehende System $v_1, v_2, \ldots, v_i, \widehat{v_i}, v_{i+1}, \cdots, v_n$ kein Erzeugendensystem von V ist. Für den Nachweis der maximalen lineare Unabhängigkeit von $v_1, v_2, \ldots, v_n$ reicht es zu zeigen, dass jedes Hinzufügen eines einzigen Vektors v zu einem linear abhängigen System $v_1, v_2, \ldots, v_n, v$ führt.

Satz 3.1 (Kennzeichnung von Basen) *Für ein System $v_1, v_2, \ldots, v_n$ von Vektoren von V sind äquivalent:*

(*a*) $(v_1, v_2, \ldots, v_n)$ *ist eine* Basis *von V.*
(*b*) $(v_1, v_2, \ldots, v_n)$ *ist ein* unverkürzbares Erzeugendensystem[11] *von V.*
(*c*) $(v_1, v_2, \ldots, v_n)$ *ist* maximal linear unabhängig *in V.*
(*d*) *Jedes $v \in V$ lässt sich aus $v_1, v_2, \ldots, v_n$ eindeutig linear kombinieren.*

Beweis. Wir orientieren den Nachweis am folgenden Schema

$$\begin{array}{ccccc} & & (b) & & \\ & & \Updownarrow & & \\ (d) & \Leftrightarrow & (a) & \Leftrightarrow & (c). \end{array}$$

Dabei ist uns die Implikation $(a) \Rightarrow (d)$ schon aus Satz 1.14 bekannt. Wir setzen nun (d) voraus. Ersichtlich ist dann $v_1, v_2, \ldots, v_n$ ein Erzeugendensystem von V. Ferner besitzt der Nullvektor wegen (d) nur die Darstellung $0 = 0.v_1 + 0.v_2 + \cdots + 0.v_n$, woraus die lineare Unabhängigkeit von $v_1, v_2, \ldots, v_n$ und dann auch die Basiseigenschaft, somit (a), folgt.

$(a) \Rightarrow (c)$: Als Basis ist $v_1, v_2, \ldots, v_n$ natürlich linear unabhängig. Ist dann v irgendein ein Vektor aus V, so können wir v als Linearkombination $v = a_1.v_1 + \cdots + a_n.v_n$ schreiben, woraus sich wegen $a_1.v_1 + \cdots + a_n.v_n + (-1).v = 0$ die lineare Abhängigkeit von $v_1, v_2, \ldots, v_n, v$ ergibt. Das System $v_1, v_2, \ldots, v_n$ ist daher maximal linear unabhängig.

[11] Auch die Bezeichnung *minimales Erzeugendensystem* ist gebräuchlich.

$(c) \Rightarrow (a)$: Wenn $v_1, v_2, \ldots, v_n$ ein maximal linear unabhängiges System in V ist, so ist zur Basiseigenschaft noch zu zeigen, dass sich jeder Vektor v aus $v_1, v_2, \ldots, v_n$ linear kombinieren lässt. Nach Voraussetzung ist $v_1, v_2, \ldots, v_n, v$ linear abhängig. Es gibt daher eine lineare Relation $a_1.v_1 + \cdots + a_n.v_n + a.v = 0$, bei der nicht alle Koeffizienten verschwinden. Dabei kann a nicht gleich Null sein, da sonst die Vektoren $v_1, v_2, \ldots, v_n$ linear abhängig wären. Wegen $a \neq 0$ lässt sich die obige Beziehung daher nach v auflösen, was die Behauptung zeigt.

$(a) \Rightarrow (b)$: Als Basis ist zunächst $v_1, v_2, \ldots, v_n$ auch ein Erzeugendensystem. Wir nehmen an, dass es verkürzbar ist, also durch Weglassen sagen wir des i-ten Mitglieds ein Erzeugendensystem $v_1, \ldots, \widehat{v_i}, \ldots, v_n$ verbleibt[12]. Wir können dann den ausgelassenen Vektor v_i als Linearkombination $v_i = a_1.v_1 + \ldots + a_{i-1}.v_{i-1} + a_{i+1}.v_{i+1} + \cdots + a_n.v_n$ der übrigen Vektoren darstellen. Bringen wir alle Terme auf eine Seite, so erschließt sich die lineare Abhängigkeit von $v_1, v_2, \ldots, v_n$, Widerspruch. Damit ist $v_1, v_2, \ldots, v_n$ ein unverkürzbares Erzeugendensystem.

$(b) \Rightarrow (a)$: Wir müssen noch die lineare Unabhängigkeit eines unverkürzbaren Erzeugendensystems $v_1, v_2, \ldots, v_n$ zeigen. Wir nehmen dazu $a_1.v_1 + \cdots + a_i.v_i + \cdots + a_n.v_n = 0$ und $a_i \neq 0$ an. Durch Auflösen dieser Gleichung nach v_i ergibt sich, dass v_i eine Linearkombination von $v_1, \ldots, \widehat{v_i}, \ldots, v_n$ ist. Die Vektoren $v_1, v_2, \ldots, v_n$ — und damit alle Vektoren aus V — liegen somit sämtlich in der linearen Hülle $\langle v_1, \ldots, \widehat{v_i}, \ldots, v_n \rangle$. Dies widerspricht der vorausgesetzten Minimalität des Erzeugendensystems $v_1, v_2, \ldots, v_n$. □

Bemerkung 3.2 Es reicht nicht, nur eine einzige der vier Kennzeichnungen von Basen zu kennen.

Je nach Sachlage ist die eine oder die andere von ihnen — oft sogar sehr viel — vorteilhafter; wir sollten daher alle vier Kennzeichnungen kennen und sachgerecht anwenden können.

Donnerstag, 18. Dezember 2003

Folgerung 3.3 *Jedes endliche Erzeugendensystem eines Vektorraums V enthält eine Basis von V.*

Beweis. Wir lassen solange Elemente eines endlichen Erzeugendensystems weg, bis ein minimales Erzeugendensystem entsteht[13]. □

Wir machen hier auf eine weit verbreitete Fehlinterpretation von Satz 3.1 aufmerksam. Die dort angeführte Kennzeichnung einer Basis $B = \{v_1, v_2, \ldots, v_n\}$ als

[12] Auch später werden wir — wo dies nicht zu Mißverständnissen führt — durch $\widehat{\ }$ das Weglassen des Elements aus einer Folge bezeichnen.

[13] Im Einklang mit den getroffenen Definitionen ist die leere Menge Basis jedes nur aus dem Nullvektor bestehenden Vektorraums $\{0\}$.

minimales (unverkürzbares) Erzeugendensystem bezieht sich nicht auf die *Anzahl der Basiselemente*, sondern auf die Minimalität von B bezüglich Inklusion $\subseteq$. Entsprechendes gilt für Eigenschaft einer Basis *maximal* linear unabhängig zu sein. Der Satz taugt daher nicht zu einer Begründung, dass je zwei Basen B und B' von V dieselbe Mitgliederzahl haben. Dieses zutreffende Faktum herzuleiten, erfordert weitere Argumente (Austauschsatz), die wir gleich im Anschluss behandeln.

Lemma 3.4 (Austauschlemma) *Sei $v_1, v_2, \ldots, v_n$ eine Basis von V und w ein von Null verschiedener Vektor aus V, den wir als Linearkombination*

$$(*) \qquad w = a_1.v_1 + \cdots + a_i.v_i + \cdots + a_n.v_n$$

der Basis $v_1, v_2, \ldots, v_n$ darstellen. Falls $a_i \neq 0$, entsteht durch Austausch von w mit v_i aus der Basis $v_1, v_2, \ldots, v_n$ eine neue Basis $v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n$ von V.

Beweis. Wegen $a_i \neq 0$ lässt sich $(*)$ nach v_i auflösen. Es ergibt sich v_i als Linearkombination von $v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n$. Folglich liegen alle v_j, $j = 1, \ldots, n$, im Unterraum $\langle v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n \rangle$ und damit auch $V = \langle v_1, v_2, \ldots, v_n \rangle$. Das System $v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n$ ist folglich ein Erzeugendensystem von V.

Um die lineare Unabhängigkeit des Systems zu zeigen, nehmen wir nun an, dass

$$c_1.v_1 + \cdots + c_{i-1}.v_{i-1} + c.w + c_{i+1}.v_{i+1} + \cdots + c_n.v_n = 0$$

ist. Falls $c = 0$ ist, erhalten wir aus der linearen Unabhängigkeit von $v_1, v_2, \ldots, v_n$ das Verschwinden aller c_j. Falls $c \neq 0$ ist, können wir w als Linearkombination $w = a'_1.v_1 + \cdots + 0.v_i + \cdots + a'_n.v_n$ darstellen, was der Eindeutigkeit der Darstellung von w in der Basis $v_1, v_2, \ldots, v_n$ widerspricht. $\square$

Satz 3.5 (Austauschsatz) *Es sei $v_1, v_2, \ldots, v_n$ sei eine Basis von V. Ist zudem $w_1, w_2, \ldots, w_r$ ein in V linear unabhängiges System, so können wir durch Umnummerierung der $v_1, v_2, \ldots, v_n$ erreichen, dass durch Austausch von $w_1, w_2, \ldots, w_r$ mit den ersten r Elementen der Basis $v_1, v_2, \ldots, v_r$ eine neue Basis*

$$w_1, w_2, \ldots, w_r, v_{r+1}, \ldots, v_n$$

von V entsteht. Insbesondere ist $r \leq n$.

Beweis durch **Induktion nach** r.
Induktionsverankerung: Der Fall $r = 0$ ist klar.
Induktionsschritt: Nun sei $r \geq 1$ und per Induktionsannahme der Satz für jedes System von $r - 1$ auszutauschenden Vektoren gültig. Wir können daher — nach Umnummerierung der v_j — das linear unabhängige System $w_1, \ldots, w_{r-1}$ zu einer

Basis der Form $w_1, \ldots, w_{r-1}, v_r, \ldots, v_n$ von V ergänzen, und insbesondere w_r in der Form

$$w_r = a_1.w_1 + \cdots + a_{r-1}.w_{r-1} + a_r.v_r + \cdots + a_n.v_n$$

darstellen. Dabei können $a_r, \ldots, a_n$ nicht alle 0 sein, da andernfalls die Vektoren $w_1, w_2, \ldots, w_r$ linear abhängig wären. Nach erneutem Umnummerieren können wir daher $a_r \neq 0$ annehmen. Per Austauschlemma können wir dann den Vektor v_r der aktuellen Basis gegen w_r austauschen. Es folgt, dass $w_1, w_2, \ldots, w_r, v_{r+1}, \ldots, v_n$ eine Basis von V ist. □

Folgerung 3.6 (Invarianz der Dimension) *Je zwei endliche Basen von V haben dieselbe Anzahl von Mitgliedern.*

Beweis. Sind $v_1, v_2, \ldots, v_n$ und $w_1, w_2, \ldots, w_m$ endliche Basen von V, so folgt aus dem Austauschsatz $n \leq m$ und $m \leq n$, also $n = m$. □

Dieses Faktum erst ermöglicht die Definition der Dimension von Vektorräumen, die sich überhaupt als wichtigste Eigenschaft eines Vektorraums herausstellen wird.

Definition 3.7 (Dimension) *Sei V ein endlichdimensionaler Vektorraum. Die Anzahl der Mitglieder einer Basis von V heißt die* ***Dimension*** *von V. Schreibweise:* $\dim_K V$ *oder* $\dim V$.

Beispiel 3.8 Im K^n bilden die "Einheitsvektoren" $e_1, e_2, \ldots, e_n$ die Standardbasis. Somit ist $\dim K^n = n$.

Einen Spezialfall der folgenden Aussage haben wir mit dem Dimensionsaxiom der anschaulichen Vektorrechnung (6.6) schon kennengelernt.

Satz 3.9 *Ist V ein Vektorraum der Dimension n, so sind je $n+1$ Vektoren $x_1, x_2, \ldots, x_n, x_{n+1}$ linear abhängig.* □

Satz 3.10 (Basisergänzungssatz) *V sei ein endlichdimensionaler Vektorraum. Jeder Unterraum U von V ist dann ebenfalls endlichdimensional. Ferner lässt sich jede Basis $w_1, w_2, \ldots, w_r$ von U zu einer Basis*

$$w_1, w_2, \ldots, w_r, v_{r+1}, \ldots, v_n$$

von V ergänzen.

Beweis. Sei $n = \dim V$, Wegen des Austauschsatzes hat jedes (endliche) linear unabhängige System $w_1, w_2, \ldots, w_r$ von U höchstens $r \leq n$ Mitglieder. Wir können daher annehmen, dass r maximal gewählt ist. Als maximal linear unabhängiges System von U ist $w_1, w_2, \ldots, w_r$ nach Satz 3.1 eine Basis von U.

Per Austauschsatz 3.5 lässt sich ferner das in V linear unabhängige System $w_1, w_2, \ldots, w_r$ zu einer Basis der Form $w_1, w_2, \ldots, w_r, v_{r+1}, \ldots, v_n$ von V ergänzen. □

Satz 3.11 (Dimension von Unterräumen) *Ist U ein Unterraum des endlichdimensionalen Vektorraums V so folgt* $\dim U \leq \dim V$. *Gleichheit* $\dim U = \dim V$ *gilt genau dann wenn* $U = V$.

Beweis. Sei $w_1, w_2, \dots, w_r$ eine Basis von U. Mittels Basisergänzungssatz können wir dieselbe zu einer Basis $w_1, w_2, \dots, w_r, v_{r+1}, \dots, v_n$ von V ergänzen. Aus $r = \dim U = \dim V = n$ folgt dann, dass $w_1, w_2, \dots, w_r$ zugleich eine Basis von U und von V ist, woraus durch Übergang zur linearen Hülle $U = \langle w_1, w_2, \dots, w_r \rangle = V$ folgt. □

Satz 3.12 (Dimensionsformel für Unterräume) *Sind U_1, U_2 Unterräume des endlichdimensionalen Vektorraums V, so gilt*

$$\dim U_1 + \dim U_2 = \dim U_1 \cap U_2 + \dim U_1 + U_2.$$

Beweis. Wir veranschaulichen die Verhältnisse durch das Hasse-Diagramm

$$\begin{tikzcd}[arrows=dash]
 & U_1 + U_2 & \\
U_1 \arrow[ur] & & U_2, \arrow[ul] \\
 & U_1 \cap U_2 \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

wobei die Verbindungslinien die vorhandenen Inklusionen angeben. Zum Beweis der Dimensionsformel starten mit einer Basis $e_1, e_2, \dots, e_s$ von $U_1 \cap U_2$, dem kleinsten der beteiligten Unterräume. Wir können dann $e_1, e_2, \dots, e_s$ sowohl zu einer Basis $e_1, e_2, \dots, e_s$, $f_1, f_2, \dots, f_p$ von U_1 als auch zu einer Basis $e_1, e_2, \dots, e_s$, $g_1, g_2, \dots, g_q$ von U_2 ergänzen. Wir behaupten nun, dass die in dieser Konstruktion insgesamt auftretenden Vektoren

$$(*) \qquad e_1, e_2, \dots, e_s,\ f_1, f_2, \dots, f_p,\ g_1, g_2, \dots, g_q$$

eine Basis von $U_1 + U_2$ bilden, woraus sich die behauptete Formel sofort ergibt.

Montag, 5. Januar 2004

Jedes Element aus U_1 (bzw. U_2) lässt sich aus $e_1, e_2, \dots, e_s$ und $f_1, f_2, \dots, f_p$ (bzw. aus $e_1, e_2, \dots, e_s$ und $g_1, g_2, \dots, g_q$) linear kombinieren. Jedes $x \in U_1 + U_2$ ist daher eine Linearkombination der Elemente $e_1, e_2, \dots, e_s$, $f_1, f_2, \dots, f_p$ und $g_1, g_2, \dots, g_q$. Damit ist $(*)$ ein Erzeugendensystem von $U_1 + U_2$.

Zum Nachweis der linearen Unabhängigkeit von $(*)$ betrachten wir eine verschwindende Linearkombination

$$\sum_{i=1}^{s} a_i.e_i + \sum_{j=1}^{p} b_j.f_j + \sum_{k=1}^{q} c_k.g_k = 0.$$

Es folgt, dass das Element

$$\sum_{i=1}^{s} a_i.e_i + \sum_{j=1}^{p} b_j.f_j = \sum_{k=1}^{q} (-c_k).g_k$$

sowohl in U_1 (betrachte dazu die linke Seite des Ausdrucks) als auch in U_2 (betrachte dazu die rechte Seite) und folglich in $U_1 \cap U_2$ gelegen ist. Daher ist zunächst $c_k = 0$ für $k = 1, \ldots, q$ und dann wegen der linearen Unabhängigkeit der Basis $e_1, e_2, \ldots, e_s, f_1, f_2, \ldots, f_p$ von U_1 auch $a_i = 0$ $(i = 1, \ldots, s)$ und $b_j = 0$ $(j = 1, \ldots, p)$. Dies zeigt die lineare Unabhängigkeit von $e_1, e_2, \ldots, e_s$, $f_1, f_2, \ldots, f_p$, $g_1, g_2, \ldots, g_q$. □

Anwendung 3.13 (Schnitt von Ebenen) **(a)** U_1 und U_2 seien zweidimensionale Unterräume des K^3, also Ebenen durch 0. Falls die beiden Ebenen nicht übereinstimmen, ist ihr Schnittgebilde eine Gerade durch 0. Wegen der Dimensionsformel sind nämlich nur die Möglichkeiten $U_1 = U_2$ bzw. $\dim U_1 \cap U_2 = 1$ möglich.

(b) In höheren Dimensionen, so dem K^4, kann es dagegen vorkommen, dass zwei Ebenen durch 0 nur den Nullpunkt gemeinsam haben. Als Beispiel betrachten wir die zweidimensionalen Unterräume $U_1 = K.e_1 + K.e_2$ und $U_2 = K.e_3 + K.e_4$ mit dem Durchschnitt $U_1 \cap U_2 = \{0\}$. Wie früher bezeichnet hier e_1, e_2, e_3, e_4 die Standardbasis des K^4.

3.4 Das kartesische (direkte) Produkt

Satz 4.1 *Sind V_1, V_2 Vektorräume, so ist das kartesische Produkt $V_1 \times V_2 = \{(v_1, v_2) \mid v_1 \in V_1, v_2 \in V_2\}$ bzgl. der komponentenweisen Definition*

$$\begin{aligned}(v_1, v_2) + (w_1, w_2) &= (v_1 + w_1, v_2 + w_2) \\ a.(v_1, v_2) &= (a.v_1, a.v_2)\end{aligned}$$

von Addition und Multiplikation mit Skalaren wieder ein Vektorraum.

Beweis. durch einfaches Nachrechnen. □

Der Vektorraum $V_1 \times V_2$ heißt das **kartesische Produkt** oder gleichbedeutend das **direkte Produkt** von V_1 und V_2.

Entsprechend bilden wir das direkte Produkt $V_1 \times V_2 \times \cdots \times V_n$ für mehrere Faktoren $V_1, V_2, \ldots, V_n$. Gilt zusätzlich noch $V_1 = V_2 = \cdots = V_n =: V$, so ist hierfür die Potenzschreibweise $V \times \cdots \times V = V^n$ gebräuchlich.

Der schon früher besprochene K^n ist somit ein Spezialfall des direkten Produkts von Vektorräumen[14].

Satz 4.2 (Dimensionsformel direktes Produkt) *Sind V_1, V_2 Vektorräume, so gilt* $\dim V_1 \times V_2 = \dim V_1 + \dim V_2$.

Beweis. Wir zeigen etwas mehr: Sind $e_1, e_2, \ldots, e_n$ und $f_1, f_2, \ldots, f_m$ Basen von V_1 (bzw. von V_2), so ist das System

$$(*) \qquad (e_1, 0), \ldots, (e_n, 0), (0, f_1), \ldots, (0, f_m)$$

eine Basis von $V_1 \times V_2$. Sind nämlich $v_1 \in V_1$ und $v_2 \in V_2$, so gibt es Darstellungen

$$\begin{aligned}v_1 &= a_1.e_1 + \cdots + a_n.e_n \\ v_2 &= b_1.f_1 + \cdots + b_m.f_m\end{aligned}$$

von v_1 und v_2 als Linearkombinationen, woraus

$$\begin{aligned}(v_1, v_2) &= (v_1, 0) + (0, v_2) \\ &= (a_1.e_1 + \cdots + a_n.e_n, 0) + (0, b_1.f_1 + \cdots + b_m.f_m) \\ &= a_1.(e_1, 0) + \cdots + a_n.(e_n, 0) + b_1.(0, f_1) + \cdots + b_m(0, f_m)\end{aligned}$$

folgt. Das System (*) ist folglich ein Erzeugendensystem von $V_1 \times V_2$.

Das System (*) ist auch linear unabhängig. Aus einer linearen Relation

$$a_1.(e_1, 0) + \cdots + a_n.(e_n, 0) + b_1.(0, f_1) + \cdots + b_m(0, f_m) = 0$$

[14] Es kommt für die jetzige Betrachtung ersichtlich nicht darauf an, ob wir Elemente als Zeilen oder Spalten schreiben.

folgt nämlich unter Verwendung der obigen Formeln

$$(a_1.e_1 + \cdots + a_n.e_n, b_1.f_1 + \cdots + b_m.f_m) = (0, 0),$$

woraus sich die beiden Gleichungen

$$\begin{aligned} a_1.e_1 + \cdots + a_n.e_n &= 0 \\ b_1.f_1 + \cdots + b_m.f_m &= 0 \end{aligned}$$

ergeben. Da $e_1, e_2, \ldots, e_n$ und $f_1, f_2, \ldots, f_m$ linear unabhängig sind, folgt dann $a_1 = a_2 = \cdots = a_n = 0$ und $b_1 = b_2 = \cdots = b_m = 0$, somit die lineare Unabhängigkeit von $(*)$. $\square$

3.5 Lineare Abbildungen

Donnerstag, 8. Januar 2004

Lineare Abbildungen benötigen wir, um Vektorräume zu vergleichen. Es sind dies diejenigen Abbildungen zwischen Vektorräumen, die die Vektorraumstruktur repräsentieren. Eigentlich sind lineare Abbildungen sehr viel wichtiger als die Vektorräume selbst! Der Gehalt dieser Aussage wird mit fortschreitender Vorlesung zunehmend klarer ins Blickfeld rücken.

Definition 5.1 *V und W seien Vektorräume. Eine Abbildung $f : V \to W$ heißt* ***linear****, wenn gilt:*

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2) \quad \text{für alle } v_1, v_2 \in V, \textit{und} \\ f(a.v) &= a.f(v) \quad \text{für alle } a \in K \text{ und } v \in V. \end{aligned}$$

Rechenregeln 5.2 *Für eine lineare Abbildung $f : V \to W$ gelten folgende Rechenregeln:*

(1) $f(0_V) = 0_W$,

(2) $f(-v) = -f(v)$, *für alle $v \in V$,*

(3) $f(a_1.v_1 + a_2.v_2) = a_1.f(v_1) + a_2.f(v_2)$, *für alle Skalare a_i und Vektoren $v_i \in V$,*

(4) $f\left(\sum_{i=1}^n a_i.v_i\right) = \sum_{i=1}^n a_i.f(v_i)$ *für alle Skalare a_i und Vektoren $v_i \in V$.*

Beweis. Zu (1): Auf die in V gültige Beziehung $0 + 0 = 0$ wenden wir f an und erhalten in W die Beziehung $f(0) + f(0) = f(0)$, aus der durch Subtraktion von $f(0)$ auf beiden Seiten die Behauptung $f(0) = 0$ folgt[15].

Zu (2): In V gilt $v + (-v) = 0$, woraus sich $f(v) + f(-v) = f(0) = 0$ ergibt. Es folgt $f(-v) = -f(v)$.

Zu (3) und (4): Dies ergibt sich sofort aus der Linearität von f. □

Beispiele 5.3 (1) Die Abbildung $f : V \to W$ mit $f(v) = 0$ für alle $v \in V$ ist linear. Schreibweise $f = 0$.

(2) Die **identische Abbildung** $1_V : V \to V$ mit $1_V(v) = v$ für alle $v \in V$ ist linear.

[15] Wir leiten hier die Eigenschaften (1) und (2) allein aus der Additivität $f(x+y) = f(x)+f(y)$ einer linearen Abbildung her. Heranziehen der Verträglichkeit $f(a.v) = a.f(v)$ mit der skalaren Multiplikation liefert wegen $0.0 = 0$ und $(-1).v = -v$ natürlich erhebliche Beweisabkürzungen für (1) und (2).

(3) Die Abbildung $f : K^3 \to K^2$, $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto \begin{pmatrix} a_2 \\ a_3 \end{pmatrix}$, ist linear.

(4) Die Abbildung $f : K^3 \to K$, $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto a_1 + a_2 + a_3$ ist linear.

(5) Die Abbildung $f : K \to K^3$, $t \mapsto \begin{pmatrix} 0 \\ t \\ t^2 \end{pmatrix}$, ist nicht linear.

(6) Die Abbildung $f : K \to K^3$, $t \mapsto \begin{pmatrix} t \\ t \\ 1+t \end{pmatrix}$, ist nicht linear.

Weitere Eigenschaften 5.4 Die Abbildung $f : V \to W$ sei linear. Dann gilt

(1) Ist U ein Unterraum von V, so ist $f(U)$ ein Unterraum von W, den wir das **Bild von U unter** f nennen.

(2) Es ist $f(V)$ ein Unterraum von W, den wir das **Bild** von f nennen und mit $\mathrm{Bild}(f)$ bezeichnen.

(3) Ist U ein Unterraum von W, so ist $f^{-1}(U)$ ein Unterraum von V. Wir nennen $f^{-1}(U)$ das **Urbild** von U unter f.

(4) Insbesondere ist $f^{-1}(\{0\})$ ein Unterraum von V, den wir als **Kern** von f bezeichnen.

(5) Sind $v_1, v_2, \ldots, v_n$ linear abhängig in V, so sind $f(v_1), \ldots, f(v_n)$ linear abhängig in W.

(6) Sind $f(v_1), \ldots, f(v_n)$ linear unabhängig in W, so sind $v_1, v_2, \ldots, v_n$ linear unabhängig in V. Die Umkehrung gilt nicht!

Beweis. Zu (1) und (2): Wegen $f(0) = 0$ liegt 0 in $f(U)$. Sind ferner zwei Elemente $x = f(u_1)$, $u_1 \in U$, und $y = f(u_2)$, $u_2 \in U$, in $f(U)$ gelegen, so gilt dies auch für $x + y = f(u_1 + u_2)$. Analog liegt für jeden Skalar a mit $x = f(u)$, $u \in U$, auch $a.x = f(a.u)$ in $f(U)$. Folglich ist $f(U)$ ein Unterraum von W.

Zu (3) und (4): Wegen $f(0) = 0 \in U$ liegt 0 in $f^{-1}(U)$. Sind ferner x und y Elemente aus V mit $f(x) \in U$ und $f(y) \in U$, so folgt $f(x+y) = f(x) + f(y) \in U$. Es ist damit $f^{-1}(U)$ gegen Bildung von Summen abgeschlossen. Schließlich gilt für jeden Skalar a und jedes $x \in f^{-1}(U)$ die Beziehung $f(a.x) = a.f(x) \in U$, so dass auch $a.x$ im Urbild von U gelegen ist. Damit ist $f^{-1}(U)$ ein Unterraum von V.

Zu (5): Sind $v_1, v_2, \ldots, v_n$ linear abhängig in V, so gibt es nicht sämtlich verschwindende Skalare $a_1, a_2, \ldots, a_n$ mit $\sum_{i=1}^n a_i.v_i = 0$. Aus der Linearität von f folgt dann

$$\sum_{i=1}^{n} a_i.f(v_i) = f\left(\sum_{i=0}^{n} a_i.v_i\right) = f(0) = 0.$$

Damit sind die Vektoren $f(v_1), f(v_2), \ldots, f(v_n)$ in W linear abhängig.

Zu (6): Bei der ersten Aussage handelt es sich um eine Umformulierung von **(5)**. Schließlich macht die Nullabbildung, die bekanntlich linear ist, aus jedem linear unabhängigen System $v_1, v_2, \ldots, v_n$ ein linear abhängiges System. Dies zeigt auch die zweite Behauptung. □

Satz 5.5 **(1)** *Sind $f : U \to V$ und $g : V \to W$ lineare Abbildungen, so ist die durch*

$$g \circ f : U \to W, \quad u \mapsto g(f(u))$$

*erklärte **Verknüpfung**[16] oder auch **Komposition** von f und g wieder eine lineare Abbildung.*

(2) *Sind $f,\ g : V \to W$ linear, so ist auch die Abbildung $f + g : V \to W$, $v \mapsto f(v) + g(v)$, linear.*

(3) *Ist $f : V \to W$ linear, so ist für jeden Skalar a auch die Abbildung $a.f : V \to W$, $v \mapsto a.f(v)$ linear.*

Beweis. Zu (1): Sei $h = g \circ f$, also $h(u) = g(f(u))$ für alle $u \in U$, so folgt

$$\begin{aligned} h(u_1 + u_2) &= g(f(u_1 + u_2)) = g(f(u_1) + f(u_2)) = h(u_1) + h(u_2) \\ h(r.u) &= g(f(r.u)) = g(r.f(u)) = r.(g(f(u)) = r.h(u) \end{aligned}$$

für alle $u, u_1, u_2 \in U$ und $r \in K$.
Zu (2): Sei $h = f + g$, also $h(v) = f(v) + g(v)$ für alle $v \in V$, so folgt

$$\begin{aligned} h(v_1 + v_2) &= f(v_1 + v_2) + g(v_1 + v_2) = f(v_1) + f(v_2) + g(v_1) + g(v_2) \\ &= (f + g)(v_1) + (f + g)(v_2) = h(v_1) + h(v_2). \end{aligned}$$

Entsprechend folgt für alle Skalare a und Vektoren $v \in V$

$$\begin{aligned} h(a.v) &= f(a.v) + g(a.v) = a.f(v) + a.g(v) \\ &= a.(f(v) + g(v)) = a.h(v). \end{aligned}$$

[16]Wir werden später an Stelle von $g \circ f$ häufig einfach gf schreiben.

Zu (3): Sei $h = a.f$, so folgt

$$\begin{aligned}
h(v_1 + v_2) &= a.f(v_1 + v_2) = a.(f(v_1) + f(v_2)) \\
&= a.f(v_1) + a.f(v_2) = h(v_1) + h(v_2) \\
h(b.v) &= a.f(b.v) = a.(b.f(v)) \\
&= (a \cdot b).f(v) = (b \cdot a).f(v) \\
&= b.(a.f(v)) = b.h(v).
\end{aligned}$$

Wir beachten, dass im letzten Argument die Vertauschbarkeit von Skalaren (Kommutativität) verwendet wurde. □

Ist $f : V \to W$ eine bijektive Abbildung, so gibt es zu jedem $w \in W$ ein eindeutig bestimmtes $v \in V$ mit $f(v) = w$; Schreibweise: $v = f^{-1}(w)$. Durch die Zuordnung $w \mapsto v$ mit $f(v) = w$ wird eine Abbildung $f^{-1} : W \to V$ erklärt, welche **Umkehrabbildung** zu f oder auch die zu f **inverse Abbildung** heißt. Wir beachten, dass die Beziehungen

$$f \circ f^{-1} = 1_W \text{ und } f^{-1} \circ f = 1_V$$

gelten[17]. Da $v = f^{-1}(w)$ die Eigenschaft $f(v) = w$ hat, folgt nämlich $w = f(f^{-1}(w)) = (f \circ f^{-1})(w)$ für alle $w \in W$ und damit $1_W = f \circ f^{-1}$. Für $v \in V$ liegt $w = f(v)$. Nach Definition von f^{-1} gilt daher $f^{-1}(w) = v$, also $f^{-1} \circ f(v) = f^{-1}(f(v)) = f^{-1}(w) = v$ für alle $v \in V$, folglich wie behauptet $f^{-1} \circ f = 1_V$.

Satz 5.6 (Linearität der inversen Abbildung) *$f : V \to W$ sei eine bijektive lineare Abbildung. Dann ist die Umkehrabbildung $f^{-1} : W \to V$ ebenfalls linear.*

Beweis. Seien $w_1, w_2 \in W$ und $v_1 = f^{-1}(w_1)$, $v_2 = f^{-1}(w_2)$. Es gilt somit $f(v_1) = w_1$ und $f(v_2) = w_2$, folglich $f(v_1 + v_2) = f(v_1) + f(v_2) = w_1 + w_2$, woraus $f^{-1}(w_1 + w_2) = v_1 + v_2 = f^{-1}(w_1) + f^{-1}(w_2)$ folgt.

Entsprechend gelten für jeden Skalar a und jedes $w \in W$ mit $v = f^{-1}(w)$ die Beziehungen $f(v) = w$, somit $f(a.v) = a.f(v) = a.w$, woraus $f^{-1}(a.w) = a.v = a.f^{-1}(w)$ folgt. □

Definition 5.7 *Eine bijektive lineare Abbildung $f : V \to W$ heißt **Isomorphismus** von V nach W. Wir nennen die beiden Vektorräume V und W in diesem Fall **isomorph**. Bezeichnung $V \cong W$.*

[17] Ist $f : V \to W$ eine bijektive Abbildung und U eine Teilmenge von W, so hat $f^{-1}(U)$ zwei mögliche Interpretationen, einerseits die als Urbild von U unter f und andererseits die als Bild von U unter f^{-1}. Beide Interpretationen liefern dasselbe Ergebnis!

Es ist dann $f^{-1} : W \to V$ ein Isomorphismus von W nach V. Für jeden Vektorraum V ist die identische Abbildung $1_v : V \to V$ ein Isomorphismus. Schließlich ist die Komposition von zwei Isomorphismen stets wieder ein Isomorphismus. Die Isomorphiebeziehung $\cong$ ist daher *reflexiv* ($V \cong V$), *symmetrisch* ($V \cong W$ impliziert $W \cong V$) und *transitiv* ($U \cong V$ und $V \cong W$ impliziert $U \cong W$).

Etwas pauschal formuliert, haben isomorphe Vektorräume übereinstimmende mathematischen Eigenschaften. Wir werden uns schrittweise von der Gültigkeit dieser Behauptung überzeugen.

Montag, 13. Januar 2004

Satz 5.8 *Isomorphe Vektorräume haben dieselbe Dimension.*

Beweis. Sei $f : V \to W$ ein Isomorphismus und $v_1, v_2, \ldots, v_n$ eine Basis von V. Wir zeigen, dass $(*)$ $f(v_1), \ldots, f(v_n)$ eine Basis von W ist, woraus durch Anzahlvergleich folgt, dass V und W dieselbe Dimension n haben.

$(*)$ **ist ein Erzeugendensystem**: Jedes w aus W hat die Form $w = f(v)$. Da $v_1, v_2, \ldots, v_n$ eine Basis von V ist, ist v eine Linearkombination von $v_1, v_2, \ldots, v_n$ und folglich $w = f(v)$ eine Linearkombination von $f(v_1), f(v_2), \ldots, f(v_n)$.

(*) ist linear unabhängig: Sei $\sum_{i=1}^n \alpha_i.f(v_i) = 0$, so folgt dass $f(\sum_{i=1}^n \alpha_i.v_i) = 0 = f(0)$ gilt. Wegen der Bijektivität von f erzwingt dies $\sum_{i=1}^n \alpha_i.v_i = 0$, woraus sich wegen der Basiseigenschaft $\alpha_1 = \cdots = \alpha_n = 0$ ergibt. $\square$

Satz 5.9 (Klassifikation endlichdimensionaler Vektorräume)

(a) *Hat ein Vektorraum V die Dimension n, so ist $V \cong K^n$.*

(b) *Es gilt $K^m \cong K^n$ genau dann, wenn $m = n$.*

Beweis. Zu (a): Sei $v_1, v_2, \ldots, v_n$ eine Basis von V. Wir betrachten die schon in Folgerung 1.15 betrachtete Abbildung

$$h : K^n \to V, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^n x_i.v_i.$$

Wir wissen, dass h bijektiv ist, da es sich bei $v_1, v_2, \ldots, v_n$ um eine Basis von V handelt. Nachrechnen zeigt ferner, dass h linear ist. Somit ist h ein Isomorphismus.

Zu (b): Falls $m = n$ ist gilt $K^m = K^n$, damit sind beide Vektorräume erst recht zueinander isomorph. Wir nehmen nun umgekehrt an, dass K^m und K^n zueinander isomorph sind. Nach Satz 5.8 haben dann beide Vektorräume dieselbe Dimension $m = n$. $\square$

Folgerung 5.10 *Die endlichdimensionalen Vektorräume V und W sind genau dann isomorph, wenn* $\dim V = \dim W$ *ist.*

Beweis. Falls V und W isomorph sind, haben sie nach Satz 5.8 dieselbe Dimension. Falls V und W andererseits dieselbe Dimension n haben, folgen $K^n \cong V$ und $K^n \cong W$ aus Satz 5.9. Unter Berücksichtigung von Symmetrie und Transitivität der Isomorphiebeziehung ergibt sich dann $V \cong W$. □

Bemerkung 5.11 (Isomorphie und Gleichheit) Wir müssen aufpassen, die Begriffe Gleichheit und Isomorphie von Vektorräumen trotz ihrer Ähnlichkeit nicht durcheinander zu bringen.

Zwei Vektorräume V und W sind gleich, wenn V und W aus denselben Elementen bestehen und zusätzlich Addition und Multiplikation mit Skalaren für V und W übereinstimmen.

Andererseits sind beispielsweise die Vektorräume

$$K^2 \text{ und } \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 + x_2 + x_3 = 0 \right\}$$

zueinander *isomorph*, da sie beide die Dimension zwei haben.

Aber natürlich sind diese Vektorräume *nicht gleich*.

3.6 Matrizen und lineare Abbildungen

Wie konstruiert man lineare Abbildungen von V nach W? Eine Antwort auf diese Frage gibt der Satz über lineare Fortsetzung. Wir sehen schon an seiner Formulierung die wichtige Rolle von Basen.

Satz 6.1 (Satz über lineare Fortsetzung) *V und W seien Vektorräume.*

(a) *Stimmen die linearen Abbildungen $f, g : V \to W$ auf einem Erzeugendensystem von V überein, so folgt $f = g$.*

(b) *Ist $b_1, b_2, \ldots, b_n$ eine Basis von V und $w_1, w_2, \ldots, w_n$ ein System von n Vektoren aus W, so gibt es genau eine lineare Abbildung $f : V \to W$ mit*

$$f(b_1) = w_1, f(b_2) = w_2, \ldots, f(b_n) = w_n.$$

Es gilt dabei $f\left(\sum_{i=1}^n r_i.b_i\right) = \sum_{i=1}^n r_i.w_i$ für jede Auswahl der Skalare r_i.

Es reicht somit völlig aus, eine lineare Abbildung $f : V \to W$ auf einer festen Basis von V zu kennen. Ferner können wir durch beliebige Wertewahl auf der gewählten Basis eine lineare Abbildung auf $f : V \to W$ festlegen.

Donnerstag, 15. Januar 2004

Beweis. Zu (a): Wir können jedes $x \in V$ als Linearkombination $x = \sum_{i=1}^n x_i.b_i$ darstellen. Aus der Linearität von f und g folgt dann

$$f(x) = \sum_{i=1}^n x_i.f(b_i) = \sum_{i=1}^n x_i.g(b_i) = g(x).$$

Es folgt damit $f = g$.

Zu (b): Wir definieren $f : V \to W$ durch die Formel

$$f\left(\sum_{i=1}^n r_i.b_i\right) = \sum_{i=1}^n r_i.w_i.$$

In der Tat definiert diese Formel eine Abbildung, da sich jedes x aus V eindeutig als Linearkombination $x = \sum_{i=1}^n r_i.b_i$ darstellen lässt. Die Darstellung eines Basiselements b_i als Linearkombination der Basiselemente $b_1, b_2, \ldots, b_n$ ist gerade $b_i = 0.b_1 + \cdots + 0.b_{i-1} + 1.b_i + 0.b_{i+1} + \cdots + 0.b_n$, so dass $f(b_i) = w_i$ folgt.

Es bleibt, dass wir uns mit der Linearität von f befassen. Sind $x = \sum_{i=1}^n x_i.b_i$ und $y = \sum_{i=1}^n y_i.b_i$ die Basisdarstellungen von x und y aus V, so folgt $x + y = \sum_{i=1}^n (x_i + y_i).b_i$. Wir erhalten damit

$$f(x + y) = \sum_{i=1}^n (x_i + y_i).b_i = \sum_{i=1}^n x_i.b_i + \sum_{i=1}^n y_i.b_i = f(x) + f(y).$$

Entsprechend zeigen wir, dass $f(a.x) = a.f(x)$ für alle Skalare a gilt. Damit haben wir die Existenz einer linearen Abbildung $f : V \to W$ gezeigt, die auf $b_1, b_2, \ldots, b_n$ der Reihe nach die Werte $w_1, w_2, \ldots, w_n$ annimmt. Die Eindeutigkeit von f folgt aus Teil (a). □

Wir befassen uns anschließend mit einer wichtigen Folgerungen des Satzes über lineare Fortsetzung, welcher die Beschreibung von linearen Abbildungen durch Matrizen betrifft. Wir bezeichnen dazu mit $\mathrm{M}_{m,n}(K)$ die Menge aller $m \times n$-Matrizen über K. Wir werden im folgenden die Mengen

$$\mathrm{M}_{m,n}(K) = (K^m)^n = (K^m)^n$$

miteinander identifizieren, indem wir eine $m \times n$-Matrix A je nach Kontext durch ihren Spaltenaufbau $[s_1, s_2, \ldots, s_n]$ als Mitglied des $(K^m)^n$ oder durch ihren Zeilenaufbau $\begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$ als Mitglied von $(K^n)^m$ auffassen. Insbesondere ist $\mathrm{M}_{m,n}(K)$ bezüglich der argumentweisen Operationen ein Vektorraum der Dimension $m \cdot n$.

Satz 6.2 (Darstellungsmatrix von $f : K^n \to K^m$) *(a) Jede lineare Abbildung $f : K^n \to K^m$ ist eindeutig durch ihre* Darstellungsmatrix

$$M(f) = [f(e_1), \ldots, f(e_n)] \in \mathrm{M}_{m,n}(K),$$

vom Format $m \times n$ bestimmt, deren Spalten in dieser Reihenfolge die Bilder der Einheitsvektoren $e_1, e_2, \ldots, e_n$ unter f sind.

(b) Sei A umgekehrt eine $m \times n$-Matrix mit den Spalten $s_1, s_2, \ldots, s_n$, d.h.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = [s_1, s_2, \ldots, s_n].$$

Dann gibt es genau eine lineare Abbildung $f : K^n \to K^m$ mit $M(f) = A$, also mit $f(e_i) = s_i$ Diese Abbildung ist formelmäßig gegeben durch

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i.s_i = \begin{pmatrix} \sum_{i=1}^{n} a_{1i}.x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi}.x_i \end{pmatrix}.$$

Beweis. Der erste Teil des Satzes über lineare Fortsetzung zeigt, dass die lineare Abbildung $f : K^n \to K^m$ durch die Folge $f(e_1), f(e_2), \ldots, f(e_n)$ der Bilder der Einheitsvektoren des K^n, also durch ihre Darstellungsmatrix $M(f)$, eindeutig bestimmt ist.

Die nächste Aussage, wie auch die Formel, folgt aus Teil (b) des Satzes über lineare Fortsetzung. □

Wir verwenden diesen Satz, um in einem zweistufigen Verfahren die Multiplikation von Matrizen einzuführen.

Definition 6.3 (Matrizenmultiplikation) *Sei A eine $m \times n$-Matrix wie oben.*

(a) ($m \times n$-Matrix) $\times$ Spalte: *Sei x eine $n \times 1$-Matrix, d.h. eine n-Spalte. Dann heißt*

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \sum_{i=1}^{n} x_i. \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1i}.x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi}.x_i \end{pmatrix}$$

*das **Produkt** der $m \times n$-Matrix A mit dem Spaltenvektor x.*

(b) ($m \times n$-Matrix) $\times$ ($n \times p$-Matrix): *Sei A eine $m \times n$-Matrix und B eine $n \times p$-Matrix mit dem Spaltenaufbau $B = [b_1, b_2, \ldots, b_p]$. Dann heißt*

$$A \cdot B = A \cdot [b_1, b_2, \ldots, b_p] := [A.b_1, \ldots, A.b_p]$$

das Matrizenprodukt von A und B.

Satz 6.4 (Berechnung Matrizenprodukt) *A sei eine $m \times n$-Matrix und B eine $n \times p$-Matrix. $e_1, e_2, \ldots, e_n$ sei die Standardbasis des K^n und x eine n-Spalte. Dann gilt*

(1) Matrix $\times$ Einheitsvektor: $A \cdot e_i = [s_1, s_2, \ldots, s_n] \cdot e_i = s_i.$

(2)(n-Zeile) $\times$ (n-Spalte): $$(a_1, a_2, \ldots, a_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_i x_i.$$

(3) ($m \times n$-Matrix) $\times$ (n-Spalte):

$$\begin{aligned} A \cdot x &= [s_1, s_2, \ldots, s_n] \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i.s_i \\ &= \begin{pmatrix} z_1 \cdot x \\ \vdots \\ z_m \cdot x \end{pmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \cdot x \end{aligned}$$

(4) ($m \times n$-Matrix) $\times$ ($n \times p$-Matrix):

$$\begin{aligned} A \cdot B &= \\ &= A \cdot [b_1, b_2, \ldots, b_p] \\ &= [A \cdot b_1, A \cdot b_2, \ldots, A \cdot b_p] \\ &= \begin{bmatrix} z_1 \cdot b_1 & \cdots & z_1 \cdot b_p \\ \vdots & \ddots & \vdots \\ z_m \cdot b_1 & \ldots & z_m \cdot b_p \end{bmatrix} \\ &= \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix} \cdot [b_1, b_2, \ldots, b_p] = \begin{bmatrix} z_1 \cdot B \\ \vdots \\ z_m \cdot B \end{bmatrix} \end{aligned}$$

Beweis. Die einzelnen Aussagen ergeben sich direkt aus der Definition. □

Fazit 6.5 (Matrizenprodukt) Das Matrizenprodukt $A \cdot B = C$ lässt sich genau dann bilden, wenn die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt, d.h. wenn A eine $m \times n$-Matrix und B eine $n \times p$-Matrix ist.

Das Produkt $C = A \cdot B$ ist dann diejenige $m \times p$-Matrix, deren Einträge durch

$$c_{ik} = (i\text{-te Zeile von } A) \cdot (k\text{-te Spalte von } B) = \sum_{j=1}^{n} a_{ij} b_{jk}$$

gegeben sind $(1 \leq i \leq m,\ 1 \leq k \leq p)$.

Mit Hilfe der Matrixmultiplikation lässt sich die zuvor behandelte Korrespondenz zwischen linearen Abbildungen und Matrizen besonders elegant ausdrücken:

Satz 6.6 (Korrespondenz: Lineare Abbildungen und Matrizen) *Es sei $e_1, e_2, \ldots, e_n$ die Standardbasis des K^n.*

(a) *Sei $f : K^n \to K^m$ linear mit Darstellungsmatrix $M(f) = [f(e_1), \ldots, f(e_n)]$. Dann ist $A = M(f)$ eine $m \times n$-Matrix und es gilt*

$$f(x) = A \cdot x \quad \text{für alle } x \in K^n.$$

(b) *Sei umgekehrt A eine $m \times n$-Matrix. Dann ist die Abbildung*

$$f : K^n \to K^m, \quad x \mapsto A \cdot x$$

linear mit Darstellungsmatrix $M(f) = A$.

Beweis. Bei **(a)** handelt es sich um eine Umformulierung des ersten Teils von Satz 6.2, während **(b)** aus dem zweiten Teil des genannten Satzes folgt. □

Wir sind damit in der Lage — unter geeigneten Voraussetzungen an das Format — Matrizen zu addieren, mit Skalaren oder auch miteinander zu multiplizieren. Gleiches können wir — unter geeigneten Voraussetzungen an Definitions- und Wertebereiche — mit linearen Abbildungen tun. Wir erinnern daran, dass für jedes $n \in \mathbb{N}$ die $n \times n$ Einheitsmatrix als

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

definiert ist.

Satz 6.7 (Darstellungsmatrizen) *Die Zuordnung $f \mapsto M(f)$, von linearen Abbildungen $f : K^n \to K^m$ zu $m \times n$-Matrizen hat die folgenden Eigenschaften.*

(1) *$M(f+g) = M(f) + M(g)$ für alle linearen Abbildungen $f, g : K^n \to K^m$.*

(2) *$M(a.f) = a.M(f)$ für alle linearen Abbildungen $f : K^n \to K^m$ und Skalare a.*

(3) *$M(g \circ f) = M(g) \cdot M(f)$ für alle linearen Abbildungen $f : K^m \to K^n$ und $g : K^n \to K^p$.*

(4) *$M(1_{K^n}) = E_n$, wobei E_n die $n \times n$-Einheitsmatrix bedeutet.*

Montag, 19. Januar 2004

Beweis. Zu (1): Es gilt $(f+g)(e_i) = f(e_i) + g(e_i)$, daher $M(f+g) = [f(e_1) + f(e_2), \ldots, f(e_n)+g(e_n)] = [f(e_1), \ldots, f(e_n)]+[g(e_1), \ldots, g(e_n)]$ und folglich $M(f+g) = M(f) + M(g)$. **(2)** beweist man analog.
Zu (3): Es seien $A = M(f)$, $B = M(g)$ und $C = M(g \circ f)$. Daher ist $A = [a_1, \ldots, a_m]$ mit $a_i = f(e_i)$. Nach Satz hat g die Form $g(x) = B \cdot x$. Es ist daher

$$(g \circ f)(e_i) = g(f(e_i)) = g(a_i) = B \cdot a_i,$$

damit wegen Satz 6.4 (4)

$$M(g \circ f) = [B \cdot a_1, B \cdot a_2, \ldots, B \cdot a_n] = B \cdot A,$$

und somit $M(g \circ f) = M(g) \cdot M(f)$.
Zu (4): Als Darstellungsmatrix $M(1_{K^n})$ der identischen Matrix erhalten wir $[e_1, e_2, \ldots, e_n]$, also die $n \times n$-Einheitsmatrix. □

Satz 6.8 (Rechenregeln) (i) *Die Menge* $\mathrm{M}_m(n)(K)$ *der* $m \times n$*-Matrizen bildet bzgl. Addition und Multiplikation mit Skalaren einen Vektorraum der Dimension* mn*.*

(2) *Es gilt für jede* $m \times n$*-Matrix* A

$$E_m \cdot A = A = A \cdot E_n$$

(3) *Die Matrixmultiplikation ist assoziativ.*

(4) *Es gelten die distributiven Gesetze* $(A+B)\cdot C = A\cdot C + B\cdot C$ *und* $A\cdot(B+C) = A \cdot B + A \cdot C$*.*

(5) *Im allgemeinen ist die Matrixmultiplikation nicht kommutativ.*

Beweis. Zu (1): Dies ergibt sich aus der Identifizierung $\mathrm{M}_{m,n}(K) = (K^m)^n$.
Zu (2): Für die lineare Abbildung $f : K^n \to K^m$, $x \mapsto A \cdot x$ gilt $1_{K^m} \circ f = f = f \circ 1_{K^n}$. Durch Übergang zu den Darstellungsmatrizen folgt unter Verwendung von Satz 6.7 (3) und (4) die Beziehung $E_m \cdot A = A = A \cdot E_n$.
Zu (3) und (4): Es ist leicht zu sehen, dass die Verknüpfung von Abbildungen assoziativ und im Fall von linearen Abbildungen auch bezüglich der Summenbildung distributiv ist. Übergang zu den Darstellungsmatrizen liefert dann wegen Satz 6.7 die Behauptung.
Zu (5): Sei $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ und $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Dann gilt

$$A \cdot B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = B \cdot A.$$

□

3.7 Invertierbare Matrizen

Definition 7.1 *Eine $n \times n$-Matrix A heißt* ***invertierbar****, wenn es eine $n \times n$-Matrix B gibt, so dass $A \cdot B = E_n = B \cdot A$ gilt.*

In diesem Fall ist B durch A eindeutig bestimmt und heißt zu A ***inverse Matrix****. Bezeichnung: A^{-1}.*

Zur Eindeutigkeit: Sind B und B' zu A invers, so folgt

$$B' = B' \cdot E_n = B' \cdot (A \cdot B) = (B' \cdot A) \cdot B = E_n \cdot B = B,$$

also wie behauptet $B' = B$. Wir werden übrigens im nächsten Satz sehen, dass wir für für $n \times n$-Matrizen, also für quadratische Matrizen, nur einseitig ein Inverses fordern müssen.

Satz 7.2 (Invertierbarkeit) *Für eine $n \times n$-Matrix A sind äquivalent:*

(1) *A ist invertierbar.*

(2) *Die Multiplikationsabbildung $f : K^n \to K^n$, $x \mapsto A \cdot x$, mit der Matrix A ist ein Isomorphismus.*

(3) *Die Spalten $s_1, s_2, \ldots, s_n$ der Matrix A bilden eine Basis des K^n.*

(4) *Es gibt eine $n \times n$-Matrix B mit $A \cdot B = E_n$.*

(5) *Es gibt eine $n \times n$-Matrix C mit $C \cdot A = E_n$.*

In diesem Fall sind die in (4) und (5) auftretenden Matrizen B und C beide gleich A^{-1}.

Donnerstag, 22. Januar 2004

Beweis. Wir zeigen zunächst die Implikationen

$$\begin{array}{ccc} (2) & \Rightarrow & (1) \\ \Uparrow & & \Downarrow \\ (3) & \Leftarrow & (5) \end{array}$$

und danach die Implikationen (1) $\Rightarrow$ (4) $\Rightarrow$ (2).
(2)$\Rightarrow$(1): Sei $g : K^n \to K^n$ die zu $f : K^n \to K^n$, $x \mapsto A \cdot x$, inverse Abbildung. Aus $f \circ g = 1_{K^n} = g \circ f$ folgt durch Übergang zu Darstellungsmatrizen , dass $M(f) \cdot M(g) = E_n = M(g) \cdot M(f)$ gilt. Daher ist $B = M(g)$ zu $A = M(f)$ invers.
(1)$\Rightarrow$(5) ist trivial. **(5)$\Rightarrow$(3):** Wegen $C \cdot A = E_n$ gilt $[C \cdot s_1, \ldots, C \cdot s_n] = [e_1, e_2, \ldots, e_n]$, somit $C \cdot s_i = e_i$ für $i = 1, \ldots, n$. Daher sind die Spalten $s_1, s_2, \ldots, s_n$

der Matrix A linear unabhängig. Aus $\sum_{i=1}^{n} \alpha_i \cdot s_i = 0$ folgt nämlich durch Multiplikation mit C (von links), dass $\sum_{i=1}^{n} \alpha_i \cdot e_i = 0$ gilt und damit alle α_i verschwinden. Nach Satz 3.11 bilden dann $s_1, s_2, \ldots, s_n$ eine Basis von K^n.
(3)⇒(2): Die Abbildung $f : K^n \to K^n$, $x \mapsto A \cdot x$, ist linear. Ferner bildet sie die Standardbasis $e_1, e_2, \ldots, e_n$ auf die aus den Spaltenvektoren $s_1, s_2, \ldots, s_n$ von A bestehende Basis des K^n ab. f ist daher ein Isomorphismus.
(1)⇒(4) ist trivial. **(4)⇒(2):** Seien $b_1, b_2, \ldots, b_n$ die Spalten von B. Wegen $A \cdot B = E_n$ gilt dann $[A \cdot b_1, A \cdot b_2, \ldots, A \cdot b_n] = [e_1, e_2, \ldots, e_n]$. Die lineare Abbildung $f : K^n \to K^n$, $x \mapsto Ax$, bildet daher die Basis $b_1, b_2, \ldots, b_n$ auf die Standardbasis $e_1, e_2, \ldots, e_n$ ab, ist daher ein Isomorphismus.

Sei nun $A \cdot B = E_n$, dann liefert Multiplikation mit A^{-1} von links, dass $B = A^{-1}$ gilt. Entsprechend folgt aus $C \cdot A = E_n$ durch Multiplikation mit A^{-1} von rechts, dass $C = A^{-1}$ gilt. □

Satz 7.3 (Inverse Matrizen: Rechenregeln) *A, B seien $n \times n$-Matrizen. Es gilt*

(1) Die Einheitsmatrix E_n ist invertierbar und $E_n^{-1} = E_n$.
(2) Ist A invertierbar, so auch A^{-1}. Es gilt dann: $(A^{-1})^{-1} = A$.
(3) Sind A und B invertierbar, so auch $A \cdot B$. Es gilt dann:

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

Wir beachten in (3) die Umkehrung der Reihenfolge!

Beweis. Zu (1): Aus $E_n \cdot E_n = E_n$ folgt, dass E_n zu sich selbst invers ist. Eindeutigkeit der Inversen liefert die Behauptung.

Zu (2): Die Beziehung $A \cdot A^{-1} = E_n = A^{-1} \cdot A$, die ausdrückt, dass A^{-1} zu A invers ist, lässt sich durch Verkehrung der Rollen beider Faktoren auch so lesen, dass A zu A^{-1} invers ist, also $A = (A^{-1})^{-1}$ ist.

Zu (3): Es gilt

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot E_n \cdot A^{-1} = A \cdot A^{-1} = E_n.$$

Entsprechend folgt $(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = E_n$ und dann die Behauptung. □

Invertierbare Matrizen und lineare Gleichungssysteme

A sei eine invertierbare $n \times n$-Matrix und b eine n-Spalte. Wir befassen uns erneut mit dem Lösen des linearen Gleichungssystems zur erweiterten Matrix $[A, b]$. Dieses Gleichungssystem können wir mit Hilfe der Matrixmultiplikation kompakt in der Form $A \cdot x = b$ darstellen.

Satz 7.4 (Gleichungssystem mit invertierbarer Matrix) *A sei eine invertierbare Matrix. Das lineare Gleichungssystem $A \cdot x = b$ hat genau eine Lösung, nämlich $x = A^{-1} \cdot b$.*

Beweis. Einsetzen von $x = A^{-1} \cdot b$ ergibt, dass x eine Lösung von $A \cdot x = b$ ist.

Sei andererseits x eine Lösung von $A \cdot x = b$, so ergibt Links-Multiplikation mit der Matrix A^{-1}, dass dann $x = A^{-1} \cdot b$ gelten muss. □

Bemerkung 7.5 (Bewertung der Lösungsformel) (1) Gegenüber der algorithmischen Lösung von $A \cdot x = b$ durch Umformen von $[A|b]$ auf Zeilenstufenform ist es vorteilhaft, dass wir die Lösung x durch eine allgemeingültige Formel $x = A^{-1} \cdot x$ angeben können.

(2) Diese Formel verdeutlicht zudem die Denkökonomie der Matrizenrechnung. Bis auf die Feinheit, dass wir auf die Reihenfolge der Faktoren zu achten haben, entspricht sie genau der Lösung der Zahlengleichung $ax = b$ für $a \neq 0$.

(3) Zu beachten ist, dass die Formel — anders als der Gauß-Algorithmus — nur für invertierbare Matrizen A gilt.

(4) Ferner ist der Rechenaufwand zu beachten: Wir erzeugen hohen Rechenaufwand durch die anschließend diskutierte Bestimmung der inversen Matrix und anschließende Matrixmultiplikation. Im Vergleich schneidet die direkte algorithmische Lösung von $A \cdot x = b$ besser ab.

Sei A eine invertierbare $n \times n$-Matrix. Wie berechnet man die zu A inverse Matrix? Nach Voraussetzung bilden die Spalten von A eine Basis des K^n. Zur Bestimmung der zu A inversen Matrix X müssen wir das Matrixproblem $A \cdot X = E_n$ mit $X = [x_1, x_2, \ldots, x_n]$ lösen. Völlig gleichbedeutend ist die Aufgabe $[A \cdot x_1, \ldots, A \cdot x_n] = [e_1, \ldots, e_n]$, somit das simultane Lösen der linearen Gleichungssysteme $A \cdot x_i = e_i$ $(1 \leq i \leq n)$. Wir wissen, dass diese Lösung durch elementare Zeilenumformungen möglich ist; dieselben können wir simultan durchführen:

Satz 7.6 (Algorithmische Berechnung der inversen Matrix) *Eine $n \times n$-Matrix A ist genau dann invertierbar, wenn ihre Zeilenstufenform die Einheitsmatrix E_n ist.*

In diesem Fall lässt sich $[A|E_n]$ durch elementare Zeilenumformungen auf die Form $[E_n|B]$ bringen. Es ist dann B die zu A inverse Matrix.

Beweis. Sei A invertierbar. Wir bringen A durch elementare Zeilenumformungen auf Zeilenstufenform A'. Da das lineare Gleichungssystem $A \cdot x = 0$ wegen der Invertierbarkeit von A nur die Nullösung besitzt, gilt dies auch für $A' \cdot x = 0$. Damit hat A' den Rang n und ist wegen der Zeilenstufenform folglich die Einheitsmatrix E_n.

Wir setzen nun voraus, dass die Zeilenstufenform von A die Einheitsmatrix ist. In diesem Fall lässt sich $[A|E_n]$ durch elementare Zeilenumformungen zu $[E_n|B]$ umformen. Nach Vorbetrachtung ist $A \cdot B = E_n$, somit $B = A^{-1}$. $\square$

Wir führen diese Überlegungen an einem einfachen Beispiel durch:

Beispiel 7.7 Gegeben sei die Matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & 1 \end{bmatrix}$. Wir bringen die erweiterte Matrix $[A|E_3]$ durch elementare Zeilenumformungen auf Zeilenstufenform:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 & 1 & 0 \\ 3 & 4 & 1 & 0 & 0 & 1 \end{array}\right] \mapsto \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & -2 & 1 & -3 & 0 & 1 \end{array}\right] \mapsto$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 2 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right] \mapsto \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 2 & 0 \\ 0 & 1 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right]$$

Es folgt damit, dass A invertierbar ist mit inverser Matrix

$$A^{-1} = \begin{bmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ 1 & -2 & 1 \end{bmatrix}.$$

3.8 Der Rangsatz

Die Eigenschaften einer linearen Abbildung $f : V \to W$ werden stark von den beiden mit ihr verbundenen Unterräumen $\mathrm{Kern}(f) = \{x \in V \mid f(x) = 0\}$ und $\mathrm{Bild}(f) = \{f(v) \mid v \in V\}$ — *und ihren Dimensionen* — bestimmt. Wir richten unser Augenmerk daher auf:

- den Kern von f, welcher ein Unterraum von V ist;
- das Bild von f, welches ein Unterraum von W ist.

Allein schon die Dimensionen von $\mathrm{Kern}(f)$ und $\mathrm{Bild}(f)$ geben wichtige Auskunft über f.

Definition 8.1 (Rang) *Für eine lineare Abbildung $f : V \to W$ heißt die Dimension von* Bild(f) *der* ***Rang*** *von f. Bezeichnung: rg(f).*

Ist A eine $m \times n$-Matrix, so erklären wir den Rang von A *als Rang der zugeordneten linearen Abbildung $f : K^n \to K^n$, $x \mapsto A \cdot x$.*

Für die durch die $m \times n$-Matrix A gegebene lineare Abbildung $f : K^n \to K^m$ wird $\mathrm{Bild}(f)$ von den Spalten von A aufgespannt. Daher ist der Rang $\mathrm{rg}(A)$ von A die maximale Anzahl eines Systems linear unabhängiger Spalten von A. Im Fall einer Matrix von Zeilenstufenform stimmt dieser mit dem früher definierten Rang überein. Wir sprechen daher genauer auch vom **Spaltenrang**[18] von A. Wir besprechen nun, wie die diskutierten Dimensionen miteinander zusammenhängen.

Satz 8.2 (Rangsatz für lineare Abbildungen) *Sei $f : V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann gilt*

$$\dim V = \dim \mathrm{Kern}(f) + \dim \mathrm{Bild}(f).$$

Da uns — in der Regel — bei gegebenem f die Dimension von V (und auch die von W) bekannt ist, bestimmen sich folglich die Dimensionen von $\mathrm{Kern}(f)$ und $\mathrm{Bild}(f)$ wechselseitig. Als Regelbezeichnung[19] hat sich hier der Rang von f durchgesetzt.

Beweis des Rangsatzes. Da V und W endlichdimensional sind, haben auch $\mathrm{Kern}(f)$ und $\mathrm{Bild}(f)$ endliche Dimension.

Seien nun $(e_1, e_2, \ldots, e_p)$ eine Basis von $\mathrm{Kern}(f)$ und $(g_1, g_2, \ldots, g_q)$ eine Basis von $\mathrm{Bild}(f)$. Im ersten Schritt wählen wir für jedes g_i $(1 \leq i \leq q)$ ein Urbild f_i aus V.

[18] Entsprechend lässt sich der **Zeilenrang** von A als Maximalzahl linear unabhängiger Zeilen von A erklären. Wir werden im nächsten Abschnitt sehen, dass Zeilen- und Spaltenrang stets übereinstimmen.

[19] Manchmal ist von der Dimension von $\mathrm{Kern}(f)$ als dem Defekt von f die Rede.

Wir behaupten, dass $B := (e_1, e_2, \ldots, e_p, f_1, f_2, \ldots, f_q)$ eine Basis von V ist, woraus dann der Rangsatz sofort folgt.
(1) B **ist linear unabhängig**: Aus

$$(a_1.e_1 + a_2.e_2 + \cdots + a_p.e_p) + (b_1.f_1 + b_2.f_2 + \cdots + b_q.f_q) = 0$$

folgt durch Anwendung von f, dass

$$b_1.g_1 + b_2.g_2 + \cdots + b_q.g_q = 0$$

und dann alle b_j verschwinden, da die g_j's eine Basis bilden. Es ergibt sich nunmehr

$$a_1.e_1 + a_2.e_2 + \cdots + a_p.e_p = 0,$$

woraus das Verschwinden auch der a_i folgt.

(2) B **ist ein Erzeugendensystem von** V: Sei $v \in V$, somit $f(v) \in \mathrm{Bild}(f) = \langle g_1, g_2, \ldots, g_q \rangle$. Wir erhalten also

$$f(v) = b_1.g_1 + \cdots + b_q.g_q = f(b_1.f_1 + \cdots + b_q.f_q)$$

mit geeigneten Skalaren b_j. Es folgt

$$f(v - (b_1.f_1 + b_2.f_2 + \cdots + b_q.f_q)) = 0,$$

somit ist

$$v - (b_1.f_1 + b_2.f_2 + \cdots + b_q.f_q)$$

in $\mathrm{Kern}(f)$ gelegen, daher von der Form $a_1.e_1 + a_2.e_2 + \cdots + a_p.e_p$ mit geeigneten Skalaren a_i. Zusammengefasst:

$$v = (a_1.e_1 + a_2.e_2 + \cdots + a_p.e_p) + (b_1.f_1 + b_2.f_2 + \cdots + b_q.f_q). \qquad \square$$

Montag, 26. Januar 2004

Satz 8.3 (Rangsatz für Matrizen) *Sei A eine $m \times n$-Matrix und H der Lösungsraum des homogenen Gleichungssystems $A \cdot x = 0$. Die Dimension von H ist gerade $n - rg(A)$.*

Beweis. Wir wenden den Rangsatz 8.2 auf die lineare Abbildung $f : K^n \to K^m$, $x \mapsto A \cdot x$ an und beachten, dass $H = \mathrm{Kern}(f)$ und $\mathrm{rg}(A) = \mathrm{rg}(f)$ gilt. $\square$

3.9 Elementarmatrizen

Definition 9.1 *Unter einer* ***Elementarmatrix*** *verstehen wir eine Matrix die aus einer $n \times n$-Einheitsmatrix E_n durch eine einzige elementare Zeilenumformung hervorgeht.*

Wir haben wir damit folgende *drei Typen von Elementarmatrizen*:

(1) Für $i \neq k$ die Matrix $E^{i,k}$, die aus E_n durch Vertauschen von i-ter und k-ter Zeile hervorgeht. Es ist $E^{i,k} = E^{k,i}$.

(2) Für $i = 1, \ldots, n$ und einen Skalar $\alpha \neq 0$ die Matrix $E^i(\alpha)$, die aus E_n durch Multiplikation der i-ten Zeile mit dem Faktor α hervorgeht.

(3) Für $i \neq k$ und einen beliebigen Skalar α die Matrix $E^{i,k}(\alpha)$, die aus E_n durch Addition des α-fachen der k-ten Zeile zur i-ten Zeile hervorgeht.

Somit gilt beispielsweise für $i < k$

$$E^{i,k} = \begin{array}{cc} \begin{array}{ccccccc} & & i & & k & & \end{array} & \\ \left[\begin{array}{ccccccc} 1 &&&&&& \\ & \ddots &&&&& \\ && 0 && 1 && \\ &&& \ddots &&& \\ && 1 && 0 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{array}\right] & \begin{array}{c} \\ \\ i \\ \\ k \\ \\ \\ \end{array} \end{array} , \quad E^{i,k}(\alpha) = \begin{array}{cc} \begin{array}{ccccccc} & & i & & k & & \end{array} & \\ \left[\begin{array}{ccccccc} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 && \alpha && \\ &&& \ddots &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{array}\right] & \begin{array}{c} \\ \\ i \\ \\ k \\ \\ \\ \end{array} \end{array}$$

Hinweis. Elementare *Spaltenumformungen* der Einheitsmatrix führen auf dieselben Typen von Matrizen.

Satz 9.2 (Elementarmatrizen und elementare Zeilenumformungen)

(a) Elementare Zeilenumformungen (bzw. Spaltenumformungen) ergeben sich gerade durch Linksmultiplikation (bzw. Rechtsmultiplikation) mit einer Elementarmatrix.

(b) Jede Elementarmatrix ist invertierbar und ihre Inverse ist wieder eine Elementarmatrix.

(c) Der Spaltenrang einer Matrix ändert sich nicht durch elementare Zeilen- oder Spaltenumformungen.

Beweis. Wir verwenden momentan $e_1, e_2, \ldots, e_n$ zur Bezeichnung der Standardbasis des K^n in *Zeilenform.*

Zu (a): Sei $A = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ eine im Zeilenaufbau gegebene $n \times p$-Matrix, so ist $e_i \cdot A = a_i$. Die Behauptung (a) folgt dann unmittelbar hinsichtlich der Zeilenumformungen. Der Beweis für elementare Spaltenumformungen verläuft analog.

Zu (b): Von einer Elementarmatrix E kommen wir durch eine einzige elementare Zeilenumformung zur Einheitsmatrix E_n zurück. Nach (a) gibt es daher eine Elementarmatrix E' mit $E' \cdot E = E_n$.

Zu (c): Wegen (b) reicht es zu zeigen, dass die $m \times n$-Matrizen A, $A \cdot B$ und $C \cdot A$ denselben Spaltenrang haben, wenn B eine invertierbare $m \times m$-Matrix und C eine invertierbare $n \times n$-Matrix ist.

Übersetzt in die Sprache der linearen Abbildungen müssen wir nämlich nachweisen, dass für eine lineare Abbildung $f : K^n \to K^m$ die linearen Abbildungen f, $f \circ g$ und $h \circ f$ denselben Rang haben, falls $g : K^m \to K^m$ sowie $h : K^n \to K^n$ Isomorphismen sind. Offensichtlich haben — wegen $g(K^m) = K^m$ — die beiden Abbildungen f und $f \circ g$ dasselbe Bild und daher denselben Rang. Da ferner h ein Isomorphismus ist, sind $\text{Bild}(f) = f(\mathbb{R}^n)$ und $\text{Bild}(h \circ f) = h(f\mathbb{R}^n)$ isomorph, somit haben auch f und $h \circ f$ denselben Rang. □

Die Zeilenstufenform einer invertierbaren Matrix ist die Einheitsmatrix. Es folgt:

Satz 9.3 (Invertierbare Matrizen und Elementarmatrizen) *Jede invertierbare $n \times n$-Matrix A ist darstellbar als Produkt von Elementarmatrizen.*

Beweis. Als invertierbare Matrix lässt sich A durch elementare Zeilenumformungen zur Einheitsmatrix umformen. Dabei entspricht jeder Umformungsschritt der Linksmultiplikation mit einer Elementarmatrix.

Es gibt daher Elementarmatrizen $F_1, F_2, \ldots, F_s$ mit

$$F_s \cdots F_2 \cdot F_1 \cdot A = E_n.$$

Folglich ist $A = F_1^{-1} \cdot F_2^{-1} \cdots F_s^{-1}$ ein Produkt von Elementarmatrizen. □

Folgerung 9.4 *Jede invertierbare $n \times n$-Matrix ist ein Produkt von Elementarmatrizen.* □

Jede invertierbare Matrix erhalten wir aus der $n \times n$-Matrix durch eine Folge elementarer Spaltenumformungen. Aus Symmetriegründen gelten entsprechende Aussagen für elementare Zeilenumformungen.

Jede $m \times n$-Matrix A lässt sich durch elementare Zeilenumformungen auf *Zeilenstufenform* und analog durch elementare Spaltenumformungen auf *Spaltenstufenform*[20] bringen.

Satz 9.5 *Zu jeder $m \times n$-Matrix A gibt es eine invertierbare Matrizen Q und R, so dass $Q \cdot A$ Zeilenstufenform und $A \cdot R$ Spaltenstufenform hat.*

[20] Dies bedeutet, dass durch Vertauschen von Spalten und Zeilen eine Matrix von Zeilenstufenform entsteht.

Beweis. Durch elementare Zeilenumformungen lässt sich A auf Zeilenstufenform bringen. Damit gibt es Elementarmatrizen $E_1, E_2, \dots, E_s$, so dass $E_1 \cdot E_2 \cdots E_s \cdot A$ Zeilenstufenform hat[21]. Das Produkt $Q = E_1 \cdot E_2 \cdots E_s$ ist eine invertierbare Matrix mit der gewünschten Eigenschaft.
Der Beweis für die Spaltenstufenform verläuft analog. □

Satz 9.6 *Jede $m \times n$-Matrix A lässt sich durch zulässige Zeilen- und Spaltenumformungen auf die Form*

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

bringen, wobei E_r die $r \times r$-Einheitsmatrix ist und die Symbole 0 für Nullmatrizen geeigneten Formats stehen.
Dabei ist r der Spaltenrang von A.

Beweis. Durch elementare Zeilenumformungen bringen wir A zunächst auf Zeilenstufenform. In den r Stufen dieser Matrix stehen die ersten r Vektoren $e_1, e_2, \dots, e_r$ der Standardbasis.

Diese Spalten bringen wir anschließend durch elementare Spaltenumformungen in die ersten r Positionen; von den verbleibenden Spalten wissen wir, dass sie Linearkombinationen von $e_1, e_2, \dots, e_r$ sind. Dieselben können daher durch weitere elementare Spaltenumformungen zu Null umgeformt werden. □

Satz 9.7 *Sei A eine $m \times n$-Matrix vom Spaltenrang r. Dann gilt:*
(a) Es gibt Elementarmatrizen $E_1, E_2, \dots, E_s$ bzw. $F_1, F_2, \dots, F_t$, so dass

$$E_1 \cdot E_2 \cdots E_s \cdot A \cdot F_1 \cdot F_2 \cdots F_t = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

(b) Es gibt invertierbare Matrizen Q und R, so dass

$$Q \cdot A \cdot R = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Beweis. (a) ergibt sich unmittelbar aus dem vorangehenden Satz. Aussage (b) folgt dann aus (a). □

Als Anwendung dieser Überlegungen erhalten wir jetzt die Übereinstimmung von Zeilen- und Spaltenrang.

Satz 9.8 (Zeilenrang=Spaltenrang) *Für jede $m \times n$-Matrix A stimmen Zeilenrang (die maximale Anzahl eines Systems linear unabhängiger Zeilen) und der Spaltenrang (die maximale Anzahl eines Systems linear unabhängiger Spalten) überein.*

[21] In der hier verwendeten Schreibweise nicht E_i mit der $i \times i$-Einheitsmatrix verwechseln!

Beweis. Durch elementare Zeilen- und Spaltenumformungen lässt sich A zu einer Matrix

$$B = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

umformen. Bei dieser Umformung bleiben Zeilen- und Spaltenrang erhalten. Ersichtlich stimmen für B Zeilen- und Spaltenrang überein, also auch für A. $\square$

Kapitel 4

Determinanten

4.1 Weshalb Determinanten?

Donnerstag, 29. Januar 2004

Es gibt drei Problemfelder, für welche Determinanten von großem Nutzen sind:

1. Die formelmäßige Überprüfung der linearen Unabhängigkeit eines Systems von n Vektoren des K^n,

2. Die formelmäßige Lösung eines linearen Gleichungssystems $A \cdot x = b$ mit **invertierbarer** Matrix A,

3. Die formelmäßige Berechnung der inversen Matrix einer invertierbaren Matrix A.

Der Fall $n = 2$

Es macht Sinn, dass wir zunächst den Fall von 2×2-Matrizen anschauen.

Satz 1.1 *Für eine 2×2-Matrix $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ sind äquivalent:*
(1) Die Spalten von A sind linear unabhängig.
(2) A ist invertierbar.
(3) Die Determinante $|A| := ad - bc$ ist von Null verschieden.
In diesem Fall ist die inverse Matrix durch die Formel

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

gegeben.

Beweis. Die Implikation (1) $\Leftrightarrow$ (2) kennen wir schon. (3) $\Rightarrow$ (2): Produktbildung liefert

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

woraus die Behauptung sofort folgt.

(2) $\Rightarrow$ (3): Wir nehmen an, dass A invertierbar ist und erhalten daher eine Matrixgleichung

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

aus der wir — beispielhaft für $c \neq 0$ — die Gleichungen

$$\begin{aligned} ae + bg &= 1 \\ ce + dg &= 0 \end{aligned}$$

extrahieren. Subtraktion des c-fachen der ersten von dem a-fachen der zweiten Gleichung liefert

$$(ad - cb)g = -c,$$

woraus $|A| \neq 0$ folgt. Ähnlich argumentieren wir, falls ein anderer Eintrag von A nicht Null ist. □

Lösungsformel für $A \cdot x = b$

Satz 1.2 (Cramersche Regel) *Seien*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \textit{ und } b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Wir setzen voraus, dass A ***invertierbar*** *ist.*

Die eindeutig bestimmte Lösung des Gleichungssystems ist dann durch

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{22} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

gegeben.

Beweis. Bilde das Produkt $A^{-1} \cdot b$ mit der diskutierten Formel für A^{-1}. □

Der Fall $n = 3$

In Kapitel 1 haben wir gesehen, wie mit Hilfe des Spatprodukts eine formelmäßige Überprüfung der linearen Unabhängigkeit eines Systems von drei Vektoren des K^3 tatsächlich möglich ist. Die Cramersche Regel hat dort gleichfalls eine Lösungsformel für lineare Gleichungssysteme mit invertierbarer 3×3-Matrix ergeben. Zum dritten Punkt verfügen wir bisher "nur" über algorithmische Lösungsverfahren. Wir erinnern ferner an die Rückführung von 3×3-Determinanten auf solche vom Format 2×2 durch "Entwicklung" der Determinante.

4.2 Existenz von Determinanten

Sinngemäß lassen sich die für $n = 2$ und $n = 3$ dargestellten Ergebnisse auf höhere Dimensionen n verallgemeinern. Hauptschwierigkeit ist dabei die **Determinantendefinition** selbst, da es nicht so klar ist, wie die Formel $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ und die entsprechende Formel für $n = 3$ auf höhere Dimensionen zu übertragen ist.

Gerade noch für $n = 3$ gibt es eine ähnlich "leichte" Formel, die **Sarrussche Regel**, welche $|A|$ als Summe von 6 Produkten (jeweils von drei Matrixeinträgen) mit zusätzlich alternierendem Vorzeichenfaktor darstellt.

Die **Leibnizsche Formel**, die wir erst am Ende dieses Kapitels behandeln, stellt generell — die Fälle $n = 2, 3$ verallgemeinernd — eine $n \times n$-Determinante als Summe von $n!$ Summanden mit alternierenden Vorzeichen dar; diese Summanden sind ihrerseits Produkte aus jeweils n Matrixeinträgen.

Satz 2.1 (Kennzeichung der Determinantenfunktion) *Es gibt genau eine Funktion*

$$\det : \mathrm{M}_n(K) \to K, \quad [a_1, a_2, \ldots, a_n] \mapsto \det(a_1, a_2, \ldots, a_n)$$

mit folgenden Eigenschaften:

(D 1) det *ist linear in jeder Spalte, d.h. es gilt stets*

$$\begin{aligned} \det(a_1, \ldots, a_i + a_i', \ldots, a_n) &= \det(a_1, \ldots, a_i, \ldots, a_n) + \det(a_1, \ldots, a_i', \ldots, a_n) \\ \det(a_1, \ldots, \lambda.a_i, \ldots, a_n) &= \lambda \cdot \det(a_1, \ldots, a_i, \ldots, a_n). \end{aligned}$$

(D 2) $\det(A) = 0$*, wenn die Spalten von A linear abhängig sind.*

(D 3) $\det(E_n) = 1$.

Diese Funktion det nennen wir **Determinantenfunktion**. Für eine $n \times n$-Matrix A nennen wir $|A| := \det(A)$ die **Determinante** von A.

Der Nachweis der Eindeutigkeit von det beruht auf folgendem einfach zu beweisenden Hilfssatz.

Lemma 2.2 (Einfache Eigenschaften) *Die Funktion $D : \mathrm{M}_n(K) \to K$ habe die Eigenschaften **(D 1)** und*

(D 2') *Hat A zwei gleiche Spalten, so ist $D(A) = 0$.*

Dann gilt

(a) *Bei Vertauschen zweier Spalten ändert $D(A)$ das Vorzeichen.*

(b) *Bei Multiplikation einer Spalte von A mit einem Faktor λ ändert sich $D(A)$ um den Faktor λ.*

(c) *Der Wert $D(A)$ ändert sich nicht, wenn wir ein Vielfaches einer Spalte zu einer* ***anderen*** *Spalte addieren.*

(d) *D erfüllt die Bedingung* ***(D 2)****.*

Wir können daher Determinantenfunktionen auch durch die Bedingungen ***(D 1)****,* ***(D 2')*** *und* ***(D 3)*** *definieren.*

Beweis. Zu (a): Wir nehmen $i \neq k$ an und werten die Formel

$$0 = D(a_1, \ldots, \boxed{a_i + a_k}, \ldots, \boxed{a_i + a_k}, \ldots, a_n)$$

unter Verwendung der Linearität (D 1) und der Bedingung (D 2') aus.

Zu (b): Dies ist durch die geforderte Linearität von D in den Spalten abgedeckt.

Zu (c): Wir nehmen $i < k$ an und erhalten wegen (1)

$$\begin{aligned} &D(a_1, a_2, \ldots, a_i, \ldots, \boxed{a_k + \lambda . a_i}, \ldots, a_n) = \\ &D(a_1, a_2, \ldots, a_i, \ldots, a_k, \ldots, a_n) + \lambda \underbrace{D(a_1, \ldots, \boxed{a_i}, \ldots, \boxed{a_i}, \ldots, a_n)}_{=0}. \end{aligned}$$

Der Fall $k < i$ wird analog behandelt.

Zu (d): Wir nehmen an, dass die Spalten $a_1, a_2, \ldots, a_n$ der Matrix A linear abhängig sind. Dann lässt sich eine von ihnen, sagen wir a_k, als Linearkombination der übrigen schreiben. Durch sukzessive Addition von Vielfachen von Spalten a_i mit $i \neq k$ lässt sich dann A zu einer Matrix A' umformen, deren k-te Spalte Null ist. Wegen **(c)** gilt $D(A) = D(A')$. Andererseits ist wegen **(b)** $D(A') = 0$. □

Nachweis der Eindeutigkeit

Wir nehmen an, dass det und det′ beide die Eigenschaften (D 1), (D 2) und (D 3) haben. Falls die Spalten $a_1, a_2, \ldots, a_n$ der Matrix A linear abhängig sind, liefern sowohl det als auch det′ auf A den Wert Null. Beide Funktionen stimmen also auf nicht invertierbaren Matrizen überein.

Sei nun A eine **invertierbare** $n \times n$-Matrix. In diesem Fall entsteht A aus der Einheitsmatrix E_n durch eine Folge elementarer Spaltenumformungen. Das Lemma 2.2 zeigt uns, dass sich det und det′ unter elementaren Zeilenumformungen identisch verhalten und folglich $\det(A) = \det'(A)$ gilt. □

Existenzbeweis und Rekursionsformel

Der folgende Existenzbeweis liefert zugleich ein rekursives Berechnungsverfahren für die Determinantenabbildung. Mit anderen Worten zeigen wir die Existenz einer Determinantenabbildung $\det : \mathrm{M}_n(K) \to K$ durch Induktion nach n.

Der Fall $n = 1$: Hier erfüllt die $\det : \mathrm{M}_1(K) \to K$, $(a) \mapsto a$ alle Anforderungen.

Der Fall $n > 1$: Wir nehmen an, dass wir schon eine Determinantenfunktion $\det_{n-1}$ für $(n-1) \times (n-1)$-Matrizen haben und konstruieren dann eine $\det_n$ für $n \times n$-Matrizen:

Für $A \in \mathrm{M}_n(K)$ sei A_{ik} die durch Streichen von i-ter Zeile und k-ter Spalte aus A entstehende $(n-1) \times (n-1)$-Matrix. Wir fixieren dann einen Zeilenindex i und setzen

$$(*) \qquad \det_n(A) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij} \cdot \det_{n-1}(A_{ij})$$

(sogenannte Entwicklung nach der i-ten Zeile).

Wegen Lemma 2.2 genügt es zu zeigen, dass $\det_n$ den Eigenschaften **(D 1)**, **(D 2')** und **(D 3)** einer Determinantenfunktion genügt.

Zu (D 3): Im Fall der Einheitsmatrix $A = E_n$ bleibt von (*) nur der Summand $(-1)^{i+i} \det(E_{n-1})$, so dass $\det_n(E_n) = 1$ folgt.

Zu (D 1): Die Linearität von $\det_n$ in den Spalten ist leicht zu verifizieren.

Der Kern des Beweises ist der Nachweis von **(D 2')**. Wir beschränken uns für den Nachweis auf den Fall von zwei gleichen benachbarten Spalten $a_k = a_l$ für $l = k+1$. Dann hat A_{ij} für $j \notin \{k, l\}$ ebenfalls zwei gleiche Spalten und $\det_{n-1}(A_{ij})$ verschwindet. Ferner ist $A_{ik} = A_{il}$ und somit nach (*)

$$\det_n(A) = (-1)^{i+k} \det_{n-1}(A_{ik}) + (-1)^{i+l} \det_{n-1}(A_{il}) = 0. \qquad \square$$

Montag, 2. Februar 2004

Explizite Berechnung der Determinante

Der gerade diskutierte Eindeutigkeitsbeweis sagt uns zugleich, wie wir $\det(A)$ explizit bestimmen können.

Sei A eine $n \times n$-Matrix. Durch **elementare Spaltenumformungen** können wir A auf Spaltenstufenform A' bringen. Falls A' von der Einheitsmatrix verschieden ist, hat A' und damit auch A einen Rang $< n$. In diesem Fall ist $\det(A) = 0$.

Andernfalls ist A invertierbar und $A' = E_n$. Somit hat A' die Determinante 1. Ferner wissen wir (Lemma!), wie elementare Spaltenumformungen den Wert der Determinante ändern. Nur das Vertauschen zweier Spalten und die Multiplikation einer Spalte mit dem Faktor $\lambda \neq 0$ liefern einen Beitrag, nämlich -1 bzw. λ.

Zusammenfassung 2.3 *Wir protokollieren die für die Umformung der invertierbaren Matrix A zur Einheitsmatrix notwendigen elementaren Spaltenumformungen, genauer notieren wir:*

1. *die Faktoren $\lambda_1, \lambda_2, \ldots, \lambda_s$ der verwendeten Multiplikationen von Spalten, und*

2. *die Anzahl t der Spaltenvertauschungen.*

Es ist dann $\det(A) = (-1)^t / \prod_{i=1}^{s} \lambda_i$.

Beispiel 2.4 (Determinantenberechnung) Zu berechnen sei die Determinante der 3×3-Matrix $\begin{bmatrix} 1 & 3 & 5 \\ 1 & 2 & 3 \\ 3 & 2 & 3 \end{bmatrix}$.

Mit $v'(i,j)$, $m'(i,\lambda)$, $a'(i,j,\alpha)$ bezeichnen wir die Spaltenumformungen "Vertauschen von i-ter und j-ter Spalte", "Multiplikation der i-ten Spalte mit $\lambda \neq 0$" bzw. "Addition des α-fachen der i-ten Spalte zur j-ten Spalte". Dann gilt

1. $a'(1,2,-3)$ und $a'(1,3,-5)$ liefern $\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & -2 \\ 3 & -7 & -12 \end{bmatrix}$.

2. $a'(2,1,1), a'(2,3,-1/2)$ und $m'(2,-1)$ liefern dann $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 7 & 2 \end{bmatrix}$.

3. $a'(3,1,2), a'(3,2,-7/2)$ und $m'(3,1/2)$ liefern schließlich die Einheitsmatrix.

Fazit. *Es gab keine Spaltenvertauschungen. Zwei Spaltenmultiplikationen mit den Faktoren -1 bzw. $1/2$ traten auf. Die Determinante von A ist demgemäß -2.*

4.3 Entwicklungssatz

Satz 3.1 (Kennzeichnung der linearen Unabhängigkeit) *Sei $a_1, a_2, \dots, a_n$ ein System von Vektoren des K^n. Äquivalent sind:*

(1) $a_1, a_2, \dots, a_n$ ist linear unabhängig.
(2) $a_1, a_2, \dots, a_n$ ist eine Basis des K^n.
(3) Es ist $\det(a_1, a_2, \dots, a_n) \neq 0$.

Beweis. Die Äquivalenz von (1) und (2) kennen wir schon. (3) $\Rightarrow$ (1): Falls $a_1, a_2, \dots, a_n$ linear abhängig ist, folgt nach Definition der Determinantenfunktion (Eigenschaft (2)), dass $\det(a_1, a_2, \dots, a_n) = 0$. Falls diese Determinante ungleich Null ist, sind daher $a_1, a_2, \dots, a_n$ linear unabhängig.

(1) $\Rightarrow$ (3): Sei umgekehrt $a_1, a_2, \dots, a_n$ linear unabhängig. Dann lässt sich die Matrix $A = [a_1, a_2, \dots, a_n]$ durch elementare Spaltenumformungen zur Einheitsmatrix E_n umformen. Bei dieser Umformung ändert sich die Determinante um einen Skalar $\neq 0$. Wegen $\det(E_n) \neq 0$ folgt dann $\det(A) \neq 0$. □

Satz 3.2 (Entwicklung nach der i-ten Zeile) *Die Determinante einer $n \times n$-Matrix A lässt sich rekursiv durch Entwicklung nach jeder Zeile (hier der i-ten) berechnen:*

$$\det\nolimits_n(A) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij} \cdot \det\nolimits_{n-1}(A_{ij}).$$

Beweis. Die Formel ist Bestandteil des Existenzbeweises für Determinantenfunktionen. □

Bemerkung 3.3 *Für kleine Werte von n empfiehlt sich die Determinantenberechnung nach dieser Methode. Für $n = 2$ ergibt sich sofort die Formel $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$. Für großes n sind wir besser bedient mit der diskutierten Vereinfachung von A durch elementare Zeilenumformungen.*

Spezialfall: Sarrussche Regel

Wir berechnen $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ durch Entwicklung nach der ersten Zeile und erhalten

$$\begin{aligned} & a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ = \;& a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} \\ - \;& (a_{11}\, a_{23}\, a_{32} + a_{12}\, a_{21}\, a_{33} + a_{13}\, a_{22}\, a_{31}), \end{aligned}$$

somit gerade die Sarrussche Regel, die wir schon in Kapitel 1 gesehen haben. □

4.4 Der Determinantenproduktsatz

Beim nächsten Satz handelt es sich um eine nützliche Vorbereitung zum Determinantenproduktsatz.

Satz 4.1 *Sei $v_1, v_2, \ldots, v_n$ ein System von Vektoren des K^n; ferner sei A eine $n \times n$-Matrix. Dann gilt*

$$\det(A \cdot v_1, A \cdot v_2, \ldots, A \cdot v_n) = \det(A) \cdot \det(v_1, v_2, \ldots, v_n).$$

Beweis. Wir fixieren A und betrachten die Abbildung

$$\begin{aligned} d : \mathrm{M}_n(K) \to K, \quad [v_1, v_2, \ldots, v_n] &\mapsto \det(A \cdot v_1, \ldots, A \cdot v_n) \\ &- \det(A) \cdot \det(v_1, v_2, \ldots, v_n). \end{aligned}$$

Diese Abbildung ist linear in den Spalten $v_1, v_2, \ldots, v_n$, sie verschwindet auf linear abhängigen Systemen und ebenfalls auf der Einheitsmatrix E_n. Eine leichte Variante des Eindeutigkeitssatzes zeigt, dass d die Nullabbildung ist. □

Satz 4.2 (Determinanten-Produktsatz) *Seien A und B beides $n \times n$-Matrizen. Dann gilt*

$$|A \cdot B| = |A| \cdot |B|.$$

Beweis. Wir beschreiben die Matrix B durch ihren Spaltenaufbau $B = [b_1, b_2, \ldots, b_n]$ und erhalten $A \cdot B = [A \cdot b_1, \ldots, A \cdot b_n]$. Der vorangehende Satz liefert die Behauptung. □

Folgerung 4.3 (Determinantenformel für inverse Matrix) *Ist A eine invertierbare $n \times n$-Matrix, so folgt*

$$|A^{-1}| = \frac{1}{|A|}.$$

Beweis. Aus $A \cdot A^{-1} = E_n$ folgt $|A| \cdot |A^{-1}| = 1$. □

4.5 Transponieren

Die Determinante ändert sich nicht beim Vertauschen von Zeilen und Spalten, dem sogenannten Transponieren.

Donnerstag, 5. Februar 2004

Satz 4.1 (Invarianz unter Transponieren) *Für jede $n \times n$-Matrix A gilt*

$$|A| = |A^{tr}|.$$

Dabei bezeichnet A^{tr} die zu $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$ transponierte Matrix $A^{tr} = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{bmatrix}$, welche aus A durch Spiegeln an der Hauptdiagonalen entsteht.

Hinweis 4.2 *(ohne Beweis). Man überlegt sich leicht $(A^{tr})^{tr} = A$ und $(A \cdot B)^{tr} = B^{tr} \cdot A^{tr}$.*

Beweis von $|A| = |A^{tr}|$: Wir müssen zeigen, dass $d : \mathrm{M}_n(K) \to K$, $A \mapsto \det(A^{tr})$, eine Determinantenfunktion ist. Für $A = [a_1, a_2, \ldots, a_n]$ hat A^{tr} den Zeilenaufbau $\begin{bmatrix} a_1^{tr} \\ \vdots \\ a_n^{tr} \end{bmatrix}$. Wir zeigen, dass d die Eigenschaften **(D 1)**, **(D 2)** und **(D 3)** einer Determinantenfunktion erfüllt.

(D 3) ist klar, da $E_n^{tr} = E_n$ gilt.

(D 1) Durch Entwicklung von $\det(A^{tr})$ nach der i-ten Zeile folgt, dass $A \mapsto \det(A^{tr})$ in der i-ten Spalte von A linear ist.

(D 2) Wir nehmen an, dass die Spalten von A linear abhängig sind. Wegen der Gleichheit von Zeilenrang und Spaltenrang sind dann die Spalten von A^{tr} linear abhängig und $d(A) = \det(A^{tr}) = 0$ folgt. □

Wir kennen schon die Entwicklung der Determinante einer Matrix nach einer Zeile. Der gerade behandelte Satz zeigt, dass wir bei der Determinantenrechnung generell Zeilen und Spalten vertauschen können. Dies liefert:

Satz 4.3 (Spaltenentwicklung der Determinante) *Ist A eine $n \times n$-Matrix, so gilt für jedes $k = 1, \ldots, n$ die Spaltenentwicklung*

$$|A| = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \cdot |A_{ik}|. \qquad \square$$

Wie früher bedeutet dabei A_{ik} die aus A durch Streichen von i-ter Zeile und k-ter Spalte hervorgehende $n \times n$-Matrix.

Eine schöne Anwendung der Entwicklungsformeln ist:

Satz 4.4 (Die Determinante einer Dreiecksmatrix) *Sei*

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

eine obere Dreiecksmatrix. Dann gilt $|A| = a_{11} \cdot a_{22} \cdots a_{nn}$.

Beweis. Wir verwenden Induktion nach n. Für $n = 1$ ist alles klar. Für $n > 1$ *entwickeln* wir $|A|$ *nach der letzten Zeile* und erhalten $|A| = a_{nn} \cdot |A_{nn}|$. Dabei ist A_{nn} eine $(n-1) \times (n-1)$-Matrix von oberer Dreiecksform, die nach Induktionsvoraussetzung die Determinante $a_{11} \cdots a_{n-1\,n-1}$ hat. □

Folgerung 4.5 *Das entsprechendes Resultat gilt auch für untere Dreiecksmatrizen.*

Beweis. Transponieren macht aus einer unteren eine obere Dreiecksmatrix. Die Determinante ändert sich dabei nicht. □

Effiziente Berechnung von Determinanten

Durch Kombination der Verfahren

- Elementare Zeilen- bzw. Spaltenumformungen
- Entwicklung nach Zeilen bzw. Spalten.

formen wir die zu berechnende Determinante solange in eine Summe von Determinantentermen um, bis für die verbleibenden Determinanten der Wert sofort ablesbar ist, etwa nur noch die Berechnung der Determinanten von Dreiecksmatrizen oder die von 2×2-Matrizen übrig bleibt oder zum Bespiel ein Induktionsargument greift.

Als Beispiel berechnen wir für gegebene Skalare $a_1, a_2, \ldots, a_n$ die Vandermondesche Determinante.

Satz 4.6 (Vandermondesche Determinante) *Es gilt*

$$V_n := \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < k \leq n} (a_k - a_i).$$

Beweis. Wir zeigen die Behauptung durch Induktion nach n. Für $n = 1$ ist die Behauptung korrekt[1]. Wir nehmen nun an, dass die Behauptung für Matrizen des Formats $(n-1) \times (n-1)$ richtig ist. Für den Induktionsschritt subtrahieren wir, rechts beginnend, jeweils wir das a_1-fache der Vorgängerspalte von der aktuellen Spalte des Determinantenausdrucks von V_n, erhalten somit ohne Änderung des Determinantenwerts als neue $i+1$-te Spalte

$$\begin{pmatrix} a_1^i \\ a_2^i \\ \vdots \\ a_n^i \end{pmatrix} - a_1 \begin{pmatrix} a_1^{i-1} \\ a_2^{i-1} \\ \vdots \\ a_n^{i-1} \end{pmatrix} = \begin{pmatrix} 0 \\ (a_2 - a_1)a_2^{i-1} \\ \vdots \\ (a_n - a_1)a_n^{i-1} \end{pmatrix}$$

für alle $i = 1, \ldots, n-1$. Die erste Spalte bleibt dabei unverändert. Es ist somit

$$V_n = \left| \begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 1 & a_2 - a_1 & (a_2 - a_1)a_2 & \cdots & (a_2 - a_1)a_2^{n-2} \\ 1 & a_3 - a_1 & (a_3 - a_1)a_3 & \cdots & (a_3 - a_1)a_3^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n - a_1 & (a_n - a_1)a_n & \cdots & (a_n - a_1)a_n^{n-2} \end{array} \right|$$

Durch Entwicklung der Determinante nach der ersten Zeile von ergibt sich

$$V_n = (a_n - a_1)(a_{n-1} - a_1) \cdots (a_2 - a_1) \cdot \begin{vmatrix} 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-2} \end{vmatrix}.$$

Nach Induktionsvorraussetzung hat die verbleibende $(n-1) \times (n-1)$-Determinante den Wert $\prod_{2 \le i < k \le n} (a_k - a_i)$, woraus der behauptete Ausdruck für V_n folgt. $\square$

[1] Hier verwenden wir die Konvention, dass ein Produkt aus null Faktoren gleich Eins ist.

4.6 Cramersche Regel, Determinantenformeln

Die Cramersche Regel gilt nicht nur für $n = 2$ (Satz 1.2) und $n = 3$ (Satz 9.4), sondern generell.

Satz 4.1 (Cramersche Regel) *Sei $A \cdot x = b$ ein lineares Gleichungssystem mit invertierbarer Matrix mit dem Spaltenaufbau $A = [a_1, a_2, \ldots, a_n]$. Die Koordinaten der eindeutig bestimmten Lösung x des Systems sind dann gegeben durch*

$$x_1 = \frac{|b, a_2, a_3, \ldots, a_n|}{|A|}, \ldots, x_i = \frac{|a_1, \ldots, b, \ldots, a_n|}{|A|}, \ldots, x_n = \frac{|a_1, \ldots, a_{n-1}, b|}{|A|}.$$

In der Formel für x_i steht b in der i-ten Spalte. Bemerkenswert ist die Einfachheit des anschließenden Beweises.

Beweis. Es ist $A \cdot x = \sum_{k=1}^n x_k.a_k = b$. Es folgt

$$\begin{aligned} \det(a_1, \ldots, a_{i-1}, \boxed{b}, a_{i+1}, \ldots, a_n) &= \sum_{k=1}^n \det(a_1, \ldots, a_{i-1}, a_k, a_{i+1}, \ldots, a_n) \\ &= x_i \cdot \det(a_1, \ldots, a_n). \end{aligned}$$ □

Definition 4.2 *Für eine $n \times n$-Matrix A heißt*

$$A^{ad} = (\alpha_{ik}) \text{ mit } \alpha_{ik} = (-1)^{i+k}|A_{ki}|$$

*die zu A **adjungierte Matrix**.*

Satz 4.3 (Adjungierte Matrix) *Für jede $n \times n$-Matrix A gilt*

$$A \cdot A^{ad} = |A|.E_n = A^{ad} \cdot A.$$

Beweis. Wir berechnen den (i, l)-Eintrag b_{il} des Produkts $A \cdot A^{ad}$.
Fall $i = l$: Nach Definition von A^{ad} erhalten wir

$$\begin{aligned} b_{ii} &= \sum_{k=1}^n a_{ik}\, \alpha_{ki} \\ &= \sum_{k=1}^n (-1)^{i+k} a_{ik}|A_{ik}| = |A| \end{aligned}$$

unter Verwendung der Entwicklung von $|A|$ nach der i-ten Zeile.
Fall $i \neq l$: Analog ist für $i \neq l$

$$\begin{aligned} b_{il} &= \sum_{k=1}^n a_{ik}\, \alpha_{kl} \\ &= \sum_{k=1}^n (-1)^{l+k} a_{ik}|A_{lk}| = |A'| \end{aligned}$$

die Zeilenentwicklung derjenigen Matrix A' nach der l-ten Zeile, die aus A durch Ersetzen der l-ten Zeile durch die i-te Zeile entsteht. Da A' zwei gleiche Zeilen hat (die i-te und die l-te), ist $|A'| = 0$ und die Behauptung bewiesen. □

Als Folgerung ergibt sich sofort:

Satz 4.4 (Determinantenformel für inverse Matrix) *Für jede invertierbare Matrix A gilt*

$$A^{-1} = \frac{1}{|A|}.A^{ad}.$$ □

4.6.1 Matrizen von Blockdreiecksform

Satz 4.5 *Für Matrizen in Blockdreiecksform (mit quadratischen Matrizen A und C) gilt*

$$\left|\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right| = |A| \cdot |C|.$$

Beweis. Sei C eine $m \times m$-Matrix. Wir fixieren A und B und betrachten die Abbildung

$$d : \mathrm{M}_m(K) \to K, \quad X \mapsto \left|\begin{array}{c|c} A & B \\ \hline 0 & X \end{array}\right| - |A| \cdot |X|.$$

Es ist klar, dass die Abbildung d in den Zeilen von X linear ist und im Fall linear abhängiger Zeilen von X verschwindet. Ferner sehen wir durch Entwicklung nach den letzten m Zeilen ein, dass $d(E_m) = |A| - |A| = 0$ gilt. Eine Variante des Eindeutigkeitsbeweises für Determinanten zeigt, dass $d = 0$ gilt. □

Folgerung 4.6 *Sei*

$$A = \begin{bmatrix} A_1 & * & * & \cdots & * \\ 0 & A_2 & * & \cdots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{n-1} & * \\ 0 & \cdots & 0 & 0 & A_n \end{bmatrix}$$

eine Matrix in Blockdreiecksform mit Diagonaleinträgen A_i, die quadratische Matrizen sind. Dann gilt

$$|A| = |A_1| \cdot |A_2| \cdots |A_n|.$$

Beweis. Dies ergibt sich aus Satz 4.5 durch Induktion nach n. □

4.7 Permutationen

Anordnungen einer endlichen Menge

Wir stellen uns n Elemente $a_1, a_2, \ldots, a_n$ vor, etwa die natürlichen Zahlen $1, 2, \ldots, n$.
Aufgabe: Bestimme Anzahl aller möglichen Anordnungen.

Für $n = 3$ ergeben sich $6 = 1 \cdot 2 \cdot 3$ Anordnungen oder **Permutationen**:

$$\begin{array}{ccc} (a_1, a_2, a_3) & (a_2, a_1, a_3) & (a_3, a_1, a_2) \\ (a_1, a_3, a_2) & (a_2, a_3, a_1) & (a_3, a_2, a_1) \end{array}$$

Generell definieren wir für eine natürliche Zahl n die Zahl $n!$ (gesprochen: n Fakultät) als Produkt aller natürlichen Zahlen von 1 bis n, also

$$n! = 1 \cdot 2 \cdots (n-1) \cdot n.$$

Satz 4.1 (Anzahl der Permutationen) *Die Anzahl aller Permutationen einer n elementigen Menge ist $n!$.*

Beweis. Wir zeigen die Behauptung durch *Induktion nach n*. Seien dazu $a_1, a_2, \ldots, a_n$ die Elemente der n-elementigen Menge M.

Induktionsverankerung: Für $n = 1$ gibt es nur die Anordnung (a_1).

Induktionsschritt von n auf $n + 1$: Unter der Voraussetzung, dass jede n-elementige Menge genau $n!$ Permutationen hat, zeigen wir, dass jede $n + 1$-elementige Menge $\{a_1, a_2, \ldots, a_{n+1}\}$ genau $(n + 1)!$ Permutationen besitzt. Wir betrachten hierzu die Menge A aller Permutationen von $\{a_1, a_2, \ldots, a_{n+1}\}$.

Für $k = 1, \ldots, n+1$ sei A_k die Menge der mit a_k beginnenden Permutationen. Es ist also, wenn wir generell mit $|A|$ die Anzahl der Elemente einer Menge A bezeichnen,

$$|A| = |A_1| + |A_2| + \cdots + |A_{n+1}|.$$

Nun besteht A_k aus denjenigen Anordnungen, die a_k an der ersten Stelle haben, gefolgt von einer Anordnung der n Elemente

$$a_1, \ldots, \widehat{a_k}, \ldots, a_{n+1} \quad \text{(Das Zeichen } \widehat{\ } \text{ bedeutet Weglassen.)}$$

Also ist $|A_k| = n!$ für jedes k und somit wie behauptet $|A| = (n + 1) \cdot n! = (n + 1)!$. $\square$

Für die folgenden Ausführungen ist ein leichter Wechsel des Standpunkts hilfreich. Unter einer Permutation einer Menge M verstehen wir eine bijektive Abbildung $\sigma : M \to M$. Mit S_M bezeichnen wir die **Menge aller Permutationen** von M. Die Komposition $\sigma \circ \rho$ zweier Permutationen σ und ρ ist wieder eine Permutation von M. Es ist die identische Abbildung 1_M eine Permutation von

M. Falls σ eine Permutation von M ist, so gilt dies auch für die zu σ inverse Abbildung σ^{-1}.

Diese Fakten können wir wie folgt formulieren: Auf der Menge $G = S_M$ aller Permutationen von M liefert die Verknüpfung eine Operation

$$G \times G \to G, \quad (\sigma, \rho) \mapsto \sigma \circ \rho,$$

welche den folgenden Bedingungen genügt:

(G 1) Es gilt $(\tau \circ \sigma) \circ \rho = \tau \circ (\sigma \circ \rho)$ für alle $\rho, \sigma, \tau \in S_M$.

(G 2) Es gibt $1_M \in S_M$ mit $1_M \circ \sigma = \sigma = \sigma \circ 1_M$ für alle $\sigma \in S_M$.

(G 3) Zu jedem $\sigma \in S_M$ gibt es ein $\tau \in S_M$ mit $\sigma \circ \tau = 1_M = \tau \circ \sigma$.

Wir drücken dies so aus, dass S_M bzgl. der Verknüpfung von Permutationen eine **Gruppe** bildet[2].

Falls $M = \{1, 2, \ldots, n\}$, so schreiben wir abkürzend S_n. Für eine Permutation σ von M schreiben wir in diesem Fall

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n-1) & \sigma(n) \end{pmatrix}.$$

Für $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$ ist somit $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$.

Definition von Transpositionen und

Satz 4.2 *Jede Permutation σ ist darstellbar als Produkt $\sigma = \tau_1 \circ \cdots \circ \tau_r$ von Transpositionen.*

Definition 4.3 *Für jede Permutation σ heißt $P(\sigma) = [e_{\sigma(1)}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}]$ die **Permutationsmatrix** von σ. Die Determinante $|P(\sigma)|$ heißt die **Signatur** oder auch das **Vorzeichen** von σ.*

Satz 4.4 *Es gelten die Eigenschaften*

(1) $P(1) = E_n$,
(2) $P(\sigma \circ \rho) = P(\sigma)P(\rho)$,
(3) $P(\sigma^{-1}) = P(\sigma)^{-1}$.

Satz 4.5 *(1) Es gilt* $\operatorname{sgn}(\sigma_1 \circ \sigma_2) = \operatorname{sgn}(\sigma_1) \circ \operatorname{sgn}(\sigma_2)$.

(2) Ist eine Permutation σ ein Produkt von r Transpositionen, so ist $\operatorname{sgn}(\sigma) = (-1)^r$.

Beweis. (1) folgt aus $P(\sigma_1 \circ \sigma_2) = P(\sigma_1) \cdot P(\sigma_2)$ durch Übergang zu Determinanten. Da die Permutationsmatrix einer Transposition aus der Einheitsmatrix durch Vertauschen zweier Spalten entsteht, hat jede Transposition die Signatur -1. Es ist dann (2) eine direkte Folge aus (1). $\square$

[2] Momentan ist dies für uns eine bequeme Redeweise, um die Eigenschaften **(G 1)** — **(G 3)** zusammenzufassen. Den Gruppenbegriff werden wir erst später eingehender untersuchen.

4.8 Die Leibnizsche Regel

Satz 5.1 (Leibnizsche Formel) *Sei $A = (a_{ik})$ eine $n \times n$-Matrix. Dann gilt*

$$|A| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdots a_{\sigma(n),n}.$$

□

Die Summe wird gebildet über alle Permutationen von $\{1, 2, \dots, n\}$. Dabei bezeichnet $\operatorname{sgn}(\sigma)$ das im vorigen Abschnitt eingeführte Vorzeichen der Permutation σ.

Beweis. A habe den Spaltenaufbau $[a_1, a_2, \dots, a_n]$, so dass

$$a_k = \sum_{i=1}^{n} a_{ik} e_i$$

gilt. Wir erhalten somit

$$\begin{aligned}
\det(a_1, a_2, \dots, a_n) &= \det\Big(\sum_{i_1=1}^{n} a_{i_1,1} e_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n,n} e_{i_n}\Big) \\
&= \sum_{i_1,\dots,i_n=1}^{n} a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} \det(e_{i_1}, \dots, e_{i_n}) \\
&= \sum_{\sigma \in S_n} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}.
\end{aligned}$$

□

Berücksichtigung von $|A| = |A^{tr}|$ liefert die ebenfalls gültige Formel

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

4.9 Determinanten von Endomorphismen, Basiswechsel

Sei $B = (b_1, b_2, \ldots, b_n)$ eine Basis des K-Vektorraums V. Damit ist die lineare Abbildung

$$h_B : K^n \to V, \quad \begin{pmatrix} x_1 \\ \ddots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i.b_i$$

ein Isomorphismus von Vektorräumen. Eine K-lineare Abbildung $f : V \to V$ von V in sich nennen wir auch einen **Endomorphismus** von V. Wir wollen erklären, was unter der Determinante von f zu erklären ist. Dazu bestimmen wir zunächst die — von der gewählten Basis B abhängige — Darstellungsmatrix $M_B(f)$ von f und erklären die Determinante von f als Determinante dieser Darstellungsmatrix.

Satz 6.1 *Sei $f : V \to V$ ein Endomorphismus des K-Vektorraums V und sei $B = (b_1, b_2, \ldots, b_n)$ eine Basis von V. Dann gibt es genau eine $n \times n$-Matrix A, so dass das Diagramm*

$$\begin{array}{ccc} K^n & \xrightarrow{h_B} & V \\ A\cdot \downarrow & & \downarrow f \\ K^n & \xrightarrow{h_B} & V \end{array}$$

kommutativ ist, also $f(h_B(x)) = h_B(A \cdot x)$ für alle $x \in K^n$ gilt.

Beweis. Es ist klar, dass es genau eine lineare Abbildung $u : K^n \to K^n$ gibt mit $h_B \circ h = f \circ h_B$, nämlich $u = h_b^{-1} \circ f \circ h_B$. Die Matrix $A = M(u)$ erfüllt dann die Bedingungen des Satzes. $\square$

Wir nennen in diesem Fall A die **Darstellungsmatrix von** f bezüglich der Basis B. Bezeichnung: $M_B(f)$. Falls — mit den verwendeten Bezeichnungen — $f(b_k) = \sum_{i=1}^{n} \alpha_{ik} c_k$ gilt, so bildet

$$\begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{nk} \end{pmatrix}$$

die k-te Spalte von A, also $M_B(f) = (\alpha_{ik})$.

Wir untersuchen nun, wie die Darstellungsmatrix $M_B(f)$ in Abhängigkeit von der Basis B verhält. Seien dazu $B = (b_1, b_2, \ldots, b_n)$ und $C = (c_1, c_2, \ldots, c_n)$ zwei Basen von V. Den Basiswechsel von B nach C bildet dann diejenige $n \times n$-Matrix $W = W_{CB}$ ab, die das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{h_B} & V \\ W\cdot \downarrow & \nearrow h_C & \\ K^n & & \end{array}$$

kommutativ macht. Wir nennen $W = W_{CB}$ die **Basiswechselmatrix** von B nach C. Mit ähnlicher Argumentation wie vorher erhalten wir $W = M(v)$, wenn $v = h_C^{-1} \circ h_B : K^n \to K^n$ gilt. Nach Konstruktion ist daher W_{CB} invertierbar. Konkret sei $c_j = \sum_{i=1}^n \alpha_{ij} b_j$ die Basisdarstellung der c_j in der Basis B. Dann ist $W_{CB} = (\alpha_{ij})$.

Mit $W = W_{CB}$ folgt

$$W_{CB} \cdot M_B(f) \cdot W_{CB}^{-1} = M_C(f),$$

gleichbedeutend

$$M_B(f) = W_{CB}^{-1} \cdot M_C(f) \cdot W_{CB}.$$

Übergang zu Determinanten zeigt unter Verwendung des Determinanten-Produktsatzes, dass $M_B(f)$ und $M_C(f)$ dieselbe Determinante haben, welche man die Determinante des Endomorphismus f nennt.

Bemerkung 6.2 *Durch Betrachtung von Determinantenfunktionen $f : V^n \to V$ lässt sich in Analogie zu Satz 4.1 gleich eine invariante Definition von Endomorphismen vornehmen.*

4.10 Kurzer Historischer Abriss

Der Determinantenbegriff wurde in etwa zeitgleich um 1683 von Leibniz und dem japanischen Mathematiker Seki entwickelt. An ihrer Ausgestaltung zur jetzigen Form waren noch viele Mathematiker beteiligt.

Der Matrizenbegriff ist dagegen sehr viel älter. Schon 200 v.Chr. haben chinesische Mathematiker die Koeffizienten linearer Gleichungssysteme in Form von Rechtecktabellen notiert und elementare Transformationen zur Lösung verwendet.

Eine ausgefeilte Matrizenrechnung (Addition, Multiplikation, Inversenbildung) gibt es andererseits erst seit Mitte des 19-ten Jahrhunderts. Hier ist besonders Cayley hervorzuheben.

4.10.1 Gottfried Wilhelm von Leibniz (1646–1716)

./Bilder/Leibniz_4.jpg

Leibniz wirkte als Universalgelehrter in vielen Disziplinen. Neben Sir Isaac Newton (1643–1727) ist Leibniz der Begründer der Differential- und Integralrechnung. Leibniz begründete die europäische Determinantentheorie zwischen 1678 und 1713.

4.10.2 Takakazu Seki (1642–1708)

./Bilder/Seki_3.jpg

Zeitgleich mit Leibniz hat Seki seine Theorie der Determinanten entwickelt, allerdings 1683 deutlich vor Leibniz veröffentlicht.

4.10.3 Arthur Cayley (1821–1895)

./Bilder/Cayley.jpg

Zu den bedeutendsten Leistungen Cayleys zählt die Entwicklung einer umfassenden Matrizentheorie.

Kapitel 5

Vektorräume mit Skalarprodukt

5.1 Elementare Eigenschaften des Skalarprodukts

Dienstag, 20. April 04

Wollen wir in einem Vektorraum — wie in der anschaulichen Vektorrechnung — auch **Längen und Winkel messen**, benötigen wir einen reellen Vektorraum mit Skalarprodukt. Das resultierende Konzept eines euklidischen Vektorraums vereinigt diejenigen Anforderungen, die wir zusätzlich an einen reellen Vektorraum stellen müssen, um erfolgreich Längen- und Winkelmessung betreiben zu können.

Definition 1.1 *Sei V ein reeller Vektorraum. Ein **Skalarprodukt** auf V ist eine Abbildung*

$$\langle -,- \rangle : V \times V \to \mathbb{R}, \quad (x,y) \mapsto \langle x,y \rangle,$$

welche die folgenden Bedingungen erfüllt:

(S 1) *$\langle -,- \rangle$ ist **bilinear**, d.h. in jeder der beiden Variablen linear.*

(S 2) *$\langle -,- \rangle$ ist **symmetrisch**, d.h. es gilt $\langle x,y \rangle = \langle y,x \rangle$ für alle $x,y \in V$.*

(S 3) *$\langle -,- \rangle$ ist **positiv definit**, d.h. es gilt $\langle x,x \rangle > 0$ für alle $0 \neq x \in V$.*

*Ein Paar $(V, \langle -,- \rangle)$ bestehend aus einem reellen Vektorraum V und einem Skalarprodukt $\langle -,- \rangle$ auf V heißt **euklidischer Vektorraum**.*

Beispiele 1.2 *(a) Im Anschauungsraum ist durch $\langle x,y \rangle = |x| \cdot |y| \cdot \cos\sphericalangle(x,y)$ ein Skalarprodukt gegeben.*

(b) Auf dem $\mathbb{R}^n$ ist durch $\langle x,y \rangle = \sum_{i=1}^n x_i y_i$ ein Skalarprodukt gegeben.

(c) Ist V der Vektorraum der auf $[0,1]$ stetigen Funktionen, so wird auf V durch*

$$\langle f,g \rangle = \int_0^1 f(x) \cdot g(x) dx$$

ein Skalarprodukt erklärt[1].

Definition 1.3 *Ist $(V, \langle -,- \rangle)$ ein euklidischer Vektorraum und $x \in V$, so heißt*

$$||x|| = \sqrt[+]{\langle x,x \rangle}$$

*die **Norm** von x.*

Wir wollen nun den Winkel $\sphericalangle(x,y)$ zwischen zwei Vektoren $x, y \neq 0$ durch einen durch die anschauliche Vektorrechnung nahegelegten Ansatz

$$\langle x,y \rangle = ||x|| \cdot ||y|| \cdot \cos\sphericalangle(x,y)$$

[1]Der Nachweis von **(S 3)** erfordert etwas Aufwand: Sei $g : [0,1] \to \mathbb{R}$ eine stetige Funktion mit $g(x) \geq 0$ für alle $x \in [0,1]$ und $b := g(x_0) > 0$ für ein $x_0 \in [0,1]$. Wegen der Stetigkeit von g gibt dann eine Umgebung $[x_0 - a, x_0 + a]$, $a > 0$, von x_0, in der $g(x) > b/2$ ist. Es folgt dann $\int_0^1 g(x)dx \geq ab > 0$.

erklären, wobei $\sphericalangle(x, y)$ durch den resultierenden Cosinus-Wert

$$\cos \sphericalangle(x, y) = \frac{\langle x, y\rangle}{||x|| \cdot ||y||}$$

dann im Intervallbereich $0 \leq \sphericalangle(x, y) \leq \pi$ eindeutig bestimmt ist.

Allerdings kann dieser Ansatz nur funktionieren, wenn (weiterhin unter der Annahme $x, y \neq 0$) stets

$$-1 \leq \frac{\langle x, y\rangle}{||x|| \cdot ||y||} \leq 1$$

gilt. Dass diese Ungleichung tatsächlich immer erfüllt ist, ist die Aussage der folgenden Cauchy-Schwarzschen Ungleichung.

Satz 1.4 (Cauchy-Schwarzsche Ungleichung) *In jedem euklidischen Vektorraum* V[2] *gilt*

$$|\langle x, y\rangle| \leq ||x|| \cdot ||y||$$

für alle x, y aus V.

Beweis. Für $y = 0$ verschwinden beide Seiten der Ungleichung und wir sind fertig. Wir nehmen daher $y \neq 0$ an. Für jedes reelle λ ist

$$\begin{aligned} 0 &\leq \langle x + \lambda y, x + \lambda y\rangle \\ &= \langle x, x\rangle + 2\lambda\langle x, y\rangle + \lambda^2\langle y, y\rangle \\ &= ||x||^2 + 2\lambda\langle x, y\rangle + \lambda^2||y||^2. \end{aligned}$$

Es folgt nach Teilen durch $||y||^2$

$$\begin{aligned} 0 &\leq \lambda^2 + 2\lambda\frac{\langle x, y\rangle}{||y||^2} + \frac{||x||^2}{||y||^2} \\ &= \left(\lambda + \frac{\langle x, y\rangle}{||y||^2}\right)^2 + \frac{||x||^2}{||y||^2} - \frac{\langle x, y\rangle^2}{||y||^4}. \end{aligned}$$

Diese Ungleichung wird für $\lambda = -\frac{\langle x,y\rangle}{||y||^2}$ am besten ausgenutzt. Es folgt dann

$$\frac{\langle x, y\rangle^2}{||y||^4} \leq \frac{||x||^2}{||y||^2}, \text{ somit } \langle x, y\rangle^2 \leq ||x||^2 \cdot ||y||^2$$

und dann durch Bilden der (positiven) Quadratwurzel die Behauptung. □

Freitag, 23. April 04

Folgerung 1.5 *Im Spezialfall des mit dem Standard-Skalarprodukt versehenen $\mathbb{R}^n$ ergibt sich:*

$$\left|\sum_{i=1}^n x_i\, y_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}.$$

□

[2]Von nun an unterdrücken wir die explizite Angabe des Skalarprodukts.

Folgerung 1.6 *Wir können für $x, y \in V$, $x, y \neq 0$, durch*

$$\cos\alpha = \frac{\langle x, y\rangle}{||x|| \cdot ||y||}$$

*den im Intervallbereich $0 \leq \alpha \leq \pi$ liegenden **Winkel** $\alpha = \sphericalangle(x, y)$ zwischen x und y erklären.* □

Wir nehmen weiterhin $x \neq 0 \neq y$ an. Es ist $\sphericalangle(x, y) = \frac{\pi}{2}$ genau wenn $\langle x, y\rangle = 0$ gilt. Auf dieses Konzept von Orthogonalität werden wir im kommenden Abschnitt näher eingehen.

Satz 1.7 (Norm) *Sei V ein euklidischer Vektorraum. Dann hat die Norm(funktion)*

$$||-|| : V \longrightarrow \mathbb{R}, \quad x \mapsto ||x||$$

die folgenden Eigenschaften:
(N 1) *Es ist $||x|| \geq 0$ für alle $x \in V$.*
(N 2) *Es ist $||x|| = 0$ genau wenn $x = 0$.*
(N 3) *Es gilt $||\lambda\, x|| = |\lambda| \cdot ||x||$ for alle $x \in V$ und $\lambda \in \mathbb{R}$.*
(N 4) *Es gilt die Dreiecksungleichung $||x + y|| \leq ||x|| + ||x||$ für alle $x, y \in V$.*

Beweis. Nur **(N 4)** bedarf eines Beweises. Es ist

$$\begin{array}{rcl} ||x + y||^2 & = & \langle x + y, x + y\rangle \\ & = & ||x||^2 + 2\langle x, y\rangle + ||y||^2 \\ & \leq & ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2 \text{ (Cauchy-Schwarz)} \\ & = & (||x|| + ||y||)^2. \end{array}$$ □

5.2 Orthogonalität

Definition 2.1 *Es sei V ein euklidischer Vektorraum. Wir nennen $x, y \in V$ zueinander **orthogonal**, wenn $\langle x, y\rangle = 0$ gilt.*

Der Nullvektor ist damit orthogonal zu jedem Vektor. Falls x und y beide nicht Null sind, bedeutet Orthogonalität nach unserer Winkelerklärung, dass der Winkel zwischen x und y gerade $\pi/2$ oder gleichbedeutend $90°$ ist.

Definition 2.2 *Es sei V ein euklidischer Vektorraum. Ein System $v_1, v_2, \ldots, v_n$ von Vektoren aus V nennen wir **Orthogonalsystem** (bzw. **Orthonormalsystem**), falls*

(1) $$v_i \neq 0 \text{ (bzw. } ||v_i|| = 1\text{) für jedes } i = 1, \ldots, r.$$

(2) $$\langle v_i, v_k\rangle = 0 \text{ für alle } i \neq k.$$

Mit Hilfe des Kronecker-Symbols können wir die Anforderung an ein Orthonormalsystem zu

$$\langle v_i, v_k\rangle = \delta_{ik} \quad i, k = 1, \ldots, r$$

zusammenfassen.

Satz 2.3 *Jedes Orthogonalsystem $v_1, v_2, \ldots, v_r$ eines euklidischen Vektorraums ist linear unabhängig.*

Gilt zusätzlich $r = \dim V$, so ist $v_1, v_2, \ldots, v_r$ eine Basis von V. Wir sprechen dann von einer **Orthonormalbasis**von V.

Beweis. Es sei

$$\lambda_1 v_1 + \cdots + \lambda_i v_i + \cdots + \lambda_r v_r = 0.$$

Wir bilden das Skalarprodukt mit v_i und erhalten

$$\begin{aligned} 0 &= \langle v_i, \sum_{j=0}^{n} \lambda_j v_j\rangle \\ &= \sum_{j=0}^{r} \lambda_j \langle v_i, v_j\rangle \\ &= \lambda_i \langle v_i, v_i\rangle. \end{aligned}$$

∎ Falls U ein Unterraum von V und $v \in V$ ist, schreiben wir $U^\perp$ für die Menge aller $v \in V$ für die $\langle v, U\rangle = 0$ gilt, also $\langle v, u\rangle = 0$ für alle $u \in U$ gilt. Wir nennen $U^\perp$ das **orthogonale Komplement** zu U in V. Immer gilt:

Lemma 2.4 *Sei U ein Unterraum eines euklidischen Vektorraums V. Dann ist $U^\perp$ ein Unterraum von V mit $U \cap U^\perp = \{0\}$.*

Beweis. Es ist klar, dass $0 \in U^\perp$ gilt. Sind ferner x und y in U gelegen, so gilt $\langle x-y,u\rangle = \langle x,u\rangle - \langle y,u\rangle = 0$ für jedes $u \in U$. Damit ist $x-y$ ebenfalls in $U^\perp$ gelegen und $U^\perp$ damit ein Unterraum von V.

Für $x \in U \cap U^\perp$ gilt $\langle x,x\rangle = 0$, also $x = 0$. □

Satz 2.5 (Projektionslemma) *Sei V ein euklidischer Vektorraum und U ein Unterraum von V mit der Orthonormalbasis $u_1, u_2, \ldots, u_r$. Dann lässt sich jedes v aus V eindeutig in der Form*

$$v = u + w \text{ mit } u \in U \text{ und } w \in U^\perp$$

schreiben. Es gilt dabei $u = \sum_{i=1}^r \langle v, u_i\rangle u_i$.

Beweis. *Existenz*: Wir setzen $u = \sum_{i=1}^r \langle v, u_i\rangle u_i$ und $w = v-u$. Es ist dann $u \in U$ und wegen der Orthogonalität $\langle u, u_j\rangle = \langle v, u_j\rangle$, somit $\langle w, u_j\rangle = 0$ für $j = 1, \ldots, n$, woraus $w \in U^\perp$ folgt.

Eindeutigkeit: Wir nehmen an, dass

$$u + w = u' + w' \text{ mit } u, u' \in U \text{ und } w, w' \in U^\perp$$

gilt. Es folgt, dass das Element $(u - u') = (w' - w)$ in $U \cap U^\perp$ liegt und damit gleich Null ist. □

Wir beantworten nun die Frage, wieweit Orthonormalbasen existieren. Das Projektionslemma spielt dabei die Rolle des Induktionsschritts.

Satz 2.6 *Jeder endlichdimensionale euklidische Vektorraum hat eine Orthonormalbasis.*

Beweis. Wir verfahren durch Induktion nach $n = \dim V$. Sei $v_1, v_2, \ldots, v_n$ eine Basis von V.

$n = 1$: Wir normieren den Basisvektor v_1 und erhalten mit $e_1 = v_1/||v_1||$ eine Orthonormalbasis von V.

$n - 1 \to n$: Nach Induktionsvoraussetzung besitzt der von $v_1, v_2, \ldots, v_{n-1}$ aufgespannte Unterraum U eine Orthonormalbasis $e_1, e_2, \ldots, e_{n-1}$. Mit Hilfe des Projektionslemmas können wir den Vektor v_n in der Form $v_n = u+w$ schreiben, wobei $w \in U^\perp$ und $u = \sum_{i=1}^{n-1} \langle v_n, e_i\rangle e_i$. Nach Voraussetzung gehört v_n nicht zu U, daher ist $w \neq 0$. Durch Normieren von w erhalten wir einen weiteren Einheitsvektor

$$e_n = \frac{v_n - \sum_{i=1}^{n-1} \langle v_n, e_i\rangle e_i}{||v_n - \sum_{i=1}^{n-1} \langle v_n, e_i\rangle ||e_i},$$

der zusammen mit $e_1, e_2, \ldots, e_{n-1}$ eine Orthonormalsystem aus $n = \dim V$ Vektoren und damit eine Orthonormalbasis von V bildet. □

Bemerkung 2.7 *Explizite Durchformung des Induktionsarguments liefert das sogenannte* ***Schmidtsche Orthonormalisierungsverfahren****, welches konstruktiv aus einer Basis $v_1, v_2, \ldots, v_n$ von V eine Orthonormalbasis von V macht.*

Dienstag, 27. April 04

Satz 2.8 *V sei ein endlichdimensionaler euklidischer Vektorraum. Für jeden Unterraum U von V gilt dann*
(a) $U + U^{\perp} = V$, $U \cap U^{\perp} = \{0\}$, $\langle U, U^{\perp} \rangle = 0$.
(b) $(U^{\perp})^{\perp} = U$.

Wir sagen in der Situation (a), dass V die **orthogonale Summe** von U und $U^{\perp}$ ist und schreiben dafür $V = U \perp U^{\perp}$.

Beweis. Zu (a): Es ist U ebenfalls endlichdimensional. Nach Satz 2.6 hat U dann eine Orthonormalbasis, so dass wir auf U das Projektionslemma anwenden können. Dasselbe zeigt $U + U^{\perp} = V$; die übrigen Eigenschaften von (a) sind offensichtlich.

Zu (b): Nach Definition ist $U \subseteq (U^{\perp})^{\perp}$. Sei $v \in (U^{\perp})^{\perp}$, so lässt sich v wegen $V = U + U^{\perp}$ in der Form $v = u + x$ mit $u \in U$ und $x \in U^{\perp}$ schreiben. Bilden des Skalarprodukts mit x führt zu $0 = \langle v, x \rangle = \langle u, x \rangle + \langle x, x \rangle = \langle x, x \rangle$, woraus $x = 0$ und damit $v = u \in U$ folgt. □

Orthonormalbasen lassen sich im Zusammenhang mit Koordinatendarstellungen sehr gut handhaben.

Satz 2.9 *Sei $e_1, e_2, \ldots, e_n$ eine Orthonormalbasis des euklidischen Vektorraums V. Dann gilt:*

(1) *Für jedes $v \in V$ gilt $v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i$.*

(2) $|| \sum_{i=1}^{n} x_i e_i || = \sqrt{\sum_{i=1}^{n} x_i^2}$.

(3) $\langle \sum_{i=1}^{n} x_i e_i, \sum_{i=1}^{n} y_i e_i \rangle = \sum_{i=1}^{n} x_i y_i$.

Beweis. Die Beweise ergeben sich durch direktes Ausrechnen. □

5.3 Orthogonale Abbildungen

Definition 3.1 *V und W seien euklidische Vektorräume. Eine lineare Abbildung $f : V \to W$ heißt **orthogonal**, wenn*

$$\langle f(v_1), f(v_2)\rangle_W = \langle v_1, v_2\rangle_V$$

für alle $v_1, v_2 \in V$ gilt.

Eine orthogonale Abbildung bewahrt insbesondere die Norm, ist daher stets injektiv. Von besonderem Interesse sind die orthogonalen Selbstabbildungen $f : V \to V$. Dieselben sind für einen endlichdimensionalen Vektorraum V stets Isomorphismen. Die folgenden Eigenschaften sind leicht zu überprüfen:

(1) Die identische Abbildung 1_V ist orthogonal.

(2) Die Komposition von orthogonalen Abbildungen ist orthogonal.

(3) Ist $f : V \to W$ eine bijektive orthogonale Abbildung, so ist auch $f^{-1} : W \to V$ orthogonal.

Insbesondere gilt:

Satz 3.2 *Die orthogonalen Selbstabbildungen eines endlichdimensionalen euklidischen Vektorraums bilden bezüglich der Verknüpfung von Abbildungen eine Gruppe $\mathbf{O}(V)$, die **orthogonale Gruppe** von V.* □

Dabei heißt ein Paar $(G, \circ)$, bestehend aus einer Menge G und einer Verknüpfung $\circ : G \times G \to G$, $(x, y) \mapsto x \circ y$ — nachfolgend oft Multiplikation genannt — eine **Gruppe**, wenn die folgenden Bedingungen erfüllt sind:

(G 1) Die Verknüpfung ist assoziativ: $a \circ (b \circ c) = (a \circ b) \circ c$.

(G 2) Es gibt ein neutrales Element e mit $a \circ e = a = e \circ a$.

(G 3) Zu jedem $x \in G$ gibt es ein Inverses y mit $x \circ y = e = y \circ x$.

Freitag, 30. April 04

Durch die Anforderungen sind neutrales Element und Inverses eindeutig bestimmt:

Nehmen wir an, dass e und e' neutrale Elemente von $(G, \circ)$ sind, so folgt $e = e \circ e' = e'$. Fortan sprechen wir also von *dem* neutralen Element der Gruppe $(G, \circ)$.

Seien ferner y und z zu x inverse Elemente, so folgt

$$y = y \circ e = y \circ (x \circ z) = (y \circ x) \circ z = e \circ z = z,$$

so dass wir künftig von *dem* zu x inversen Element x^{-1} aus G sprechen können. Es gelten die folgenden Rechenregeln:

$$e^{-1} = e, \quad (x^{-1})^{-1} = x, \quad (x \circ y)^{-1} = y^{-1} \circ x^{-1}.$$

Die erste Behauptung folgt aus $e \circ e = e$, die zweite folgt aus $x \circ x^{-1} = e = x^{-1} \circ x$. Für die dritte sehen die Gleichung $(x \circ y) \circ (y^{-1} \circ x^{-1}) = x \circ (y \circ y^{-1}) \circ x^{-1} = x \circ x^{-1} = e$ und die entsprechend nachzuweisende Gleichung $(y^{-1} \circ x^{-1}) \circ (x \circ y) = e$ an, aus denen auch hier die Behauptung folgt.

Die Gruppenaxiome **(G 1)—(G 3)** sind offensichtlich im Fall $\mathbf{O}(V)$ bzgl. der üblichen Verknüpfung von Abbildungen erfüllt. Insbesondere ist die Assoziativität erfüllt, weil das Nacheinanderausführen von Abbildungen dem Assoziativgesetz genügt. Ferner übernimmt 1_V die Rolle des neutralen Elements und die zu u inverse Abbildung u^{-1} die Rolle des zu u in $\mathbf{O}(V)$ inversen Elements.

Weitere Beispiele von Gruppen sind uns, ohne das wir sie als solche bezeichnet hätten, schon früher begegnet:

(1) Die Gruppe $(\mathbb{Z}, +)$ bezüglich der Addition.

(2) Sei K ein Körper. Dann ist $(K, +)$ eine Gruppe, aber auch $(K \setminus \{0\}, \cdot)$. Insbesondere gilt dies für $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$.

In diesen Gruppen gilt stets $x \circ y = y \circ x$. Sie sind **kommutativ** oder auch **abelsch**[3].

Ein weiteres Beispiel einer Gruppe wird durch die Gruppe $\mathbf{GL}(V)$ der K-Automorphismen $f : V \to V$ eines K-Vektorraums gebildet. Diese Gruppe heißt die **allgemeine lineare Gruppe** oder auch **volle lineare Gruppe** von V[4].

Ist zusätzlich V ein euklidischer Vektorraum, so ist $\mathbf{O}(V)$ in $\mathbf{GL}(V)$ gelegen; ferner stimmen beide Gruppenmultiplikationen für Elemente aus $\mathbf{O}(V)$ überein: es ist $\mathbf{O}(V)$ eine Untergruppe von $\mathbf{GL}(V)$.

Sei generell $(G, \circ)$ eine Gruppe. Wir nennen eine Teilmenge U von G **Untergruppe** von G, falls U gegen die Verknüpfung von G abgeschlossen ist und bezüglich der resultierenden Verknüpfung $\circ_U : U \times U \to U$, $(x, y) \mapsto x \circ y$, selbst eine Gruppe ist.

Es ist nicht schwer nachzuweisen, dass U genau dann eine Untergruppe von G ist, wenn die folgenden drei Eigenschaften erfüllt sind:

(U 1) Für $x, y \in U$ gehört auch $x \circ y \in U$.

(U 2) Das neutrale Element e von G gehört zu U.

(U 3) Für jedes $x \in U$ gehört auch das in G gebildete x^{-1} zu U.

[3] Nach dem norwegischen Mathematiker Niels Henrik Abel (1802–1829)

[4] Dabei steht GL für general linear.

Definition 3.3 *Eine $n \times n$-Matrix A heißt* ***orthogonal****, wenn die lineare Abbildung*

$$f : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto A \cdot x$$

orthogonal ist, d.h. wenn $\langle Ax, Ay \rangle = \langle x, y \rangle$ für alle $x, y \in \mathbb{R}^n$ gilt.

Aus den Eigenschaften von $\mathbf{O}(\mathbb{R}^n)$ folgt sofort, dass die Einheitsmatrix zu den orthogonalen $n \times n$-Matrizen gehört, dass dieselben ferner gegen Produktbildung abgeschlossen sind und schließlich die zu einer orthogonalen Matrix inverse Matrix wieder orthogonal ist. Kurz: die orthogonalen $n \times n$-Matrizen bilden bezüglich der Matrixmultiplikation eine Gruppe $\mathbf{O}(n)$, die **orthogonale Gruppe** des Grades n.

Ganz entsprechend bilden die *invertierbaren* $n \times n$-Matrizen über einem Körper K hinsichtlich Matrixmultiplikation eine Gruppe $\mathbf{GL}_n(K)$, die als allgemeine oder **volle lineare Gruppe** des Grades n bezeichnet wird.

> Sind G und H Gruppen, so heißt eine Abbildung $u : G \to H$ **Gruppenhomomorphismus**, wenn $u(x \circ y) = u(x) \circ u(y)$ für alle $x, y \in G$ gilt. Für Gruppen sind diese Abbildungen daher die strukturbewahrenden Abbildungen.
>
> Falls u zusätzlich bijektiv ist, sprechen wir von einem **Isomorphismus** $u : G \to H$ von Gruppen. Es ist nicht schwer zu sehen, dass in diesem Fall auch die Umkehrabbildung $u^{-1} : H \to G$ ein Gruppenhomomorphismus und dann ebenfalls ein Isomorphismus ist. Falls es einen Isomorphismus $u : G \to H$ gibt, nennen wir die Gruppen G und H isomorph, Schreibweise $G \cong H$. Wir verbinden mit diesem Begriff die Vorstellung, dass isomorphe Gruppen über "dieselben" mathematischen Eigenschaften verfügen.

Es sind $\mathbf{GL}_n(K)$ und $\mathbf{GL}(K^n)$ isomorphe (aber nicht gleiche) Gruppen. Es ist nämlich durch

$$\varphi : \mathbf{GL}_n(K) \to \mathbf{GL}(K^n), \quad A \mapsto (x \mapsto A \cdot x)$$

ein Isomorphismus definiert mit Umkehrabbildung

$$\psi : \mathbf{GL}(K^n) \to \mathbf{GL}_n(K), \quad f \mapsto [f(e_1), \ldots, f(e_n)].$$

Mit gleichlautenden Abbildungsvorschriften sind für den mit Standardskalarprodukt versehenen $\mathbb{R}^n$ die orthogonalen Gruppen $\mathbf{O}(n)$ und $\mathbf{O}(\mathbb{R}^n)$ ebenfalls isomorph.

Der nächste Satz gibt eine konkrete Beschreibung der orthogonalen Matrizen.

Satz 3.4 *Für eine reelle $n \times n$-Matrix A sind äquivalent:*

(1) *A ist orthogonal.*

(2) *$A^{tr} \cdot A = E_n$ (bzw. $A \cdot A^{tr} = E_n$).*

(3) *Die Spalten (bzw. Zeilen) von A bilden ein Orthonormalsystem.*

Beweis. Wir betrachten die Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto A \cdot x$. Für die folgenden Argumente merken wir zunächst an, dass $A^{tr} \cdot A = E_n$ zu $A^{-1} = A^{tr}$ und somit zu $A \cdot A^{tr} = E_n$ äquivalent ist. Wir können uns beim Nachweis der Äquivalenz von (2) und (3) daher anschließend auf die Spaltenversion von (3) beschränken.

(1) ⇒ (3): Es ist $A = [f(e_1), f(e_2), \ldots, f(e_n)]$, wenn $e_1, e_2, \ldots, e_n$ die Standardbasis von A bezeichnet. Da f als orthogonale Abbildung Skalarprodukte bewahrt, ist mit $e_1, e_2, \ldots, e_n$ auch $f(e_1), f(e_2), \ldots, f(e_n)$ ein Orthonormalsystem.

(3) ⇔ (2): Ist $A = [a_1, a_2, \ldots, a_n]$, so ist $A^{tr} = \begin{bmatrix} {a_1}^{tr} \\ \vdots \\ {a_n}^{tr} \end{bmatrix}$ und daher

$$A^{tr} \cdot A = \begin{bmatrix} \langle a_1, a_1 \rangle & \cdots & \langle a_1, a_n \rangle \\ \vdots & & \vdots \\ \langle a_n, a_1 \rangle & \cdots & \langle a_n, a_n \rangle \end{bmatrix}.$$

(3) ⇒ (1): Sei wieder $A = [a_1, a_2, \ldots, a_n]$. Für $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$ ist $f(x) = \sum_{i=1}^n x_i a_i$ und $f(y) = \sum_{i=1}^n y_i a_i$. Da $a_1, a_2, \ldots, a_n$ ein Orthonormalsystem ist, folgt die Orthogonalität von f aus

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = \langle f(x), f(y) \rangle.$$

Zur letzten Gleichheit haben wir Satz 2.9 (3) benutzt. □

Es gibt somit eine natürliche Bijektion zwischen (geordneten) Orthonormalbasen $(a_1, a_2, \ldots, a_n)$ des $\mathbb{R}^n$ einerseits und orthogonalen $n \times n$-Matrizen $a_1, a_2, \ldots, a_n$ andererseits.

Jede orthogonale Matrix $A \in \mathbf{O}(n)$ erfüllt wegen $A \cdot A^{tr} = E_n$ die Bedingung $|A|^2 = 1$, hat also die Determinante ± 1. Hier gehen ein der Determinantenproduktsatz und die Invarianz der Determinante unter Transponieren.

Die orthogonalen Matrizen mit Determinante 1 bilden — wegen des Determinantenproduktsatzes — eine Untergruppe

$$\mathbf{SO}(n) = \{ A \in \mathbf{O}(n) \mid |A| = 1 \},$$

von $\mathbf{O}(n)$, die **spezielle orthogonale Gruppe** des Grades n.

Beispiel 3.5 (Drehmatrizen) *Für jeden Winkel α liegt*

$$D(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

in $\mathbf{SO}(2)$. Es gelten die folgenden Eigenschaften:

(1) $D(\alpha + \beta) = D(\alpha) \cdot D(\beta)$.

(2) $D(0) = E_2$.

(3) $D(-\alpha) = D(\alpha)^{-1}$.

(4) $|D(\alpha)| = 1$.

Ferner ist $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ *in* $\mathbf{O}(2) \setminus \mathbf{SO}(2)$ *gelegen. Allgemeiner sind alle Matrizen der Form*

$$A \cdot D(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & -\cos\alpha \end{bmatrix}$$

orthogonale Matrizen mit Determinante -1.

Es ist nicht schwer zu zeigen, siehe das Argument im nachfolgenden Beispiel, dass jede orthogonale 2×2-Matrix entweder eine Drehmatrix $D(\alpha)$ oder eine Matrix von der Form $A \cdot D(\alpha)$ ist. Mit anderen Worten

$$\mathbf{SO}(2) = \{D(\alpha) \mid \alpha \in \mathbb{R}\}$$

und mit obiger Bedeutung von A gilt

$$\mathbf{O}(2) = \mathbf{SO}(2) \cup \{A \cdot X \mid X \in \mathbf{O}(2)\}.$$

Wie ihr Name sagt, bewirken die Matrizen $D(\alpha)$ in der Ebene $\mathbb{R}^2$ eine Drehung $\mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto D(\alpha) \cdot x$, um den Winkel α. Dagegen bewirken die Matrizen $A \cdot D(\alpha)$ eine Drehspiegelung, d.h. eine Drehung verknüpft mit einer Spiegelung.

Beispiel 3.6 *Die komplexen Zahlen* $\mathbf{U} = \{z \in \mathbb{C} \mid |z| = 1\}$ *vom Betrag* 1 *bilden bezüglich der Multiplikation komplexer Zahlen eine Gruppe. Hierzu berücksichtigen wir, dass* $\mathbf{U}$ *ersichtlich gegen Multiplikation abgeschlossen ist, dass* 1 *in* $\mathbf{U}$ *liegt und dass schließlich für jede komplexe Zahl* z *vom Betrag* 1 *auch* $1/z$ *den Betrag* 1 *hat.*

Die Gruppen $\mathbf{U}$ *und* $\mathbf{SO}(2)$ *sind isomorph, denn die Abbildung*

$$\varphi : \mathbf{U} \to \mathbf{SO}(2), \quad a + b\,i \mapsto \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

ist ein Isomorphismus mit Umkehrabbildung

$$\psi : \mathbf{SO}(2) \to \mathbf{U}, \quad \begin{bmatrix} a & c \\ b & d \end{bmatrix} \mapsto a + b\,i.$$

Zu zeigen ist hier nur, dass $\begin{bmatrix} a & c \\ b & d \end{bmatrix} \in \mathbf{SO}(2)$ impliziert, dass $c = -b$ und $a = d$ ist. Den Nachweis lassen wir als Übungsaufgabe.

Index