Lineare Algebra

B. Herzog, Universität Leipzig, Institut für Mathematik und Informatik,
Vorlesung des ersten Studienjahrs im Herbstsemester 2012

Mo. 13-15 Uhr in Hs. 2
Mi. 9-11 Uhr in Hs. 2
Klausur: am 4.2.13, 10.00-12.00 Uhr im Auditorium Maximum (Augusteum)
Nachklausur: am 4.4.13, 10.00-12.00 Uhr im Auditorium Maximum (Augusteum)

Warnung

Dies ist ein Manuskript, also nicht fehlerlos. Sie sollten nicht alles kritiklos hinnehmen, was hier steht !

Hinweise

Übungsaufgaben

Am Anfang jeder Woche werden jeweils 3 Aufgaben ins Netz gestellt (www.math.uni-leipzig.de).

Die Lösungen dieser Aufgaben sind am Anfang der Montagsvorlesung der nachfolgenden Woche abzugeben.

Für die Lösung einer Aufgabe werden bis zu 4 Punkte vergeben.

Am Ende jedes Semesters findet eine Klausur statt. Um zur Klausur zugelassen zu werden, müssen sie mindestens 60% der Gesamtpunktzahl für die Lösung der Aufgaben erhalten haben.

Sie sollten versuchen alle Aufgaben zu lösen, auch wenn ihre aktuelle Punktzahl oberhalb von 60% des Maximums liegt, denn die Aufgaben werden nicht leichter werden.

Für das Verständnis der Vorlesung ist es für Sie von entscheidender Bedeutung, daß Sie die Aufgaben selbständig lösen. Es ist nichts dagegen einzuwenden, wenn Sie mit anderen Studententen über die in den Aufgaben formulierten Probleme diskutieren. Aber vorher sollten Sie allein darüber nachdenken, auch wenn das unangenehm, anstrengend und vielleicht sogar frustierend sein sollte. Jede Idee, die Sie beim Nachdenken über die Aufgaben bekommen, wird Ihnen das Lösen zukünftiger Aufgaben und das Verständnis der Vorlesung erleichtern. Jede, die Sie nicht bekommen, wird Ihre Chancen bei zukünftigen Aufgaben und beim Bestehen der Abschlußprüfungen verringern.

Auch wenn Sie die Aufgaben nicht selbständig lösen können und sich die Lösungen bei anderen holen, sollten Sie Ihre Lösung eigenständig zu formulieren versuchen. Falls Sie immer wieder feststellen, daß Sie die Aufgaben trotz intensiver Bemühungen nicht lösen können, sollten Sie darüber nachdenken, ob Sie nicht die Studienrichtung wechseln sollten, denn Sie werden spätestens im zweiten Studienjahr, die notwendigen Prüfungen nicht mehr bestehen.

Die Korrektoren der Aufgaben sind nicht angewiesen, aufwendig nach Abschreibern zu suchen. Wenn ihnen jedoch identische Formulierungen auffallen, werden sie die für die Aufgabe vergebene Punktzahl durch die Anzahl der Exemplare teilen, in denen die identischen Formulierungen vorkommen, d.h. bei einer Fünfergruppe mit identischen Formulierungen bekommt jeder der Fünf nur ein Fünftel der Punktzahl.

Zum Vorlesungsbesuch

Niemand wird kontrollieren, ob Sie eine Vorlesung besuchen oder nicht. Bei den Klausuren und Prüfungen wird niemanden interessieren, ob Sie die Vorlesung besucht haben oder nicht. In den vergangenen 30 Jahren habe ich allerdings niemanden kennengelernt, der die Prüfungen bestanden hat ohne die Vorlesung zu besuchen.

Zu Beginn jeder Vorlesung sollten Sie die vergangenen Vorlesungen so gut kennen, daß Sie der aktuellen Vorlesung relativ mühelos folgen können. Können Sie das nicht, so sollten Sie das als Alarmzeichen betrachten und sich um Hilfe bemühen. Falls Sie das nicht rechtzeitig tun, werden Ihre Chancen, mit dem Studium zurechtzukommen, sehr schnell sinken.

Die Übungsassistenten und ich sind bereit, Ihnen nach Kräften zu helfen. Ich werde nach jeder Vorlesung für Fragen zur Verfügung stehen. Falls das nicht reicht, können wir auch gerne einen gesonderten Termin vereinbaren.

Sie sollten sich auch nicht scheuen, während der Vorlesung Fragen zu stellen.

Da ihr mathematisches Vorwissen sehr unterschiedlich ist, werde ich sehr einfach beginnen. Einige von Ihnen werden sich deshalb zunächst langweilen, und es besteht die Gefahr, daß sie den Moment verpassen, an dem sich das ändert. Ich muß Sie deshalb warnen: es gibt regelmäßig mathematisch talentierte Studenten, die aus solchen Gründen ihr Studium nicht bestehen.

Vorlesungsmanuskript

Diese Vorlesung wird einer Vorlesung, die ich im Wintersemester 2007/2008 gehalten habe, sehr ähnlich sein. Ein Manuskript dieser letzteren Vorlesung könne sie sich herunterladen unter

www.math.uni-leipzig.de/ ~herzog/Manuskripte/Manuskripte.html
(zur Webseite ‘www.math.uni-leipzig.de’ gehen, ‘Herzog’ klicken, ‘Vorlesungsmanuskripte’ klicken)

Bezeichnungen

Abb(M, V)	Vektorraum der Abbildungen der Menge M mit Werten im Vektorraum V, vgl. 3.2.3
$\mathrm{Aut}_K(V)$	Vektorraum der K-linearen Automorphismen des K-Vektorraums V, vg. 3.1
char(K)	Charakteristik des Körpers K, vgl. 4.4.2.
$\mathbb{C}$	Körper der komplexen Zahlen, vgl. 2.8.1
χ_f	das charakteristische Polynome eines linearen Endomorphismus f, vgl. 5.1.3
χ_A	das charakteristische Polynom einer quadratischen Matrix A, vgl. 5.1.3
det(A)	Determinante der quadratischen Matrix A, vgl. 4.3.1
dim V	Dimension des Vektorraums V, vgl. 3.3.8
δ_m	charakteristische Funktion der einelementigen Menge {m}, vgl. 3.2.10
End(V)	Endomorphismenring des Vektorraums V, vgl. 3.1 und 5.3.2
f_A	die zur Matrix A gehörige lineare Abbildung $x \mapsto Ax$, vgl. 3.4.2

$F_K(M)$	der von der Menge M frei erzeugte K-Vektorraum, vgl. 3.2.10
$\mathbb{F}_2$	der Körper mit zwei Elementen, vgl. 2.6.3
$\mathbb{F}_n$	der Körper mit n Elementen, vgl. 2.6.3
$GL_n(\mathbb{R})$	die Gruppe der umkehrbaren n×n-Matrizen über den reellen Zahlen $\mathbb{R}$, vgl. 2.6.1
$GL(n, R)$	Gruppe der umkehrbaren n×n-Matrizen mit Einträgen aus dem kommutativen Ring R mit Eins, vgl. 2.6.2
$GL_K(V)$	Vektorraum der K-linearen Automorphismen des K-Vektorraums V, vg. 3.1
$Hom_K(V',V'')$	Vektorraum der K-linearen Abbildungen $V' \to V''$, vgl 3.1
$\mathbb{H}$	Schiefkörper der Quaternionen, vgl. 2.8.2
i	imaginäre Einheit, vgl. 2.8.1
i, j, k	imaginäre Einheiten von $\mathbb{H}$, vgl. 2.8.2
Im(z)	Imaginärteil der komplexen Zahl z, vgl. 2.8.1
im(f)	Bild der Abbildung f, vgl. 3.2.9
$J_d(c)$	d×d-Jordan-Block zum Eingenwert c, vgl. 5.2.4
ker(f)	Kern der linearen Abbildung f, vgl. 3.2.9
K	ein Körper (beginnend mit 3.1).
K^n	K-Vektorraum der n-zeiligen Spalten mit Einträgen aus K, vgl. 3.2.1
$K^{m\times n}$	K-Vektorraum der m×n-Matrizen mit Einträgen aus K, vgl. 3.2.2
$L_K(V', V'')$	Vektorraum der K-linearen Abbildungen $V' \to V''$, vgl. 3.1
#(V)	die äußere Algebra des K-Vektorraums V, vgl. 6.6.20
$M_n(\mathbb{R})$	Ring der n×n-Matrizen mit Einträgen aus dem Körper $\mathbb{R}$ der reellen Zahlen, vgl. 2.6.2
$M_{m,n}(K)$	K-Vektoraum der m×n-Matrizen mit Einträgen aus K, vgl 3.2.2
$M^v_w(f)$	die Matrix der linearen Abbildung $f:V\to W$ bezüglich der Basen v und w von V bzw. W, vgl. 3.4.1
$M_v(b)$	die Matrix der Bilinearform b bezüglich der Basis v, vgl. 6.1.3
$\mu_j(f)$	die geometrischen Vielfachheiten eines Endomorphismus f, vgl. 5.1.5
$\mu_j(A)$	die geometrischen Vielfachheiten einer Matrix A, vgl. 5.1.5
$\nu_j(f)$	die algebraischen Vielfachheiten eines Endomorphismus f, vgl. 5.1.5
$\nu_j(A)$	die algebraischen Vielfachheiten einer Matrix A, vgl. 5.1.5
ord(f)	Ordnung des Endomorphismus f, vgl. 5.2.9.
$\mathbb{P}(V)$	die Projektivierung des K-Vektorraums V, vgl. 7.1
$\mathbb{P}^n_K$	der n-dimensionale projektive Raum über dem Körper K, vgl. 7.1
φ_v	die zur Basis v gehörige Koordinaten-Abbildung, vgl. 3.4.2
$\mathbb{Q}$	Körper der rationalen Zahlen, vgl. 2.6.2
Re(z)	Realteil der komplexen Zahl z, vgl. 2.8.1
rk A	Rang der Matrix A, vgl. 3.4.9
rk' A	Zeilenrang der Matrix A, vgl. 3.4.9

rk f	Rang der linearen Abbildung f, vgl. 3.4.9
$\rho_k(f)$	Anzahl der zyklischen Räume der Dimension k in der Jordan-Zerlegung des nilpotenten Endomorphismus f, vgl. 5.2.9.
$\rho_k(f, c)$	Anzahl der Jordan-Blöcke des Typs (k, k) zum Eigenwert c in der Jordan-Zerlegung des Endomorphismus f, vg. 5.3.9.
$\rho_k(A,c)$	Anzahl der Jordan-Blöcke des Typs (k, k) zum Eigenwert c in der Jordan-Zerlegung der Matrix A, vgl. 5.3.10.
$\mathbb{R}$	Körper der rationalen Zahlen, vgl. 2.6.2 und 2.6.3
$\mathbb{R}^*$	die multiplikative Gruppe der reellen Zahlen, vgl. 2.6.1
ρ_V	die natürliche Abbildung des Vektorraums V in dessen doppeltes Dual, vgl. 3.4.8
S(V)	die symmetrische Algebra des K-Vektorraums V, vgl. 6.6.19
T(V)	die Tensor-Algebra des K-Vektorraums V, vgl. 6.6.13
Z(R)	Zentrum des Rings R, vgl. 2.7.2.
$\mathbb{Z}$	Ring der ganzen Zahlen, vgl. 2.6.2
$\lvert A \rvert$	Determinante der quadratischen Matrix A, vgl. 4.3.1
$\lvert z \rvert$	Betrag der komplexen Zahl z, vgl. 2.8.1
$\lvert q \rvert$	Betrag des Quaternions q, vgl. 2.8.2
$\{v_i^*\}$	die zur Basis $\{v_i\}$ duale Basis, vgl. 3.3.17.
$\overline{z}$	die zur komplexen Zahl z konjugierte komplexe Zahl, vgl. 2.8.1
$\overline{q}$	das zum Quaternion q konjugierte Quaternion, vgl. 2.8.2
$\oplus$	direkte Summe von Vektorräumen, vgl. 3.2.4 und 3.2.5
$\oplus$	direkte Summe von Matrizen, vgl. 5.2.5
$\otimes$	Tensorprodukt von Vektorräumen, vgl. 6.6.2
$\times$	direktes Produkt von Vektorräumen, vgl. 3.2.4
$\langle M \rangle$	der von der Teilmenge M eines Vektorraums erzeugte Vektorraum, vgl. 3.2.6
$\prod_{i\in I} V_i$	direktes Produkt der Vektorräume V_i, vgl. 3.2.4
A^T	die zur Matrix A transponierte Matrix, vgl. 2.7.1
f*	die zur linearen Abbildung duale Abbildung, vgl. 3.4.7.1
R*	multiplikative Gruppe der Einheiten des Rings R mit Einselement, vgl. 2.6.2
R^{op}	der zu R entgegengesetzte Ring, vgl. 2.6.2
$R^{n\times n}$	Ring der n×n-Matrizen mit Einträgen aus dem kommutativen Ring R mit Eins, vgl. 2.6.2
V*	der zum Vektorraum V duale Vektorraum, vgl. 3.4.7.1
$V_c=V_c(f)$	der Eigenraum zum Eigenwert c des linearen Endomorphismus $f:V\to V$, vgl. 5.1.1
$V'_c=V'_c(f)$	der Hauptraum zum Eigenwert c des linearen Endomorphismus $f:V\to V$, vgl. 5.3.5
V/W	der Faktorraum des Vektorraums V bezüglich des linearen Unterraums W, vgl. 3.2.11
v+W	der um v verschobene Unterraum W, vgl. 3.2.8

Literatur

Fischer G.: Lineare Algebra, Vieweg-Verlag, Braunschweig 2003
Keller, O.-H.: Analytische Geometrie und lineare Algebra, VEB Deutscher Verlag der Wissenschaften, Berlin 1963
Brieskorn, E.: Lineare Algebra und analytische Geometrie I+II, Vieweg-Verlag, Braunschweig 1983
Herzog, B.: Lineare Algebra, www.mathematik.uni-leipzig.de/~herzog/Manuskripte/Manuskripte.html

In dieser Vorlesung werden wir uns weitgehend an dem Buch von Fischer orientieren. Das Buch von Keller führen wir hier an, weil es den geometrischen Aspekt der linearen Algebra besonders betont und sehr viele Sätze der klassischen Geometrie behandelt, die in den meisten modernen Büchern zu diesem Gegenstand fehlen.

1. Lineare Gleichungssysteme

Lineare Algebra ist im wesentlichen die Theorie der linearen Gleichungssysteme. Sehr viele explizite Berechungen der Mathematik beruhen darauf, daß man diese auf lineare Algebra zurückführt.

1.1 Eine Lösungsformel

Wie löst man das Gleichungssystem

$$\begin{aligned} x + 2y &= 3 \\ 4x + 5y &= 6 \end{aligned}$$

oder allgemeiner

$$\begin{aligned} ax + by &= u \\ cx + dy &= v \end{aligned}$$

Um x zu eliminieren, multiplizieren wir die erste Gleichung mit c, die zweite Gleichung mit a und bilden die Differenz:

$$(ad\text{-}bc)y = av - cu$$

Analog gehen wir vor, um y zu eliminieren:

$$(ad - bc)x = ud - vb$$

Wir erhalten also

$$x = \frac{\begin{vmatrix} u & b \\ v & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}},\ y = \frac{\begin{vmatrix} a & u \\ c & v \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \tag{1}$$

mit

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{2}$$

Bemerkungen

(i) Die Formeln (1) heißen Cramersche Regel.
(ii) Der Rechenausdruck (2) heißt Determinante.
(iii) Wir werden einen erhebliche Teil unserer Zeit damit zubringen, die Determinanten für allgemeinere Systeme zu definieren (und die allgemeine Cramersche Regel zu beweisen).

Fakten

1. Die Determinante ist das wichtigse Objekt, welches Sie in dieser Vorlesung kennenlernen werden. Einen erheblichen Teil der modernen Mathematik würde es ohne die Determinante nicht geben.

2. Zum Lösen von linearen Gleichungssystemen braucht man die Determinante nicht. In vielen Fällen (nicht in allen - Hauptachsentransformation) ist sie ein theoretisches Objekt, welches uns hilft, mathematische Phänome zu verstehen. Ausrechnen sollte man Determinanten nur in Notfällen, da dies viel zu aufwendig ist.

1.2 Ein Algorithmus zum Lösen von Gleichungssystemen

Der Algorithmus heißt Gauß-Algorithmus (obwohl Gauß wohl kaum der erste war, der ihn verwendet hat).
Das Problem:

$$
(1)\qquad \begin{array}{l} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{array}
$$

Grundidee:
Man ersetze schrittweise das gegebene Gleichungssystem durch ein System mit denselben Lösungen. Dabei versuche man das System in jedem Schritt zu vereinfachen, bis man ein System bekommt, dessem Lösungen man direkt ablesen kann.

Was ist eine Lösung von (1)?
Eine Lösung von (1) ist eine Folge von Zahlen $c_1,\ldots,c_n$ mit der Eigenschaft, daß man nach nach Ersetzen der $x_1,\ldots,x_n$ durch die $c_1,\ldots,c_n$ in (1) lauter Identitäten bekommt. Man sagt dann in dieser Situation,

$$(c_1,\ldots,c_n)$$

ist eine Lösung von (1). Eine Lösung eines Gleichungssystems in n Variablen ist also ein n-Tupel.

Wann ist die Lösung eines Systems offensichtlich?
Zum Beispiel ist das der Fall für ein System der folgenden Gestalt:

$$
\begin{array}{llll} a_1 \cdot x_1 & & & = b_1 \\ & a_2 \cdot x_2 & & = b_2 \\ & & \ldots & \\ & & a_n x_n & = b_n \end{array}
$$

Bemerkungen

(i) Diese System besitzt die einzige Lösung

$$\left(\frac{b_1}{a_1}, \ldots, \frac{b_n}{a_n}\right),$$

falls sämtliche a_i von Null verschieden sind.

(ii) Ist ein a_i gleich Null und ist das zugehörige b_i ungleich Null, so besitzt dieses System keine Lösung.

(iii) Sind ein oder mehrere a_i gleich Null und sind die zugehörigen b_i jeweils auch Null, so kann die entsprechende Koordinate x_i beliebig sein. Insbesondere gibt es in diesem Fall mehr als eine Lösung.

Wie vereinfacht man ein Gleichungssystem?

Beispiel

Die beiden folgenden Systeme haben dieselben Lösungen:

$$A \equiv 1x + 2y - 3 = 0$$
$$B \equiv 4x + 5y - 6 = 0$$

und

$$A \equiv 1x + 2y - 3 = 0$$

$$A + 17B \equiv (4x + 5y - 3)+17(1x + 2y - 6) = 0$$

Statt das 17-fache könnte man natürlich auch ein beliebiges anderes Vielfaches der ersten zur zweiten Gleichung addieren,

$$1x + 2y = 3$$

$$(4x + 5y)+\lambda\cdot(1x + 2y)= 6 + \lambda\cdot 3$$

wobei man λ möglichst geschickt wählen sollte. Vorschlag:

$$\lambda = -4$$

liefert:

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

<u>Allgemeiner Fall</u>

Um die Idee im allgemeinen Fall deutlich zu machen, ändern wir die Bezeichnungen

$$\begin{aligned} a_{11}x_1 + ... + a_{1n}x_n - b_1 &= 0 \\ \dots \\ a_{m1}x_1 + ... + a_{mn}x_n - b_m &= 0 \end{aligned} \qquad (1)$$

Bezeichnung für die i-te Gleichung des Systems:

$$f_i(x) = a_{i1}x_1 + ... + a_{in}x_n - b_i$$

Wir benutzen hier die Bezeichnung x für das Tupel $x = (x_1, ... , x_n)$. Das Gleichungssystem bekommt dann die Gestalt

$$\begin{aligned} f_1(x) &= 0 \\ f_2(x) &= 0 \\ \dots \\ f_m(x) &= 0 \end{aligned} \qquad (2)$$

<u>Grundlegener Fakt</u>:

Die Lösungen des Systems bleiben unverändert, wenn man

1. ein Vielfaches einer Gleichung zu einer anderen addiert.
2. eine Gleichung mit einem von Null verschiedenen Faktor multipliziert.
3. die Reihenfolge der Gleichungen ändert.

Beispiel

Folgende Systeme haben dieselben Lösungen:

$$\begin{aligned} f_1(x) &= 0 \\ f_2(x) &= 0 \end{aligned}$$

und

$$\begin{aligned} f_1(x) &= 0 \\ f_2(x) + 3f_1(x) &= 0 \end{aligned}$$

und

$$\begin{aligned} f_1(x) &= 0 \\ 25\cdot f_2(x) &= 0 \end{aligned}$$

1.3 Beispiele

Beispiel 1

Wir übersetzen das System in eine Tabelle, in der die Unbekannten nicht mehr vorkommen.
Statt

$$\begin{aligned} 1x + 2y + 3z &= 4 \\ 2x + 3y + 4z &= 5 \\ 3x + 4y + 6z &= 7 \end{aligned}$$

schreiben wir

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 6 & 7 \end{pmatrix}$$

Die so angeordneten Daten des gegeben Gleichungsystems nennt man auch <u>erweiterte Koeffizientenmatrix</u> des Gleichungssystems. Die analoge Matrix ohne die letzte Spalte der Absolutglieder, d.h. die Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}$$

heißt <u>Koeffizientenmatrix</u> des Gleichungssystems.

Wir führen jetzt die oben beschriebenen Operationen, welche die Lösungsmenge des Gleichungssystems nicht ändern, mit den Zeilen der Matrix, welche für die Gleichugen des Systems stehen, aus.

Subtraktion der zweiten Zeile von der dritten liefert:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 1 & 1 & 2 & 2 \end{pmatrix}$$

Subtration der ersten Zeile von der zweiten liefert:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{pmatrix}$$

Subtraktion der zweiten Zeile von der dritten liefert:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

(An dieser Stelle wissen wir bereits $z = 1$)
Subtraktion der ersten Zeile von der zweiten, Multiplikation der zweiten Zeile mit -1:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Subtraktion des Doppelten bzw. dreifachen der letzten Zeile von zweiten bzw. ersten:

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

Subtraktion der zweiten Zeile von der ersten:

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

Neues Gleichungssystem:

$$\begin{aligned} x \qquad\qquad &= -1 \\ y \qquad &= 1 \end{aligned}$$

$$z = 1$$

Lösung: (-1,1,1).
Lösungsmenge: { (-1,1,1) }

Beispiel 2

Das Gleichungssystem

$$\begin{aligned} 1x + 2y + 3z &= 4 \\ 2x + 3y + 4z &= 5 \\ 3x + 4y + 5z &= 6 \end{aligned}$$

entspricht der Matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}$$

Subtraktion des 2-fachen bzw. 3-fachen der ersten Zeile von der zweiten bzw. dritten Zeile:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & -2 & -4 & -8 \end{pmatrix}$$

Multiplikation der zweiten Zeile mit -1 und Addition des Doppelten zur dritten:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Addition des (-2)-fachen der zweiten Zeile zur ersten:

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

Neues Gleichungssystem:

$$\begin{aligned} x \qquad\quad - z &= -2 \\ y + 2z &= 3 \end{aligned}$$

Äquivalent dazu ist das System:

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

An dieser Gestalt des Systems lesen wir ab, daß man z = t beliebig wählen kann und für jede Wahl von z eindeutig bestimmte x und y deart bekommt, daß (x,y,z) eine Lösung ist.
Lösungsmenge:

$$\{ (t-2,-2t+3,t) \mid t \text{ beliebig} \}$$

Läßt man t die natürlichen Zahlen durchlaufen, so bekommt man insbesondere unendlich viele Lösungen.

Beispiel 3

Das Gleichungssystem

$$\begin{aligned} 1x + 2y + 3z &= 4 \\ 2x + 3y + 4z &= 5 \\ 3x + 4y + 5z &= 7 \end{aligned}$$

entspricht der Matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 7 \end{pmatrix}$$

Subtraktion der zweiten Zeile von der dritten:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

Subtration der ersten Zeile von der zweiten:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

Subtraktion der zweiten Zeile von der dritten:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Die dritte Gleichung des zugehörigen Gleichungssystems hat die Gestalt

$$0 \cdot z = 1.$$

Da es kein z gibt, für welches diese Bedingung erfüllt ist, hat das Gleichungssystem keine Lösung.
Lösungsmenge:

$$\varnothing$$

1.4 Allgemeine Beschreibung des Algorithmus

1.4.1 Eine weitere zulässige Operation

Um den Algorithmus etwas bequemer beschreiben zu können, wollen wir eine weitere Operation zulassen, die man in der Praxis nicht benutzen sollte, da sie sehr leicht Fehler verursachen kann. Für theoretische Zwecke, d.h. für das Verständnis des Algorithmus führt die Operation aber zu einer Vereinfachung.

Beispiel

Für die Lösung des Gleichungssystems

$$\begin{aligned} 1x + 2y &= 3 \\ 4x + 5y &= 6 \end{aligned}$$

ist es ohne Belang, ob man es wie eben oder in der Gestalt

$$\begin{aligned} 2y + 1x &= 3 \\ 5y + 4x &= 6 \end{aligned}$$

aufschreibt, d.h. ob man auf den rechten Seiten erst die Vielfachen von x und dann die von y aufschreibt oder umgekehrt. Für die zugehörigen Matrizen bedeutet dies, wir können Spalten der Matrix (mit Ausnahme der letzten) vertauschen, ohne daß sich etwas Wesentliches ändert.
Wenn man dies in der Praxis tatsächlich tut, sollte man über jede Spalte die Bezeichnung der Unbekannten schreiben, zu welcher die Spalte gehört, damit keine Unbekannten, beim letztendlichen Aufschreiben der Lösung verwechselt werden. Oder man sollte die Vertauschung ganz vermeiden.

Als zulässige Operationen, die die Lösungsmenge des Gleichungssystems unverändert lassen erhalten wir damit die folgenden:

1. Addition zu einer Zeile der Matrix ein Vielfaches einer anderen.
2. Multiplikation einer Zeile der Matrix mit einer von Null verschiedenen Zahl.
3. Vertauschen von zwei Zeilen der Matrix.
4. Vertauschen von zwei Spalten der Matrix, die beide von der letzten Spalte verschieden sind (wobei man gleichzig die Reihenfolge der verwendeten Unbestimmten in derselben Weise ändern muß)[1].

[1] Man kann dies zum Beispiel dadurch automatisch erreichen, indem man zur erweiterten Koeffizientenmatrix eine weitere Zeile hinzufügt, in der die Bezeichnungen der Unbestimmten stehen. Die Spalten der so ergänzten Matrix kann man vertauschen, ohne daß die Gefahr einer Verwechslung bei der Wahl der Unbestimmten eintreten kann.

Diese Operationen nennt man auch elementarte Operationen. Die vierte Operation spielt dabei eine Sonderrolle: sie dient nur theoretischen Zwecken und wird beim pratischen Lösen von Gleichungssystemen vermieden. Die obigen Operationen ohne die letzte nennt man auch elementare Zeilenoperationen.

1.4.2 Der Algorithmus

Wir beschreiben jetzt einen Algorithmus, der die erweiterte Koeffizientenmatrix in endlich vielen Schritte in eine Gestalt bringt, bei der man die Lösungen des zugehörigen Gleichungssystems direkt ablesen kann. Die erweiterte Koeffizientenmatrix des Gleichungssystems habe die Gestalt

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b2 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

d.h. wir haben ein Gleichungssystem aus m Gleichungen in n Unbekannten zu lösen.

1. Falls in allen Spalten der Matrix (außer eventuel in der letzten) Nullen stehen, brauchen keine Umformungen ausgeführt werden.

   1. Fall: Stehen auch in der letzten Spalte Nullen, so ist jedes n-Tupel eine Lösung.

   2. Fall: Steht in der letzten Spalte an einer Stelle ein von Null verschiedener Eintrag, so hat das System keine Lösung.

2. Wir können jetzt annehmen, es gibt, außer eventuell in der letzten Spalte, weitere von Null verschiedenen Einträge.

   Wir vertauschen in geeigneter Weise Spalten und erreichen so, daß es in der ersten Spalte einen von Null verschiedenen Eintrag gibt.

   Weiter vertauschen wir in geeigneter Weise Zeilen und erreichen so, daß der Eintrag a_{11} ungleich Null ist,

$$a_{11} \neq 0.$$

   Anschließend addieren wir Vielfache der ersten Zeile zu den anderen Zeilen in einer Weise, daß alle Einträge der ersten Spalte mit Ausnahme des ersten gleich Null werden. Die erweiterte Koeffizientenmatrix bekommt dadurch die Gestalt

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22} & \dots & a_{2n} & b2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \tag{1}$$

   Bemerkung: die Lösung des Gleichungssystems ist jetzt auf die Lösung des Systems zur Matrix ohne die erste Zeile und erste Spalte zurückgeführt:

$$\begin{pmatrix} a_{12} & \dots & a_{1n} & b_1 \\ a_{22} & \dots & a_{2n} & b2 \\ \dots & \dots & \dots & \dots \\ a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \tag{2}$$

Der Wert der ersten Unbekannten ergibt sich aus denen der übrigen in eindeutiger Weise durch

$$x_1 = -\frac{1}{a_{11}}(b_1 - a_{12}x_2 - \ldots - a_{1n}x_n).$$

3. Wir wiederholen jetzt die eben ausgeführten Opertionen mit der verkleinerten Matrix (2) anstelle von (1), d.h. wir solgen dafür, daß der Eintrag a_{22} ungleich Null wird und danach dafür, daß alle anderen Einträge der zweiten Spalte Null werden. Danach gehen wir analog mit der dritten Spalte vor. Wir erreichen nach endlich vielen Schritten,daß die Matrix die folgende Gestalt bekommt:

   In den Spalten 1 bis k ist genau der Eintrag

   $$a_{ii} \quad (i=1,\ldots,k) \tag{3}$$

   von Null verschieden. Alle anderen Einträge sind Null.

4. Sollte in den verbleibenden Spalten (ausgenommen die letzte) noch irgendwo ein von Null verschiedener Eintrag stehen, so kann man dafür sorgen daß in einer weiteren Spalte nur der Eintrag a_{ii} ungleich Null ist, d.h. man kann in (3) den Wert von k um 1 erhöhen. Da die Zahl der Unbestimmten endlich ist, kann das nur endlich oft geschehen. Nach endlich vielen Schritten hat die Matrix also die Gestalt:

$$\begin{pmatrix} a_{11} & 0 & \ldots & 0 & \ldots & b_1 \\ 0 & a_{22} & \ldots & 0 & \ldots & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{kk} & \ldots & b_k \\ 0 & 0 & 0 & 0 & \ldots & b_{k+1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & & & & \end{pmatrix}$$

mit $a_{11} \neq 0, \ldots, a_{kk} \neq 0$. Außerdem sind alle Einträge der Zeilen k+1 bis m (falls k < m ist und diese Zeilen tatsächlich existieren) mit eventueller Ausnahme der letzten gleich Null. Genauer:

$$\left(\begin{array}{cc|cc|l} a_{11} & \mathbf{0} & * & a_{1n} & b_1 \\ \mathbf{0} & a_{kk} & * & a_{kn} & b_k \\ \hline \mathbf{0} & & \underbrace{\mathbf{0}}_{\text{kann fehlen (n=k}} & & \left.\begin{matrix} b_{k+1} \\ b_m \end{matrix}\right\} \text{kann fehlen (m=k} \end{array}\right)$$

Der beschriebene Algorithmus heißt <u>Gauß-Algorithmus</u>.

1.4.3 Das Lösungsverhalten

Für die Lösungsmenge L eines linearen Gleichungssystems in n Unbestimmten gibt es nur folgende drei Möglichkeiten.

1. L ist leer.
2. L besteht aus genau einem n-Tupel.
3. L besteht aus unendlich vielen n-Tupeln (im reellen Fall).

Beweis. Betrachten wir die obige Matrix, die als Ergebnis des Gauß-Algorithmus entsteht. Betrachten wir den Fall, daß die Spalten in der Mitte tatsächlich vorhanden sind, d.h. den Fall $n > k$.

<u>Fall 1</u>: $n > k$.

<u>Fall 1a</u>: Eines der Absolutglieder $b_{k+1}, \ldots, b_m$ ist von Null verschieden.

In diesen Fall beschreibt eine der letzten Gleichungen des Systems einen Widerspruch. Das System hat keine Lösung, d.h. die Lösungsmenge ist leer.

<u>Fall 1b</u>: Die Absolutglieder $b_{k+1}, \ldots, b_m$ sind sämtlich Null.

Sie letzten Gleichungen des Systems haben dann die Gestalt

$$0 = 0.$$

Sie schränken also die Lösungen des Systems in keiner Weise ein und können deshalb weggelassen werden. Wir können somit annehmen,

$$m = k.$$

Das Gleichungssystem bekommt damit die Gestalt

$$\begin{gathered} a_{11}x_1 = b_1 - a_{1,k+1}x_{k+1} - \ldots - a_{1n}x_n \\ a_{22}x_2 = b_2 - a_{2,k+1}x_{k+1} - \ldots - a_{2n}x_n \\ \ldots \\ a_{kk}x_k = b_k - a_{k,k+1}x_{k+1} - \ldots - a_{kn}x_n \end{gathered}$$

Die Werte der

$$x_{k+1}, \ldots, x_n$$

können also beliebig gewählt werden. Die obigen Gleichgungen liefern für diese beliebig gewählten Werte entsprechende Werte für die

$$x_1, \ldots, x_k$$

sodaß man jeweils eime Lösung des Systems erhält. Die Lösungsmenge ist unendlich (im Fall der reellen Zahlen).

<u>Fall 2</u>: $n = k$.

<u>Fall 2a</u>: Eines der Absolutglieder $b_{k+1}, \ldots, b_m$ ist von Null verschieden.

Wie im Fall 1a ist die Lösungsmenge leer.

<u>Fall 2b</u>: Die Absolutglieder $b_{k+1}, \ldots, b_m$ sind sämtlich Null.

Wie im Fall 1b reduziert man auf den Fall $m = k$. Das Gleichungssystem bekommt die Gestalt

$$\begin{gathered} a_{11}x_1 = b_1 \\ a_{22}x_2 = b_2 \\ \ldots \\ a_{nn}x_n = b_n \end{gathered}$$

Es gibt genau eine Lösung.

QED.

1.5 Matrizenmultiplikation

1.5.1 Ein etwas komplizierteres Beispiel

Unser letztes Beispiel in diesem Abschnitt soll etwas komplizierter sein als die bisherigen. Wir wollen nämlich annehem, daß die rechten Seiten des Systems ebenfalls Variable sind, die erst noch zu bestimmen sind.

Beispiel:

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

Die Variablen u, v, w sollen ihrerseits durch ein lineares Gleichungssystem festgelegt sein.

$$\begin{aligned} u + v + w &= 16 \\ u - v + w &= 6 \\ u \quad\;\; - w &= -3 \end{aligned}$$

Es gibt eine naheliegende Art, dieses Problem zu lösen, indem man einfach die Gleichungen des ersten Systems in die des zweiten einsetzt.

$$\begin{aligned} (1x + 2y + 3z) + (2x + 3y + 4z) + (3x + 4y + 6z) &= 16 \\ (1x + 2y + 3z) - (2x + 3y + 4z) + (3x + 4y + 6z) &= 6 \\ (1x + 2y + 3z) \qquad\qquad\qquad\qquad - (3x + 4y + 6z) &= -3 \end{aligned}$$

Wir erhalten das System

$$\begin{aligned} 6x + 9y + 13z &= 16 \\ 2x + 3y + 5z &= 6 \\ -2x - 2y - 3z &= -3 \end{aligned}$$

das man wieder nach dem Gaußschen Algorithmus lösen kann.
Formulieren wir die eben beschriebene Situation mit Hilfe von Matrizen. Es sind zwei Gleichungssysteme gegeben mit den Matrizen

$$\begin{pmatrix} 1 & 2 & 3 & u \\ 2 & 3 & 4 & v \\ 3 & 4 & 6 & w \end{pmatrix}$$

und

$$\begin{pmatrix} 1 & 1 & 1 & 16 \\ 1 & -1 & 1 & 6 \\ 1 & 0 & 1 & -3 \end{pmatrix}$$

und wir haben aus diesen beiden Systemen ein neues gewonnen mit der Matrix

$$\begin{pmatrix} 6 & 9 & 13 & 16 \\ 2 & 3 & 5 & 6 \\ -2 & -2 & -3 & -3 \end{pmatrix}$$

Wir werden ziemlich oft mit einer Situation wie dieser konfrontiert werden und wollen deshalb gleich in voller Allgemeinheit überlegen, wie man hier vorzugehen hat, d.h. wir wollen uns eine Formel überlegen, mit der man aus den ersten beiden Matrizen die dritte berechnen kann.

1.5.2 Verallgemeinerung

Seien die beiden folgenden Systeme gegeben.

$$\begin{gathered} a_{11}x_1 + \ldots + a_{1n}x_n = u_1 \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = u_m \end{gathered}$$

und

$$\begin{gathered} b_{11}u_1 + \ldots + b_{1m}u_m = v_1 \\ \ldots \end{gathered}$$

$$b_{p1}u_1 + ... + b_{pm}u_m \;= v_p$$

Die zugehörigen Matrizen sind die folgenden

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & u_1 \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} & u_m \end{pmatrix} \text{ und } \begin{pmatrix} b_{11} & \cdots & b_{1m} & v_1 \\ \cdots & \cdots & \cdots & \cdots \\ b_{p1} & \cdots & b_{pm} & v_p \end{pmatrix}$$

<u>Problem</u>: welches ist die Matrix des zugehörigen gewöhnlichen Systems.

Um bequemer rechnen zu können schreiben wir die beiden Systeme mit Hilfe des Summenzeichens auf. Für das erste System erhalten wir

$$\sum_{j=1}^{n} a_{ij}x_j = u_i \;, \; i = 1, ... , m$$

und für das zweite

$$\sum_{k=1}^{m} a_{\ell k}u_k = v_\ell \;, \; \ell = 1, ... , p$$

Durch Einsetzen erhalten wir

$$\sum_{k=1}^{m} a_{\ell k} \sum_{j=1}^{n} a_{kj} = v_\ell \;, \; \ell = 1, ... , p$$

d.h.

$$\sum_{k=1}^{m} \sum_{j=1}^{n} a_{\ell k}a_{kj}x_j = v_\ell \;, \; \ell = 1, ... , p$$

Anders ausgedrückt,

$$\sum_{j=1}^{n} c_{\ell j}x_j = v_\ell \;, \; \ell = 1, ... , p$$

mit $c_{\ell j} = \sum_{j=1}^{n} a_{\ell k}a_{kj}$. Damit haben wir unsere Aufgabe gelöst. Die neue Matrix hat die Gestalt

$$\begin{pmatrix} c_{11} & \cdots & c_{1n} & v_1 \\ \cdots & \cdots & \cdots & \cdots \\ c_{p1} & \cdots & c_{pn} & v_p \end{pmatrix}$$

mit

$$c_{\ell j} = \sum_{k=1}^{n} a_{\ell k}a_{kj}.$$

Wir werden später auf diese Formel zurückkommen. Um die Eigenschaften von Gleichungssystemen besser verstehen zu können, müssen wir uns zunächst etwas genauer mit Matrizen beschäftigen.

2. Matrizen und Vektoren

2.1 Summe und Vielfache von Matrizen

Sei eine Matrix

$$M = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

gegeben. Wir haben dabei die Bezeichnungen so gewählt, daß die Zahlen ganz rechts keine besonder Rolle mehr spielen. Die Zahl n heißt in dieser Situation Spaltenzahl der Matrix M und die Zahl m ihre Zeilenzahl. Von der Matrix selbst werden wir sagen, das es eine $m\times n$- Matrix ist. Das Zahlenpaar (m,n) heißt auch Typ der Matrix M.

Die Zahlen a_{ij} heißen Einträge der Matrix. Die Zahl a_{ij} der i-ten Zeile und j-ten Spalte heißt auch Eintrag in der Position (i,j). Eine Matrix mit nur einer Zeile heißt auch Zeilenvektor oder einfach Zeile. Eine Matrix mit nur einer Spalte heißt auch Spaltenvektor oder einfach Spalte.

Sei jetzt eine zweite Matrix vom selben Typ gegeben.

$$N = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{m1} & \dots & b_{mn} \end{pmatrix}$$

Dann definieren wir die Summe von M und N durch die folgende Formel.

$$M+N := \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \dots & \dots & \dots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

Es ist wichtig zu beachten, die Summe ist nur für Matrizen desselben Typs definiert.
Kurzschreibweise:

$$M = (a_{ij})_{i=1,\dots,m,j=1,\dots,n} = (a_{ij})$$
$$N = (b_{ij})_{i=1,\dots,m,j=1,\dots,n} = (b_{ij})$$
$$M+N = (a_{ij}+b_{ij})_{i=1,\dots,m,j=1,\dots,n} = (a_{ij} + b_{ij})$$

Zwei Matrizen werden Addiert, indem man alle Einträge in denselben Positionen addiert.
Eine Nullmatrix ist eine Matrix, deren sämtliche Einträge Null sind. Wir werden Nullmatrizen oft mit dem Symbol 0 bezeichnen.

Ist a eine Zahl, so heißt die Matrix

$$aM = Ma = \begin{pmatrix} aa_{11} & \dots & aa_{1n} \\ \dots & \dots & \dots \\ aa_{m1} & \dots & aa_{mn} \end{pmatrix}$$

das a-fache der Matrix M oder auch Produkt von a mit M. Anstelle von (-1)M schreibt im allgemeine auch -M und nennt diese Matrix Negatives von M. Für M+(-N) schreibt man auch
M - N
und nennt diese Matrix Differenz von M und N.

2.2 Eigenschaften der Matrizenaddition

Für Matrizen M, N, P und Zahlen α, β gilt:
(i) M+N = N + M
(ii) (M+N) + P = M + (N+P)
(iii) M + 0 = 0 + M = M falls 0 die Nullmatrix ist.

(iv) $M + (-M) = 0$.

(v) $(\alpha+\beta)M = \alpha M+\beta M$, $\alpha(M+N) = \alpha M+\alpha N$.

(vi) $\alpha M = M\alpha$, $1\cdot M = M$.

Die Identitäten sind dabei so zu interpretieren, daß die linke Seite genau dann definiert ist, wenn es die rechte ist, und daß in diesem Fall die beiden Seiten gleich sind.

(Eine oder zwei Eigenschaften beweisen).

2.3 Das Produkt von Matrizen

Seien zwei Matrizen gegeben des Typs (m,n) bzw. (p,q) gegeben.

$$M = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

$$N = \begin{pmatrix} b_{11} & \dots & b_{1q} \\ \dots & \dots & \dots \\ b_{p1} & \dots & b_{pq} \end{pmatrix}$$

Voraussetzung

$$n = p,$$

d.h. die Spaltenzahl von M sei gleich der Zeilenzahl. In dieser Situation sagt man, M ist <u>verkettet</u> mit N. Man beachte, daß hier die Reihenfolge wichtig ist, Wenn M verkettet ist mit N, so muß N in keiner Weise verkettet sein mit M.

Sei also M verkettet mit N. In dieser Situation definieren wir eine dritte Matrix

$$P := \begin{pmatrix} c_{11} & \dots & c_{1q} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mq} \end{pmatrix} \text{ mit } c_{ij} = \sum_{k=1}^{n} a_{ik}a_{kj}$$

des Typs (m, q), d.h. P hat dieselbe Zeilenzahl wie M und dieselbe Spaltenzahl wie N. Diese Matrix heißt Produkt von M und N und wird mit

$P = M\cdot N$

bezeichnet.

Bemerkungen

(i) In der betrachteten Situation muß das Produkt MN im allgemeinen nicht definiert sein.

(ii) Für quadratische Matrizen (Zeilenzahl = Spaltenzahl) ist mit MN automatisch auch NM definiert.

(iii) Eine quadratische Matrix der Gestalt

$$Id = Id_n = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

mit $a_{ij} = \begin{cases} 1 \text{ falls } i=j \\ 0 \text{ sonst} \end{cases}$ heißt <u>Einheitsmatrix.</u>

und wird mit Id_n oder auch Id bezeichnet.

2.4 Eigenschaften der Matrizenmultiplikation

Für Matrzen M, N, P gilt:

(i) a) Im allgemeinen sind die Produkte MN und NM nicht beide definiert. Selbst wenn sie beide definiert sind, kann MN ≠ MN gelten.
b) Das Produkt von Null verschiedener Matrizen kann Null sein.

(ii) (MN)P = M(NP).

(iii) M(N+P) = MN + MP.

(iv) (M+N)P = MP + NP

(v) Id_m•M = M und M•Id_n = M, falls M den Typ (m, n) hat.

Das Assoziativgesetzt gilt auch, wenn einer oder mehrere Faktoren Zahlen sind.

Die Identitäten sind dabei so zu interpretieren, daß die linke Seite genau dann definiert ist, wenn es die rechte ist, und daß in diesem Fall die beiden Seiten gleich sind.

Zu (i). Seien M = $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ und N = $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Dann gilt

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

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

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

(Einige der Eigenschaften beweisen).

2.5 Transposition von Matrizen

2.5.1 Transponierte Matrizen

Seien K A = (a_{ij}) eine Matrix mit Einträgen a_{ij}. Dann heißt die Matrix

$$A^T = (a'_{ij})$$

mit den Einträgen

$$a'_{ij} = a_{ji} \text{ für } i=1,...,n,\ j = 1,...,m$$

die zu A transponierte Matrix.

Beispiel

Sei

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix}$$

Dann ist die zu A transponierte Matrix gleich

$$A^T = \begin{pmatrix} 1 & 6 & 11 \\ 2 & 7 & 12 \\ 3 & 8 & 13 \\ 4 & 9 & 14 \\ 5 & 10 & 15 \end{pmatrix}$$

2.5.2 Eigenschaften transponierter Matrizen

(i) $(A^T)^T = A$ für beliebige Matrizen A.

(ii) $(A+B)^T = A^T + B^T$ für Matrizen A, B desselben Typs.

(iii) $(\lambda A)^T = \lambda A^T$ für beliebige Matrizen über K und beliebige $\lambda\in K$.

(iv) $(AB)^T = B^T A^T$.

Beweis. Seien $A = (a_{ij})$ und $B = (b_{ij})$. Dann gilt

$A^T = (a'_{ij})$, $B^T = (b'_{ij})$ mit $a'_{ij} = a_{ji}$ und $b'_{ij} = b_{ji}$.

<u>Zu (i)</u>. Trivial.

<u>Zu (ii).</u> Wir setzen $c_{ij} := a_{ij} + b_{ij}$ und $c' := a'_{ij} + b'_{ij}$ und erhalten

$(A+B)^T = (c_{ij})^T = (c'_{ij}) = (a'_{ij} + b'_{ij}) = (a'_{ij}) + (b'_{ij}) = A^T+B^T$.

<u>Zu (iii)</u>. $(\lambda A)^T = (\lambda a'_{ij}) = \lambda(a'_{ij}) = \lambda A^T$.

<u>Zu (iv)</u>. Es gilt

$(AB)^T = (d'_{ij})$

mit $d'_{ij} := d_{ji}$, wobei die Einträge der Produktmatrix AB bezeichnen sollen, d.h.

$$d_{ij} := \sum_k a_{ik} b_{kj} = \sum_k a'_{ki} b'_{jk} = \sum_k b'_{jk}\, a'_{ki}.$$

Mit anderen Worten, es ist

$$d'_{ij} = \sum_k b'_{ik}\, a'_{kj}.$$

Nach Definition der Matrizenmultiplikation folgt

$(AB)^T = (b'_{ij})(a'_{ij}) = B^T A^T$.

QED.

2.6 Anwendung auf lineare Gleichungssysteme

2.6.1 Lineare Gleichungssysteme in Matrizenschreibweise

Sei ein lineares Gleichungssystem gegeben.

$$\begin{aligned} a_{11}x_1 + ... + a_{1n}x_n &= b_1 \\ & \dots \qquad (1) \\ a_{m1}x_1 + ... + a_{mn}x_n &= b_m \end{aligned}$$

Dann heißt die Matrix der a_{ij},

$$A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

<u>Koeffizientenmatrix</u> des Gleichungssystems. Die früher betrachtete Matrix

$$A' := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} & b_n \end{pmatrix}$$

heißt erweiterte Koeffizientenmatrix. Der Spaltenvektor

$$b := \begin{pmatrix} b_1 \\ \dots \\ b_m \end{pmatrix}$$

heißt Spaltenvektor der rechten Seiten des Gleichungssystems und

$$x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$$

Spaltenvektor der Unbestimmten. Das Gleichungssystem selbst kann man mit Hilfe der Matrizenmultiplikation schreiben als

$$Ax = b.$$

Mit Hilfe der Matrizenmultiplikation bekommt ein lineares Gleichungssystem die Gestalt einer einzelnen Gleichung.

2.6.2 Vereinbarung

Von jetzt ab wollen wir die Lösungen eines Gleichungssystems als Spaltenvektor und nicht wie bisher als Zeilenvektor schreiben.

2.6.3 Homogene und inhomogen Gleichungssysteme

Ein lineare Gleichungssystem

$$\begin{matrix} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{matrix} \qquad (1)$$

heißt homogen, wenn sämtliche rechten Seiten Null sind, und ansonsten inhomogen. In Matrizenschreibweise bedeutet das, auf der rechten Seite steht der Nullvektor
$Ax = 0$.

Satz

Seien

$$Ax = b \qquad (2)$$

ein lineares Gleichungssystem in Matrizenschreibweise und

$$Ax = 0 \qquad (3)$$

das zugehörige homogene Gleichungssystem. Dann gilt:

(i) Sind u und v zwei Lösungen des homogenen Systems (3) und α und β zwei beliebige Zahlen, so ist auch die Linearkombination

$$\alpha u + \beta v$$

eine Lösung des homogenen Systems.

(ii) Ist u eine Lösung des homogenen Systems und v eine des inhomogenen Systems, so ist

$$u+v$$

eine Lösung des inhomogenen Systems. Fixiert man v und läßt u die Lösungen des homogenen Systems durchlaufen, so bekommt man auf diese Weise alle Lösungen des inhomogenen Systems.

(iii) Je zwei Lösungen des inhomogenen Systems unterscheiden sich um eine des homogenen.

Beweis. Zu (i). Es gilt

$$A(\alpha u + \beta v) = \alpha Au + \beta Av = \alpha 0 + \beta 0 = 0.$$

Zu (iii). Seien v und v' zwei Lösungen des inhomogenen Systems. Dann gilt

$$A(v-v') = Av - Av' = b - b = 0,$$

d.h. v-v' ist eine Lösung des homogenen Systems.
Zu (ii). Es gilt

$$A(u+v) = Au + Av = 0 + b = b.$$

Sei v' eine beliebige Lösung des inhomogenen Systems, so gilt

$$v' = v + (v'-v),$$

wobei der erste Summand rechts die gegebene Lösung des inhomogenen Systems ist und der zweite Summand v'-v nach (iii) eine Lösung des homogenen Systems.
QED.

Beispiel

Betrachten wir das lineare Gleichungssystem

$$2x + 3y = 5$$

Es hat offensichtlich die Lösung $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Das zugehörige homogene System

$$2x + 3y = 0$$

hat die allgemeine Lösung

$$t\cdot\begin{pmatrix} 3 \\ -2 \end{pmatrix}, \text{ t beliebig.}$$

Also hat nach dem obigen Satz das inhomogene System die allgemeine Lösung

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + t\cdot\begin{pmatrix} 3 \\ -2 \end{pmatrix}, \text{ t beliebig.}$$

Bemerkung

(i) Um das Lösungsverhalten der linearen Gleichungssystem noch besser zu verstehen, brauchen wir den Begriff des Vektorraums.
(ii) Bevor wir diesen Begriff einführen können, haben wir jedoch noch einiges anderes zu klären. Zum Beispiel haben wir bisher ganz undifferenziert von "Zahlen" gesprochen. Um unsere Theorie exakter zu machen, müssen wir klarstellen, was wir darunter verstehen wollen.
(iii) Wir werden aus diesem Anlaß gleich alle wichtigen algebraischen Strukturen einführen, die wir später benötigen.

2.7 Gruppen, Ringe, Körper

Bisher haben wir über Zahlen gesprochen, ohne genauer zu präzisieren, was wir darunter verstehen wollen. Diese Lücke wollen wir in diesem Abschnitt schließen.

Bezeichnung

Sind M und N zwei Mengen, so bezeichne

$$M\times N$$

die Menge aller Paare (m,n), deren erste Koordinate m in M und deren zweite Koordinate n in N liegt,

$$M\times N := \{(m,n) \mid m\in M \text{ und } n\in N\}.$$

Die Menge M×N heißt auch Produktmenge.

2.7.1 Begriff der Gruppe

Eine Gruppe G ist eine eine Menge zusammen mit einer Abbildung

$$G\times G \to G, (a,b) \mapsto ab,$$

(genannt Gruppenoperation oder Gruppenmultiplikation), wobei die folgenden Bedinungen erfüllt sind.

(i) Es gilt das Assoziativgesetz, d.h. $a(bc) = (ab)c$ für $a,b,c\in G$.
(ii) Es gibt ein neutrales Element e in G, d.h. ein Element e mit $ae = ea = a$ für jedes $a\in G$.
(ii) Es gibt zu jedem Element $a\in G$ ein inverses Element, d.h. ein Element $a'\in G$ mit $aa' = a'a = e$.

Bezeichnung: $a^{-1} := a'$.

Gilt außerdem $ab = ba$ für $a,b\in G$, so heißt die Gruppe auch kommutativ oder abelsch.

Ein Gruppen-Homomorphismus ist eine Abbildung $h: G \to G'$ einer Gruppe G in eine Gruppe G' mit $h(g'\cdot g'') = h(g')\cdot h(g'')$ für alle $g',g''\in G$.

Beispiel 1

Die Menge $\mathbb{R}^*$ der von Null verschiedenen reellen Zahlen ist mit der gewöhnlichen Multiplikation

$$\mathbb{R}^*\times\mathbb{R}^* \to \mathbb{R}^*, (a,b) \mapsto a\cdot b,$$

eine Gruppe. Wir sagten in dieser Situation $(\mathbb{R}^*, \cdot)$ ist eine Gruppe. Das neutrale Element ist die Eins, das zu $r\in \mathbb{R}^*$ inverse Element ist $1/r$. Die Gruppe ist abelsch.

Beispiel 2

Die Menge $\mathbb{R}$ der reellen Zahlen ist mit der Addition

$$\mathbb{R}\times\mathbb{R} \to \mathbb{R}, (a,b) \mapsto a+b,$$

eine Gruppe. Wir sagen in dieser Situation auch $(\mathbb{R}, +)$ ist eine Gruppe. Das neutrale Element ist die Null, das zu $r\in\mathbb{R}$ 'inverse' Element ist dessen Negatives $-r$. Die Gruppe ist abelsch.

Beispiel 3

Sei n eine fest vorgegeben natürliche Zahl. Dann ist die Menge

$$GL_n(\mathbb{R}) = GL(n, \mathbb{R})$$

der $n\times n$-Matrizen A mit reellen Einträgen, für welche es eine $n\times n$-Matrix B mit reellen Einträgen gibt mit $AB = BA = Id$, eine Gruppe bezüglich der Matrizenmultiplikation

$$GL_n(\mathbb{R})\times GL_n(\mathbb{R}) \to GL_n(\mathbb{R}), (A',A'') \mapsto A'A''.$$

Das neutrale Element ist gerade die Einheitsmatrix Id. Das zu A inverse Element ist die nach Voraussetzung existierende Matrix B. Diese Gruppe ist nicht abelsch (außer im Fall $n = 1$). Sie heißt allgemeine lineare Gruppe (General Linear Group).

Beispiel 4 (Restklassen ganzer Zahlen)

Sei $n \in \mathbb{Z}$ eine ganze Zahl. Für jede weitere ganze Zahl $i \in\mathbb{Z}$ bezeichne

$$\overline{i} := \{k\cdot n + i \mid k\in\mathbb{Z}\}$$

die Restklasse von i modulo n, d.h. die Menge der ganzen Zahlen, die bei Division mit n den Rest n lassen. Da sich jede ganze Zahl g in der Gestalt

$$g = k\cdot n + i \text{ mit } i \in \{0, 1, \dots , n-1\}$$

schreiben läßt, gibt es gerade n verschiedene Restklassen Modulo n. Sei

$$\mathbb{Z}/(n) := \{\overline{0}, \overline{1}, \overline{2}, \dots , \overline{n-1}\}$$

die Menge aller Restklassen modulo n. Für je zwei Mengen von ganzen Zahlen

$$M \subseteq \mathbb{Z} \text{ und } N \subseteq \mathbb{Z}$$

definieren wir

$$M + N := \{a+b \mid a\in M, b\in N\}$$

Dann gilt

$$\overline{i} + \overline{j} = \overline{i+j} \qquad (1)$$

Die Rechengesetze für ganze Zahlen gelten deshalb auch für die Restklassen modulo n, d.h. die Menge

$$\mathbb{Z}/(n)$$

ist mit der Operation (1) eine kommutative Gruppe und die Abbildung

$$\rho:\mathbb{Z} \to \mathbb{Z}/(n),\ i \mapsto \overline{i}$$

ein Gruppen-Homomorphismus.
Beweis. Beweis von (1). Sei

$$\alpha \in \text{LHS}.$$

Dann gibt es ganze Zahlen k und k' mit

$$\alpha = (k\cdot n + i) + (k'\cdot n + j) = (k+k')n + (i+j)$$

d.h. es gilt $\alpha \in$RHS. Wir haben gezeigt

$$\text{LHS} \subseteq \text{RHS}.$$

Sie umgekehrt

$$\alpha \in \text{RHS}.$$

Dann gibt es eine ganze Zahl k mit

$$\alpha = k\cdot n + (i+j) = (k\cdot n + i) + (0\cdot n + j),$$

d.h. es gilt $\alpha \in$LHS. Wir haben gezeigt die beiden Seiten von (1) sind gleich.
Die Gruppengesetze: werden mit Hilfe von Formel (1) bewiesen.
Assoziativgesetz: für beliebige i, j, k$\in\mathbb{Z}$ gilt

$$(\overline{i} + \overline{j}) + \overline{k} = \overline{i+j} + \overline{k} = \overline{i+j+k} = \overline{i} + \overline{j+k} = \overline{i} + (\overline{j}+\overline{k}).$$

Existenz des neutralen Elements: für jedes i$\in\mathbb{Z}$ gilt

$$\overline{i} + \overline{0} = \overline{i+0} = \overline{i} = \overline{0+i} = \overline{0} + \overline{i},$$

d.h.die Restklasse $\overline{0}$ spielt die Rolle des neutralen Elements.
Existenz des negativen Elements: für jedes i$\in\mathbb{Z}$ gilt

$$\overline{i} + \overline{-i} = \overline{i+(-i)} = \overline{0},$$

d.h. $\overline{-i}$ ist das Negative von $\overline{i}$.
Kommutativgesetz: für beliebige i, j $\in\mathbb{Z}$ gilt

$$\overline{i} + \overline{j} = \overline{i+j} = \overline{j+i} = \overline{j} + \overline{i}.$$

Die Eigenschaft der angegebenen Abbildung, ein Homomorphismus zu sein, ergibt sich aus Formel (1).
QED.

2.7.2 Begriff des Rings

Ein Ring ist eine Menge R zusammen mit zwei Abbildungen

$$R\times R \to R,\ (a,b) \mapsto a+b, \text{ und } R\times R \to R,\ (a,b) \mapsto ab,$$

(genannt Ringaddition bzw. Ringmultiplikation), wobei die folgenden Bedingungen erfüllt sind.

(i) (R , +) ist eine abelsche Gruppe.
(ii) Die Multiplikation ist assoziativ, d.h. es gilt a(bc) = (ab)c für a,b,c$\in$R.
(iii) Es gelten die Distributivgesetze, d.h. es gilt

$$a(b+c) = ab + ac \text{ für } a,b,c\in R$$

und

$$(a+b)c = ac + bc \text{ für } a,b,c\in R.$$

Falls die Multiplikation außerdem kommutativ ist, d.h. ab = ba für a,b$\in$R, so heißt R auch kommutativer Ring. Das neutrale Element der Addition von R heißt Nullelement von R und wird im allgemeinen mit 0 bezeichnet. Ein Elemnt e$\in$R mit der Eigenschaft

$$ea = ae = a$$

für jedes $a \in R$ heißt Einselement von R und wird oft auch mit 1 bezeichnet. Falls ein Einselement existiert, so heißt R auch Ring mit 1. Ein Element a aus einem Ring R mit 1 mit der Eigenschaft, daß es ein Element $b \in R$ gibt mit $ab = ba = 1$ (d.h. es gibt ein zu a "inversers" Element) heißt Einheit von R.

Die Menge der Einheiten von R wird mit R^* bezeichnet. Die Menge

$$Z(R) := \{x \in R \mid xy = yx \text{ für jedes } y \in R\}$$

der Elemente von R, die mit allen anderen Elementen von R kommutieren, heißt Zentrum von R.

Ein Ring-Homomorphismus ist eine Abbildung $f: R \to R'$, wobei R und R' Ringe sind und außerdem gilt

$$f(x+y) = f(x) + f(y)$$
$$f(xy) = f(x)f(y)$$

für beliebige $x, y \in R$. Falls anstelle der zweiten Identität

$$f(xy) = f(y)f(x)$$

gilt, so spricht man auch von einem Ring-Antihomomorphismus.

Ein Ringhomomorphismus $f: R \longrightarrow R'$ heißt Homomorphismus von Ringen mit 1, wenn R und R' Ringe mit 1 sind und außerdem

$$f(1) = 1$$

gilt. Falls eine Umkehrung von f existiert, so sagt man auch f ist ein Isomorphismus von Ringen (mit 1).

Sei R ein kommutativer Ring mit 1. Eine R-Algebra (mit 1) ist ein Ring mit 1 zusammen mit einem Homomorphismus

$$f: R \to S$$

von Ringen mit 1, wobei zusätlich gefordert wird, daß

$$f(r) \in Z(S)$$

gilt für jedes $r \in S$.

Seien R ein kommuativer Ring mit 1, S und S' zwei R-Algebren und

$$f: R \longrightarrow S \text{ und } f': R \longrightarrow S'$$

die zugehörigen Homomorphismen. Ein R-Algebra-Homomorphismus $S \longrightarrow S'$ ist ein Ring-Homomorphismus $h: S \longrightarrow S'$ mit der Eigenschaft, daß das folgende Diagramm kommutativ ist.

$$\begin{tikzcd} S \arrow[rr, "h"] & & S' \\ & R \arrow[ul, "f"] \arrow[ur, "f'"'] & \end{tikzcd}$$

Mit anderen Worten, es gelte

$$f' = h \circ f. \qquad (1)$$

Der Begriff der R-Algebra-Anti-Homomorphismus ist auf analoge Weise definiert.

Gilt $R \subseteq S$ und $R \subseteq S'$ und sind die Homomorphjismen f und f' gerade die natürlichen Einbettungen $R \hookrightarrow S$ und $R \hookrightarrow S'$, so bedeutet (1) gerade es ist

$$h(r) = r \text{ für jedes } r \in R.$$

Beispiel 1: $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$
Die Menge der ganzen Zahlen $\mathbb{Z}$ ist mit der gewöhnlichen Addition und der gewöhnlichen Multiplikation ein kommutativer Ring mit 1. Dasselbe gilt für die Menge $\mathbb{Q}$ der rationalen und die Menge $\mathbb{R}$ der reellen Zahlen. Außerdem sind

$$\mathbb{Z}, \mathbb{Q}, \mathbb{R}$$

(auf genau eine Weise) $\mathbb{Z}$-Algebren

$$\mathbb{Q}, \mathbb{R}$$

(auf genau eine Weise) $\mathbb{Q}$-Algebren und

$$\mathbb{R}$$

ist eine $\mathbb{R}$ -Algebra bezüglich der identischen Abbildung $\mathbb{R} \to \mathbb{R}$.
Beispiel 2 (Ring ohne 1)
Die Menge $2\mathbb{Z}$ geraden Zahlen ist mit den gewöhnlichen Operationen + und $\cdot$ ein kommutativer Ring (ohne 1).
Beispiel 3 (Matrizenring)
Die Menge

$$M_n(\mathbb{R}) = \mathbb{R}^{n\times n}$$

der n×n-Matrizen mit reellen Einträgen ist mit den oben definierten Operationen Matrizenaddition und Matrizenmultiplikation ein Ring mit 1. Dieser Ring ist nicht kommutativ. Das Einelement ist gerade die Einheitsmatrix Id. Das Nullelement 0 ist die Nullmatrix, d.h. die n×n-Matrix, deren sämtliche Einträgen 0 sind.

Beispiel 4 (für eine Gruppe, die Einheitengruppe)
Sei R ein Ring mit 1. Dann ist die Menge

$$R^* := \{ r\in R \mid r \text{ ist Einheit von } R \}$$

der Einheiten von R eine Gruppe bezüglich der Multiplikation von R, welche Einheitengruppe von R heißt.
Die Einheitengruppe des Rings $M_n(\mathbb{R})$ ist gerade die Gruppe $GL(n,\mathbb{R})$.

Beispiel 5: GL(n, R)
Sei R ein kommutativer Ring mit 1. Die Menge

$$R^{n\times n}$$

der n×n-Matrizen mit Einträgen aus R ist mit den gewöhnlichen Matrizenoperationen ein Ring und heißt voller Matrizenring über R. Der Ring ist nicht kommutativ (außer im Fall n = 1). Die zugehörigen Einheitengruppe wird mit

$$(R^{n\times n})^* = GL(n, R)$$

bezeichnet und heißt allgemeine lineare Gruppe über R.
Beispiel 6: R^{op}
Sei R ein Ring. Der entgegengesetzte Ring R^{op} besteht aus denselben Elementen wie R und ist mit derselben Additon versehen. Die Multplikation $\circ$ ist wie folgt definiert:

$$a\circ b := b\cdot a,$$

wobei der Punkt auf der rechten Seite die Multiplikation von R bezeichnen soll. Der Ring

$$R^{op}$$

ist ein Ring mit 1, falls R ein Ring mit 1 ist.
Beispiel 7: direktes Produkt
Seien R und S zwei Ringe (mit 1). Dann ist die Menge

$$R\times S := \{(r,s) \mid r\in R \text{ und } s\in S\}$$

mit den Operationen

$$(r, s) + (r', s') := (r+r', s+s')$$
$$(r, s) \cdot (r', s') := (rr', ss')$$
ein Ring (mit dem Einselement (1,1)), welcher <u>direktes Produkt</u> der Ringe R und S heißt.

Beispiel 8 (Ringe von Restklassen ganzer Zahlen)

Seien $n \in \mathbb{Z}$ eine ganze Zahl. Für je zwei Mengen von ganzen Zahlen
$$M \subseteq \mathbb{Z} \text{ und } N \subseteq \mathbb{Z}$$
definieren wir
$$M \cdot N := \{a\cdot b + c\cdot n \mid a\in M, b\in N, c\in\mathbb{Z}\}.$$
Dann gilt

(2) $$\bar{i} \cdot \bar{j} = \overline{i\cdot j}$$

Insbesondere ist die Gruppe der Restklassen modulo n,
$$\mathbb{Z}/(n)$$
ein kommutativer Ring mit 1 bezüglich der Multiplikation (2) und die Abbildung
$$\rho:\mathbb{Z} \to \mathbb{Z}/(n), i \mapsto \bar{i}$$
ein Homomorphismus von Ringen mit 1.

Beweis. <u>Beweis von (2)</u>. Sei
$$\alpha\in \text{LHS}.$$
Dann hat α die Gestalt
$$\alpha = (k\cdot n + i) \cdot (k'\cdot n + j) + \ell n = kk'n^2 + knj + ik'n + ij + \ell n = (kk'n+kj+ik'+\ell)n + ij,$$
d.h. es gilt $\alpha\in$RHS. Wir haben gezeigt,
$$\text{LHS} \subseteq \text{RHS}.$$
Sie umgekehrt
$$\alpha\in\text{RHS}.$$
Dann hat α die Gestalt
$$\alpha = kn + ij = (k\cdot n + i)(0\cdot n + j) + kn - knj = (k\cdot n + i)(0\cdot n + j) + (k-kj)n,$$
d.h. es gilt $\alpha\in$LHS. Wir haben gezeigt, die beiden Seiten von (2) sind gleich.

Die Ringaxiome für $\mathbb{Z}/(n)$ ergeben sich aus denen von $\mathbb{Z}$ mit Hilfe von Formel (2). Zum Beispiel gilt für beliebige i, j, k$\in\mathbb{Z}$:
$$(\bar{i} \cdot \bar{j})\cdot \bar{k} = \overline{i\cdot j} \cdot \bar{k} = \overline{ijk} = \bar{i}\cdot \overline{jk} = \bar{i} \cdot (\bar{j}\cdot \bar{k}),$$
d.h. es gilt das Assoziativgesetz. Die Kommutativität von $\mathbb{Z}/(n)$ ergibt sich aus der von $\mathbb{Z}$:
$$\bar{i} \cdot \bar{j} = \overline{i\cdot j} = \overline{j\cdot i} = \bar{j}\cdot \bar{i}$$

Der Ring $\mathbb{Z}/(n)$ hat ein Einselement, denn es gilt
$$\bar{1}\cdot \bar{j} = \overline{1\cdot j} = \bar{j}.$$
Die Homomorphie-Eigenschaft der obigen Abbildung ergibt sich aus Formel (1) von 2.6.1 und Formel (2). Es ist ein Homomorphismus von Ringen mit 1, weil $\bar{1}$ das Einselement des Restklassenrings ist.

QED.

2.7.3 Begriff des Körpers

Ein <u>Schiefkörper</u> K ist ein Ring mit 1, in dem jedes von 0 verschiedene Element eine Einheit ist. Ein <u>Körper</u> ist ein Schiefkörper, der als Ring kommutativ ist.

Sei R ein kommutativer Ring mit 1. Eine R-Algebra S heißt Divisionsalgebra über R, wenn S ein Schiefkörper ist. Eine Divisionsionsalgebra S über R heißt zentral, wenn Z(S) = R gilt.

Beispiel 1
Der Ring $\mathbb{R}$ der reellen Zahlen (mit den gewöhnlichen Operationen) ist ein Körper.
Beipiel 2
Der Ring $\mathbb{Z}$ der ganzen Zahlen ist kein Körper. Zum Beispiel ist $2\in\mathbb{Z}$ keine Einheit in $\mathbb{Z}$.
Beispiel 3
Die Menge

$$\mathbb{F}_2 := \{0,1\}$$

der Restklassen modulo 2 ist mit den Operationen + und • versehen, welche durch die folgenden beiden Tabellen beschrieben werden könnnen.

+	**0**	**1**
0	0	1
1	1	0

.	**0**	**1**
0	0	0
1	0	1

Das einzige von 0 verschiedene Element ist 1, welches trivialerweise eine Einheit ist. Also ist $\mathbb{F}_2$ ein Körper.

Beispiel 4
Die Menge

$$\mathbb{Z}/(6) := \{0,1,3,4,5\}$$

der Restklassen modulo 6 ist mit den Operationen + und • versehen, welche durch die folgenden beiden Tabellen beschrieben werden könnnen.

+	**0**	**1**	**2**	**3**	**4**	**5**
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

.	**0**	**1**	**2**	**3**	**4**	**5**
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Dieser kommutative Ring mit 1 ist kein Körper. Es gilt nämlich

$$2\cdot 3 = 0.$$

Gäbe es ein zu 2 inverses Element, so könnte man diese Identät damit multiplizieren und erhielte

$$3 = 0,$$

was offensichtlich falsch ist.

Beispiel 5: die Restklassen modulo einer Primzahl

Der Ring der Restklassen modulo n,

$$\mathbb{Z}/(n)$$

ist genau dann ein Körper, wenn n eine Primzahl ist.

Beweis. Ist n keine Primzahl, sagen wir

$$n = a\cdot b,$$

so gilt in $\mathbb{Z}/(n)$,

$$\bar{a}\cdot \bar{b} = \bar{n} = \bar{0},$$

d.h. das Produkt zweier von Null verschiedener Elemente ist Null. Dann kann aber der Ring kein Körper sein.

Sei jetzt n eine Primzahl. Wir haben zu zeigen,

$$\mathbb{F}_n := \mathbb{Z}/(n)$$

ist ein Körper. Zum Beweis ist es sinnvoll, den Begriff des <u>größten gemeinsamen Teilers</u> und den <u>Euklidischen Algorithmus</u> zur Berechnung des größten gemeinsamen Teilers einzuführen.

Der <u>größte gemeinsame Teiler</u> zweier ganzer Zahlen a und b ist eine ganze Zahl d mit den folgenden beiden Eigenschaften:

(i) d teilt die Zahlen a und b:

$$d \mid a \text{ und } d \mid b.$$

(ii) Jede ganze Zahl g, welche a und b teilt, teilt auch d:

$$g \mid a \text{ und } g \mid b \Rightarrow g \mid d.$$

Der größte gemeinsame Teiler von a und b ist, falls er existiert, bis aufs Vorzeichen eindeutig bestimmt. Den nicht negativen der beiden möglichen Werte bezeichnen wir mit

$$\mathrm{ggT}(a,b).$$

Der Euklidische Algorthmus zweier von Null verschiedener ganzer Zahlen a und b besteht in der Berechung einer endlichen echt absteigenden Folge von ganzen Zahlen

$$(a_0 \geq)\ a_1 > a_2 > \ldots > a_r\ ,$$

die sich wie folgt ergeben. O.B.d.A. sei

$$a \geq b \geq 0$$

(ansonsten vertauschen wir a und b bzw. ändern die Vorzeichen).

$$a_0 := a$$

$$a_1 := b$$

<u>Berechnung von</u> a_{i+1} <u>aus</u> a_i <u>und</u> a_{i-1}: wir teilen a_{i-1} mit Rest durch a_i und schreiben

(1) $a_{i-1} = q_i a_i + r_i$ mit $0 \leq r_i < a_i$ $\quad (q_i, r_i \in \mathbb{Z})$

Falls $r_i = 0$ ist, bricht der Algorithmus ab (d.h. es gibt kein a_{i+1}). Andernfalls setzen wir

$$a_{i+1} := r_i.$$

Da die Glieder der Folge positiv sind und echt absteigen, muß die Folge endlich sein, d.h. der Algorithmus bricht nach endlich vielen Schritten ab, d.h. man findet stets ein r mit

$$a_{r-1} = q_r a_r$$

<u>Eigenschaften</u> der a_i:

(iii) a_{i-1} und a_i haben denselben ggT wie a_i und a_{i+1} (insbesondere existiert der ggT der ersten beiden Zahlen genau dann, wenn der ggT der zweiten beiden Zahlen existiert).

(iv) der ggT von a_{r-1} und a_r existiert und ist gleich a_r

Folgerung 1
Für je zwei ganze von Null verschiedene Zahlen existiert der ggT. Er ist gleich dem letzten Glied des Euklidischen Algorithmus angewandt auf diese ganzen Zahlen.

Folgerung 2
Für je drei von Null verschiedene ganze Zahlen a, b, d mit d > 0 sind die folgenden beiden Bedingungen äquivalent.

(v) d = ggT(a,b).

(vi) Es gibt ganze Zahlen u, v mit d = ua + vb, d.h. d ist eine $\mathbb{Z}$-Linearkombination von a und b.

Folgerung 3
Für jede Primzahl n ist $\mathbb{Z}/(n)$ ein Körper.

Die Beweise. Eigenschaft (iii). Nach Definition von a_{i+1} gilt

$$a_{i-1} = q_i a_i + a_{i+1}$$

Damit gilt aber für jede ganze Zahl g:

$$g \mid a_{i-1} \text{ und } g \mid a_i \Leftrightarrow g \mid a_i \text{ und } g \mid a_{i+1},$$

d.h. ein ggT der ersten beiden Zahlen ist auch ein ggT der zweiten beiden und umgekehrt.

Eigenschaft (iv). Wegen

$$a_{r-1} = q_r a_r$$

ist a_r ein gemeinsamer Teiler von a_{r-1} und a_r, und jeder gemeinsame Teiler der letzten beiden Zahlen ist trivialerweise ein Teiler von a_r. Mit anderen Worten:

$$a_r = \mathrm{ggT}(a_{r-1}, a_r).$$

Zu Folgerung 1. Die Existenz des ggT folgt unmittelbar aus (iii) und (iv). Außerdem gilt

$$\mathrm{ggT}(a, b) = \mathrm{ggT}(a_0, a_1) = \ldots = \mathrm{ggT}(a_{i-1}, a_i) = \ldots = \mathrm{ggT}(a_{r-1}, a_r) = a_r .$$

Zu Folgerung 2. Falls (vi) gilt, so ist jeder gemeinsame Teiler von a und b auch ein Teiler von d, d.h. es gilt (v). Es reicht also die Implikation

$$(v) \Rightarrow (vi).$$

zu beweisen. Sei also

$$d = \mathrm{ggT}(a, b).$$

Wir haben zu zeigen,

d ist eine $\mathbb{Z}$-Linearkombination von a und b.

Wir wenden den Euklidischen Algorithmus auf a und b an und erhalten eine Folge

$$a_0 \geq a_1 > a_2 > \ldots > a_r ,$$

mit $a = a_0$, $b = a_1$ und $d = a_r$. Es reicht deshalb wenn wir zeigen:

1. d ist eine $\mathbb{Z}$-Linearkombination von a_r und a_{r-1}.
2. Ist d eine $\mathbb{Z}$-Linearkombination a_i und a_{i+1}, so ist d auch eine von a_{i-1} und a_i (falls $i \geq 1$ ist).

Zu 1: es gilt

$$d = a_r = 0\cdot a_{r-1} + 1\cdot a_r.$$

Zu 2: Sei

$$d = u\cdot a_i + v\cdot a_{i+1}.$$

Wegen $a_{i-1} = q_i a_i + a_{i+1}$, d.h. $a_{i+1} = a_{i-1} - q_i a_i$ folgt

$$\begin{aligned} d &= u\cdot a_i + v\cdot(a_{i-1} - q_i a_i) \\ &= v\cdot a_{i-1} + (u - q_i)a_i . \end{aligned}$$

Zu Folgerung 3. Sei $\overline{i} \in \mathbb{Z}/(n)$ von Null verschieden, d.h. die ganze Zahl ist kein Vielfaches von n. Weil n eine Primzahl ist, gilt

$$ggT(n, i) = 1,$$

d.h. es gibt ganze Zahlen u, v mit

$$u\cdot n + v\cdot i = 1.$$

Wir wenden den natürlichen Homomorphismus $\mathbb{Z}\to \mathbb{Z}/(n)$ an und erhalten:

$$\overline{un} + \overline{v}\cdot \overline{i} = \overline{1}$$

Weil un bei Division durch n den Rest Null ergibt, gilt $\overline{un} = \overline{0}$, also

$$\overline{v}\cdot \overline{i} = \overline{1}.$$

Wir haben gezeigt, jedes von Null verschiedene Element $\overline{i}$ von $\mathbb{Z}/(n)$ besitzt ein Inverses, ist also eine Einheit. Mit anderen Worten $\mathbb{Z}/(n)$ ist ein Körper.
QED.

2.8 Weitere Anwendungen

2.8.1 Der Körper der komplexen Zahlen

Die folgende Teilmenge von $\mathbb{R}^{2\times 2}$ heißt Körper der komplexen Zahlen.

$$\mathbb{C} := \{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a,b\in\mathbb{R}\}$$

Produkt und Summe zweier Matrizen aus $\mathbb{C}$ sind wieder Matrizen aus $\mathbb{C}$. wie man durch direktes Nachrechnen überprüft. Zum Beispiel ist das Produkt zweier komplexer Zahlen wieder eine komplexe Zahl.

(1) $$\begin{pmatrix} a' & b' \\ -b' & a' \end{pmatrix}\cdot\begin{pmatrix} a'' & b'' \\ -b'' & a'' \end{pmatrix} = \begin{pmatrix} a'a''-b'b'' & a'b''+b'a'' \\ -a'b''-b'a'' & a'a''-b'b'' \end{pmatrix} = \begin{pmatrix} u & v \\ -v & u \end{pmatrix}$$

mit $u := a'a''-b'b''$ und $v := a'b''+b'a''$.

Bemerkungen

(i) Die Menge $\mathbb{C}$ ist mit der Addition von Matrizen und der Multiplikation von Matrizen ein Ring mit 1.[2]

(ii) Es gilt das Kommutativgesetz.[3]

---

[2] $\mathbb{C}$ ist mit der Addition eine Gruppe: das Assoziativgesetz gilt weil es für die Addition von beliebigen Matrizen (gleichen Typs) gilt. Die Nullmatrix spielt die Rolle des Nullelements und die Matrix $\begin{pmatrix} -a & -b \\ b & -a \end{pmatrix}$ die Rolle des Negaiven der Matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Das Assoziativgesetz der Multiplikation und die Distributivgesetze gelten, da sie für beliebige n×n-Matrizen gelten. Also ist $\mathbb{C}$ ein Ring. . Es ist sogar ein Ring mit 1, die Einheitsmatrix spielt die Rolle des Einselements.

(iii) Die Menge der reellen Zahlen kann man mit einer Teilmenge von $\mathbb{C}$ identifizieren vermittels der Abbildung

$$\mathbb{R} \to \mathbb{C},\ a \mapsto \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},$$

Man beachte, die Abbildung ist injektiv. Es ist sogar ein Homomorphismus von Ringen mit 1. Der Ring $\mathbb{C}$ wird auf diese Weise eine $\mathbb{R}$-Algebra.

(vi) Die komplexen Zahlen unterscheiden sich von den reellen insbesondere dadurch, daß sie negative Quadrate haben können. Zum Beispiel hat $i := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{C}$ das Quadrat

$$i^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -1 \text{ (als Element von } \mathbb{R}\text{)}.$$

Die komplexe Zahl i heißt <u>imaginäre Einheit</u>.

(v) Eine etwas üblichere Schreibweise erhält man, wenn man schreibt

$$z = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ -b & a \end{pmatrix} = a + bi \text{ für } a,b \in \mathbb{R}.$$

Die Zahl a heißt <u>Realteil</u> der komplexen Zahl z und wird mit

$$a = \mathrm{Re}(z)$$

bezeichnet. Die Zahl b heißt <u>Imaginärteil</u> von z und wird mit

$$b = \mathrm{Im}(z)$$

bezeichnet.

Die beiden Operationen + und · bekommen dann die Gestalt

$$(a'+b'i) + (a''+b''i) = (a'+a'') + (b'+b'')i$$
$$(a'+b'i) \cdot (a''+b''i) = (a'a'' - b'b'') + (a'b'' + b'a'')i$$

Die Formel für die Multiplikation ergibt sich dabei aus den Distributivgesetzen und $i^2 = -1$.

(vi) Die Abbildung

$$\mathbb{C} \to \mathbb{C},\ a+ib \mapsto \overline{a+ib} := a-ib,$$

überführt Summen in Summen und Produkte in Produkte[4],

$$\overline{z'+z''} = \overline{z'} + \overline{z''} \text{ und } \overline{z'\cdot z''} = \overline{z'}\cdot\overline{z''} \text{ für } z',z''\in\mathbb{C}$$

Sie heißt <u>komplexe Konjugation</u>. Die komplexe Zahl $\overline{a+ib}$ heißt konjugiert komplex zu a+ib. Mit Hilfe der komplexen Konjugation kann man Real- und Imaginärteil einer komplexen Zahl ausdrücken:

$$\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z}),\ \mathrm{Im}(z) =^{5} \frac{1}{2i}(z - \overline{z}).$$

Außerdem gilt

$$z\bullet\overline{z} = \overline{z}\bullet z = a^2 + b^2$$

für jede komplexe Zahl $z = a + ib$.[6] Die komplexe Konjugation ist ein Beispiel für einen $\mathbb{R}$-Algebra-Homomorphismus.

---

[3] Die rechte Seite von (1) ändert sich nicht, wenn man die einfach gestrichenen Größen durch die doppelt gestrichenen und die doppelt gestrichenen durch die einfach gestrichenen ersetzt.

[4] Denn es ist gerade die Matrizen-Transposition.

[5] Man beachte, wegen

$$(2i)\bullet(-\frac{1}{2}i) = i\bullet(-i) = 1$$

ist 2i eine Einheit, besitzt also ein Inverses $\frac{1}{2i} = -\frac{1}{2}i$.

[6] Auf Grund der letzten Formel von (v).

(vii) Der Betrag einer komplexen Zahl $z = a + ib$ ist definiert als die nicht-negative reelle Zahl

$$|z| = \sqrt{z\cdot\bar{z}} = \sqrt{a^2+b^2}\ (\geq 0).$$

Nach Definition gilt[7]

1. $|z| = 0 \Leftrightarrow z = 0$.
2. $|z+z'| \leq |z| + |z'|$ (Dreiecksungleichung)
3. $|z\cdot z'| = |z|\cdot|z'|$

(viii) Jede von Null verschieden komplexe Zahl $z = a + bi$ ist eine Einheit. Mit

$$z' := \frac{1}{a^2+b^2}\cdot\bar{z} = \frac{1}{|z|^2}\bar{z}$$

gilt nämlich $z\cdot z' = \frac{1}{a^2+b^2}\cdot z\cdot\bar{z} = 1$.

2.8.2 Die Divisionsalgebra der Quaternionen

Die folgende Teilmenge von $\mathbb{C}^{2\times 2}$ heißt Ring der Quaternionen.

$$\mathbb{H} := \{\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \mid a,b\in\mathbb{C}\}$$

Durch direktes Nachrechnen stellt man fest, daß Produkte und Summen von je zwei solchen Matrizen wieder Matrizen dieser Art sind. Zum Beispiel ist das Produkt von zwei Quaternionen wieder ein Quaternion.

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}\cdot\begin{pmatrix} a' & b' \\ -\bar{b}' & \bar{a}' \end{pmatrix} = \begin{pmatrix} aa'-b\bar{b}' & ab'+b\bar{a}' \\ -a'\bar{b}-\bar{a}\bar{b}' & \bar{a}\bar{a}'-b'\bar{b} \end{pmatrix} \qquad (1)$$

Bemerkungen

(i) Mit der Addition und der Multiplikation von Matrizen ist $\mathbb{H}$ ein Ring mit 1. Man zeigt das auf dieselbe Art und Weise wie im Fall der komplexen Zahlen $\mathbb{C}$.

(ii) Der Ring der Quaternionen ist nicht kommutativ.[8] Jedes Quaternion kommutiert allerding mit jeder reellen Zahl,[9]

$$q\cdot r = r\cdot q \text{ für } q\in\mathbb{H} \text{ und } r\in\mathbb{R}.$$

(iii) Die Menge der komplexen Zahlen läßt sich mit einer Teilmenge von $\mathbb{H}$ identifizieren vermittels der Abbildung

$$\mathbb{C} \to \mathbb{H},\ a \mapsto \begin{pmatrix} a & 0 \\ 0 & \bar{a} \end{pmatrix},$$

Man beachte, die Abbildung ist injektiv und ein Homomorphismus von Ringen mit 1.[10] Wie wir gleich sehen werden, wird $\mathbb{H}$ auf diese Weise nicht zu einer $\mathbb{C}$-Algebra sondern nur zu einer $\mathbb{R}$-Algebra, nämlich durch die Zusammensetzung

---

[7] Außer der Dreiecksungleichung sind diese Ausagen trivial. Zum Beweis der Dreieckungleichung beachte man zunächst, es gilt

(2) $\quad Re(z) \leq |z|$ und $Im(z) \leq |z|$.

Weiter gilt

$$|z+z'|^2 = (z+z')(\bar{z} + \bar{z}') = z\bar{z} + z\bar{z}' + z'\bar{z} + z'\bar{z}' = |z|^2 + 2Re(z\bar{z}') + |z'|^2$$

Wegen (2) erhält man damit

$$|z+z'|^2 = (z+z')(\bar{z} + \bar{z}') \leq |z|^2 + 2|z\bar{z}'| + |z'|^2 = |z|^2 + 2|z|\cdot|z'| + |z'|^2 = (|z|+|z'|)^2$$

also gilt die Behauptung.

[8] siehe zum Beispiel Bemerkung (iii): für die Quaternionen i und j gilt zum Beispiel $i\bullet j = - j\bullet i$.

[9] siehe (ii): Die reellen Zahlen entsprechen gerade den Skalar-Matrizen mit reellen Einträgen. Solche Matrizen liegen aber sogar im Zentrum des gesamten Matrizen-Rings (d.h. sie kommutieren mit jeder Matrix).

$$\mathbb{R} \hookrightarrow \mathbb{C} \hookrightarrow \mathbb{H}.$$

(iv) Eine besondere Rolle spielen die folgenden Quaternionen.

$$i := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \; j := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; k := \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Das Rechnen im Ring der Quaternionen wird durch die folgenden Identitäten bestimmt. Wir schreiben im folgenden für die Einheitsmatrix Id auch einfach nur 1, denn sie in des Einheitselement des Rings $\mathbb{H}$.

$$i^2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\,\mathrm{Id} = -1.$$

$$j^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\,\mathrm{Id} = -1.$$

$$k^2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\,\mathrm{Id} = -1.$$

$$ij = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = k.$$

$$ji = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -k.$$

Weiter ist damit

$ik = iij = -j$
$ki = iji = -jii = j$
$jk = jij = -\,jji = i$
$kj = ijj = -i$

(v) Eine etwas üblichere Schreibweise erhält man, wenn man schreibt

$$\begin{pmatrix} a+bi & c+id \\ -c+id & a-ib \end{pmatrix} = a+ib + jc + kd \text{ für } a,b,c,d \in \mathbb{R}$$

Das Rechnen mit Quaternioen in dieser Schreibweise erfolgt dann mit Hilfe der Distributivgesetze und der in (iii) angegebenen Identitäten.

(vi) Die Abbildung

$$\mathbb{H} \to \mathbb{H},\ a+ib+jc+kd \mapsto \overline{a+ib+jc+kd} := a-ib-jc-kd,$$

ist zu sich selbst invers und ein $\mathbb{R}$-Algebra-Antihomomorphismus.[11] Insbesondere überführt sie Summen und Produkte in Summen und Produkte, wobei sich bei den Produkten die Reihenfolge der Faktoren umkehrt.

---

[10] Sind in Formel (1) alle b's gleich Null, so sind auf der rechten Seite alle Einträge außerhalb der Hauptdiagonalen gleich Null. Die komplexe Zahl 1 wird in die Einheitsmatrix abgebildet.

[11] Für die zugehörigen Matrizen entspricht die Konjugation von Quaternionen der Komposition der zwei folgenden Operationen.

1. Konjugation aller (komplexewertigen) Einträge der Matrix.
2. Transposition der Matrix.

$$\overline{\begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}}^T = \begin{pmatrix} a-ib & c-id \\ -c-id & a+ib \end{pmatrix}^T = \begin{pmatrix} a-ib & -c-id \\ c-id & a+ib \end{pmatrix}$$

Die erste Operation ist ein Homomorphismus und die zweite kehrt die Reihenfolge der Faktoren um. Alternativ kann man die Quaternionen auch als 4×4-Matrizen reeller Zahlen schreiben, d.h. für

$$q = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}, \; z = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \; w = \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \text{ mit } a,b,c,d \in \mathbb{R}$$

schreibt man

$$\overline{q'+q''} = \overline{q'} + \overline{q''} \text{ und } \overline{q'\cdot q''} = \overline{q''}\cdot\overline{q'} \text{ für } q',q''\in\mathbb{H}$$

Die Abbildung läßt sich auch als $\mathbb{R}$-Algebra-Isomorphismus

$$\mathbb{H} \to \mathbb{H}^{op}, q \mapsto \overline{q},$$

auffassen. Sie heißt <u>Konjugation</u>. Das Quaternion $\overline{a+ib+jc+kd}$ heißt das zu a+ib+jc+kd konjugierte Quaternion. Die Konjugation von Quaternionen ist ein Beispiel für einen $\mathbb{R}$-Algebra-Antihomomorphismus.

(vii) Der Betragt eines Quaternions q:= a+ib + jc + kd ist in Analogie zum Fall der komplexen Zahlen definiert als

$$|q| := \sqrt{q\cdot\overline{q}} = \sqrt{a^2+b^2+c^2+d^2}.$$

Dies ist eine nicht-negative reelle Zahl, welche nur für q=0 Null ist. Man beachte, es gilt tatsächlich[12]

$$q\cdot\overline{q} = a^2+b^2+c^2+d^2.$$

Insbesondere ist

$$|q\cdot r| = \sqrt{q\cdot r\cdot\overline{q\cdot r}} = \sqrt{q\cdot r\cdot\overline{r}\cdot\overline{q}} = \sqrt{q\cdot|r|^2\cdot\overline{q}} = \sqrt{q\cdot\overline{q}}\cdot|r| = |q|\cdot|r|,$$

Direkt aus der obigen Identität für $q\cdot\overline{q}$ liest man ab,

$$|q|^2 = q\cdot\overline{q} = \overline{q}\cdot q = |\overline{q}|^2$$

(viii) Wie im Fall der komplexen Zahlen sieht man, daß jedes von Null verschiedene Element

$$q \in \mathbb{H} - \{0\}$$

ein Inverses besitzt, d.h. man kann durch von Null verschiedene Quaternionen teilen:

$$q\cdot(\frac{1}{|q|^2}\cdot\overline{q}) = \frac{1}{|q|^2}\cdot(q\cdot\overline{q}) = \frac{1}{|q|^2}\cdot|q|^2 = 1.$$

und

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

Die Konjugationsabbildung bekommt dann die Gestalt

$$\mathbb{H}\to\mathbb{H}, \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix} \mapsto \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & -a \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}^T,$$

d.h. es ist gerade der Übergang zur transponierten Matrix. Von der Transposition wissen wir aber, daß es sich um einen Algebra-Anti-Homorphismus handelt, der zu sich selbst invers ist.

[12] Es gilt $q = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}$ mit z = a+bi und w = c+di und $\overline{q} = \begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix}$ also

$$q\cdot\overline{q} = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}\cdot\begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix} = \begin{pmatrix} z\overline{z}+w\overline{w} & 0 \\ 0 & \overline{z}z+\overline{w}w \end{pmatrix} = |z|^2+|w|^2 = a^2 + b^2 + c^2 + d^2$$

$$(\frac{1}{|q|^2}\cdot \bar{q})\cdot q = \frac{1}{|q|^2}\cdot (\bar{q}\cdot q) = \frac{1}{|q|^2}\cdot |\bar{q}|^2 = 1$$

Die Quaternionen bilden somit eine Divisionsalgebra über den reellen Zahlen.

(viii) Für ein Quternion

$$q = a + bi + cj + dk$$

gilt

$$iq = ai - b + ck - dj \qquad jq = aj - bk - c + di$$
$$qi = ai - b - ck + dj \qquad qj = aj + bk - c - di$$

Das Quaternion kommutiert genau dann mit i und j,

$$iq = qi,$$

wenn $b = c = d = 0$ gilt. Insbesondere ist $Z(\mathbb{H}) \subseteq \mathbb{R}$. Umgekehrt kommutiert jede reelle Zahl mit allen Quaternionen (vgl. (ii)). Es gilt also

$$Z(\mathbb{H}) = \mathbb{R}.$$

Mit anderen Worten, die Quaternionen bilden eine zentrale Divisionsalgebra über den reellen Zahlen.

Problem

Gibt es außer $\mathbb{H}$ und $\mathbb{R}$ weitere zentrale Divisionsalgebren über den reelen Zahlen ?

Die Antwort lautet: Nein. Der Beweis geht weit über die Möglichkeiten dieser Vorlesung hinaus. Die moderne Mathematik bietet die Möglichkeit, alle zentralen Divisionsalgebren über jeden Körper zu bestimmen (Satz von Merkurjev).

3. Vektorräume

In diesem Abschnitt wollen wir die Art und Weise, wie wir bisher mit linearen Gleichungen und Matrizen gerechnet haben, axiomatisieren. Wir werden dadurch in der oft Lage sein, in derselben Weise mit Objekten umzugehen, wie wir es von Matrizen gewohnt sind, obwohl diese Objekte von ihrer Natur her zunächst wenig mit Matrizen zu tun zu haben scheinen.

3.1 Vektorräume, Unterräume und lineare Abbildungen

Sei K ein Körper. Ein K-Vektoraum ist eine Menge, deren Elemente man addieren und mit den Elementen von K multiplizieren. Genauer: ein K-Vektorraum ist eine Menge V zusammen mit zwei Abbildungen

$$+: V\times V \to V,\ (v', v'') \mapsto v'+v'',$$

$$K\times V \to V,\ (a,v) \mapsto a\cdot v,$$

wobei die folgenden Bedingungen erfüllt sind.

(i) V ist bezüglich Operation + eine abelsche Gruppe.

(ii) Die Multiplikation ist assoziativ, d.h. es gilt

$$a(a'v) = (ab)v \text{ für } a,a'\in K \text{ und } v\in V.$$

(ii) Die Multiplikation und Addition verhalten sich distributiv, d.h.

$$a(v'+v'') = av' + av''$$
$$(a+a')v = av + a'v$$

für $a,a'\in K$, $v',v'',v\in V$.

(iii) Die Multiplikation mit dem Einselement von K induziert die identische Abbildung auf V, d.h.

$$1\cdot v = v \text{ für } v\in V.$$

Seien V,V' Vektorräume über dem Körper K. Eine K-lineare Abbildung

$$f: V' \to V''$$

ist eine Abbildung mit

$$f(\lambda v+\lambda' v') = \lambda\cdot f(v) + \lambda'\cdot f(v')$$

für $\lambda, \lambda'\in K$ und $v,v'\in V$. Die Menge aller K-linearen Abbildungen $V'\to V''$ wird mit

$$L_K(V', V'') \text{ oder } \mathrm{Hom}_K(V',V'')$$

bezeichnet. Eine lineare Abbildung, die eine Umkehrabbildung besitzt, welche ebenfalls linear ist, heißt auch linearer Isomorphismus. Ein linearer Isomorphismus $V \to V$ heißt auch linearer Automorphismus. Eine lineare Abbildung $V\to V$ heißt auch linearer Endomorphismus. Die Menge der lineare Automorphismen eines Vektorraums V wird auch mit

$$\mathrm{Aut}_K(V) \text{ oder } GL_K(V)$$

bezeichnet. Die der Endomorphismen mit

$$\mathrm{End}_K(V).$$

Ein K-linearer Unterraum des K-Vektorraums V ist eine Teilmenge von V, welche mit den Operationen von V wieder ein K-Vektorraum ist.

Bemerkungen

(i) Kriterium für Isomorphie de bijektive lineare Abbildung ist ein Isomorphismus.

(ii) Die Komposition von K-linearen Abbildungen ist K-linear.

(iii) Die linearen Automorphimen eines K-Vektorraums V bilden bezüglich der Hintereinanderausführung von Abbildungen eine Gruppe, welche Automorphismengruppe oder auch allgemeine lineare Gruppe von V heißt.

(iv) Ein K-linearer Unterraum von V ist ein K-Vektorraum W mit folgenden Eigenschaften.

1. $W \subseteq V$ als Menge.

2. Die Abbildung $W \to V$, $w \mapsto w$, ist K-linear.

(v) Unterraumkriterium. Eine Teilmenge $W\subseteq V$ eines K-Vektorraums ist genau dann ein K-linearer Unterraum, wenn die folgenden drei Bedingungen erfüllt sind.

a) Der Nullvektor von V liegt in W,

$$0\in W.$$

b) Mit je zwei Vektoren von V liegt auch deren Summe in W,

$$v,v'\in W \Rightarrow v+v'\in W.$$

c) Das K-Vielfache eines Vektors von W liegt in W,

$$\lambda\in K,\ v\in W \Rightarrow \lambda\cdot v\in W.$$

Beweis. Zu (i). Sei $f:V\to V$ eine bijektive K-lineare Abbildung. Dann gilt

$$f(\lambda v+\lambda' v') = \lambda\cdot f(v) + \lambda'\cdot f(v')$$

und die Umkehrabbildung $f^{-1}:W\to V$ ist wohldefiniert. Insbesondere können wir in diese Identität

$$v = f^{-1}(w) \text{ und } v' = f^{-1}(w')$$

mit $w,w'\in W$ einsetzen. Wir erhalten

$$f(\lambda\, f^{-1}(w)+\lambda'\, f^{-1}(w')) = \lambda\cdot w + \lambda'\cdot w'.$$

Wir wenden auf beide Seiten f^{-1} an und erhalten

$$\lambda\, f^{-1}(w)+\lambda'\, f^{-1}(w') = f^{-1}(\lambda\cdot w + \lambda'\cdot w').$$

Mit anderen Worten f^{-1} ist ebenfalls linear.

Zu (ii). Sind $f:U\to V$ und $g:V\to W$ K-lineare Abbildungen, so gilt

$$(g\circ f)(\lambda v+\lambda' v') = g(f(\lambda v+\lambda' v'))$$

$= g(\lambda\, f(v)+\lambda'\, f(v'))$ (da f linear ist)

$= \lambda g(f(v))+\lambda'\, g(f(v'))$ (da g linear ist)

$= \lambda(g\circ f)\,(v)+\lambda'(g\circ f)(v')$

Also ist auch $g\circ f$ linear.

Zu (iii). Nach (ii) ist die Komposition von linearen Automorphismen eine wohldefinierte Abbildung

$$\mathrm{Aut}_K(V)\times \mathrm{Aut}_K(V) \to \mathrm{Aut}_K(V),\ (f, g) \mapsto f\circ g.$$

Diese Komposition ist assoziativ: für beliebige drei Abbildungen f, g und h, die man zusamensetzen kann, gilt

$$(f\circ g)\circ h = f\circ(g\circ h).$$

Für jedes Element x aus dem Definitionsbereich von h gilt nämlich

$$\begin{aligned}(f\circ g)\circ h\,(x) &= (f\circ g)(h(x)) = f(g(h(x)))\\ &= f((g\circ h)(x))\\ &= f\circ(g\circ h)\,(x).\end{aligned}$$

Die identische Abbildung hat offensichtlich die Eigenschaften eines neutralen Elements. Schließlich ist die Umkehrabbildung jedes Automorphismus wieder ein Automorphismus, der die Eigenschaften eines inversen Elements besitzt.

Zu (iv). Bedingung 2 bedeutet gerade, daß die auf W definierten Operationen + und · dieselben sind wie die von V (genauer : die Einschränkungen der entsprechenden Operationen von V).

Zu (v). Für jeden linearen Unterraum sind die Bedingungen erfüllt. Sei jetzt umgekehrt W eine Teilmenge, für welche die drei Bedingungen erfüllt sind.

Wir haben zu zeigen, die Vektorraumaxiome sind für die Menge W erfüllt bezüglich der auf V definierten Operationen. Wegen Bedingung a) ist W eine nicht-leere Menge. Wegen b) definiert die Addition von V eine Abbildung

$$+\colon W \times W \to W,\ (w',w'') \mapsto w'+w''.$$

Wegen c) definiert die Multiplikation der Vektoren von V mit Elementen aus K eine Abbildung

$$\cdot\colon K\times W \to W,\ (\lambda, w) \mapsto \lambda w.$$

Wir haben zu zeigen, (W,+) ist eine abelsche Gruppe. Das Assoziativgesetzt ist erfüllt, da es sogar für alle Vektoren aus der größeren Menge V gilt. Die Existenz des neutralen Elements ist gesichert wegen a). Die Existenz des negativen Vektors ergibt sich aus c):

$$w\in W \mapsto -w = (-1)w \in W.$$

Die übrigen Vektorraumaxiome gelten ebenfalls, da sie sogar für die Elemente aus der größeren Menge V gelten.

QED.

3.2 Beispiele

3.2.1 Der Vektorraum K^n

Sei K ein Körper. Wir versehen die Menge

$$K^n := \{ \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} \mid a_1,\dots,a_n \in K\}$$

der n-zeiligen Spalten mit Einträgen aus K mit den üblichen Matrizen-Operationen.

$$\begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ \dots \\ a_n+b_n \end{pmatrix}$$

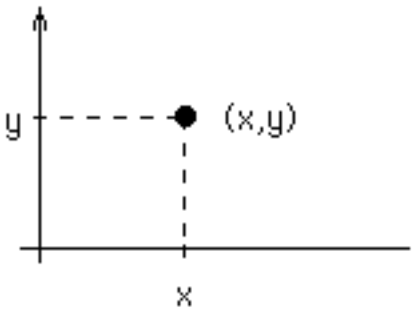
$$\lambda \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \dots \\ \lambda a_n \end{pmatrix}$$

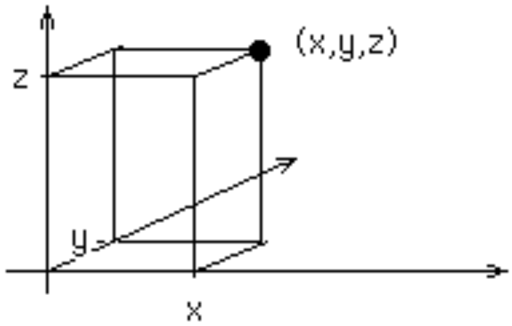
für a_i, b_i, $\lambda \in K$. Mit diesen Operationen ist K^n ein K-Vektorraum.

Bemerkungen

(i) Im Fall n=1 betrachtet man die Elemente von K^n auch als Punkte eines Raumes, die durch eine Koordinate gegeben sind. Mit dieser Menge K^1 verbindet man oft die Vorstellung einer Geraden, obwohl das nicht immer korrekt ist. Im Fall $K = \mathbb{C}$ wäre die Vorstellung einer Ebene angemessener.

(ii) Im Fall n=2 betrachtet man die Elemente von K^n auch als Punkte eines Raumes, die durch zwei Koordinaten gegeben sind. Mit der Menge K^2 selbst verbindet man oft die Vorstellung einer Ebene.

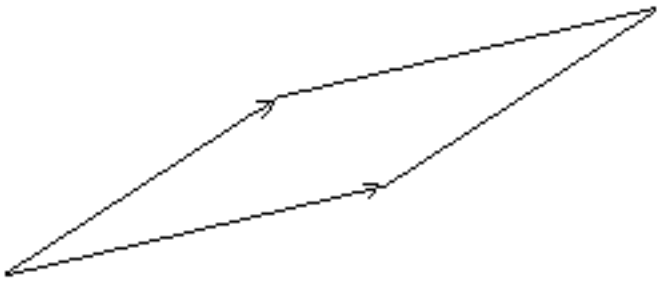




(iii) Im Fall n=3 betrachtet man die Elemente von K^n auch als Punkte eines Raumes, die durch drei Koordinaten gegeben sind. Mit der Menge K^3 selbst verbindet man die Vorstellung eines dreidimensionalen Raumes.

(v) Statt als Punkte stellt man sich die Vektoren des K^n auch als Pfeile vor, wobei man Pfeile derselben Richtung und Länge als gleich ansieht. Zum Beispiel kann man sich vorstellen, daß der Pfeil eine Bewegung beschreibt in der durch den Pfeil gegebenen Richtung und in einer durch die Länge des Pfeils gegebenen Ausdehnung. Den Pfeil mit dem Angriffspunkt im Ursprung $\begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix}$ und der Spitze in $\begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}$ bezeichnet man dabei einfach mit $\begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}$.

(iv) Die Summe von zwei Vektoren in den beschriebenen Modellen entspricht der Diagonale eines Parallelogramms, dessen Seiten die beiden Vektoren bilden.



(iiv) Die Multiplikation eines Vektors mit einem Körperelement entspricht dem Ersetzen eines Pfeils durch einen dadurch parallelen Pfeil, dessen Länge sich um den gegebenen Faktor verändert hat..

(iiiv) Die obigen mehr heuristischen Betrachtungen kann man im Fall $K = \mathbb{R}$ korrekt machen. Für beliebiges K können sie, wenn man mit der nötigen Vorsicht mit ihnen umgeht, immer noch hilfreich sein.

3.2.2 Der Vektorraum $K^{m\times n}$

Seien K ein Körper und m,n natürliche Zahlen. Dann ist die Menge

$$M_{m,n}(K) = K^{m\times n}$$

der Matrizen vom Typ (m,n) mit Einträgen aus K zusammen mit der gewöhnlichen Matrizen-Multipliktion und der gewöhnlichen Multiplikaition mit Elememten aus K ein K-Vektorraum. (Man beweise dies).

3.2.3 Abbildungen mit Werten in einem Vektorraum

Seien M eine Menge, K ein Körper und V ein K-Vektorraum. Wir betrachen die Menge der auf M definierten Abbildungen $M \longrightarrow V$ mit Werten in V,

$$\mathrm{Abb}(M, V) = \mathrm{Hom}_{\mathrm{Ens}}(M, V)$$

zusammen mit zwei wie folgt definierten Abbildungen.
Die erste Abbildungen bezeichnen wir mit '+'.

$$+: \mathrm{Abb}(M,V)\times\mathrm{Abb}(M,V) \longrightarrow \mathrm{Abb}(M,V),\ (f, g) \mapsto f+g.$$

Dabei bezeichne f+g die Abbildung mit

$$(f+g)(m) \quad := f(m) + g(m)$$

für jedees $m \in M$.

Die zweite Abbildung bezeichnen wir mit '•'.

$$\bullet: K\times\mathrm{Abb}(M,V) \longrightarrow \mathrm{Abb}(M,V),\ (a, f) \mapsto a\bullet f.$$

Dabei bezeichne a•f die Abbildung mit

$$(a\cdot f)(m) \quad := a\cdot f(m)$$

für jedes $m \in M$.

Die Menge Abb(M, V) bildet zusammen mit diesen beiden Operationen einen K-Vektorraum. (Man beweise dies).

3.2.4 Lineare Abbildungen

Seien K ein Körper und V', V" zwei K-Vektorräume. Dann ist die Menge

$$\mathrm{Hom}_K(V',V'') = \mathrm{Hom}_{K\text{-}\mathrm{Mod}}(V', V'')$$

der K-linearen Abbildungen $V' \to V''$ mit denselben Operation wie in 3.2.3 ein K-Vektorraum. Mit anderen Worten, $Hom_K(V',V'')$ ist ein K-linearer Unterraum von

$$Abb(V', V'').$$

(Man beweise dies).

3.2.5 Direkte Produkte von Vektorräumen

Seien K ein Körper und V',V'' zwei K-Vektorräume. Dann ist die Menge

$$V'\times V'' := V'\oplus V'' := \{(v',v'') \mid v'\in V,\ v''\in V''\}$$

zusammen mit den folgenden Operationen ein K-Vektorraum.

$$(v',v'') + (w',w'') := (v'+w',\ v''+w'')$$
$$a\cdot(v',v'') := (a\cdot v',\ a\cdot v'')$$

Dieser K-Vektorraum heißt <u>direktes Produkt</u> oder auch <u>direkte Summe</u> von V' und V''.

Sei jetzt $\{V_i\}_{i\in I}$ eine beliebige Familie von K-Vektorräumen. Dann ist die Menge

$$\prod_{i\in I} V_i := \{ (v_i)_{i\in I} \mid v_i\in V_i \}$$

aller Familien $(v_i)_{i\in I}$, deren i-tes Glied im Vektorraum V_i liegt mit den folgenden Operationen ein K-Vektorraum.[13]

$$(v_i) + (v'_i) := (v_i+v'_i)$$
$$a\cdot(v_i) := (a\cdot v_i)$$

Bemerkungen

(i) Ist die Index-Menge I = {1,...,n} die Menge der ersten n natürlichen Zahlen und $V_i = K$ für jedes i, so gilt

$$\prod_{i\in I} V_i = K^n$$

---

[13] Sei $V := \prod_{i\in I} V_i$. Die beiden Operationen definieren Abbildungen

$$+\colon V \times V \longrightarrow V,\ ((v_i), (v'_i)) \mapsto (v_i) + (v'_i),$$

und

$$\times\colon K \times V \longrightarrow V,\ (a, (v_i)) \mapsto a\cdot(v_i).$$

Die K-Vektorraum-Eigenschaften der V_i übersetzen sich dabei in die Vektorraum-Eigenschaften von V. Zum Beispiel gilt für die Elemente (v_i) von V das Assoziativgesetz, weil für deren Koordinaten das Assoziativgesetz gilt:

$$\begin{aligned}(u_i) + ((v_i) + (w_i)) &= (u_i) + ((v_i+w_i)) = (u_i +(v_i+w_i))\\ &= ((u_i + v_i) +w_i) \qquad \text{(in } V_i \text{ gilt das Assoziativgesetz)}\\ &= (u_i + v_i) + (w_i)\\ &= ((u_i) + (v_i)) + (w_i)\end{aligned}$$

(ii) Für jeden direkten Faktor V_j eines direkten Produkts $\prod_{i\in I} V_i$ ist die Abbildung

$$\prod_{i\in I} V_i \to V_j\ ,\ \{v_i\} \mapsto v_j,$$

welche jede Familie von $\prod_{i\in I} V_i$ auf ihre j-te Koordinate abbilded, eine K-lineare Abbildung. Sie heißt <u>Projektion auf den j-ten Faktor</u>.

3.2.6 Direkte Summe von Familien

Sei $\{V_i\}_{i\in I}$ eine beliebige Familie von K-Vektorräumen. Dann ist die Menge

$$\oplus_{i\in I} V_i := \{\ (v_i)_{i\in I} \mid v_i \in V_i\ ,\ v_i = 0 \text{ für fast alle } i\in I\}$$

mit den folgenden Operationen ein K-Vektorraum.

$$(v_i) + (v'_i) := (v_i + v'_i)$$

$$a\cdot(v_i) := (a\cdot v_i)$$

Dabei bedeute “fast alle” dasselbe wie “alle bis auf endlich viele Ausnahmen”.

Bemerkungen

(i) Für jede Familie $\{V_i\}_{i\in I}$ von K-Vektorräumen ist deren direkte Summe ein K-linearer Unterraum des direkten Produkts,

$$\oplus_{i\in I} V_i \subseteq \prod_{i\in I} V_i$$

(ii) Ist die Menge I unendlich und sind die Vektorräume V_i sämtlich von Null verschieden, so ist diese Inklusion echt.

(iii) Jeder der direkten Summanden V_j einer direkten Summe $\oplus_{i\in I} V_i$ kann mit einen K-linearen Unterraum von $\oplus_{i\in I} V_i$ identifiziert werden vermittels der Abbildung

(1) $$V_j \to \oplus_{i\in I} V_i\ ,\ v \mapsto \{\delta_{ij}\cdot v\}_{i\in I},$$

welche jedem Vektor $v\in V_j$ die Familie $\{\delta_{ij}\cdot v\}_{i\in I}$ zuordnet, deren einziges von Null verschiedenes Glied das j-te Glied ist, welches seinerseits mit v übereinstimmt. Das Symbol

$$\delta_{ij} := \begin{cases} 1 \text{ falls } i=j \\ 0 \text{ falls } i\neq j \end{cases}$$

heißt <u>Kronecker-Symbol</u>. Die K-lineare Abbildung (1) heißt <u>natürliche Einbettung</u> des j-ten direkten Summanden in die direkte Summe $\oplus_{i\in I} V_i$.

3.2.7 Durchschnitte von Unterräumen

Seien K ein Körper und U’, U” $\subseteq$ V zwei K-lineare Unterräume von V. Dann ist

$$U' \cap U''$$

ein K-linearer Unterraum von V.

Allgemeiner sei $\{U_i\}_{i\in I}$ eine Familie von K-linearen Unterräumen von V. Dann ist der Durchschnitt

$$\bigcap_{i\in I} U_i$$

ein K-linearer Unterraum von V.

Beweis. Es reicht die allgemeinere zweite Aussage zu beweisen. Dazu reicht es zu zeigen der Durchschnitt der U_i genügt den drei Bedingungen des Unterraum-Kriteriums.

Zu Bedingung 1:
Da der Null-Vektor in jedem U_i liegt, liegt er auch in deren Durchschnitt.

Zu Bedingung 2:
Für je zwei Vektoren u und v aus dem Durchschnitt der U_i gilt $u+v \in U_i$ für jedes i. Also liegt u + v im Durchschnitt der U_i.

Zu Bedingung 3:
Für jeden Vektor u aus dem Durchschnitt der U_i und jedes $\lambda\in K$ gilt $\lambda\cdot u \in U_i$ für jedes i. Also liegt $\lambda\cdot u$ im Durchschnitt der U_i.

QED.

3.2.8 Der von einer Teilmenge erzeugte Unterraum

Seien K ein Körper, V ein K-Vektorraum und $M\subseteq V$ eine beliebige Teilmenge. Dann ist die Menge

$$<M> := \{c_1v_1 + ... + c_kv_k \mid c_1,...,c_k\in K,\ v_1,...,v_k\in M,\ k=1,2,3,...\}$$

aller (endlichen) Linearkombinationen von Vektoren aus M mit den Operationen von V selbst wieder ein K-Vektorraum. Dieser heißt der von M erzeugte Unterraum von V oder auch kurz das Erzeugnis von M. Die Menge M heißt Erzeugendensystem von V, wenn

$$<M> = V$$

gilt, d.h. wenn jeder Vektor von V als K-Linearkombination von jeweisl endlich vielen Vektoren aus M geschrieben werden kann.

Bemerkungen

(i) Jeder K-lineare Unterraum $W\subseteq V$, welcher die Menge M als Teilmenge enthält, enthält auch das Erzeugnis von M,

$$M \subseteq W,\ W \text{ linearer Unterraum} \Rightarrow\ <M> \subseteq W.$$

(ii) Insbesondere liegt M im Durchschnitt aller linearen Unterräume von V, welche die Menge M enthalten, ist also “kleiner” als alle diese Unterräume.

(iii) Der Raum $<M>$ ist selbt ein linearer Unterraum, welcher die Menge M enthält. Der Durchschnitt aller linearen Unterräume mit dieser Eigenschaft, ist also gleich $<M>$,

$$<M> = \bigcap_{\substack{M\subseteq W \\ W\subseteq V \text{ linear}}} W$$

(iv) Die Eigenschaft von $<M>$, gleich dem Durchschnitt aller linearen Unterräume zu sein, die die Menge M enthalten, werden wir im folgenden oft durch die etwas laxe Formulierung ausdrücken, daß $<M>$ der kleinste Unterraum ist, welcher die Menge M enthält.

$<M>$:= kleinster linearer Unterraum, welcher die Menge M enthält.

3.2.9 Erzeugendensysteme und lineare Abbildungen

Seien V ein K-Vektorraum mit dem Erzeugendensystem

$$\{v_i\}_{i\in I}$$

und

$$f, g: V \to W$$

zwei K-lineare Abbildungen mit

$$f(v_i) = g(v_i) \text{ für jedes } i\in I.$$

Dann gilt

$$f = g.$$

Beweis. Sei $v\in V$ ein vorgegebener Vektor. Wir haben zu zeigen,

$$f(v) = g(v).$$

Weil $\{v_i\}_{i\in I}$ ein Erzeugendensystem ist, gibt es eine Familie $\{c_i\}_{i\in I}$ von Elementen aus K, die fast alle Null sind, mit

$$v = \sum_{i\in I} c_i v_i.$$

Weil f und g lineare Abibldungen sind, folgt

$$f(v) = \sum_{i\in I} c_i f(v_i) = \sum_{i\in I} c_i f(v_i) = g(v).$$

QED.

3.2.10 Der von einer Menge frei erzeugte Vektorraum

Seien M eine beliebige Menge und K ein Körper. Für jedes $m\in M$ bezeichne

$$\delta_m: M \to K$$

die <u>charakteristische Abbildung</u> von $\{m\}$, d.h. δ_m ist gleich 1 an der Stelle m und sonst gleich Null,

$$\delta_m(x) := \begin{cases} 1 \text{ falls } x=m \\ 0 \text{ sonst} \end{cases}$$

Dann ist die folgende Abbildung injektiv,

$$\delta: M \to \mathrm{Abb}(M, K),\ m \mapsto \delta_m,$$

denn für $m_1 \neq m_2$ sind δ_{m_1} und δ_{m_2} verschiedene Abbildungen (die eine ist an der Stelle m_1 gleich 1 und die andere gleich 0). Mit anderen Worten, wir können die Menge M mit der Teilmenge $\delta(M)$ des K-Vektorraums Abb(M) identifizieren, indem wir m mit δ_m identifizieren,

$$M \subseteq \mathrm{Abb}(M,K).$$

Es hat also Sinn, von Erzeugnis

$$\langle M\rangle$$

der Menge M zu sprechen. Dieses ist ein K-Vektorraum und heißt der von M <u>frei erzeugte K-Vektorraum</u>. Er wird mit

$$F_K(M)$$

bezeichnet.

3.2.11 Der Faktorraum

Seien V ein K-Vektorraum und $W \subseteq V$ ein K-linearer Unterraum. Dann heißt für jedes $v \in V$ die Menge

$$v+W := \{v+w \mid w \in W\}$$

affiner Unterraum von V oder auch Verschiebung von W (mit v) in V. Die Menge aller Verschiebungen von W in V wird mit

$$V/W := \{v+W \mid v \in V\} \text{ (gesprochen "V modulo W")}$$

bezeichnet und heißt Faktorraum von V modulo W.

Bemerkungen

(i) Kriterium für die Gleichheit zweier Versschiebungen. Zwei Vektoren $v,v' \in V$ können durchaus dieselbe Verschiebung definieren, auch wenn sie verschieden sind. Genauer gilt

$$v + W = v' + W \Leftrightarrow v-v' \in W.$$

(ii) Vektorraumeingenschaft von V/W. V/W ist mit den folgenden Operationen ein K-Vektorraum.

$$(v'+W) + (v''+W) := (v'+v'') + W$$
$$c\cdot(v+W) := (cv) + W$$

für $v,v',v'' \in V$ und $c \in K$.

(iii) Linearität der natürlichen Abbildung. Die Abbildung

$$\rho: V \to V/W,\ v \mapsto v+W,$$

ist K-linear (und heißt natürliche Abbildung von V auf den Faktorraum V/W).

Beweis. Zu (i). Beweis von $\Rightarrow$. Mit $v + W = v' + W$ gilt

$$v = v + 0 \in v + W = v' + W,$$

d.h. v kann in der Gestalt $v = v' + w$ mit $w \in W$ geschrieben werden. Also gilt

$$v - v' = w \in W.$$

Beweis von $\Leftarrow$. Sei $w := v - v' \in W$. Wir haben zu zeigen

1. $v + W \subseteq v' + W$
2. $v' + W \subseteq v + W$.

Da mit $v - v' \in W$ auch $v' - v \in W$ gilt, ist unsere Voraussetzung symmetrisch in v und v', d.h. es genügt eine der beiden Inklusionen zu beweisen. Beweisen wir zum Beispiel die erste. Mit $x \in v + W$ gilt

$$x = v + w'$$

für ein $w' \in W$. Damit ist aber

$$x = v' + (v-v') + w'.$$

die letzten beiden Summanden liegen in W, also liegt auch ihre Summe in W, d.h. es gilt $x \in v' + W$.

Zu (ii). Im wesentlichen ist zu zeigen, daß die in (i) beschriebenen Operationen korrekt definiert sind. Alle weiteren Aussagen sind dann leicht einzusehen. Zum Beweis der Korrektheit der Operationen beweisen wir zunächst ein Kriterium für die Gleichheit zweier Verschiebungen von W.

1. Schritt: Korrektheit der Definiton der Operationen.

Wir haben zu zeigen, sind x, x', x'' $\in$ Vektoren mit

$$x + W = v + W,\ x'+W = v'+W,\ x''+W=v''+W$$

so ist

$$(*)\qquad (x'+x'') + W = (v'+v'')+W \text{ und } (cx) + W = (cv) + W.$$

Nach (i) gilt auf Grund der Voraussetzungen

$$x-v \in W,\ x'-v' \in W,\ x''-v'' \in W.$$

Weil W ein K-Vektorraum ist, folgt

$$cx - cv = c(x-v) \in W \text{ und } (x'+x'') - (v'+v'') = (x'-v')+(x''-v'')\in W.$$
Auf Grund des Kriteriums (i) gilt dann aber (*).

<u>2. Schritt</u>: die Vektorraumeigenschaft von V/W.
Die Vektorraumaxiome für V/W ergeben sich aus den entsprechenden Axiomen für V und der Definiton Operationen von V/W. Zum Beispiel erhält man für das Assoziativgesetz der Addition:

$$((a+W) + (b+W)) + (c+W)= ((a+b)+W) + (c+W) = ((a+b)+c) + W$$
$$(a+W) + ((b+W) + (c+W))= (a+W) + ((b+c)+W) = (a+(b+c)) + W$$

Das Assoziativgesetz der Addition auf V/W folgt somit aus dem Assoziativgesetz der Addition auf V. Für die anderen Vektorraumaxiome ist die Situation analog.
...
<u>Zu (iii)</u>. Folgt unmittelbar aus der Definition der Vektorraumoperationen auf V/W.
QED.

3.2.12 Bild und Kern einer linearen Abbildung

Seien K ein Körper und f: V → V' eine K-lineare Abbildung. Dann gilt:
(i) Das Bild von f,

$$\mathrm{im}(f) := \{f(v) \mid v\in V\}$$

ist ein K-linearer Unterraum von V'.
(ii) Die Menge

$$\ker(f) := \{v\in V \mid f(v) = 0\}$$

aller Elemente von V, die bei f in die Null abgebildet werden, ist ein K-linearer Unterraum von V (und heißt <u>Kern</u> der Abbildung f).

Beweis. <u>Zu (i)</u>. <u>1. Schritt</u>: Addition und Multiplikation von V' induzieren Abbildungen

$$\mathrm{im}(f)\times \mathrm{im}(f) \to \mathrm{im}(f),\ (v',w') \mapsto v'+w',$$
$$K\times\mathrm{im}(f) \to \mathrm{im}(f),\ (c,v') \mapsto c\cdot v'.$$

Seien v',w'∈im(f) und c∈K. Wir haben zu zeigen,

(1) $$v+w\in\mathrm{im}(f)$$
(2) $$c\cdot v\in\mathrm{im}(f)$$

Nach Voraussetzung gibt es Vektoren v, w∈V mit

$$v' = f(v),\ w' = f(w).$$

Also gilt

$$v'+w' = f(v) + f(w) = f(v+w) \in \mathrm{im}(f),$$
$$c\cdot v' = c\cdot f(v) = f(c\cdot v) \in \mathrm{im}(f),$$

d.h. es gelten (1) und (2).
<u>2. Schritt</u>. im(f) ist eine abelsche Gruppe bezüglich der Addition.
Das Assoziativgesetz der Adddition gilt für die Elemente von im(f), da es sogar für die Elemente der größerene Menge V' gilt.
im(f) hat eine neutrales Element bezüglich der Addition, da das neutrale Element der Addition 0∈V' in im(f) liegt: 0 = f(0)∈im(f).
Jedes Element v'∈im(f) ⊆ V' hat die Gestalt v' = f(v) für ein v∈V. Da V ein Vektorraum ist, existiert das zu v negative Element -v. Es gilt

$$0 = v + (-v) = (-v) + v$$

also

$$0 = f(0) = f(v+(-v)) = f((-v)+v) = f(v) + f(-v) = f(-v) + f(v).$$

Also ist f(-v)∈im(f) das zu v' = f(v) negative Element.
Schließlich gilt für die Element von im(f) das Kommutativgesetz der Addition, weil es sogar für die Elemente der größeren Menge V' gilt.
<u>3. Schritt</u>. Es gelten auch die übrigen Vektorraum-Axiome.

Für die Elemente von im(f)($\subseteq$ V') gelten das Assoziativitätsgesetzt der Multiplikation, die Distributivgesetze und die Regel für die Multiplikation mit 1, weil diese Gesetze sogar für die Elemente aus der größeren Menge V' gelten.
Zu (ii). 1. Schritt: Addition und Multiplikation von V induzieren Abbildungen

$$\ker(f)\times\ker(f) \to \ker(f),\ (v,w) \mapsto v+w,$$

$$K\times\ker(f) \to \ker(f),\ (c,v) \mapsto c\cdot v.$$

Seien $v,w\in\ker(f)$ und $c\in K$. Wir haben zu zeigen, $v+w\in\ker(f)$ und $c\cdot v\in\ker(f)$. Nach Voraussetzung gilt

$$f(v) = 0 \text{ und } f(w) = 0.$$

Also gilt auch

$$f(v+w) = f(v) + f(w) = 0 + 0 = 0$$

$$f(c\cdot v) = c\cdot f(v) = c\cdot 0 = 0$$

also $v+w\in\ker(f)$ und $c\cdot v\in\ker(f)$.
2. Schritt. ker(f) ist bezüglich der Addition eine abelsche Gruppe.
Es reicht zu zeigen, ker(f) ist eine Untergruppe von V (bezüglich der Addition). Wegen

$$f(0) = 0$$

gilt $0\in\ker(f)$, d.h. ker(f) ist nicht leer. Weiter gilt mit $v,w\in\ker(f)$ auch

$$f(v-w) = f(1\cdot v + (-1)\cdot w) = 1\cdot f(v) + (-1)\cdot f(w) = 1\cdot 0 + (-1)\cdot 0 = 0$$

also $v-w\in\ker(f)$. Also ist ker(f) eine Untergruppe.
3. Schritt. Die übrigen Vektorraum-Axiome sind ebenfalls erfüllt.
Für die Elemente von ker(f)($\subseteq$ V) gelten das Assoziativitätsgesetzt der Multiplikation, die Distributivgesetze und die Regel für die Multiplikation mit 1, weil diese Gesetze sogar für die Elemente aus der größeren Menge V gelten.
QED.

3.3 Die Dimension eines Vektorraums

3.3.1 Lineare Abhängigkeit

Seien K ein Körper, V ein K-Vektorraum und

$$v_1,\dots,v_k \in V$$

endlich viele Vektoren aus V. Diese Vektoren heißen (K-) linear abhängig, wenn es solche Körperelemente $c_1,\dots,c_k\in K$ gibt, daß die folgenden beiden Bedingungen erfüllt sind

1. $c_1v_1 + \dots + c_kv_k = 0$
2. Mindestens ein c_i ist ungleich Null.

Andernfalls heißen die Vektoren $v_1,\dots,v_k$ (K-) linear unabhängig. Eine beliebige Menge von Vektoren $M\subseteq V$ heißt linear unabhängig, wenn je endlich viele Vektoren aus M linear unabhängig sind. Die Menge M heißt linear abhängig, wenn sie nicht linear unabhängig ist.

Bemerkungen

(i) Ein Ausdruck der Gestalt

$$c_1v_1 + \dots + c_kv_k$$

mit $v_i\in V$ und $c_i\in K$ für alle i heißt Linearkombination der Vektoren $v_1,\dots,v_k$ mit Koeffizienten aus K. Die Linearkombination heißt trivial, wenn alle Koeffizienten c_i gleich Null sind.

(ii) Die Vektoren $v_1,...,v_k$ von V sind nach Definition genau dann linear abhängig, wenn es eine nicht-triviale Linearkombination gibt, welche gleich Null ist.

Beispiel 1

Die Vektoren $\begin{pmatrix}2\\4\\6\end{pmatrix}$ und $\begin{pmatrix}3\\6\\9\end{pmatrix}$ sind linear abhängig, denn es gilt

$$3 \cdot \begin{pmatrix}2\\4\\6\end{pmatrix} + (-2)\cdot \begin{pmatrix}3\\6\\9\end{pmatrix} = 0$$

Beispiel 2

Der Vektor $\begin{pmatrix}0\\0\\0\end{pmatrix}$ ist linear abhängig, denn es gilt

$$1 \cdot \begin{pmatrix}0\\0\\0\end{pmatrix} = 0$$

Beispiel 3

Die Vektoren $\begin{pmatrix}1\\2\\3\end{pmatrix}$ und $\begin{pmatrix}3\\2\\1\end{pmatrix}$ des $\mathbb{R}^3$ sind linear unabhängig, denn aus

$$(*) \qquad x \cdot \begin{pmatrix}1\\2\\3\end{pmatrix} + y\cdot \begin{pmatrix}3\\2\\1\end{pmatrix} = 0$$

folgt, x und y sind Lösungen des linearen homogenen Gleichungssystems mit der Koeffizientenmatrix

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

Durch elementare Zeilenoperationen erhalten wir äquivalente homogene Systeme mit den Koeffizientenmatrizen

$$\begin{pmatrix}1&3\\0&-4\\0&-8\end{pmatrix} \sim \begin{pmatrix}1&3\\0&1\\0&0\end{pmatrix} \sim \begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}.$$

Mit anderen Worten, aus (*) folgt x = y = 0, d.h. die Vektoren sind linear unabhängig.

Beispiel 4

Seien K ein Körper, M eine beliebige Menge und

$$V := \langle M \rangle$$

der von M frei erzeugte K-Vektorraum. Dann ist M eine linear unabhängige Teilmenge von V.

Beweis. Seien $m_1,...,m_k \in M$ endlich viele paarweise verschiedene Elemente von M. Wir haben zu zeigen, die zugehörigen Elemente

$$\delta_{m_1},...,\delta_{m_k} \in \mathrm{Abb}(M,K)$$

des Vektorraums der Funktionen sind linear unabhängig. Seien $c_1,...,c_k \in K$ beliebige Körperelemente mit

$$c_1\delta_{m_1} + ... + c_k\delta_{m_k} = 0.$$

Wir haben zu zeigen, die Linearkombination auf der rechten Seite ist trivial, d.h. sämtliche c_i sind gleich Null. Nach Voraussetzung ist die Funktion

$$f := c_1\delta_{m_1} + ... + c_k\delta_{m_k} : M \to K$$

die Nullfunktion, d.h. für jedes $m\in M$ ist

$$0 = f(m) = c_1\delta_{m_1}(m) + ... + c_k\delta_{m_k}(m).$$

Speziell für $m = m_i$ ist höchstens ein Summand auf der rechten Seite ungleich Null, nämlich der i-te. Es gilt also

$$0 = f(m_i) = c_i\delta_{m_i}(m_i) = c_i\cdot 1 = c_i.$$

Da dies für beliebiges i gilt, sind sämliche $c_i = 0$.

QED.

3.3.2 Charakterisierung der linearen Unabhängigkeit endlicher Mengen

Seien K ein Körper, V ein K-Vektorraum und

$$v_1,...,v_k \in V$$

endlich viele Vektoren aus V. Dann sind folgende Aussagen äquivalent.

(i) $v_1,...,v_k$ sind K-linear unabhängig.

(ii) Keines der v_i ist K-Linearkombination der übrigen.

(iii) Die lineare Abbildung

$$\varphi:K^k \to V, (c_1,...,c_k) \mapsto c_1v_1 + ... + c_kv_k,$$

ist injektiv.

Beweis. <u>(i) ⇒ (iii)</u>. Angenommen, die Abbildung von (iii) ist nicht injektiv. Dann gibt es zwei verschiedene k-Tupel

$$(c'_1,...,c'_k), (c''_1,...,c''_k) \in K^k$$
$$(c'_1,...,c'_k) \neq (c''_1,...,c''_k),$$

mit $\varphi(c'_1,...,c'_k)=\varphi(c''_1,...,c''_k)$, .d.h. mit

$$\begin{aligned} 0 &= \varphi(c'_1,...,c'_k)-\varphi(c''_1,...,c''_k) \\ &= (c'_1v_1 + ... + c'_kv_k) - (c''_1v_1 + ... + c''_kv_k) \\ &= (c'_1-c''_1)v_1 + ... + (c'_k- c''_k)v_k \end{aligned}$$

Da die Vektoren $v_1,...,v_k$ nach Voraussetzung linear unabhängig sind, muß die Linearkombination rechts trivial sind, d.h. es gilt

$$\begin{aligned} 0 &= c'_1-c''_1 \\ &\dots \\ 0 &= c'_k- c''_k \end{aligned}$$

Das bedeutet aber, die beiden k-Tupel sind gleich,

$$(c'_1,...,c'_k) = (c''_1,...,c''_k),$$

im Widerspruch zu unserer Annahme.

<u>(iii) ⇒ (ii)</u>. Angenommen, eines der v_i ist Linearkombination der anderen,

$$v_i = c_1v_1 + ... + c_{i-1}v_{i-1} + c_{i+1}v_{i+1} + ... +c_kv_k.$$

Dann gilt

$$0 = c_1v_1 + ... + c_{i-1}v_{i-1}+(-1)v_i+ c_{i+1}v_{i+1} + ... +c_kv_k$$

$$= \varphi(c_1, \dots, c_{i-1}, -1, c_{i+1}, \dots, c_k)$$

Mit anderen Worten, das k-Tupel $(c_1, \dots, c_{i-1}, -1, c_{i+1}, \dots, c_k)$ hat bei φ dasselbe Bild wie das k-Tupel $(0,\dots,0)$. Das steht aber im Widerspruch zur Injektivität von φ.

(ii) $\Rightarrow$ (i). Angenommen, $v_1,\dots,v_k$ sind nicht linear unabhängig. Dann gibt es eine nicht-triviale Linearkombination der v_i, welche Null ist,

$$c_1 v_1 + \dots + c_k v_k = 0.$$

O.B.d.A. sei $c_1 \neq 0$ (andernfalls ändern wir die Bezeichnungen geeignet). Dann gilt aber,

$$v_1 = \frac{-c_2}{c_1} v_2 + \dots + \frac{-c_k}{c_1} v_k,$$

d.h. v_1 ist Linearkombination der übrigen Vektoren. Das steht aber im Widerspruct zur Annahme (ii).

QED.

3.3.3 Charakterisierung der linearen Unabhängigkeit beliebiger Mengen

Seien K ein Körper, V ein K-Vektorraum und

$$\{v_i\}_{i\in I}$$

eine Familie von paarweise verschiedenen Vektoren aus V. Dann sind folgende Aussagen äquivalent.

(i) $(v_i)_{i\in I}$ ist K-linear unabhängig.

(ii) Keines der v_i ist K-Linearkombination der übrigen.

(iii) Die lineare Abbildung

$$\varphi: \oplus_{i\in I} K \to V,\ \{c_i\}_{i\in I} \mapsto \sum_{i\in I} c_i v_i,$$

injektiv.

Beweis. Übungsaufgabe.

QED.

3.3.4 Basen eines Vektorraumes

Seien K ein Körper und V ein K-Vektorraum. Eine Basis von V (über K) ist eine Familie

$$(v_i)_{i\in I}$$

von Elementen aus V mit folgenden Eigenschaften.

1. Die v_i sind K-linear unabhängig.
2. Die v_i erzeugen den Vektorraum V.

Bemerkungen

(i) Bedingung 1 besagt, nur die triviale Linearkombination der v_i ist Null, d.h. aus

$$c_{i_1} v_{i_1} + \dots + c_{i_k} v_{i_k} = 0 \text{ mit } c_{i_1},\dots,c_{i_k} \in K$$

folgt stets

$$c_{i_1} = ...= c_{i_k} = 0$$

(ii) Bedingung 2 besagt, der von den v_i erzeugte Unterraum ist gleich dem gesamten Raum V. Mit anderen Worten, diese Bedingung ist äquivalent zu der folgenden Aussage.

2'. Jeder Vektor von V ist K-Linearkombination gewisser endlich vieler v_i.

(iii) Die beiden Bedingungen der obigen Definition lassen sich auch zu der folgenden Bedingung zusammenfassen.

(1+2)'. Jeder Vektor von V läßt sich auf genau eine Weise als K-Linearkombination endlich vieler Vektoren v_i schreiben

Beweis von (iii). 1. Schritt. Bedingung (1+2)' ist notwendig.
Seien die Bedingungen 1 und 2 erfüllt. Wir haben zu zeigen, es gilt dann auch (1+2)'. Wegen 2 läßt sich jeder Vektor von V als K-Linearkombination der v_i schreiben. Wir haben nur noch zu zeigen, daß verschiedene Linearkombinationen unmöglich dasselbe Element darstellen. Angenommen $v \in V$ ließe sich auf zwei Weisen als Linearkobination schreiben,

$$v = c_{i_1} v_{i_1} + ... + c_{i_k} v_{i_k} = c'_{i_1} v_{i_1} + ... + c'_{i_{k'}} v_{i_{k'}},$$

mit $c_{i_1},...,c_{i_k}, c'_{i_1},...,c'_{i_{k'}} \in K$. Indem wir bei beiden Linearkombination weitere Summanden mit dem Koeffizienten hinzufügen, können wir annehmen, k = k'. Es gilt

$$0 = v - v = (c_{i_1} - c'_{i_1})v_{i_1} + ... + (c_{i_k} - c'_{i_k})\, v_{i_k}\ .$$

Wegen 1 muß dies die triviale Linearkombination sein, d.h.

$$c_{i_1} - c'_{i_1} = ... = c_{i_k} - c'_{i_k} = 0.$$

Je zwei Linearkombinationen, die v darstellen, sind also identisch.
2. Schritt. Bedinngung (1+2)' ist hinreichend.
Sei Bedingung (1+2)' erfüllt. Wir haben zu zeigen, daß dann auch 1 und 2 erfüllt sind. Nach Voraussetztung ist jeder Vektorraum von V Linearkombination der v_i. Deshalb erzeugen die v_i den gesamten Raum V, d.h. es gilt 2. Der Vektor $0 \in V$ läßt sich nach Vorraussetzung auf genau eine Weise als Linearkombination schreiben, d.h. nur die triviale Linearkombination ist Null. Deshalb sind die v_i linear unabhängig, d.h. es gitl 1.
QED.

3.3.5 Charakterisierung der endlichen Basen eines Vektorraumes

Seien K ein Körper, V ein K-Vektorraum und

$$v_1,...,v_n \in V$$

paarweise verschiedene Vektoren aus V. Dann sind folgende Aussagen äquivalent.

(i) $v_1,...,v_n$ ist eine Basis von V über K.

(ii) Die Vektoren $v_1,...,v_n$ bilden eine maximale linear unabhängige Familie, d.h. für beliebiges $v \in V$ sind die Vektoren $v_1,...,v_n, v$ nicht mehr linear unabhängig.

(iii) Die Vektoren $v_1,...,v_n$ bilden ein minimales Erzeugendensystem, d.h. durch Weglassen eines beliebigen Vektors geht die Eigenschaft, den Raum V zu erzeugen, verloren.

(iv) Die lineare Abbildung

$$\varphi: K^n \to V, (c_1, \dots, c_n) \mapsto c_1v_1 + \dots + c_nv_n,$$

ist ein Isomorphismus.

Sind diese Bedingungen erfüllt, so nennt man die Abbilung φ von Bedingung (iv) auch <u>Koordinatenabbildung</u> zur Basis $v_1,\dots,v_n$.

Beweis. <u>(i) ⇒ (iv)</u>. Sei $v_1,\dots,v_n$ eine Basis. Dann ist jedes Element von V Linearkombination der $v_1,\dots,v_n$, d.h. die Abbildung ist surjektiv. Auf Grund unserer Charakterisierung der linearen Unabhängigkeit (vgl. 4.3.3) ist die Abbildung auch injektiv.

<u>(iv) ⇒ (iii)</u>. Auf Grund der Surjektivität der Abbildung φ bilden die Vektoren $v_1,\dots,v_n$ ein Erzeugendensystem von V. Angenommen, der Raum V wird auch von den Vektoren $v_1,\dots,v_{i-1},v_{i+1},\dots,v_n$ erzeugt. Dann ist v_i eine Linearkombination dieser Vektoren, sagen wir

$$v_i = c_1v_1+\dots+c_{i-1}v_{i-1}+c_{i+1}v_{i+1}+\dots+c_nv_n.$$

Dann ist aber $(c_1,\dots,c_{i-1}, -1, c_{i+1},\dots,c_n)\in K^n$ ein von Null verschiedener Vektor mit demselben Bild wir der Nullvektor:

$$\begin{aligned}\varphi(c_1,\dots,c_{i-1}, -1, c_{i+1},\dots,c_n) &= c_1v_1+\dots+c_{i-1}v_{i-1}+(-1)v_i+c_{i+1}v_{i+1}+\dots+c_nv_n\\ &= v_i - v_i\\ &= 0\\ &= \varphi(0,\dots,0).\end{aligned}$$

Dies steht aber im Widerspruch zur Injektivität von φ, d.h. die Vektoren $v_1,\dots,v_{i-1},v_{i+1},\dots,v_n$ erzeugen nicht den ganzen Raum V.

<u>(iii) ⇒ (ii)</u>. Sei $v_1,\dots,v_n$ ein minimales Erzeugendensystem. Zeigen wir zunächst, diese Vektoren sind linear unabhängig. Wären sie es nicht, so könnte man einen dieser Vektoren, sagen wir v_1, als Linearkombination der übrigen schreiben:

(1) $$v_1 = c_2v_2 + \dots + c_nv_n.$$

Sei jetzt $v\in V$ ein beliebiger Vektor. Da die $v_1,\dots,v_n$ den Raum V erzeugen, gilt

$$v = c'_1v_1+\dots+c'_nv_n$$

mit gewissen $c'_1,\dots,c'_n\in K$. Durch Einsetzen von (1) erhalten wir

$$\begin{aligned}v &= c'_1(c_2v_2 + \dots + c_nv_n)+c'_2v_2\dots+c'_nv_n\\ &= (c'_1c_2+c'_2)v_2 + (c'_1c_3+c'_3)v_3+ \dots + (c'_1c_n+c'_n)v_n.\end{aligned}$$

Wir haben gezeigt, jeder Vektor $v\in V$ läßt sich als Linearkombination der Vektoren $v_2, \dots, v_n$ schreiben. Das stehe aber im Widerspruch zu unserer Voraussetzung, daß die Vektoren $v_1,\dots,v_n$ ein minimales Erzeugendensystem bilden sollen. Also ist die Annahme, daß sie linear abhängig sind, falsch. Wir haben damit gezeigt,

(2) $v_1,\dots,v_n$ sind linear unabhängig.

Wir haben noch zu zeigen, durch Hinzufügen eines beliebigen Vektors $v\in V$ geht die Eigenschaft, linear unabhängig zu sein, verloren. Da die Vektoren $v_1,\dots,v_n$ ein

Erzeugendensystem bilden, läßt sich der Vektor v als Linearkombination der $v_1,\ldots,v_n$ schreiben. Dann sind die Vektoren

$$v_1,\ldots,v_n, v$$

aber linear abhängig.

(ii) ⇒ (i). Sei jetzt $v_1,\ldots,v_n$ eine maximale linear unabhängige Familie. Wir haben zu zeigen, die Vektoren $v_1,\ldots,v_n$ erzeugen den ganzen Raum V. Sei $v\in V$ beliebig. Dann sind die Vektoren

$$v_1,\ldots,v_n, v$$

nicht mehr linear unabhängig, d.h. eine nicht-triviale Linearkombination dieser Vektoren ist Null,

(3) $$c_1v_1+\ldots+c_nv_n+ cv = 0.$$

Dabei kann der Koeffizienten c nicht gleich Null sein, denn andernfalls wären die Vektoren $v_1,\ldots,v_n$ linear abhängig. Wir können die Identität (3) mit $\frac{1}{c}$ multiplizien und somit ohne Beschränkung der Allgemeinheit annehmen, daß

$$c = 1$$

gilt. Dann gilt aber

$$v =(- c_1)v_1+\ldots+(-c_n)v_n.$$

Wir haben gezeigt, jeder Vektor $v\in V$ läßt sich als Linearkombination der Vektoren v_1 $,\ldots,v_n$ schreiben. Diese Vektoren $v_1,\ldots,v_n$ erzeugen also den Raum V.

QED.

3.3.6 Charakterisierung beliebiger Basen

Seien K ein Körper, V ein K-Vektorraum und

$$\{v_i\}_{i\in I}$$

eine Familie von paarweise verschiedenen Vektoren aus V. Dann sind folgende Aussagen äquivalent.

(i) $\{v_i\}_{i\in I}$ ist eine Basis von V über K.

(ii) $\{v_i\}_{i\in I}$ ist eine maximale Familie von K-linear unabhängigen Vektoren von V.

(iii) $\{v_i\}_{i\in I}$ ist ein minimales Erzeugendensystem von V.

(ii) Die lineare Abbildung

$$\varphi:\oplus_{i\in I}K \to V, \{c_i\}_{i\in I} \mapsto \sum_{i\in I} c_iv_i,$$

ist ein Isomorphismus.

Den Beweis überlassen wir dem Leser.

3.3.7 Die Existenz von Basen

3.3.7.1 Vorbemerkung: Auswahlaxiom, Zornsches Lemma, Wohlordnungssatz

Das Ziel diese Abschnitts ist der Beweis der Aussage, daß jeder Vektorraum eine (im allgemeinen unendliche) Basis besitzt. Dieser Satz ist eine Konsequenz des sogenannten Auswahlaxioms der Mengenlehre.

Auswahlaxiom

Sei $\mathbf{M} := \{M_i \mid i\in I\}$ eine Familie von nicht-leeren Mengen. Dann kann man aus jeder der Mengen von $\mathbf{M}$ ein Element auswählen, d.h. es gibt eine Abbildung

$$\varphi: \mathbf{M}\to\cup\mathbf{M} := \cup_{i\in I} M_i$$

mit der Eigenschaft, daß für jedes $M\in\mathbf{M}$ das Bild $\varphi(M)$ ein Element der Menge M ist,

$$\varphi(M)\in M \text{ für jedes } M\in\mathbf{M}.$$

Bemerkungen

(i) Man lasse sich nicht durch die Plausibilität der Formulierung des Auswahlaxioms täuschen. Es kann nicht aus den übrigen Axiomen der Mengenlehre abgeleitet (also nicht "bewiesen") werden.

(ii) Das Auswahlaxiom ist äquivalent zu einer Aussage, welche Zornsches Lemma heißt. Wir werden im folgenden die Aussage, daß jeder Vektorraum eine Basis besitzt aus dem Zornschen Lemma ableiten.

(iii) Der Beweis der Aussage, daß das Zornsche Lemma äquivalent zum Auswahlaxiom ist, ist nicht sehr schwer, sprengt aber etwas den Rahmen dieser Vorleseung. Wir werden deshalb den Beweis nicht angeben. Man kann ihn zum Beispiel in dem Buch von Kurosch finden.

Kuroš: Vorlesungen über allgemeine Algebra, S. 13-15

Es werden dort gleichzeitig mehrere zum Auswahlaxiom äquivalente Aussagen (darunter der Wohlordnungssatz) behandelt.

(iv) Um das Zornsche Lemma formulieren zu können, müssen wir den Begriff der Halbordnung und einige damit zusammenhängende Begriffe einführen.

3.3.7.2 Halbgeordnete Mengen

Sei M eine Menge. Eine Relation auf M ist eine beliebige Teilmenge

$$R \subseteq M\times M$$

der Produktmenge von M mit sich selbst. Falls $(x,y)\in R$ gilt, so sagt man, das Element x steht zum Element y in der Relation R und schreibt

$$xRy.$$

Die Relation R heißt reflexiv, wenn gilt

$$xRx$$

für jedes $x\in M$. Sie heißt symmetrisch, wenn die folgende Implikation besteht.

$$xRy \Rightarrow yRx.$$

Sie heißt antisymmetrisch, wenn die folgende Implikation besteht.

$$xRy \text{ und } yRx \Rightarrow x=y.$$

Sie heißt transitiv, wenn die folgende Implikation besteht.

$$xRy \text{ und } yRz \Rightarrow xRz.$$

Zwei Elemente $x,y\in M$ heißen vergleichbar bezüglich R, wenn

$$xRy \text{ oder } yRx$$

gilt.

Eine Äquivalenzrelation auf M ist eine reflexive, symmetrische und transitive Relation auf M. Eine Halbordnung auf M ist eine reflexive, antisymmetrische und transitive Relation auf M. Eine Halbordnung R auf M heißt linear, wenn je zwei Elemente von M vergleichbar bezüglich R sind. Die Menge M heißt in diesem Fall (bezüglich R) linear geordnet oder auch Kette. Eine halbgeordnete Menge ist eine Menge M zusammen mit einer Halbordnung auf M.

Beispiele

1. Die Relation

$$R_1 := \{(x,x) \mid x \in M\}$$

heißt <u>Gleichheit</u> auf M und wird mit = bezeichnet, d.h.

$$x = y$$

bedeutet, die Elemente x und y sind gleich. Diese Relation ist eine Äquivalentzrelation, eine Halbordnung und ist nicht linear (es sei denn M enthält höchstens ein Element).

2. Bezeichne $\mathbb{R}_+$ die Menge der positiven reellen Zahlen. Die Relation

$$R_2 := \{(x,y) \in \mathbb{R} \times \mathbb{R} \mid x\text{-}y \in \mathbb{R}_+\}$$

heißt <u>Größerrelation</u> auf $\mathbb{R}$ und wird mit > bezeichnet, d.h.

$$x > y$$

bedeutet, die reelle Zahl x ist größer als die reelle Zahl y. Diese Relation ist nicht reflexiv, nicht symmetrisch, nicht anti-symmetrisch, aber sie ist transitiv.

3. Die Relation

$$R_3 := \{(x,y) \in \mathbb{R} \times \mathbb{R} \mid y\text{-}x \in \mathbb{R}_+ \cup \{0\}\}$$

heißt <u>Kleinergleichrelation</u> auf $\mathbb{R}$ und wird mit ≤ bezeichnet, d.h.

$$x \le y$$

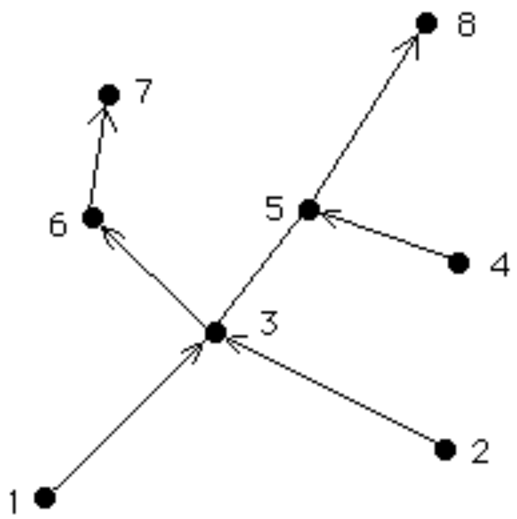
bedeutet, die reelle Zahl x ist kleiner oder gleich der reellen Zahl y. Diese Relation ist reflexiv, anti-symmetrisch, nicht symmetrisch, aber sie ist transitiv. Es handelt sich also um keine Äquivalenzrelation aber um eine Halbordnung, und zwar sogar um eine lineare Halbordnung.

4. Die Relation

$$R_4 := \{(x,y) \in \mathbb{R} \times \mathbb{R} \mid |y\text{-}x| = 1\}$$

ist nicht reflexiv, nicht anti-symmetrisch, nicht transitiv, aber symmetrisch.

5. Sei M die Menge der Ecken des folgenden “gerichteten Graphen”.



Für zwei Elemente $x,y \in M$ gelte genau dann

$$x \le y,$$

wenn x=y gilt oder man von x nach y gelangen kann, indem man eine Folge von Kanten in Pfeilrichtung durchläuft. Die so definierte Relation auf M ist eine Halbordnung, keine Äquivalenzrelation und sie ist nicht linear.

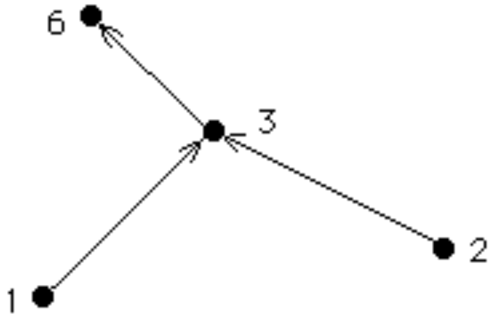
3.3.7.3 Obere Schranken und maximale Elemente

Seien M eine Menge, ≤ eine Halbordnung auf M und $S \subseteq M$ eine Teilmenge von M. Ein Element $m \in M$ heißt <u>obere Schranke</u> von S in M, wenn

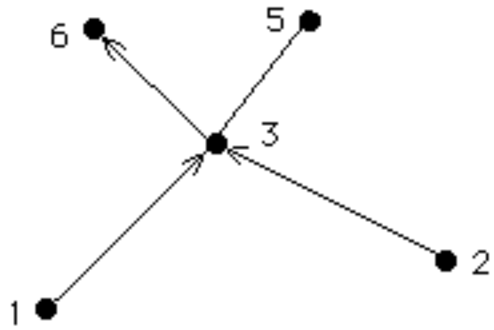
$$x \le m$$

gilt für jedes $x \in S$. Ein Element $m \in M$ heißt <u>maximal</u>, wenn für jedes $x \in M$ mit $m \le x$ automatisch sogar $m = x$ gilt.

Beispiel
Im Beispiel 5 von 4.3.7.2 sind die Elemente 7 und 8 die maximalen Elemente der Menge M. Die Teilmenge S={1,2,3,6} der Ecken des Teilgraphen



hat die Elemente 6 und 7 als obere Schranken. Die Teilmenge S={1,2,3,5,6} der Ecken von



besitzt keine obere Schranke.

3.3.7.4 Zornsches Lemma

Sei eine halbgeordnete Menge mit der Eigenschaft, daß jede linear geordnete Teilmenge von M eine obere Schranke in M besitzt. Dann gibt es in M ein maximales Element.

Zur Ableitung aus verschiedenen Axiomen der Mengenlehre siehe

A. G. Kuroš: Allgemeine Algebra, §6, Satz von Kuratowski-Zorn.

3.3.7.5 Die Exisitenz Vektorraumbasen

Seien K ein Körper und V ein K-Vektorraum. Dann gibt es eine Familie $\{v_i\}_{i\in I}$ von Elementen aus V, welche eine Basis von V bilden.

Beweis. Wir werden diesen Satz aus dem Zornschen Lemma ableiten. Es genügt zu zeigen, in V gibt es eine maximale linear unabhängige Menge von Vektoren. Bezeichne

$$M := \{S \mid S \text{ ist K-linear unabhängige Teilmenge von V}\}$$

die Menge aller linear unabhängigen Teilmengen des Vektorraumes V. Für je zwei Elemente

$$S', S''\in M$$

schreiben wir

$$S' \le S''$$

wenn $S'\subseteq S''$ gilt. Auf diese Weise ist auf der Menge M eine Halbordnung ≤ definiert. Ein maximales Element bezüglich dieser Halbordnung ist nach unserer Charakterisierung der Basen (vgl.3.3.6) gerade eine Basis von V. Es reicht also, wenn wir zeigen, die Menge M besitzt ein maximales Element. Auf Grund des Zornschen Lemmas genügt es zu zeigen, jede linear geordnete Teilmenge besitzt in M eine obere Schranke. Sei also eine linear geordnete Teilmenge

$$\mathcal{S} := \{S_i \mid i\in I\}$$

von M. Die Eigenschaft, linear geordnet zu sein, bedeutet gerade, daß je zwei (linear unabhängige) Mengen $S_{i_1}, S_{i_2} \in \mathcal{S}$ ineinander enthalten sind, d.h. es gilt

$$S_{i_1} \subseteq S_{i_2} \text{ oder } S_{i_1} \supseteq S_{i_2}.$$

Wir setzen

$$S := \cup_{i\in I} S_i.$$

Nach Konstruktion enthält S jedes Element von S_i von $\mathcal{S}$. Wenn wir also zeigen können, daß S ein Element von M ist, so ist S eine obere Schranke von $\mathcal{S}$ und der Beweis der Behauptung ist abgeschlossen.

Es genügt also zu zeigen, $S\in M$, d.h. S besteht aus linear unabhängigen Vektoren. Seien

$$v_1,\dots,v_n\in S$$

endlich viele vorgegebene paarweise verschiedene Vektoren aus S. Es genügt zu zeigen, diese Vektoren sind linear unabhängig. Nach Konstruktion von S gibt es für jedes v_j eine Menge $S_{i_j} \in \mathcal{S}$ mit

$$v_j\in S_{i_j} \quad (j=1,\dots,n).$$

Für je zwei Vektoren $v_{j'}$ und $v_{j''}$ gilt

$$S_{i_{j'}} \subseteq S_{i_{j''}} \text{ oder } S_{i_{j'}} \subseteq S_{i_{j''}}.$$

Indem wir die Reihenfolge der Vektoren (und ihre Bezeichnung) geeignet abhändern, können wir erreichen, daß für j'<j" stets die erste Inklusion besteht. Wir können also ohne Beschränkung der Allgemeinheit annehmen, es gilt

$$S_{i_1} \subseteq S_{i_2} \subseteq \dots \subseteq S_{i_n}.$$

Dann gilt aber

$$v_1,\dots,v_n\in S_{i_n} \quad (\in\mathcal{S}\subseteq M).$$

Als Element von M besteht die Menge S aus linear unabhängigen Vektoren. Insbesondere sind die Vektoren $v_1,\dots,v_n$ also linear unabhängig.

QED.

3.3.8 Die Dimension eines Vektorraums

Sei V ein K-Vektorraum. Falls V eine Basis aus endlich vielen Vektoren besitzt, so heißt V endlich-dimensional und die Anzahl der Basisvektoren heißt Dimension von V und wird mit

$$\dim V = \dim_K V$$

bezeichnet.

Bemerkungen

(i) Die Dimension eines Vektorraumes läßt sich auch im nicht-endlich-dimesionalen Fall definieren als Anzahl der Basiselemente. In diesem Fall ist die Dimension eine unendliche Kardinalzahl. Da wir hier die Bekanntschaft mit den letzteren nicht voraussetzen wollen, werden wir im folgenden diesen Fall etwas vernachlässigen und alle wesentlichen Betrachtungen an endlich-dimensionalen Vektorräumen vornehmen. Im unendlich-dimensionalen Fall schreiben wir

$$\dim V = \dim_K V = \infty.$$

Letztere Identität soll also nur bedeuten, daß V keine Basis aus endlich vielen Elementen besitzt.

(ii) Die obige Definition der Dimension ist nur sinnvoll, wenn jede Basis eines Vektorraums aus derselben Anzahl von Elementen besteht. Dies ist eine Folgerung des nachfolgenden Satzes.

3.3.9 Satz von Steinitz

Seien V ein K-Vektorraum, $v_1,\dots,v_n$ ein Erzeugendensystem von V und $w_1,\dots,w_m$ beliebige linear unabhängige Vektoren von V. Dann gibt es n-m Vektoren

$$v_{i_1},\dots,v_{i_{n-m}}$$

des gegebenen Erzeugendensystems derart, daß

$$v_{i_1},\dots,v_{i_{n-m}}, w_1,\dots,w_m$$

ein Erzeugendensystem von V ist. Insbesondere gilt $n \ge m$.

Bemerkungen

(i) Der Satz sagt aus, in einem Erzeugendensystem kann man gewisse Vektoren durch die Vektoren einer linear unabhängigen Menge so ersetzen, daß die Eigenschaft, Erzeugendensystem zu sein erhalten bleibt. Man spricht deshalb auch oft vom Austauschsatz von Steinitz.

(ii) Eine analoge Variante des Austauchsatzes für unendliche Mengen $\{v_i\}$ und $\{w_j\}$ ist ebenfalls gültig und kann mit Hilfe des Zornschen Lemmas bewiesen werden.

Beweis des Satzes von Steinitz durch Induktion nach m.

1. Schritt. Der Fall m = 1.

Weil $v_1,\dots,v_n$ ein Erzeugendensystem ist, läßt sich der Vektor $w = w_1$ als Linearkombination der v_i schreiben,

$$w = c_1 v_1 + \dots + c_n v_n.$$

Nach Voraussetzung ist $w \ne 0$, also muß auch einer der Koeffizienten c_i von Null verschieden sein,

$$c_i \ne 0.$$

Dann gilt

$$v_i = (-\frac{c_1}{c_i})v_1 + \dots + (-\frac{c_{i-1}}{c_i})v_{i-1} + \frac{1}{c_i}w + (-\frac{c_{i+1}}{c_i})v_{i+1} + \dots + (-\frac{c_n}{c_i})v_n.$$

Mit anderen Worten, v_i liegt in dem von w und den $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ erzeugten Unterraum. Es gilt damit

$$v_1, v_2, \dots, v_n \in \langle w, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\rangle,$$

also

$$(V =)\langle v_1, v_2, \dots, v_n\rangle \subseteq \langle w, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\rangle \ (\subseteq V),$$

also

$$V = \langle w, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\rangle$$

Mit anderen Worten, die Vektoren $w, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ bilden ein Erzeugendensystem von V.

Bemerkung

Die obige Argumentation gibt uns einen Hinweis darauf, welchen Vektor des Erzeugendensystems wir gegen den neuen Vektor w austauschen können: man schreibe w als Linearkombination des Erzeugendensystems

$$w = c_1v_1 + \dots + c_nv_n.$$

Dann kann man jedes v_i gegen w eintauschen, welches in der Linearkombination mit einem von Null verschiedenen Koeffizienten auftritt.

2. Schritt. Der Induktionsschritt.

Nach Induktionsvorausetzung gibt es Vektoren

(1) $$v_{i_1}, \dots, v_{i_{n-m+1}}$$

derart, daß

(2) $$v_{i_1}, \dots, v_{i_{n-m+1}}, w_1, \dots, w_{m-1}$$

ein Erzeugendensystem von V bilden. Wir haben zu zeigen, einer der ersten n-m+1 Vektoren des Systems (2) läßt sich gegen den Vektor w_m austauschen. Um den Austauch vorzunehme, schreiben wir w_m als Linearkombination der Vektoren des Erzeugengensystems:

$$w_m = c_1v_{i_1} + \dots + c_{n-m+1}v_{i_{n-m+1}} + c'_1w_1 + \dots + c'_{m-1}w_{m-1}.$$

Dabei können die Koeffizienten $c_1, \dots, c_{n-m+1}$ nicht sämtlich Null sein, denn andernfalls wären die Vektoren $w_1, \dots, w_m$ linear abhängig. Nach der Bemerkung am Ende des ersten Schritts läßt sich also w_m gegen einen der Vektoren $v_{i_1}, \dots, v_{i_{n-m+1}}$ austauschen.

QED.

3.3.10 Unabhängigkeit der Dimension von der Wahl der Basis

Sei V ein K-Vektorraum. Dann bestehen je zwei Basen aus derselben Anzahl von Vektoren.

Beweis. Seien $\{v_i \mid i\in I\}$ und $\{w_j \mid j\in J\}$ zwei Basen von V.

1. Schritt. Die Mengen $\{v_i \mid i\in I\}$ und $\{w_j \mid j\in J\}$ sind entweder beide endlich oder beide unendlich.

Angenommen die eine Menge, sagen wir

(1) $$\{v_i \mid i\in I\} = \{v_{i_1}, \dots, v_{i_n}\},$$

wäre endlich und die andere unendlich. Dann können wir aus der unendlichen Menge n+1 Vektoren auswählen.

(2) $$w_{j_1}, \dots, w_{j_{n+1}}.$$

Da die Vektoren von (1) ein Erzeugendensystem bilden und die Vektoren (2) linear unabhängig sind, folgt nach dem Satz von Steinitz, daß

$$n+1 \le n$$

gilt, was offensichtlich nicht möglich ist. Also kann nicht eine der Basis endlich und die andere unendlich sein.

<u>2. Schritt</u>. Gleichheit der Elementzahl im endliche Fall.
Seien also beide Basen endlich,

$$\{v_i \mid i\in I\} = \{v_{i_1},\ldots,v_{i_n}\},$$

$$\{\{w_j \mid j\in J\} = \{w_{j_1},\ldots,w_{j_m}\}.$$

Da beide Basen insbesondere Erzeugendensysteme sind und aus linear unabhängigen Vektoren bestehen, gilt nach dem Satz von Steinitz sowohl $n\le m$ also auch $m\le n$.

QED.

Bemerkung

Bijektive lineare Abbildungen $f:V\to V'$ überführen Basen in Basen. Isomorphe Vektorräume haben deshalb dieselbe Dimension.

3.3.11 Existenz von linearen Abbildungen mit vorgegebenen Werten auf einer Basis

Seien V und W zwei K-Vektorräume,

$$\{v_i\}_{i\in I}$$

eine Basis von V und

$$\{w_i\}_{i\in I}$$

eine beliebige Familie von Vektoren aus W. Dann gibt es genau eine K-lineare Abbildung

$$f: V \to W \text{ mit } f(v_i) = w_i.$$

Beweis. Die Eindeutigkeitsaussage folgt aus 3.2.7, denn Basen sind insbesondere Erzeugendensysteme. Sei

$v \in V$

ein vorgegebener Vektor. Weil $\{v_i\}_{i\in I}$ eine Basis ist, gibt es eindeutig bestimmte Elemente $c_i\in K$, die fast alle gleich Null sind, mit

$$v = \sum_{i\in I} c_i v_i.$$

Wir setzten

$$f(v) := \sum_{i\in I} c_i w_i .$$

Auf diese Weise ist eine Abbildung

$$f: V \to W$$

definiert mit $f(v_i) = w_i$ für jedes $i\in I$. Wir haben noch zu zeigen, diese Abbildung ist linear.

Nach Konstruktion gilt für jedes $\lambda\in K$

$$f(\lambda v) = \sum_{i\in I} c_i \lambda w_i = \lambda f(v).$$

Ist

$$v' = \sum_{i\in I} d_i v_i$$

ein zweiter Vektor aus V, so gilt

$$v + v' = \sum_{i\in I} (c_i+d_i)v_i$$

also

$$f(v+v') = \sum_{i\in I} (c_i+d_i)w_i = f(v) + f(v').$$

Die Abbildung f: V → W ist somit linear.

QED.

3.3.12 Die Dimension von Kern und Bild einer linearen Abbildung

Sei f: V → V' eine K-lineare Abbildung. Weiter seien Familien

$$\{v_i\}_{i\in I} \text{ und } \{w_j\}_{j\in J}$$

von Elementen aus V bzw. ker(f) gegeben. Es gelte

1. Die w_j bilden eine Basis von ker(f).
2. Die $f(v_i)$ bilden eine Basis von im(f).

Dann bilden die Vektoren v_i und w_j zusammen eine Basis von V. Insbesondere gilt

$$\dim \ker(f) + \dim \operatorname{im}(f) = \dim V.$$

Dabei sei die Summe aus ∞ und einer beliebigen nicht-negativen ganzen Zahl gleich ∞.

Beweis. haben zu zeigen, die Vektoren v_i und w_j sind linear unabhängig und erzeugen den Raum V.

<u>1. Schritt</u>. Die v_i und w_j erzeugen den Raum V.

Sei v∈V ein beliebiges Element. Wir haben zu zeigen, v läßt sich als (endliche) Linearkombination der v_i und w_j schreiben. Da die $f(v_i)$ ein Basis von im(f) bilden, läßt sich f(v) als Linearkombination von endlich vielen der $f(v_i)$ schreiben,

$$f(v) = c_1\cdot f(v_{i_1}) + ... + c_r\cdot f(v_{i_r})$$

mit geeigneten $c_1,...,c_r\in K$. Es gilt

$$f(v - c_1\cdot v_{i_1} + ... + c_r\cdot v_{i_r}) = f(v) - c_1\cdot f(v_{i_1}) - ... - c_r\cdot f(v_{i_r}) = 0,$$

d.h. der Vektor

(1) $$w := v - c_1\cdot v_{i_1} - ... - c_r\cdot v_{i_r}$$

liegt im Kern von f. Da die w_j ein Basis von ker(f) bilden, kann man w als Linearkombination von endlich vielen der w_j schreiben,

(2) $$w = d_1 w_{j_1} + ... + d_s w_{j_s}$$

mit geeigneten $d_1,...,d_s\in K$. Wir setzen (2) in (1) ein und erhalten

$$v = c_1 \cdot v_{i_1} + \dots + c_r \cdot v_{i_r} + d_1 w_{j_1} + \dots + d_s w_{j_s}.$$

Wir haben gezeigt, jeder Vektor $v \in V$ ist Linearkombination der Vektoren v_i, w_j, d.h. diese Vektoren bilden ein Erzeugendensystem von V.

<u>2. Schritt</u>. Die v_i und w_j sind linear unabhängig.

Wir haben zu zeigen, ist eine Linearkombination der Vektoren v_i, w_j gleich Null, so handelt es sich um die triviale Linearkombination. Sei also

(3) $$c_1 \cdot v_{i_1} + \dots + c_r \cdot v_{i_r} + d_1 w_{j_1} + \dots + d_s w_{j_s} = 0$$

mit $c_1, \dots + c_r, d_1, \dots, d_s \in K$. Wir haben zu zeigen, sämtliche Koeffizienten sind Null. Wir wenden f aus die Identität (3) an und erhalten

$$\begin{aligned} 0 &= c_1 \cdot f(v_{i_1}) + \dots + c_r \cdot f(v_{i_r}) + d_1 f(w_{j_1}) + \dots + d_s f(w_{j_s}) \\ &= c_1 \cdot f(v_{i_1}) + \dots + c_r \cdot f(v_{i_r}). \end{aligned}$$

Man beachte, die letzten s Summanden der ersten Zeile sind Null, da die Vektoren w_j nach Voraussetzung im Kern von f liegen. Ebenfalls nach Vorausetzung sind die $f(v_i)$ linear unabhängig (da sie eine Basis von im(f) bilden). Der letzte Ausdruck kann also nur dann Null sein, wenn sämliche Koeffizienten Null sind,

(4) $$c_1 = \dots = c_r = 0.$$

Wir setzen (4) in die Ausgangsidentität (3) ein und erhalten

(5) $$d_1 w_{j_1} + \dots + d_s w_{j_s} = 0.$$

Nach Voraussetzung bilden die w_j eine Basis (von ker(f)) sind also linear unabhängig. Mit (5) gilt also

(6) $$d_1 = \dots = d_s = 0.$$

Wir haben gezeigt, mit (3) gilt (4) und (6), d.h. nur die triviale Linearkombination der Vektoren v_i, w_j ist Null. Die Vektoren sind also linear unabhängig.

QED.

3.3.13 Die Dimension eines Faktorraums

Seien V ein K-Vektorraum und $W \subseteq V$ ein K-linearer Unterraum. Dann gilt

$$\dim W + \dim V/W = \dim V.$$

Beweis. Wir betrachten die natürliche Abbildung

$$\rho: V \to V/W, v \mapsto v+W.$$

Nach dem Satz über die Dimension von Kern und Bild (3.3.11) gilt

$$\ker(\rho) + \operatorname{im}(\rho) = \dim W.$$

Es reicht also zu zeigen,

(1) $$\ker(\rho) = W$$

(2) $$\operatorname{im}(\rho) = V/W.$$

Beweis von (1). Es gilt

$$v \in \ker(\rho) \Leftrightarrow \rho(v) = 0 \Leftrightarrow v+W = 0+W \Leftrightarrow v = v-0 \in W.$$

Beweis von (2). Es gilt

$$\operatorname{im}(\rho) = \{\rho(v) \mid v \in V\} = \{v+W \mid v \in V\} = V/W.$$

QED.

3.3.14 Basis-Ergängzungssatz

Seien V ein K-Vektorraum und $W \subseteq V$ ein K-linearer Unterraum. Dann läßt sich jede Basis von W zu einer Basis von V ergänzen.

Insbesondere gilt

$$\dim W \le \dim V \text{ und } \dim V/W \le \dim V.$$

Beweis: Seien

$$\{w_j\}_{j\in J} \text{ eine Basis von } W$$

$$\{\bar{v}_i\}_{i\in I} \text{ eine Basis von } V/W.$$

Für jedes $i \in I$ wählen wir ein Urbild $v_i \in V$ des Vektors $\bar{v}_i$ bei der natürlichen Abbildung $V \longrightarrow V/W$. Das ist möglich, weil die natürliche Abbildung surjektiv ist. Nach 3.3.12 bilden dann die w_j zusammen mit den v_i eine Basis von V.

QED.

3.3.15 Die Dimension von Unterräumen und Faktorräumen

Seien V ein endlich-dimensionaler Vektorraum K-Vektorraum und $W \subseteq V$ ein K-linearer Unterraum. Dann sind folgende Aussagen äquivalent.

(i) $\dim W = \dim V$.

(ii) $\dim V/W = 0$.

(iii) $W = V$.

Beweis. Trivialerweise bestehen die Implikationen

$$\text{(iii)} \Rightarrow \text{(i) und (iii)} \Rightarrow \text{(ii)}.$$

Zum Beweis der umgekehrten Implikationen

$$\text{(i)} \Rightarrow \text{(iii) und (ii)} \Rightarrow \text{(iii)}$$

verwenden wir die Bezeichnungen wie im Beweis von 3.3.14. Weil V endlich-dimensional ist, sind die Familien

$$\{w_j\}_{j\in J}, \{\bar{v}_i\}_{i\in I} \text{ und } \{v_i\}_{i\in I}$$

endlich.

Bedingunen (i) und (ii) bedeuten dann gerade,

$$\# J = \# J + \# I \text{ bzw. } \# I = 0.$$

In beiden Fällen ist also $\# I = 0$, d.h. die Index-Menge I ist leer. Dann ist aber die Basis $\{w_j\}_{j\in J}$ von W auch eine Basis von V, d.h. es gilt

$$W = V.$$

QED.

3.3.16 Exakte Sequenzen

Eine Folge von K-linearen Abbildungen

$$\dots \to V_{i-1} \xrightarrow{f_{i-1}} V_i \xrightarrow{f_i} V_{i+1} \to \dots$$

heißt <u>Komplex</u>, wenn die Zusammensetzung von je zwei benachbarten Abbildungen Null ist,

$$f_i \circ f_{i-1} = 0 \text{ für jedes } i\in\mathbb{Z}.$$

Diese Bedingungen sind äquivalent zu den Inklusionen

$$\mathrm{im}(f_{i-1}) \subseteq \ker(f_i) \text{ für jedes } i\in\mathbb{Z}.$$

Der Komplex heißt exakt an der Stelle V_i, wenn anstelle der Inklusion sogar das Gleichheitszeichen gilt,

$$\mathrm{im}(f_{i-1}) = \ker(f_i).$$

Eine exakte Sequenz oder auch exakter Komplex ist ein Komplex, der an allen Stellen exakt ist. Ein komplex heißt beschränkt oder auch von endlicher Länge, wenn es nur endlich viele V_i gibt , die vom trivialen Vektorraum $\{0\}$ verschieden sind.

Beispiel

Seien V ein K-Vektorraum und $W \subseteq V$ ein K-linearer Unterraum. Wir bezeichnen mit

$$i: W \to W, w \mapsto w,$$

die natürliche Einbettung von W in V und mit

$$\rho: V \to V/W, v \mapsto v + W,$$

die natürliche Projektion auf den Faktorraum. Dann ist

$$0 \to W \xrightarrow{i} V \xrightarrow{\rho} V/W \to 0$$

eine exakte Sequenz (von den Null-Vektorräumen die links von W und rechts von V/W stehen lassen wir jeweils alle bis auf einen weg).

3.3.17 Beschränkte exakte Sequenzen endlich-dimensionaler Vektorräume

Sei

$$\ldots \to V_i \xrightarrow{f_i} V_{i+1} \xrightarrow{f_{i+1}} V_{i+2} \to \ldots$$

eine exakte Sequenz endlich-dimensionaler K-Vektorräume, die beschränkt ist (als Komplex). Dann gilt

$$\sum_{i\in\mathbb{Z}} (-1)^i \dim V_i = 0.$$

Beispiel

Auf Grund der exakten Sequenz des Beispiels von 3.3.13 ist

$$\dim W - \dim V + \dim V/W = 0.$$

Diese Formel kennen wir bereits von 3.3.12.

Beweis. Sei n die Anzahl der von Null verschiedenen V_i. Wir führen den Beweis durch Induktion nach n. Im Fall n = 0 sind alle Vektorräumen gleich $\{0\}$, also alle Dimensionen Null, und die Behauptung gilt trivialerweise. Sei jetzt n > 0. Ohne Beschränkung der Allgemeinheit können wir annehmen, der am meisten links stehende von Null verschiedene Vektorraum ist der Vektorraum V_1. Die exakte Sequenz hat also die Gestalt

(1) $$0 \to V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} V_3 \xrightarrow{f_3} \ldots$$

Wir betrachten anstelle des Komplexes (1) den folgenden Komplex.

(2) $$0 \to 0 \xrightarrow{\alpha} \mathrm{im}(f_2) \xrightarrow{i} V_3 \xrightarrow{f_3} \ldots$$

Dabei sei $i(v) = v$ für jedes $v\in \mathrm{im}(f_2)$. Dieser Komplex ist an allen Stellen mit eventueller Ausnahmen der Stellen $\mathrm{im}(f_2)$ und V_3 exakt, denn links von diesen Stellen ist er trivialerweise exakt (weil Null) und rechts davon stimmt er mit dem Komplex (1) überein, der nach Voraussetzung überall exakt ist. Zeigen wir, (2) ist auch an den verbleibenden beiden Stellen exakt.

Exaktheit an der Stelle $\mathrm{im}(f_2)$: Nach Definition von i ist

$$\ker(i) = 0 = \mathrm{im}(\alpha).$$

Exaktheit an der Stelle V_3: Es gilt

$$\mathrm{im}(i) = \{i(v) \mid v\in \mathrm{im}(f_2) \} = \{ v \mid v \in \mathrm{im}(f_2) \} = \mathrm{im}(f_2) = \ker(f_3).$$

Das letzte Gleichheitszeichen gilt, weil (1) überall exakt ist.

Wir haben damit gezeigt, (2) ist eine exakte Sequenz. Nach Konstruktion ist die Anzahl der von Null verschiedenen Vektorräume von (2) kleiner als n. Nach Induktionsvoraussetzung ist deshalb

(3) $$\dim \mathrm{im}(f_2) + \sum_{i=3}^{\infty} (-1)^i \dim V_i = 0.$$

Weiter ist

$$\begin{aligned}\dim \mathrm{im}(f_2) &= \dim V_2 - \dim \ker(f_2) && \text{(nach 3.3.11)}\\ &= \dim V_2 - \dim \mathrm{im}(f_1) && \text{(weil (1) exakt ist)}\\ &= \dim V_2 - (\dim V_1 - \dim \ker(f_1)) && \text{(nach 3.3.11)}\\ &= \dim V_2 - \dim V_1,\end{aligned}$$

denn $\ker(f_1) = \mathrm{im}\,(0 \to V_1) = 0$. Wir setzen das Ergebnis dieser Berechnung in (3) und erhalten die Behauptung:

$$\sum_{i=1}^{\infty} (-1)^i \dim V_i = 0.$$

QED.

3.3.18 Die Dimension einer direkten Summe

Seien V' und V" zwei K-Vektorräume. Dann gilt

$$\dim V'\oplus V'' = \dim V' + \dim V''.$$

Beweis. Wir betrachten die Abbildung

$$f: V'\oplus V'' \to V', (v',v'') \mapsto v'.$$

Diese Abbildung ist linear (wie wir früher im Zusammenhang mit der Einführung des direkten Produkts gesehen haben). Nach dem Satz über die Dimension von Kern und Bild folgt

$$\ker(f) + \mathrm{im}(f) = \dim V'\oplus V''.$$

Es reicht also zu zeigen,

(1) $\ker(f) = V''$

(2) $\mathrm{im}(f) = V'$

Beweis von (1). Man beachte im Zusammenhang mit (1), daß man V" als Unterraum der direkten Summe auffassen kann, in dem man den Raum mit der Menge

$$V'' = \{(0,v'') \mid v''\in V''\}$$

aller Paare mit verschwindender erster Koordinate identifiziert. Mit dieser Identifikation gilt

$$(v',v'')\in\ker(f) \Leftrightarrow f(v',v'') = 0 \Leftrightarrow v'=0 \Leftrightarrow (v',v'') \in V''.$$

Beweis von (2). Es gilt

$$\mathrm{im}(f) = \{f(v',v'') \mid v'\in V' \text{ und } v''\in V''\} = \{v' \mid v'\in V'\} = V'.$$

QED.

3.3.19 Dimension von Durchschnitt und Summe zweier Unterräume

Seien V ein endlich-dimensionaler K-Vektorraum und W', W" zwei K-lineare Unterräume von V. Dann gilt

$$\dim W' + \dim W'' = \dim W' \cap W'' + \dim W'+W''.$$

Dabei sollen $W' \cap W''$ und $W'+W''$ die folgenden linearen Unterräume von V.

$$W' \cap W'' = \{w \mid w \in W' \text{ und } w \in W''\}$$

$$W'+W'' = \{w'+w'' \mid w' \in W' \text{ und } w'' \in W''\}$$

Beweis. Es reicht eine exakte Sequenz der Gestalt

(1) $\quad 0 \to W' \cap W'' \xrightarrow{f} W' \otimes W'' \xrightarrow{g} W'+W'' \to 0$

zu finden, denn dann ist nach 3.3.14

$$\dim W' \oplus W'' = \dim W' \cap W'' + \dim W'+W''.$$

Wegen 3.3.15 ist das aber gerade die Behauptung. Wir setzen

$$f(w) = (w, -w) \text{ für } w \in W' \cap W''$$

und

$$g(w', w'') = w'+w'' \text{ für } w' \in W' \text{ und } w'' \in W''.$$

Dann sind f ung g lineare Abbildung und es gilt

$$\ker(f) = \{w \in W' \cap W'' \mid f(w) = 0\} = \{w \in W' \cap W'' \mid (w, -w) = (0,0)\} = 0,$$

d.h. (1) ist an der Stelle $W' \cap W''$ exakt. Weiter gilt

$$\begin{aligned}
\ker(g) &= \{(w', w'') \mid w' \in W', w'' \in W'', g(w', w'') = 0\} \\
&= \{(w', w'') \mid w' \in W', w'' \in W'', w'+ w'' = 0\} \\
&= \{(w, - w) \mid w \in W', w \in W''\} \\
&= \{(w, -w) \mid w \in W' \cap W''\} \\
&= \{f(w) \mid w \in W' \cap W''\} \\
&= \mathrm{im}(f),
\end{aligned}$$

d.h. (1) ist an der Stelle $W' \cap W''$ exakt. Weiter gilt

$$\begin{aligned}
\mathrm{im}(g) &= \{g(w',w'') \mid w' \in W', w'' \in W''\} \\
&= \{w'+w'' \mid w' \in W', w'' \in W''\} \\
&= W'+W'' \\
&= \ker(W'+W'' \to 0).
\end{aligned}$$

Damit ist die Exaktheit von (1) bewiesen.
QED

3.3.20 Dimension des dualen Vektorraums Hom(V,K)

Sei V ein K-Vektorraum. Dann hat der Raum der Linearformen von V, d.h. der Raum der K-linearen Abbildungen $V \to K$,

$$V^* := \mathrm{Hom}_K(V, K),$$

dieselbe Dimension wie V,

$$\dim V^* = \dim V.$$

Beweis. Wir wählen eine Basis

$$\{v_i\}_{i \in I}$$

des Vektorraumes V. Jedes Element $v \in V$ läßt sich dann auf genau eine Weise als Linearkombination der v_i schreiben,

$$v = \sum_{i \in I} f_i v_i \tag{1}$$

mit eindeutig bestimmten Koeffizienten

$$f_i = f_i(v) \in K.$$

Betrachten wir die Abbildungen

$$f_i: V \to K,\ v \mapsto f_i(v). \tag{2}$$

Nach Konstruktion sind die f_i: die eindeutig bestimmten Abbildungen mit

$$v = \sum_{i\in I} f_i(v) v_i \tag{3}$$

für jedes $v \in V$.

<u>1. Schritt</u>. Die Abbildungen (2) sind K-linear.

Für $v,v'\in V$ und $c,c'\in K$ gilt

$$\begin{aligned}\sum_{i\in I} f_i(cv+c'v')\cdot v_i &= cv+c'v' \\ &= c\cdot \sum_{i\in I} f_i(v)\cdot v_i + c'\cdot \sum_{i\in I} f_i(v')\cdot v_i \\ &= \sum_{i\in I} (cf_i(v) + c'f_i(v'))\cdot v_i\end{aligned}$$

Da die v_i eine Basis bilden, sind zwei Linearkombinationen genau dann gleich, wenn alle einander entsprechenden Koeffizienten gleich sind. Also gilt

$$f_i(cv+c'v') = cf_i(v) + c'f_i(v')$$

für beliebige $v,v'\in V$, beliebige $c,c'\in K$ und beliebige $i\in I$. Mit anderen Worten,die Abbildungen f_i sind K-linear,

$$f_i \in V^* = \operatorname{Hom}_K(V,K).$$

<u>2.Schritt</u>: Die $f_i \in V^*$ sind linear unabhängig. Insbesondere gilt im Fall $\dim V = \infty$ auch $\dim V^* = \infty$.

Wir haben zu zeigen, nur die triviale Linearkombination der Abbildungen f_i ist gleich der Nullabbildung. Sei also

$$\sum_{i\in I} c_i\cdot f_i = 0$$

die Nullabbildung für gewisse $c_i\in K$. Dann gilt für jedes $v\in V$,

$$0 = (\sum_{i\in I} c_i\cdot f_i)(v) = \sum_{i\in I} c_i\cdot f_i(v).$$

Speziell für $v = v_j$ erhalten wir

$$0 = \sum_{i\in I} c_i\cdot f_i(v_j). \tag{4}$$

Wegen (3) ist nun

$$f_i(v_j) = \begin{cases} 1 \text{ für } i=j \\ 0 \text{ sonst}\end{cases} \tag{5}$$

Auf der rechten Seite von (4) steht also höchsten ein von Null verschiedener Summand, d.h. es gilt

$$0 = c_j \cdot 1 = c_j$$

Da j beliebig gewählt war, ergibt sich, daß sämtliche Koeffizienten c_j Null sein müssen, d.h. f_i sind linear unabhängig.

<u>3. Schritt</u>: Abschluß des Beweises.

Wir haben noch zu zeigen, im Fall dim V < ∞ gilt

$$\dim V = \dim V^*.$$

Dazu reicht es zu zeigen, daß die im zweiten Schritt konstruierte Familie von Linearformen

$$f_i \in V^*$$

eine Basis von V* bilden. Im 2. Schritt haben wir bereits gezeigt, die f_i sind linear unabhängig. Es reicht also zu zeigen, sie bilden ein Erzeugendensystem von V*.

Sei also $f \in V^*$ ein beliebiges Element. Wir setzen

$$g := f - \sum_{i\in I} f(v_i) f_i \in V^* \tag{6}$$

Es reicht zu zeigen, daß g die Nullabbildung ist, denn dann ist

$$f = \sum_{i\in I} f(v_i) f_i$$

d.h. jede Linearform f auf V ist eine Linearkombination der f_i, d.h. die f_i erzeugen den Raum der Linearformen.

Wegen (5) gilt für jedes $j \in I$,

$$g(v_j) = f(v_j) - \sum_{i\in I} f(v_i) f_i(v_j) = f(v_j) - f(v_j)\cdot 1 = 0.$$

d.h. g ist auf den Elementen einer Basis von V gleich Null. Dann muß aber die lineare Abbildung g identisch Null sein,

$$g = 0.$$

(nach 3.2.9 oder 3.3.11).

QED.

Bemerkungen

(i) Die Aussage, daß die Linearformen f_i den Raum Hom(V,K) erzeugen, ist falsch im unendlich-dimensionalen Fall. Unser Beweis versagt in diesen Fall, weil dann die Summe auf der rechten Seite von (5) aus unendlich vielen von Null verschiedenen Summanden bestehen kann, also überhaupt nicht definiert ist.

(ii) Im Fall $\dim_K V < \infty$ haben wir gezeigt, daß es zu jeder Basis

$$\{v_i\}_{i\in I}$$

von V einen Basis $\{f_i\}_{i\in I}$ von V* gibt mit

$$f_j(v_i) = \delta_{ij} = \begin{cases} 1 \text{ für } i = j \\ 0 \text{ für } i \neq j \end{cases}$$

Diese Basis heißt die zur Basis der v_i <u>duale Basis</u>. Ihre Vektoren werden wir oft mit

$$v_j^* := f_j$$

bezeichnen.

3.3.21 Dimension von Hom(V,V')

Seien V und V' zwei K-Vektorräume. Dann hat der Raum der K-linearen Abbildungen $V \to V'$ die Dimension

$$\dim \mathrm{Hom}_K(V,V') = \dim V \cdot \dim V'.$$

Dabei sei das Produkt auf der rechten Seite Null, wenn einer der Vektorräume V oder V' die Dimension Null hat (selbst wenn die Dimension des anderen Raums unendlich ist).

Beweis. Ist einer der beiden Räume V, V' von der Dimension Null, also gleich dem Nullvektorraum

$$\{0\},$$

so ist die einzige lineare Abbildung $V \to V'$ die Nullabbildung, d.h. Hom(V,V') ist 0-dimensional. Seien jetzt die Dimensionen von V und V' beide ungleich Null.

Wir wählen eine Basis

$$\{v'_i\}_{i\in I}$$

von V' und betrachten die Abbildung

$$(1) \qquad \varphi: \oplus_{i\in I}\mathrm{Hom}(V,K) \to \mathrm{Hom}_K(V,V'),\ \{f_i\}_{i\in I} \mapsto (v \mapsto \sum_{i\in I} f_i(v)v'_i),$$

d.h. nach Definition ist

$$\varphi(\{f_i\}_{i\in I})(v) = \sum_{i\in I} f_i(v)v'_i\ .$$

<u>1. Schritt</u>. Die Abbildung φ ist K-linear.

Für beliebige Familien

$$\{f'_i\}_{i\in I} \text{ und } \{f''_i\}_{i\in I}$$

aus der direkten Summe auf der linken Seite von (1), beliebige $c', c'' \in K$ und beliebige $v\in V$ gilt

$$\begin{aligned}
\varphi(c'\cdot\{f'_i\}_{i\in I} + c''\cdot\{f''_i\}_{i\in I})(v) &= \varphi(\{c'\cdot f_i + c''\cdot f''_i\}_{i\in I})(v) \\
&= \sum_{i\in I} (c'\cdot f_i + c''\cdot f''_i)(v)v'_i \\
&= \sum_{i\in I} (c'\cdot f_i(v) + c''\cdot f''_i(v))v'_i \\
&= c' \sum_{i\in I} f_i(v)v'_i + c''\cdot \sum_{i\in I} f''_i(v)v'_i \\
&= c'\cdot\varphi(\{f'_i\}_{i\in I}) + c''\cdot\varphi(\{f''_i\}_{i\in I}).
\end{aligned}$$

<u>2. Schritt</u>. $\ker(\varphi) = 0$.

Sei

$$\varphi(\{f_i\}_{i\in I}) = 0$$

die Nullabbildung, d.h. für jedes $v\in V$ gelte

$$0 = \varphi(\{f_i\}_{i\in I})(v) = \sum_{i\in I} f_i(v)v'_i .$$

Da $\{v'_i\}_{i\in I}$ eine Basis des K-Vektorraumes V' ist, folgt damit

$f_i(v) = 0$

für alle $v\in V$, d.h.die Abbildungen f_i sind Null. Dann ist aber $\{f_i\}_{i\in I}$ Nullvektor der direkten Summe auf der linken Seite von (1).

<u>3. Schritt</u>. Eine lineare Abbildung f ist genau dann injektiv, wenn ihr Kern der Nullraum 0 ist.

Wenn die Abbildung injektiv ist, kann nur der Nullvektor in den Nullvektor abgebildet werden, d.h. es gilt ker(f) = 0. Sei jetzt umgekehrt,

$$\ker(f) = 0$$

und seien v',v" zwei Vektoren aus dem Definitionsbereich von f mit f(v') = f(v"). Wir haben zu zeigen, dann gilt v' = v".

Es gilt

$$0 = f(v') - f(v'') = f(v'-v'')$$

also

$$v'-v''\in\ker(f) = 0,$$

als0 v'-v" = 0, also v' = v".

<u>4. Schritt</u>. Der Fall $\dim V' = \infty$

Auf Grund des 2. und 3. Schritts ist die Abbildung (1) injektiv, d.h. die direkte Summe

(2) $$\oplus_{i\in I}\mathrm{Hom}(V,K)$$

kann mit einem Unterraum von $\mathrm{Hom}_K(V,V')$ identifiziert werden. Dasselbe gilt, wenn man in (2) die unendliche Indexmenge I durch eine n-elementige Teilmenge I' ersetzt (n eine beliebige natürliche Zahl). Also gilt

$$\begin{aligned}\dim \mathrm{Hom}_K(V,V') &\ge \dim \oplus_{i\in I'}\mathrm{Hom}(V,K)\\ &= \dim \mathrm{Hom}(V,K) + ... + \dim \mathrm{Hom}(V,K) \text{ (n-mal)}\\ &= n\cdot\dim \mathrm{Hom}(V,K)\\ &= n\cdot \dim V \text{ (nach 4.3.14).}\end{aligned}$$

Da n beliebig war und dim V ungleich Null, muß

$$\dim \mathrm{Hom}(V,V'') = \infty = \dim V\cdot\dim V'$$

gelten.

<u>5. Schritt</u>. Der Fall $\dim V' < \infty$.

Nach Voraussetung ist die Menge I endlich. Es reicht zu zeigen, daß dann die Abbildung (1) ein Isomorphismus ist, denn dann gilt

$$\begin{aligned}\dim \mathrm{Hom}(V,V') &= \dim \oplus_{i\in I}\mathrm{Hom}(V,K)\\ &= (\#I)\cdot\dim \mathrm{Hom}(V,K)\\ &= \dim V'\cdot \dim \mathrm{Hom}(V,K)\\ &= \dim V'\cdot \dim V \qquad \text{(nach 4.3.14).}\end{aligned}$$

Da bereits gezeigt wurde, daß φ injektiv ist, reicht es zu zeigen, φ ist surjektiv. Sei

$$f\in\mathrm{Hom}(V,V')$$

beliebig. Unter Verwendung der oben eingeführten Basis von V' können wir für jedes Element $v\in V$ das Element f(v) als Linearkombination der v'_i schreiben,

$$f(v) = \sum_{i\in I} f_i \cdot v'_i$$

mit eindeutig bestimmten (von v abhängigen)

$$f_i = f_i(v) \in K.$$

Wir erhalten damit Abbildungen

$$f_i: V \to K$$

die durch die Bedingung

(3) $$f(v) = \sum_{i\in I} f_i(v)\cdot v'_i$$

eindeutig festgelegt sind. Es reicht zu zeigen, diese Abbildungen sind K-linear, denn dann gilt

$$\{f_i\}_{i\in I} \in \oplus_{i\in I} \mathrm{Hom}(V,K)$$

und

$$\varphi(\{f_i\}_{i\in I})(v) = \sum_{i\in I} f_i(v)\cdot v'_i = f(v)$$

für alle v, d.h.

$$\varphi(\{f_i\}_{i\in I}) = f.$$

Es ist noch die Linearität der f_i zu beweisen. Für $c',c''\in K$ und $v', v''\in V$ gilt

$$\begin{aligned}\sum_{i\in I} f_i(c'v'+c''v'')\cdot v'_i &= f(c'v'+c''v'')\\ &= c'\cdot f(v') + c''\cdot f''(v'')\\ &= c'\cdot \sum_{i\in I} f_i(v')\cdot v'_i + c''\cdot \sum_{i\in I} f_i(v'')\cdot v'_i\\ &= \sum_{i\in I} (c'\cdot f_i(v') + c''\cdot f_i(v''))\cdot v'_i\end{aligned}$$

Da die v'_i eine Basis von V' bilden, folgt für jedes i,

$$f_i(c'v'+c''v'') = c'\cdot f_i(v') + c''\cdot f_i(v''),$$

d.h. die f_i sind lineare Abbildungen.

QED.

3.4 Lineare Abbildungen

3.4.1 Die Matrix einer linearen Abbildung

Sei f: V→W eine K-lineare Abbildung von endlich-dimensionalen Vektorräumen. Weiter seien

$$v_1,\dots,v_m \in V$$

$$w_1,\dots,w_n \in W$$

Basen von V bzw. W. Dann läßt sich jeder der Vektoren $f(v_i)$ auf genau eine Weise als Linearkombination der Basisvektoren $w_1,...,w_n$ schreiben,

$$f(v_i) = c_{1i}w_1+...+c_{ni}w_n \qquad (i=1,...,m)$$

mit eindeutig bestimmten Koeffizienten $c_{ij}\in K$, i=1,...,m,j=1,...,n. Diese Koeffizienten bilden eine Matrix

$$\begin{pmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nm} \end{pmatrix} \in K^{n\times m}$$

Diese Matrix heißt <u>Matrix</u> der Abbildung $f:V\to W$ bezüglich der Basen $\{v_i\}$ und $\{w_j\}$ und wird mit

$$M(f) := M^{v_1,...,v_m}_{w_1,...,w_n}(f) := \begin{pmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nm} \end{pmatrix}$$

Beispiel

Seien a,b,c eine Basis von V und r,s,t,u eine Basis von W und sei $f:V\to W$ die lineare Abbildung mit

$$\begin{aligned} f(a) &= r + 2s + 3t + 4u \\ f(b) &= 5r + 6s + 7t + 8u \\ f(c) &= 9r + 8s + 7t + 6u \end{aligned}$$

Dann gilt

$$M(f) = M^{a,b,c}_{r,s,t,u}(f) = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 8 \\ 3 & 7 & 7 \\ 4 & 8 & 6 \end{pmatrix}$$

Bemerkungen

(i) Die Abbildung $f:V\to W$ ist durch ihre Matrix M(f) bezüglich der gegebenen Basen $\{v_i\}$ und $\{w_j\}$ eindeutig festgelegt. Gilt nämlich

(1) $$M(f) = \begin{pmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nm} \end{pmatrix},$$

so ist für jedes i

$$f(v_i) = \sum_{j=1}^{n} c_{ji}f_j$$

also

$$f(\sum_{i=1}^{m} x_ie_i) = \sum_{i=1}^{m} x_if(e_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_ic_{ji}f_j$$

(2) $$f(\sum_{i=1}^{m} x_ie_i) = \sum_{j=1}^{n} (x_ic_{ji}x_i)f_j$$

Mit anderen Worten, das Bild jedes beliebigen Vektors von V ist bereits eindeutig festgelegt.

(ii) Umgekehrt ist für beliebig vorgegebene $c_{ij} \in K$ durch (2) eine lineare Abbildung $f: V \to W$ definiert, deren Matrix bezüglich der vorgegebenen Basen von V und W gerade die Matrix (1) ist.

3.4.2 Ein kommutatives Diagramm

Seien K ein Körper,

$$f: V \to W$$

eine K-lineare Abbildung endlich-dimensionaler K-Vektorräume,

$$v_1,\dots,v_m \in V \text{ bzw. } w_1,\dots,w_n \in W$$

Basen von V bzw. W und

$$A = (a_{ij}) := M(f) = M^v_w(f)$$

die Matrix der Abbildung f bezüglich der gegebenen Basen.

Dann ist das folgende Diagramm kommutativ,

$$\begin{array}{ccc} K^m & \xrightarrow{\varphi_v} & V \\ f_A \downarrow & & \downarrow f \\ K^n & \xrightarrow{\varphi_w} & W \end{array} \tag{1}$$

d.h. es gilt

$$f \circ \varphi_v = \varphi_w \circ f_A$$

Dabei seien φ_v und φ_w die zu den Basen der v_i bzw. w_j gehörigen Isomorphismen von 3.3.5, d.h.

$$\varphi_v\left(\begin{pmatrix} x_1 \\ \dots \\ x_m \end{pmatrix}\right) := \sum_{i=1}^m x_i v_i \text{ und } \varphi_w\left(\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}\right) := \sum_{j=1}^n x_j w_j\ .$$

und f_A die zur Matrix A gehörige lineare Abbildung, d.h.
$f_A(x) = Ax$.

Beweis. Für i = 1, ... , m gilt

$$\begin{aligned} f \circ \varphi_v(e_i) &= f(v_i) \\ &= \sum_{j=1}^n a_{ji} w_j \\ &= \sum_{j=1}^n a_{ji} \varphi_w(e_j) \\ &= \varphi_w\left(\sum_{j=1}^n a_{ji} e_j\right) \\ &= \varphi_w(\text{ i- te Spalte vn A}) \\ &= \varphi_w(Ae_i) \end{aligned}$$

$$= \varphi_W(f_A(e_i))$$

Die beiden K-linearen Abbildungen $f\circ\varphi_V$ und $\varphi_W\circ f_A$ haben also and der Stelle e_i für i = 1,...,m denselben Wert. Da die e_i ein EZS von K^m bilden, sind die beiden Abbildungen an allen Stellen gleich, d.h. es gilt

$$f\circ\varphi_V = \varphi_W\circ f_A.$$

QED.

Bemerkungen

(i) Betrachten wir den Spezielfall, daß die beteiligten Vektorräume von der Gestalt K^N und die gewählten Basen aus den Standard-Einheitsvektoren bestehen:

$$v_i = e_i \in K^m =: V \text{ und } w_j = e_j \in K^m =: W,$$

Dann sind die horizontalen Abbildungen des Diagramms (1) gerade die identischen Abbildungen.
Insbesondere gilt

$$M_e^e(f_A) = A$$

(Die Matrix der Multiplikation mit A, $x \mapsto Ax$, ist gleich A).

(ii) Im allgemeinen Fall sind die horizontalen Abbildungen Isomorphismen. Für jedes f gibt es genau ein f_A (also genau ein A)[14], nämlich

$$f_A := \varphi_W^{-1}\circ f\circ\varphi_V$$

und umgekehrt gibt es für jede Matrix A genau ein f, nämlich

$$f := \varphi_W\circ f_A\circ\varphi_V^{-1}.$$

Diese Aussage läßt sich auch wie folgt formulieren:

(iii) Die Abildung

$$K^{m\times n} \to \mathrm{Hom}_K(V, W),\ A \mapsto \varphi_W\circ f_A\circ\varphi_V^{-1} \tag{2}$$

ist bijektiv und besitzt als Umkehrung die Abbildung

$$\mathrm{Hom}_K(V, W) \to K^{m\times n},\ f \mapsto M_e^e(\varphi_W^{-1}\circ f\circ\varphi_V). \tag{3}$$

Diese Abbildungen gestatten es uns, die beiden Mengen zu identifizieren (wobei dei Art der Identifikation von der Wahl der Basen abhängt). Man beachte φ_V, φ_W, φ_V^{-1}, φ_W^{-1} sind K-lineare Abildungen, sodaß (2) und (3) K-lineare Isomorphismen sind.

(iv) Die Aussage kann man auch wie folgt interpretieren: benutzt man die zu den gegebenen Basen gehörigen <u>Koordinaten-Abbildungen</u> φ_V und φ_W um V mit dem K^m und W mit dem K^n zu identifizieren,

$$K^m \overset{\varphi_V}{\cong} V$$

$$K^n \overset{\varphi_W}{\cong} W$$

[14] Die Matrix A definiert die Abbildung $f_A: x \mapsto Ax$, und umgekehrt definiert f_A die Matrix A, denn $f_A(e_i) = Ae_i$ ist gerade die i-te Spalte von A.

so identifizieren sich die linearen Abbildungen $V \to W$ mit den Matrize(-Multiplikationen der Matrizen) von $K^{m\times n}$.

3.4.3 Komposition von Abbildungen

Seien $U \xrightarrow{f} V$ und $V \xrightarrow{g} W$ zwei K-lineare Abbildungen endlich-dimensionaler Vektorräumen und u,v,w K-Vektorraumbasen von U, V bzw. W. Dann gilt

$$M^u_w(g\circ f) = M^v_w(g)\circ M^u_v(f).$$

Beweis. Seien

$$u_1,\dots,u_\ell \in U$$
$$v_1,\dots,v_m \in V$$
$$w_1,\dots,w_n \in W$$

die Vektoren der mit u, v bzw. w bezeichneten Basen.

Nach 3.4.2 haben wir kommutive Diagramme

$$\begin{array}{ccc} K^\ell & \xrightarrow{\varphi_u} & U \\ f_A\downarrow & & \downarrow f \quad \text{mit } A := M^u_v(f) \\ K^m & \xrightarrow{\varphi_v} & V \end{array}$$

und

$$\begin{array}{ccc} K^m & \xrightarrow{\varphi_v} & V \\ f_B\downarrow & & \downarrow g \quad \text{mit } B := M^v_w(g) \\ K^n & \xrightarrow{\varphi_w} & W \end{array}$$

und

$$\begin{array}{ccc} K^\ell & \xrightarrow{\varphi_u} & U \\ f_C\downarrow & & \downarrow g\circ f \quad \text{mit } C := M^u_w(g\circ f) \\ K^n & \xrightarrow{\varphi_w} & W \end{array}$$

Durch Zusammensetzen der ersten beiden Diagramme erhalten wir ein kommuatatives Diagramm

$$\begin{array}{ccc} K^\ell & \xrightarrow{\varphi_u} & U \\ f_B\circ f_A\downarrow & & \downarrow g\circ f \\ K^n & \xrightarrow{\varphi_w} & W \end{array}$$

Durch Vergleich der letzten beiden Diagramme sehen wir (da die linken vertikalen Abbildungen durch die rechten vertikalen Abbildungen eindeutig festgelegt sind)

$$(f_{BA} =)\ f_B\circ f_A = f_C,$$

also

$$BA = C.$$

Das ist aber gerade die Behauptung.

QED.

Beispiel

Sei $f:U\to V$ die lineare Abbildung von Beispiel 3.4.1, d.h. es gelte

$$
\begin{aligned}
f(a) &= r + 2s + 3t + 4u\\
f(b) &= 5r + 6s + 7t + 8u\\
f(c) &= 9r + 8s + 7t + 6u
\end{aligned}
$$

wobei a,b,c und r,s,t,u Basen von U bzw. V bezeichnen sollen. Weiter sei $g:V\to W$ die lineare Abbildung mit

$$
\begin{aligned}
g(r) &= x + 2y\\
g(s) &= x - 2y\\
g(t) &= 2x+3y\\
g(u) &= 2x-3y
\end{aligned}
$$

wobei x,y eine Basis von W bezeichne. Wir wenden g auf die definierenden Gleichungen von f an und erhalten

$$
\begin{aligned}
gf(a) &= g(r) + 2g(s) + 3g(t) + 4g(u) &&= 17x - 6y\\
gf(b) &= 5g(r) + 6g(s) + 7g(t) + 8g(u) &&= 41x - 5y\\
gf(c) &= 9g(r) + 8g(s) + 7g(t) + 6g(u) &&= 43x + 5y
\end{aligned}
$$

Für die Matrizen der Abbildungen f, g und gf ergibt sich

$$
M(f) = \begin{pmatrix} 1 & 5 & 9\\ 2 & 6 & 8\\ 3 & 7 & 7\\ 4 & 8 & 6 \end{pmatrix} \qquad M(g) = \begin{pmatrix} 1 & 1 & 2 & 2\\ 2 & -2 & 3 & -3 \end{pmatrix}
$$

und

$$
M(gf) = \begin{pmatrix} 17 & 41 & 43\\ -5 & -5 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 & 2\\ 2 & -2 & 3 & -3 \end{pmatrix} \begin{pmatrix} 1 & 5 & 9\\ 2 & 6 & 8\\ 3 & 7 & 7\\ 4 & 8 & 6 \end{pmatrix}
$$

Bemerkungen

(i) Die Matrix der identische Abbildung Id: $V \to V$ ist offensichtlich die Einheitsmatrix Id, wenn man für Definitions- und Bildraum dieselbe Basis verwendet,

$$M^v_v(Id) = Id$$

für beliebige Basen v von V.

(ii) Sind $f: V\to W$ und $g:W\to V$ zueinander inverse lineare Abbildungen, d.h. gilt

$$f\circ g = Id \text{ und } g\circ f = Id,$$

so folgt aus der eben bewiesenen Formel und Bemerkung (i), daß

$$M(f)M(g) = Id \text{ und } M(g)M(f) = Id \text{ gilt,}$$

genauer,

$$M^v_w(f)M^w_v(f^{-1}) = Id = M^w_v(f^{-1})M^v_w(f)$$

für beliebige Basen v von V und beliebige Basen w von W.

3.4.4 Verhalten bei Basiswechsel

Sei $f: V\to W$ eine lineare Abbildung von endlich-dimensionalen Vektorräumen. Weiter seien für jeden der beiden Vektorräume zwei Basen

v, v’ von V
w,w’ von W

gegeben. Dann gilt

$$M^v_w(f) = M^{w'}_w(Id)M^{v'}_{w'}(f)M^v_{v'}(Id).$$

Die Matrix $M^v_{v'}(Id)$ heißt <u>Basiswechselmatrix</u> für den Übergang von der Basis v zur Basis v'.

Bemerkungen

(i) Die behauptete Formel ergibt sich unmittelbar aus der Formel für die Komposition von Abbildungen.

(ii) Die beiden Matrizen $M^v_{v'}(Id)$ und $M^{w'}_w(Id)$ heißen <u>Basiswechselmatrizen</u>. Sie beschreiben die Beziehung zwischen den beiden Basen v, v' von V bzw. w,w' von W. Genauer sie geben an, wie die Vektoren der einen Basis mit Hilfe anderen als Linearkombination geschrieben werden.

(iii) Ist zum Beispiel

$$M^v_{v'}(Id) = (c_{ij})$$

so gilt für den i-ten Vektor der Basis v,

$$v_i = \sum_{j=1}^{n} c_{ji} v'_j$$

(mit n = dim V).

(iv) Stimmen die beiden Basen überein, so gilt $c_{ij} = \delta_{ij}$, d.h. $M^v_v(Id)$ ist die Einheitsmatrix,

$$M^v_v(Id) = Id$$

für jede Basis v von V.

(v) Speziell für W=V, f=Id, w=v, w'=v' erhalten wir die Identität

$$Id = M^{v'}_v(Id)M^v_{v'}(Id),$$

d.h. die Basiswechselmatrizen eines endlich-dimensionalen Raumes sind umkehrbar und die Übergänge v $\mapsto$ v' und v' $\mapsto$ v gehören zu zueinander inversen Matrizen,

$$M^{v'}_v(Id) = M^v_{v'}(Id)^{-1}.$$

(vi) Seien f:V→V ein linearer Endomorphismus des (endlich-dimensionalen Vektorraums) und v,v' zwei Basen von V. Dann gilt

$$M^{v'}_{v'}(f) = M^{v'}_v(Id)^{-1}M^v_v(f)M^{v'}_v(Id),$$

d.h.

$$A' = S^{-1}AS$$

mit $A' := M^{v'}_{v'}(f)$, $A := M^v_v(f)$ und $S := M^{v'}_v(Id)$.

3.4.5 Eine Anwendung auf Matrizen: Kriterium für die Umkehrbarkeit

Sei A und B quadratische Matrizen mit Einträgen aus dem Körper K,

$$A,B\in K^{n\times n}.$$

Es gelte

(1) $$AB = Id.$$

Dann sind A und B umkehrbare Matrizen, d.h. es gilt auch

$$BA = Id.$$

Beweis. Wir betrachten die zu den Matrizen gehörigen K-linearen Abbildungen

$$f_A: K^n \to K^n, x \mapsto Ax$$

$$f_B: K^n \to K^n, x \mapsto Bx$$

(bezüglich der Standardbasen des K^n). Wegen AB = Id gilt

$$f_A \circ f_B = f_{AB} = f_{Id} = Id.$$

Diese Zusammensetzung ist injektiv, also ist auch f_B injektiv und wir können den Definitionsbereich von f_B mit dem Bild identifizieren. Insbesondere gilt

$$\dim Im(f_B) = \dim K^n = n.$$

Wir haben damit ineinander liegende Räume derselben Dimension:

$$Im(f_B) \subseteq K^n.$$

Also gilt (nach 3.3.15)

$$Im(f_B) = K^n,$$

d.h. f_B ist nicht nur injektiv, sondern sogar bijektiv. Die zu f_B inverse Abbildung ist linear, also von der Gestalt f_C mit einer Matrix C. Wegen

$$f_B \circ f_C = Id = f_C \circ f_B$$

d.h.

$$f_{BC} = f_{Id} = f_{CB}$$

gilt

$$BC = Id = CB.$$

Mulitplikation der linken Identität von links mit A liefert wegen (1)

$$C = A,$$

also

$$BA = Id = AB.$$

QED.

Bemerkung

Aus dem Beweis ergibt sich insbesondere auch, daß eine Matrix A genau dann umkehrbar ist, wenn dies für die zugehörige Abbildung f_A gilt.

3.4.6 Fortsetzbarkeit von linearen Abbbildungen auf Unterräume

Seien V ein K-Vektoraum, $V' \subseteq V$ ein K-linearer Unterraum und

$$f': V' \to W$$

eine lineare Abbildung. Dann gibt es eine lineare Abbildung

$$f: V \to W,$$

welche auf dem Unterraum V' mit f' übereinstimmt,

$$f(v') = f'(v') \text{ für } v' \in V'.$$

Beweis. Sei

$$\{v_i\}_{i\in I}$$

eine Basis von V'. Wir ergänzen diese zu einer Basis

$$\{v_j\}_{j\in J}$$

von V (vgl. 3.3.14). Insbesondere gilt $I \subseteq J$. Auf Grund der Basiseigenschaft gibt es genau eine lineare Abbildung $f: V \longrightarrow W$ mit

$$f(v_i) = \begin{cases} f'(v_i) & \text{für } i\in I \\ 0 & \text{für } i \in J-I \end{cases}$$

(nach 3.3.11). Da f auf einer Basis von V' mit f' übereinstimmt, gilt

$$f(v) = f'(v) \text{ für alle } v \in V',$$

d.h. f ist die gesuchte Fortsetzung von f' zu einer linearen Abbildung von V.
QED.

3.4.7 Die duale Abbildung

3.4.7.1 Definition

Seien $f:V\to W$ eine K-lineare Abbildung und

$$V^* := \mathrm{Hom}(V,K)$$
$$W^* := \mathrm{Hom}(W,K)$$

die zu V bzw. W dualen Vektorräume.

(i) Dann ist die folgende Abbildung wohldefiniert und K-linear und heißt zu f duale Abbildung.

$$f^*: W^* \to V^*, \ell \mapsto \ell\circ f.$$

Nach Definition ist also $f^*(\ell) := \ell\circ f$.

(ii) Für je zwei K-lineare Abbildungen $f:U\to V$ und $g:V\to W$ gilt

$$(g\circ f)^* = f^*\circ g^*$$

(iii) Die zur identischen Abbildung $Id:V\to V$ duale Abbildung ist die identische Abbildung von V^*,

$$Id_V^* = Id_{V^*}.$$

Die zur Nullabbildung duale Abbildung ist die Nullabbildung.

(iv) Zueinander inverse Abbildungen gehen beim Dualisieren in zueinander inverse Abbildungen über.

(v) Das Dual einer surjektiven Abbildung ist injektiv.

(vi) Das Dual einer injektiven Abildung ist surjektiv.

Beweis. Zu (i). Die Zusammensetzung $\ell\circ f$ ist als Komposition linearer Abbildungen wieder linear, d.h. es gilt

$$\ell\circ f\in V^*$$

für jedes $\ell\in W^*$. Mit anderen Worten, die Abbildung f^* ist wohldefiniert. Wir haben noch zu zeigen, daß sie linear ist. Für je zwei Linearformen ℓ', $\ell''\in W^*$, beliebige $c',c''\in K$ und beliebige $v\in V$ gilt

$$\begin{aligned} f^*(c'\ell'+c''\ell'')(v) &= ((c'\ell'+c''\ell'')\circ f)(v) \\ &= (c'\ell'+c''\ell'')(f(v)) \\ &= c'\ell'(f(v))+c''\ell''(f(v)) \\ &= c'(\ell'\circ f)(v))+c''(\ell''\circ f)(v)) \\ &= c'f^*(\ell')(v) + c''f^*(\ell'')(v) \\ &= (c'f^*(\ell') + c''f^*(\ell''))(v) \end{aligned}$$

Da dies für beliebige $v\in V$ so ist, gilt

$$f^*(c'\ell'+c''\ell'') = c'f^*(\ell') + c''f^*(\ell'').$$

Mit anderen Worten , f^* ist linear.

Zu (ii). Für jede Linearform $\ell:W\to K$ gilt

$$\ell\circ(g\circ f) = (\ell\circ g)\circ f,$$

also

$$(g\circ f)^*(\ell) = f^*(\ell\circ g) = f^*(g^*(\ell)) = (f^*\circ g^*)(\ell).$$

Da dies für alle Linearformen ℓ auf V richtig ist, folgt

$$(g\circ f)^* = f^*\circ g^*.$$

Zu (iii). Die Verpflanzung mit der identischen Abbildung bildet trivialerweise jede Linearform auf sich selbst ab. Die Verpflanzung mit der Nullabbildung bildet jede Linearform auf die Null ab.

Zu (iv). Gilt $g\circ f = \mathrm{Id}$, so gilt nach (ii) und (iii) auch $f^*\circ g^* = \mathrm{Id}^* = \mathrm{Id}$.

Zu (v). Sei $f:V\to W$ surjektiv. Wir haben zu zeigen,

$$\ker(f^*) = 0.$$

Sei also $\ell\in\ker(f^*)$. Da $f:V\to W$ surjektiv ist, gibt es zu jedem $w\in W$ ein $v\in V$ mit $w=f(v)$. Damit ist

$$\ell(w) = \ell(f(v)) = (\ell\circ f)(v) = f^*(\ell)(v) = 0.$$

Wir haben gezeigt, für jedes $w\in W$ gilt $\ell(w) = 0$, d.h. es gilt $\ell = 0$. Der Kern von f^* ist somit trivial, d.h. f^* ist injektiv.

Zu (vi). Sei $f:V\to W$ injektiv und sei $\ell\in V^*$ beliebig. Wir haben zu zeigen, es gibt ein

$$\ell'\in W^* \text{ mit } f^*(\ell') = \ell.$$

Da f injektiv ist, ist die zugehörige Abbildung

$$g:V\to f(V) \text{ mit } g(v) = f(v) \text{ für alle } v\in V$$

bijektiv, also ein Isomorphismus.

$$\begin{tikzcd} & K & \\ V \arrow[ur, "\ell"] \arrow[r, "g", "\cong"'] & f(V) \arrow[u, "\ell''"] & W \arrow[ul, "\ell'"'] \arrow[l, phantom, "\supseteq"] \end{tikzcd}$$

Es gibt also eine lineare Abbildung

$$\ell'' = \ell\circ g^{-1}: f(V) \to K,$$

welche auf dem Unterraum f(V) von W definiert ist mit

(1) $$\ell''\circ g = \ell\circ g^{-1}\circ g = \ell$$

Wenn wir eine Fortsetzung $\ell':W\to K$ der Linearform $\ell'':f(V)\to K$ zu einer Linearform auf dem gesamten Raum W finden können, so gilt für diese Fortsetzung die Identität,

$$\ell'\circ f = \ell''\circ g = \ell, \text{ d.h. } f^*(\ell') = \ell.$$

Jede solche Fortsetzung ist eine Linearform der gesuchten Art. Die Behauptung folgt daher aus dem Fortsetzungssatz 3.4.6..

QED.

3.4.7.2 Die duale Basis

Sei V ein endlich-dimensionaler K-Vektoraum und sei

(1) $$v_1,\dots,v_n\in V$$

eine Basis von V. Wir erinnern daran (vgl. die Bemerkung von 3.3.18), die zu dieser Basis duale Basis besteht dann aus Vektoren

$$v_1^*,\dots,v_n^*\in V^*$$

des dualen Raums V* mit

$$v_i^*(v_j) = \delta_{ij}$$

für alle i,j= 1,...,n. Diese bilden eine Basis von V*.

$$v_1^*,\dots,v_n^*\in V^*$$

Bemerkung
Unser nächstes Ziel besteht darin, zu zeigen, daß zwischen der Matrix einer Abbildung und der Matrix der dualen Abbildung ein Zusammenhang besteht.

3.4.7.3 Die Matrix der dualen Abbildung bezüglich der dualen Basis

Seien V und W ein endlich-dimensionale K-Vektorräume,

$$f: V \to W$$

eine K-lineare Abbildung,

$$v_1,\dots,v_m \in V \text{ und } w_1,\dots,w_n \in W$$

Basen von V bzw. W und

$$v^*_1,\dots,v^*_m \in V \text{ und } w^*_1,\dots,w^*_n \in W$$

die zugehörigen dualen Basen. Dann gilt

$$M^{w^*}_{v^*}(f^*) = M^v_w(f)^T,$$

d.h. die Matrix der dualen Abbildung ist transponiert zur Matrix der Ausgangsabbildung (bezüglich der dualen Basen).

Beweis. Sei

$$M^v_w(f) = A := (a_{ij}),$$

d.h. es gelte

$$f(v_i) = \sum_{\alpha=1}^{n} a_{\alpha i} w_\alpha \quad \text{für } i = 1,\dots,m.$$

Nach Definition der dualen Abbildung $f^*: W^* \to V^*$ ist

$$\begin{aligned}
f^*(w^*_j)(v_i) &= (w^*_j \circ f)(v_i)\\
&= w^*_j(f(v_i))\\
&= w^*_j\left(\sum_{\alpha=1}^{n} a_{\alpha i} w_\alpha\right)\\
&= \sum_{\alpha=1}^{n} a_{\alpha i}\, w^*_j(w_\alpha)\\
&= a_{ij}\\
&= \sum_{\alpha=1}^{n} a_{i\alpha} v^*_\alpha(v_i).
\end{aligned}$$

Mit anderen Worten, die beiden linearen Abbildungen $f^*(w^*_j)$ und $\sum_{\alpha=1}^{n} a_{i\alpha} v^*_\alpha$ haben im Basiselement v_i denselben Wert. Da i beliebig ist, stimmen die Abbildungen auf den Elementen einer Basis überein, und sind als lineare Abbildungen damit überhaupt gleich,

$$f^*(w^*_j) = \sum_{\alpha=1}^{n} a_{i\alpha} v^*_\alpha.$$

Mit anderen Worten, die Matrix von f* ist gerade A^T.

QED.

3.4.8 Anwendung: Das doppelte Dual eines endlich-dimensionalen Vektorraums

(i) Sei V ein K-Vektorraum. Dann ist die Abbildung[15]

$$\rho := \rho_V : V \to V^{**},\ v \mapsto (\ell \mapsto \ell(v)),$$

(von V ins doppelte Dual von V mit $\rho_V(v)(\ell) = \ell(v)$) wohldefiniert und injektiv.

Insbesondere ist ρ im Fall $\dim V < \infty$ ein Isomorphismus. Dieser heißt natürlicher Isomorphismus von V ins doppelte Dual.

(ii) Sei f: V→V eine K-lineare Abbildung. Dann ist das folgende Diagramm kommutativ.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\rho_V"'] & W \arrow[d, "\rho_W"] \\ V^{**} \arrow[r, "f^{**}"] & W^{**} \end{tikzcd},$$

d.h. für jedes $v \in V$ gilt $\rho_W(f(v)) = f^{**}(\rho_V(v))$.

Beweis. Zu (i). Die Abbildung ist wohldefiniert. Wir haben zu zeigen, für jedes $v \in V$ liegt die Abbildung

$$f_v := \rho(v): \ell \mapsto \ell(v)$$

in V^{**}, d.h. wir haben zu zeigen, f_v ist eine lineare Abbildung $V^* \to K$. Seien also ℓ, $\ell' \in V^*$ und $c, c' \in K$. Dann gilt

$$\begin{aligned} f_v(c\ell + c'\ell) &= (c\ell + c'\ell)(v) \\ &= c\ell(v) + c'\ell(v) \\ &= cf_v(\ell) + c'f_v(\ell'). \end{aligned}$$

Mit anderen Worten, f_v ist linear.

Die Abbildung ρ ist linear. Seien $v, v' \in V$ und $c, c' \in K$. Dann gilt für jedes $\ell \in V^*$:

$$\rho(cv + c'v')(\ell) = \ell(cv + c'v') = c\ell(v) + c'\ell(v') = c\rho(v)(\ell) + c'\rho(v')(\ell).$$

Da dies für beliebige $\ell \in V^*$ gilt, folgt

$$\rho(cv + c'v') = c\rho(v) + c'\rho(v').$$

Mit anderen Worten, die Abbildung ρ ist linear.

Die Abbildung ρ ist injektiv. Es reicht zu zeigen, der Kern von ρ ist trivial. Sei $v \in V$ ein Element mit $\rho(v) = 0$. Wir haben zu zeigen, daß dann v selbst schon Null ist. Nach Voraussetzung gilt für jedes $\ell \in V^*$:

$$0 = \rho(v)(\ell) = \ell(v),$$

d.h. es ist

(1) $\qquad \ell(v) = 0$ für jedes $\ell \in V^*$.

Angenommen v ist ungleich Null. Dann ist v eine Basis des Teilvektorraums $Kv \subseteq V$. Insbesondere ist die lineare Abbildung

---

[15] $\rho(v)$ ist die Auswertungsabbildung ℓ a $\ell(v)$ an der Stelle v.

$$g: K \to Kv,\ c \mapsto cv,$$

ein Isomorphismus. Die Abbildung

$$g^{-1}: Kv \to K,\ cv \mapsto c,$$

ist somit wohldefiniert (und linear). Insbesondere gilt $g^{-1}(v) = 1$. Nach 4.4.6.2 läßt sich g^{-1} fortsetzen zu einer linearen Abbildung

$$\ell: V \to K.$$

Es gilt $\ell(v) = g^{-1}(v) = 1$. Wir haben damit ein Element $\ell \in V^*$ gefunden, für welches die Aussage (1) falsch ist. Damit ist aber unsere Annahme $v\neq 0$ falsch und es muß v gleich Null sein. Wir haben gezeigt, ρ hat den Kern $\{0\}$, ist also injektiv.

Im Fall $\dim V < \infty$ ist ρ ein Isomorphismus. Angenommen ρ wäre nicht surjektiv. Dann wäre $\mathrm{im}(\rho)$ ein echter Unterraum von V^{**}, also

$$\dim \mathrm{im}(\rho) < \dim V^{**} = \dim V.$$

Da aber ρ, wie wir gesehen haben, injektiv ist, muß das Bild von ρ isomorph zum Definitionsbereich sein,

$$\rho: V \to \mathrm{im}(\rho)$$

ist ein Isomorphismus, d.h. es gilt $\dim \mathrm{im}(\rho) = \dim V$.

Zu (ii). Wir haben zu zeigen, für jedes $\ell \in W^*$ gilt

$$\rho_W(f(v))(\ell) = f^{**}(\rho_V(v))(\ell).$$

Für die linke Seite erhalten wir

$$\rho_W(f(v))(\ell) = \ell(f(v)) = (\ell\circ f)(v)$$

Zur Berechnung der rechten Seite beachten wir, f^{**} ist die Verpflanzung mit f^*, d.h. es ist

$$f^{**}(\rho_V(v)) = (\rho_V(v))\circ f^*$$

Für die rechte Seite erhalten wir damit

$$f^{**}(\rho_V(v))(\ell) = (\rho_V(v))(f^*(\ell)) = (\rho_V(v))(\ell\circ f) =^{16} (\ell\circ f)(v)$$

Die Abbildungen $\rho_W(f(v))$ und $f^{**}(\rho_V(v))$ haben für beliebiges $\ell\in W^*$ denselben Wert. Sie sind also gleich.

QED.

3.4.9 Zeilenrang und Spaltenrang von Matrizen

Seien K ein Körper und $A\in K^{m\times n}$ eine Matrix mit Einträgen aus K. Wir denken uns A als einzeilige Matrix geschrieben, deren Einträge die Spalten von A sind,

$$A = (a_1, \dots , a_n) \text{ mit } a_i \in K^{m\times 1}.$$

Der Rang oder genauer der Spaltenrang von A ist dann definiert als die Maximalzahl linear unabhängiger Spalten a_i in der Menge aller a_i. Mit anderen Worten, der Spaltenrang von A ist gerade

$$\mathrm{rk}\, A := \mathrm{srk}\, A = \dim \langle a_1, \dots , a_n\rangle$$

die Dimension des von den Spalten von A erzeugten linearen Unterraums von $K^{m\times 1}$.

---

[16] $\rho_V(v)$ ist die Auswertung an der Stelle v.

In analoger Weise können wir die Matrix A auch als einspaltige Matrix geschrieben denken, deren Einträge die Zeilen von A sind,

$$A = \begin{pmatrix} a^1 \\ \dots \\ a^m \end{pmatrix} \text{ mit } a^j \in K^{1\times n}.$$

Der <u>Zeilenrang</u> von A ist dann definiert als die Maximalzahl linear unabhängiger Zeilen a^j in der Menge aller a^j . Mit anderen Worten, der Zeilenring von A ist gerade

$$\text{zrk } A = \dim \langle a^1, \dots , a^m\rangle$$

die Dimension des von den Zeilen von A erzeugten linearen Unterraums von $K^{1\times n}$.

Für jede lineare Abbildung $f: V \longrightarrow W$ definieren wir schließlich deren <u>Rang</u> als die Dimension ihres Bildes,

$$\text{rk } f := \dim \text{Im}(f).$$

Beispiel

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 9 \\ 6 & 7 & 13 \end{pmatrix}$$

Die ersten beiden Spalten dieser Matrix sind nicht proportional, d.h. sie sind linear unabhängig. Die dritte Spalte ist gerade die Summe der beiden ersten. Daher gilt

$$\text{rk } A = 2.$$

Die ersten beiden Zeilen von A sind ebenfalls nicht proportional. Der Zeilenrang sind somit mindestens 2. Weiter gilt

$$2\cdot(1,2,3) - 5\cdot(4,5,9) + 3\cdot(6,7,13) = 0.$$

Der Zeilenrang ist also ebenfalls 2,

$$\text{zrk } A = 2.$$

Diese Übereinstimmung ist, wie wir später sehen werden, kein Zufall.

3.4.10 Eigenschaften des Rangs

Seien K ein Körper, A eine Matrix mit Einträgen aus K und

$$f: V \longrightarrow W$$

eine K-lineare Abbildung von Vektorräume endlicher Dimension. Dann gilt:

(i) zrk A = srk A^T und srk A = zrk A^T.
(ii) rk A = rk f_A
(iii) rk f = rk $M^v_w(f)$ für lineare Abbildungen $f: V \longrightarrow W$ und Basen v von V und w von W.
(iv) Ist f die Zusammensetzung aus einer linearen Surjektion und einer linearen Injektion,

$$f = i\circ\tilde{f} : V \overset{\tilde{f}}{\twoheadrightarrow} U \overset{i}{\hookrightarrow} W,$$

so gilt rk f = dim U.

Beweis.
<u>Zu (i)</u>.Beim Transponieren vertauschen die Zeilen und Spalten einer Matrix ihre Rollen. Dasselbe gilt somit auch für Zeilenrang und Spaltenrang.

Zu (ii). Wir schreiben A als einzeilige Matrix mit den Spalten a_i,

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

Dann gilt

$$\begin{aligned}\mathrm{Im}(f_A) &= \{ f_A(x) \mid x \in K^n \} \\ &= \{ Ax \mid x \in K^n \} \\ &= \{a_1x_1 + \ldots + a_nx_n \mid x_1, \ldots, x_n \in K\} \\ &= < a_1, \ldots a_n >\end{aligned}$$

also

$$\mathrm{rk}\, f_A = \dim < a_1, \ldots a_n > = \mathrm{rk}.$$

Zu (iii). Wir betrachten das kommutative Diagramm von 3.4.2,

$$\begin{array}{ccc} K^m & \xrightarrow{\varphi_V} & V \\ f_A\downarrow & & \downarrow f \\ K^n & \xrightarrow{\varphi_W} & W \end{array} \qquad \text{mit } A := M_W^V(f).$$

Es gilt

$$\begin{aligned}\mathrm{rk}\, f &= \dim \mathrm{Im}(f) && \text{(nach Definition des Rangs)} \\ &= \dim \mathrm{Im}(f_A) && \text{(weil } \varphi_V \text{ und } \varphi_W \text{ Isomorphismen sind)} \\ &= \mathrm{rk}\, f_A && \text{(nach Definition des Rangs)} \\ &= \mathrm{rk}\, A && \text{(nach (ii))}\end{aligned}$$

Zu (iv). Weil $\tilde{f}$ surjektiv ist, gilt

$$\mathrm{Im}(f) = \mathrm{Im}(i).$$

Weil i injektiv ist, gilt

$$\mathrm{Im}(i) \cong U.$$

Zusammen erhalten wir

$$\mathrm{rk}\, f = \dim \mathrm{Im}(f) = \dim \mathrm{Im}(i) = \dim U.$$

QED.

3.4.11 Das Verhalten des Rangs einer Abbildung beim Dualisieren

Seien

$$f: V \to W$$

eine K-lineare Abbildung endlich-dimensionaler Vektorräume und

$$f^*: W^* \to V^*$$

die zugehörige duale Abbildung. Dann gilt für die Ränge der beiden Abbildungen

$$\mathrm{rk}(f) = \mathrm{rk}(f^*).$$

Insbesondere ist für jede Matrix A der Zeilenrang gleich dem Spaltenrang von A,

$$\mathrm{rk}\, A = \mathrm{rk}\, A^T$$

Beweis.

1. Schritt: Reduktion auf den Fall linearer Abbildungen.

Der zweite Teil der Behauptung folgt aus dem ersten, denn für

$$f = f_A: K^n \longrightarrow K^m$$

ist gilt

$$M(f) \quad = A$$

bezüglich der Standard-Einheitsbasen von K^n und K^m und

$$M(f^*) \quad = A^T$$

bezüglich der zugehörigen dualen Basen (nach 3.4.7.3). Also gilt

$$\text{rk } A = \text{rk } M(f) = \text{rk } f$$

$$\text{rk } A^T = \text{rk } M(f^*) = \text{rk } f^*.$$

Aus der Gleichheit der Ränge von f und f* folgt also die Gleichheit der Ränge von A und A^T.

2. Schritt: Der Fall lineare Abbildungen.

Wir zerlegen die lineare Abbildung in die Zusammensetzung aus einer Surjektion und einer Injektion:

$$f = i \circ \tilde{f} : V \xrightarrow{\tilde{f}} \!\!\!\rightarrow U \xhookrightarrow{i} W.$$

Dabei sei U = Im(f), $\tilde{f}$ genüge derselben Abbildungsvorschrift wie f und die Abbildung i sei gerade die natürliche Einbettung von U in W, d.h. i(u) = u für u ∈ U. Es gilt

$$\text{rk } f = \dim U. \qquad (1)$$

(nach 3.4.10). Betrachten wir das Dual von f. Es gilt

$$f^* = (i \circ \tilde{f})^* = \tilde{f}^* \circ i^*: W^* \xrightarrow{i^*} \!\!\!\rightarrow U^* \xhookrightarrow{\tilde{f}^*} V^*.$$

Man beachte, das Dual einer Injektion ist eine Surjektion und das Dual einer Surjektion ist eine Injektion (nach 3.4.7.1 (v) und (vi)). Damit ist

$$\text{rk } f^* = \dim U^* = \dim U.$$

(nach 3.4.10 und 3.3.20). Zusammen mit (1) folgt die Behauptung,

$$\text{rk } f^* = \text{rk } f.$$

QED.

3.4.12 Rangkriterium für die Umkehrbarkeit einer Matrix

Sei $A \in K^{n \times n}$ eine n-reihige Matrix mit Einträgen aus K. Dann sind folgende Bedingungen äquivalent.

(i) A ist umkehrbar.

(ii) rk(A) = n.

Beweis. (i) ⇒ (ii). Nach Voraussetzung existiert A^{-1}. Mit AA^{-1} = Id gilt auch

$$f_A \circ f_{A^{-1}} = f_{Id} = Id.$$

Insbesondere ist f_A surjektiv,

$$f_A: K^n \twoheadrightarrow K^n.$$

also

$$\text{rk } A = \text{rk } f_A = \dim \text{Im}(f_A) = \dim K^n = n.$$

(ii) ⇒ (i). Nach Voraussetzung gilt $\dim f_A(K^n) = n$, d.h. die Abbildung

$$f_A: K^n \to K^n$$

ist surjektiv. Weiter ist

$$\dim \ker(f_A) = \dim K^n - \dim \text{Im}(f_A) = n - n = 0,$$

d.h. der Kern von f_A ist trivial, d.h. f ist injektiv. Wir haben gezeigt, f_A ist bijektiv, also ein Isomorphismus. Dann ist A aber umkehrbar (vgl. die Bemerkung zu 3.4.5).
QED.

4. Determinanten

Für jede quadratische Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in K^{n\times n}$$

über einem Körper K ist die Determinante von A durch die Formel von Leibniz

$$\det A := |A| := \sum_{\sigma\in S_n} \mathrm{sign}(\sigma)\cdot a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)}$$

definiert. Wir beginnen damit, die Bestandteile dieser Formel zu erklären.

4.1 Permutationen

Dieser und der nachfolgende Abschnitt haben vorbereitenden Charakter.

4.1.1 Gruppen von Abbildungen

Sei M eine beliebige Menge. Wir führen folgende Bezeichnungen ein

Abb(M) := Abb(M,M) = Menge der Abbildungen $M \to M$
S(M) := $\{f\in \mathrm{Abb}(M) \mid f \text{ bijektiv}\}$
(Menge der bijektiven Abbildungen $M \to M$)

Satz
Die Menge S(M) ist bezüglich der Komposition von Abbildungen eine Gruppe.

Die Gruppe S(M) heißt auch symmetrische Gruppe der Menge M.
Beweis. Wir führen den Beweis wieder in mehreren Schritten. Abkürzend schreiben wir wieder

$$G := S(M)$$

1. Die Abbildung $G\times G \to G$, $(f,g) \mapsto f\circ g$, ist wohldefiniert,
denn die Zusammensetzung zweier bijektiver Abbildungen ist bijektiv.
2. Die Komposition von Abbildungen ist assozialtiv
Seien f, g, $h:M\to M$ drei bijektive Abbildungen. Wir haben zu zeigen,

$$(f\circ g)\circ h = f\circ(g\circ h), \qquad (1)$$

d.h. zu zeigen ist,

$$((f\circ g)\circ h)(m) = (f\circ(g\circ h))(m) \text{ für beliebiges } m\in M \qquad (2)$$

Es gilt
LHS von (2) = $(f\circ g)(h(m)) = f(g(h(m)))$
RHS von (2) = $f((g\circ h)(m))$ 0 $f(g(h(m)))$
Beide Seiten von (2) sind also gleich.
3. Es gibt in G ein Element mit den Eigenschaften des Einselements,
nämlich die identische Abbildung.
4. Zu jedem $f\in G$ gibt es ein Element in G mit den Eigenschaften des Inversen
Das ist so nach Definition der Bijektivität (vgl. Bemerkung 3.1 (i)).
QED.

4.1.2 Symmetrische Gruppen endlicher Mengen

4.1.2.1 Bezeichnung von Permutationen

Sei M die folgende n-elementige Menge.

$$M := \{1,2,...,n\}$$

Für die symmetrischen Gruppe von M verwendet man dann die Bezeichnung

$$S_n := S(M).$$

Die Elemente von S_n, d.h. die bijektiven Abbildungen

$$f: M \to M$$

heißen auch Permutationen der Zahlen 1,...,n. Diese Elemente wollen wir durch eine Art zweireihige Matrizen beschreiben, in deren erster Zeile die Elemente von M und darunter in der zweiten Zeile deren Bilder stehen,

$$f = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ f(1) & f(2) & f(3) & \dots & f(n) \end{pmatrix}.$$

Äußerlich sieht das Symbol auf der rechten Seite der Gleichung aus wie eine zweireihige Matrix. Wenn wir Permutationen im Auge haben, soll dies jedoch keine Matrix bezeichnen sondern die Abbildung, welche 1 in f(1), 2 in f(2),..., n in f(n) abbildet.

Beispiel 1

Mit

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$$

werde die Abbildung

$$f:\{1,2,3,4\} \to \{1,2,3,4\}$$

bezeichnet mit

$$f(1) = 4,\ f(2) = 3,\ f(3) = 1,\ f(4) = 2.$$

Beispiel 2

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}$$

4.1.2.2 Zyklenschreibweise

Ein Zyklus ist eine Permutation $f: M \to M$ mit der Eigenschaft, daß es paarweise verschiedene Elemente $e_1, ... , e_r \in M$ gibt mit

$$f(e_1) = e_2,\ f(e_2) = e_3,\ ... ,\ f(e_{r-1}) = e_r,\ f(e_r) = e_1$$

und

$$f(e) = e \text{ für } e \in M - \{e_1, ... , e_r\}.$$

Für einen solchen Zyklus verwendet man die Bezeichnung

$$f = (e_1, ... , e_r) = (e_2, ..., e_r, e_1) = ...$$

Zyklen, deren zugehörige Mengen $\{e_1, ... , e_r\}$ disjunkt sind, heißen elementfremd. Zyklen, deren zugehörige Menge $\{e_1, ... , e_r\}$ aus zwei Elementen bestehen, heißen Transpositionen. Eine Permutation der Gestalt

$$(i, i+1) \in S_n$$

heißt Nachbartausch.

Bemerkungen

(i) Nicht jede Permutation ist ein Zyklus.

(ii) Jeder Zyklus ist ein Produkt von Transpositionen,

$$(e_1, \ldots , e_r) = (e_1,e_2)\circ(e_2,e_3)\circ\ldots\circ(e_{r-2},e_{r-1})\circ(e_{r-1},e_r)$$

(iii) Aus (ii) und (iii) ergibt sich, daß jede Permutation Produkt von Transpositionen ist.

Beispiele

$$\begin{pmatrix} 1\,2\,3\,4 \\ 4\,3\,1\,2 \end{pmatrix} = (1,4,2,3)$$

$$\begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 4\,3\,7\,1\,6\,8\,2\,5 \end{pmatrix} = (1,4)(2,3,7)(5,6,8)$$

$$(1,3,5)(2,4,6) = \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 3\,4\,5\,6\,1\,2 \end{pmatrix}$$

Multiplikation mit einer Transposition

$$\begin{pmatrix} 1 & \ldots & i & \ldots & j & \ldots & n \\ f(1) & \ldots & f(i) & \ldots & f(j) & \ldots & f(n) \end{pmatrix}\circ(i,j) = \begin{pmatrix} 1 & \ldots & i & \ldots & j & \ldots & n \\ f(1) & \ldots & f(j) & \ldots & f(i) & \ldots & f(n) \end{pmatrix}$$

Konjugation eines Zyklus mit einer Permutation f

$$f\circ(e_1, \ldots , e_r)\circ f^{-1} = (f(e_1), \ldots , f(e_r))$$

$$\begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 4\,3\,7\,1\,6\,8\,2\,5 \end{pmatrix}\circ(1,3,5)(2,4,6)\circ\begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 4\,3\,7\,1\,6\,8\,2\,5 \end{pmatrix}^{-1} = (4,7,6)(3,1,8)$$

4.1.2.3 Inversionen und Vorzeichen einer Permutation

Seien M die Menge der ersten n natürlichen Zahlen

$$M = \{1,\ldots,n\}$$

und f: M → M eine Permutation von M,

$$f = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ f(1) & f(2) & f(3) & \ldots & f(n) \end{pmatrix}.$$

Eine Inversion von f ist dann ein Paar $(i,j)\in M\times M$ mit folgenden Eigenschaften.

1. $i > j$
2. $f^{-1}(i) < f^{-1}(j)$.

Bezeichne

$$\mathrm{Inv}(f)$$

die Menge der Inversionen der Permutation f und

$$\#\,\mathrm{Inv}(f)$$

die Anzahl der Elemente der Menge Inv(f). Dann heißt

$$\mathrm{sign}(f) := (-1)^{\#\mathrm{Inv}(f)}$$

<u>Vorzeichen</u> der Permutation f. Permutationen mit positiven Vorzeichen heißen <u>gerade</u>, solche mit negativen Vorzeichen <u>ungerade</u>.

Beispiel

Die Inversionen der Permutation

$$f = \begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 4\,3\,7\,1\,6\,8\,2\,5 \end{pmatrix}$$

sind gerade die Paare natürlicher Zahlen aus der unteren Zeile, für welche die rechts stehende Zahl kleiner ist als die links stehende. Wir erhalten die folgenden Inversionen.

(3,1), (3,2)
(4,1), (4,2), (4,3)
(6, 2), (6,5)
(7, 1), (7,2), (7,5), (7,6)
(8,2), (8,5)

Das Vorzeichen von f ist also

$$\text{sign}(f) = (-1)^{13} = -1.$$

4.1.2.4 Zerlegung in Transpositionen

Jede Permutation ist ein Produkt von endlich vielen Transpositionen.
Beweis. Die Permutation

$$f = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ f(1) & f(2) & f(3) & \dots & f(n) \end{pmatrix} \qquad (1)$$

ist genau dann die identische Abbildung, wenn

$$f(1) < f(2) < ... < f(n)$$

gilt, d.h. wenn f keine Inversionen bestitzt. Im Fall $f \neq \text{Id}$ gibt es also ein $i \in [1,n-1]$ mit

$$f(i) > f(i+1).$$

Durch Vertauschen von f(i) und f(i+1) in der unteren Zeile wird also die Anzahl der Inversionen verkleinert. Mit anderen Worten

$$f\circ(i,i+1)$$

hat weniger Inversionen als f. Durch Wiederholen der Argumentation erhalten wir duch wiederholte Multiplikation mit Nachbartauschen eine Permutation ohne Inversionen, d.h. die identische Abbildung. Es gibt also eine Folge von Zahlen $i_1, \dots , i_r \in [1, n-1]$ mit

$$f\circ(i_1,i_1+1)\circ...\circ(i_r,i_r+1) = \text{Id}$$

Wir multiplizieren diese Identität von rechts mit $(i_r,i_r+1)\circ(i_{r-1},i_{r-1}+1)\circ...\circ(i_1,i_1+1)$ und erhalten

$$f = (i_r,i_r+1)\circ(i_{r-1},i_{r-1}+1)\circ...\circ(i_1,i_1+1).$$

QED.

4.1.2.5 Das Vorzeichen bei Multiplikation mit einer Transposition

Für jede Permutation $f \in S_n$ und jede Transposition (i,j) aus S_n gilt

$$\text{sign}(f\circ(i,j)) = - \text{sign}(f)$$

Beweis. Es gilt

$$\begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ f(1) & \dots & f(i) & \dots & f(j) & \dots & f(n) \end{pmatrix}\circ(i,j) = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ f(1) & \dots & f(j) & \dots & f(i) & \dots & f(n) \end{pmatrix}.$$

<u>1. Fall</u>. j=i+1.
Mit anderen Worten, (i,j) = (i,i+1) ist ein Nachbartausch. Wir haben die Anzahl der Inversionen von

(1) $$\begin{pmatrix} 1 & \dots & i & i+1 & \dots & n \\ f(1) & \dots & f(i) & f(i+1) & \dots & f(n) \end{pmatrix}$$

und

(2) $$\begin{pmatrix} 1 & \dots & i & i+1 & \dots & n \\ f(1) & \dots & f(i+1) & f(i) & \dots & f(n) \end{pmatrix}$$

zu vergleichen.

Ein Paar (a,b) mit $a \neq f(i)$ oder $b \neq f(i+1)$ ist genau dann eine Inversion von (1) wenn es eine von (2) ist.[17]
Das Paar (f(i), f(i+1)) ist genau dann eine Inversion von (1), wenn das Paar (f(i+1), f(i)) keine Inversion von (2) ist.
Zusammen ergibt sich, die Inversionsmengen von (1) und (2) unterscheiden sich um genau ein Element. Das Vorzeichen von (1) und (2) ist also entgegengesetzt.
2. Fall. i und j beliebig.
O.B.d.A sei $i < j$. Es gilt

(3) $$(i,j) =^{18} (i,i+1)\circ(i+1,i+2)\circ\ldots\circ(j-2,j-1)\circ(j,j-1)\circ(j-1,j-2)\circ\ldots\circ(i+2,i+1)\circ(i+1,i)$$

Statt f mit der Permutation (i,j) zu multiplizieren, können wir f auch nacheinander mit den Nachbartauschen auf der rechten Seite von (3) multiplizieren. Bei jeder Multiplikation mit einem Nachbartausch ändert die Permutation ihr Vorzeichen. Die Zahl der Nachbartausche auf der rechten Seite von (3) ist aber ungerade (bis auf den Faktor in der Mitte kommt jeder Faktor zweimal vor). Nach einer ungeraden Anzahl von Nachbartauschen ist aber das endgültige Vorzeichen dem ursprünglichen entgegengesetzt.
QED.

Bemerkungen

(i) Wendet man die obige Formel auf die identische Permutation an, so sieht man, das Vorzeichen einer Transposition ist negativ.
$$\text{sign}\,(a,b) = -1.$$
(ii) Ist die Permutation von f das Produkt eine geraden Anzahl von Transpositionen, so gilt auf Grund der obigen Formel
$$\text{sign}(f) = +1,$$
ist Produkt einer ungeraden Anzahl von Transpositionen, so gilt entsprechend
$$\text{sign}(f) = -1.$$
(iii) Allgemein, ist f das Produkt von r Transpositionen, so gilt
$$\text{sign}(f) = (-1)^r$$

Beispiel

$\text{sign}\,(1,2)\circ(2,5)\circ(3,4) = -\,\text{sign}\,(1,2)\circ(2,5) = +\,\text{sign}\,(1,2) = -1.$

4.1.2.6 Das Vorzeichen eines Produktes von Permutationen

$$\text{sign}\, f\circ g = \text{sign}\, f \cdot \text{sign}\, g$$
Mit anderen Worten, die Abbildung
$$\text{sign}: S_n \to \{+1, -1\},\ f \mapsto \text{sign}(f),$$
ist ein Gruppenhomomorphismus.
Beweis. Wir schreiben f und g als Produkte von Transpositionen,
$$f = (a_1,a_2)\circ(a_3,a_4)\circ\ldots\circ(a_r,a_{r+1})$$

---

[17] Denn dann steht b in der unteren Zeile von (1) genau dann rechts von a, wenn dasselbe bezüglich der unteren Zeile von (2) gilt.
[18] Wendet man die rechte Seite auf i an, so wird i nacheinander auf i+1,i+2,...,j-1, j abgebildet und bleibt bei den letzten j-i-1 Operationen unverändert.
Wendet man die rechte Seite auf j an, so bleibt j bei den ersten j-i-1 Operationen unverändert und wird dann nacheinander auf j-1, j-2, ... , i+1, i abgebildet.
Wendet man die rechte Seite auf ein a außerhalb des Intervalls [i, j] an, so bleibt a unverändert, da dieser Wert auf der rechten Seite nicht vorkommt.
Wendet man die rechte Seite auf ein a im Innern des Intervalls [i, j] an, so wird a zuerst auf auf a+1 abgebildet (bei den ersten j-i-1 Operationen) und anschließend auf wieder auf a (bei den übrigen Operationen).

$$g = (b_1,b_2)\circ(b_3,b_4)\circ\ldots\circ(b_s,b_{s+1})$$

Dann gilt

$$f\circ g = (a_1,a_2)\circ(a_3,a_4)\circ\ldots\circ(a_r,a_{r+1})\circ(b_1,b_2)\circ(b_3,b_4)\circ\ldots\circ(b_s,b_{s+1}).$$

Nach Bemerkung (iii) von 3.3.8.4 gilt damit

$$\mathrm{sign}(f) = (-1)^r$$
$$\mathrm{sign}(g) = (-1)^s$$
$$\mathrm{sign}(f\circ g) = (-1)^{r+s} = (-1)^r\cdot(-1)^s = \mathrm{sign}(f)\cdot\mathrm{sign}(g).$$

QED.

4.1.3 Untergruppen

4.1.3 .1 Definition

Sei G eine Gruppe. Eine Untergruppe von G ist eine Teilmenge $H\subseteq G$, welche zusammen mit der Operation von G eine Gruppe ist.

Genauer: Eine Teilmenge $H\subseteq G$ heißt Untergruppe von G, wenn die Einschränkung der Gruppenoperation

$$G\times G \to G,\ (a,b) \mapsto a\cdot b,$$

von G auf $H\times H$ eine Abbildung mit Werten in H ist,

$$H\times H \to H,\ (a,b) \mapsto a\cdot b,$$

und die Menge H ist zusammen mit dieser Abbildung eine Gruppe ist.

16.01.13 (Mi)19.30 Uhr:
Im Audomax (Augusteum) läuft der Film (Eintriff frei)
"Colors of Math"
von E. Eremenko (Mathe-Diplom in Moskau 1990).
Regisseurin antwortet im Anschluß auf Fragen

Bemerkungen

(i) Sind G und H Gruppen, so wird die Relation $H\subseteq G$ im allgemeinenen sogar bedeuten, daß H eine Untergruppe von G ist. Im Zweifelsfall werden wir zusätzlich darauf hinweisen, ob es sich um ein Enthaltensein als Mengen oder als Gruppen handelt.

(ii) Alternativ kann man den Begriff der Untergruppe wie folgt definieren: Eine Untergruppe der Gruppe G ist eine Gruppe H mit folgenden Eigenschaften.

1. Als Menge ist H in G enthalten, $H\subseteq G$.

2. Die natürliche Einbettung $H \hookrightarrow G$, $h \mapsto h$, die jedes Element auf sich selbst abbildet, ist ein Homomorphismus.

4.1.3.2 Beispiel: Abbildungen und lineare Abbildungen

Seien

$M \quad := \mathbb{R}^n = \mathbb{R}^{n\times 1}$ der Vektorraum der n-zeiligen reellen Spaltenvektoren

G := S(M) die Gruppe der bijektiven Abbildungen $M\to M$

H := $\{f_A:M\to M \mid A\in GL(n, \mathbb{R})\}$

die Gruppe der linearen Abbildungen $M \longrightarrow M$

Dann ist H eine Untergruppe von G.

Beweis. Die identifikation von H mit der Gruppe der linearen bijektiven Abbildung M $\longrightarrow$ M ergibt sich aus dem kommutativen Diagramm 3.4.2 und der Bemerkung zu 3.4.5, nach welcher die umkehrbaren Matrizen gerade den umkehrbaren linearen Abbildungen entsprechen.

Insbesondere ist H ein Gruppe (die isomorph ist zu $GL(n, \mathbb{R})$).

QED.

4.1.3 .3 Untergruppenkriterium

Seien G eine Gruppe und $H\subseteq G$ eine Teilmenge von G. Dann sind die folgenden Aussagen äquivalent.

(i) H ist eine Untergruppe von G.

(ii) (a) Das Einselement von $e\in G$ liegt in H: $e\in H$.

(b) Das Produkt von je zwei Elementen von H liegt in H: $a,b\in H \Rightarrow ab\in H$.

(c) Das Inverse jedes Elements von H liegt in H: $a\in H \Rightarrow a^{-1}\in H$

(iii) H ist nicht leer und mit $a,b\in H$ gilt stets $ab^{-1}\in H$.

Beweis. <u>(i) ⇒ (iii)</u>. trivial.

<u>(iii) ⇒ (ii)</u>. Da H nicht leer ist, gibt es ein $a\in H$, welches in H liegt,

$$a\in H.$$

Dann liegt aber auch der “Quotient” von a und a in H,

$$e = aa^{-1} \in H,$$

d.h. die erste Bedingung von (ii) ist erfüllt. Wegen $e\in H$ liegt für jedes $a\in H$ auch der Quotient

$$a^{-1} = ea^{-1}\in H$$

in H. Mit anderen Worten, die dritte Bedingung von (ii) ist erfüllt. Seien schließlich $a,b\in H$ beliebige Elemente. Dann gilt auf Grund der bereits bewiesenen Aussage (ii)(c) auch

$$b^{-1}\in H.$$

Mit a, $b^{-1}\in H$ ist aber auf Grund der Voraussetzung (iii) auch der folgende Quotient eine Element von H,

$$H\ni a\cdot(b^{-1})^{-1} = ab.$$

Wir haben gezeigt, die dritte Bedingung von (ii) ist erfüllt.

<u>(ii) ⇒ (i)</u> r haben zu zeigen, eine Teilmenge H von G, die den Bedingungen (ii) genügt ist mit der Multiplikation von G eine Gruppe. Es gilt:

1. Da für die Elemente von G das Assoziativgesetz gilt, gilt es auch für die von H.
2. Nach Bedingung (ii)(b) definiert die Multiplikation von G eine Abbildung $H\times H\to H$ mit Werten in H.
3. Nach Bedingung (ii)(a) besitzt H ein Element mit den Eigenschaften des Einselements (nämlich e).
4. Nach Bedingung (ii)(c) besitzt jedes Element von H ein (in H liegendes))Inverses.

Mit anderen Worten, H ist mit der Multiplikation von G eine Gruppe, d.h. H ist eine Untergruppe von G.

QED.

4.1.3.4 Eine Anwendung des Untergruppenkriteriums: die alternierde Gruppe

Sei n eine natürliche Zahl. Dann ist die Menge

$$A_n := \{f\in S_n \mid \mathrm{sign}(f) = +1\}$$

der geraden Permutationen von S_n eine Untergruppe von S_n. Diese Gruppe heißt alternierende Gruppe der n-elementigen Menge $\{1,...,n\}$.

Beweis. Wir zeigen, A_n genügt den Bedingungen (ii) von 5.1.3.4.Das Vorzeichen der identischen Permutation ist positiv (weil die Anzahl der Inversionen gleich Null ist),

$$\mathrm{Id}\in A_n.$$

Seien $f,g \in A_n$ zwei gerade Permutationen. Dann gilt

$$\mathrm{sign}(f\circ g) = \mathrm{sign}(f)\cdot\mathrm{sign}(g) = (+1)\cdot(+1) = +1,$$

d.h. $f\circ g\in A_n$. Schließlich gilt

$$\mathrm{sign}(f)\cdot\mathrm{sign}(f^{-1}) = \mathrm{sign}(f\circ f^{-1}) = \mathrm{sign}(\mathrm{Id}) = 1,$$

also

$$\mathrm{sign}(f^{-1}) = \mathrm{sign}(f).$$

Mit f liegt also auch f^{-1} in A_n.

QED.

4.2 Elementarmatrizen

4.2.1 Bezeichnungen

Wir führen folgende Bezeichnungen für quadratische Matrizen aus $K^{n\times n}$ ein.

$$E_{ij} := \begin{array}{cc} \begin{array}{c} \quad\quad\quad j \end{array} & \\ \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} & \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} \end{array}$$

Bezeichne E_{ij} die Matrix, deren Eintrag in der Position (i,j) gleich 1 ist, und deren sämtliche anderen Einträge Null sind.

$$M_i(c) := \mathrm{Id} + (c-1)E_{ii} = \begin{array}{cc} \begin{array}{c} \quad\quad\quad i \end{array} & \\ \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & c & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} & \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} \end{array}$$

Die <u>Multiplikationsmatrix</u> $M_i(c)$ sei die <u>Diagonalmatrix</u>[19], deren Eintrag in der Position (i,i) gleich c ist, und deren übrige Hauptdiagonal-Einträge gleich Eins sind.

---

[19] d.h. die einzigen von Null verschiedenen Einträge befinden sich auf auf der Hauptdiagonalen.

$$Q_{ij}(c) = \mathrm{Id} + c \cdot E_{ij} = \begin{array}{cl} \begin{array}{ccccc} & & & j & \end{array} & \\ \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & c & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} & \begin{array}{l} \\ i \\ \\ \\ \\ \\ \\ \end{array} \end{array} \quad (i \neq j)$$

sei die Matrix, deren Hauptdiagonaleinträge sämtlich 1 sind, deren Eintrag in der Position (i,j) gleich c ist, und deren übrige Einträge Null sind.

$$P_{ij} = \mathrm{Id} - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$

Die <u>Permutationsmatrix</u> P_{ij} sei diejenige Matrix, die aus der Einheitsmatrix durch Vertrauschen von i-ter und j-ter Spalte entsteht.

4.2.2 Definition

Die Matrizen der Gestalt

$$M_i(c), Q_{ij}(c), P_{ij} \in K^{n\times n} \text{ mit } c \in K^*,\ i \neq j,$$

heißen <u>Elementarmatrizen</u>.

4.2.3 Elementarmatrizen und elementare Umformungen

Die Elementarmatrizen zeichnen sich dadurch aus, daß man mit ihrer Hilfe die üblichen Zeilen- und Spaltenoperationen als Multiplikation mit geeigneten Matrizen beschreiben kann.

<u>Multiplikation der i-ten Spalte mit c</u>:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto A \cdot M_i(c).$$

<u>Addition des c-fachen der i-ten Spalte zur j-ten Spalte</u>:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto A \cdot Q_{ij}(c).$$

<u>Vertauschen von i-ter und j-ter Spalte</u>:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto A \cdot P_{ij}.$$

Durch Multiplikation von links erhält man die entsprechenden Zeilenoperationen.

<u>Multiplikation der i-ten Zeile mit c</u>:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto M_i(c) \cdot A.$$

<u>Addition des c-fachen der j-ten Zeile zur i-ten Zeile</u>:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto Q_{ij}(c) \cdot A.$$

<u>Vertauschen von i-ter und j-ter Zeile</u>:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto P_{ij}\cdot A.$$

4.2.4 Eigenschaften von Elementarmatrizen

(i) Die Elementarmatrizen $Q_{ij}(c)$ und P_{ij} sind Produkte von Elementarmatrizen der Gestalt $M_i(c)$ und $Q_{ij}(1)$. Genauer gilt:

$$Q_{ij}(c) = M_j(\tfrac{1}{c})Q_{ij}(1)M_j(c)$$
$$P_{ij} = Q_{ji}(1)Q_{ij}(-1)Q_{ji}(1)M_j(-1).$$

(ii) Die Elementarmatrizen sind umkehrbar und ihre Inversen sind wieder Elementarmatrizen. Genauer gilt:

$$M_i(c)^{-1} = M_i(\tfrac{1}{c})$$
$$Q_{ij}(c)^{-1} = Q_{ij}(-c)$$
$$P_{ij}^{-1} = P_{ij}$$

(iii) Jede umkehrbare Matrix ist das Produkt von endlich vielen Elementarmatrizen.

Beweis. Beweis-Idee für (i): Man multipliziere die Einheitsmatrix (zum Beispiel von Rechts) mit den beiden Seiten der behaupteten Identitäten und zeige mit Hilfe der Beschreibungen von 4.2.3, daß das Ergebnis in beiden Fällen dasselbe ist.
Beweis-Idee für (ii): Man untersuche, wie man die Wirkung der Multiplikation mit den betrachteten Matrizen wieder rückgängig machen kann.

Zu (i). Die Identitäten überprüft man zum Beispiel durch direktes Ausrechnen. Alternativ kann man auch die zu den Produkten gehörigen Zeilen- oder Spaltenoperationen ermitteln. Zum Beispiel hat man für die Matrix

$$A = (\dots,a_i,\dots,a_j,\dots)$$

mit den Spaltenvektoren $\dots,a_i,\dots,a_j,\dots$ die folgende Situation:

$$AQ_{ji}(1) = (\dots,a_i+a_j,\dots,a_j,\dots)$$
$$AQ_{ji}(1)Q_{ij}(-1) = (\dots,a_i+a_j,\dots,-\ a_i,\dots)$$
$$AQ_{ji}(1)Q_{ij}(-1)Q_{ji}(1) = (\dots,a_j,\dots,-\ a_i,\dots)$$

Beachtet man noch, wie die Multiplikation von rechts mit)$M_j(-1)$ wirkt, so ergibt sich die zweite der zu beweisenden Identitäten.
Weiter erhält man

$$A\,M_j(\tfrac{1}{c}) = (\dots,a_i,\dots,\tfrac{1}{c}\cdot a_j,\dots)$$
$$A\,M_j(\tfrac{1}{c})Q_{ij}(1) = (\dots,a_i,\dots,\tfrac{1}{c}\cdot a_j + a_i,\dots)$$
$$A\,M_j(\tfrac{1}{c})Q_{ij}(1)\,M_j(c) = (\dots,a_i,\dots,a_j + c\cdot a_i,\dots)$$

Damit gilt auch die erste Identität.
Zu (ii). Die Identitäten ergeben sich ebenfalls direkt aus den zu den Matrizen gehörigen Operationen. Die Multiplikation der i-ten Spalte mit c kann man wieder rückgängig machen, indem man anschließend mit $\frac{1}{c}$ multipliziert. Die Addition des c-fachen der j-ten

Spalte zur i-ten macht man rückgängig, indem man das c-fach dieser Spalte anschließend wieder abzieht. Schließlich erhält man durch zweimaliges Vertauschen von i-ter und j-ter Spalte die Ausgangsmatrix.
Zu (iii) i eine umkehrbare Matrix. Durch elementare Zeilen- und Spaltenoperationen läßt sich A in eine obere Dreieckmatrix überführen, d.h. es gibt eine obere Dreiecksmatrix

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

und Elementarmatrizen $B_1,...,B_r$, $C_1,...,C_s$ mit

$$A' = B_1 \cdot\ldots\cdot B_r \bullet A \bullet C_1 \cdot\ldots\cdot C_s$$

Da das Produkt umkehrbarer Matrizen umkehrbar ist, ist mit A auch A' umkehrbar. Also hat A' den Rang n. Insbesondere ist

(1) $$a_{11} \neq 0,$$

denn anderenfalls wären die Spalten von A' linear abhängig. Wegen (1) können wir aber durch weitere elementare Spaltenoperationen erreichen, daß a_{11} der einzige Eintrag $\neq 0$ in der ersten Zeile ist. Dann gilt aber

$$a_{22} \neq 0.$$

Indem wir mit dieser Argumentation fortfahren, erreichen wir, daß A' Diagonalgestalt bekommt, wobei sämtliche Einträge der Hauptdiagonalen $\neq 0$ sind. Durch weitere Multiplikation mit Multiplikationsmatrizen erreichen wir schließlich A' = Id. Es gilt also Elementarmatrizen $B_1,...,B_r$, $C_1,...,C_s$ mit

$$Id = B_1 \cdot\ldots\cdot B_r \bullet A \bullet C_1 \cdot\ldots\cdot C_s .$$

Wir mutiplizieren von links nacheinander mit den Inversen von B_1, ..., B_r und von rechts mit den Inversen von C_s, ... , C_1 und erhalten

$$B_r^{-1} \bullet\ldots\bullet B_1^{-1} \bullet C_s^{-1} \cdot C_{s-1}^{-1} \bullet \ldots \bullet C_1^{-1} = A.$$

Da die Inversen von Elementarmatrizen wieder Elementarmatrizen sind, haben wir damit A als Produkt von Elementarmatrizen dargestellt.
QED.
Bemerkung
Durch geeignete Interpretation der letzten Identität im Beweis von (iii) sieht man, daß man umkehrbare Matrizen A allein durch Zeilenoperationen bzw. allein durch Spaltenoperationen in die Einheitsmatrix überführen kann.

4.3 Die Determinanten-Definition von Leibniz

4.3.1 Definition

Sei

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in K^{n\times n}$$

eine quadratische n-reihige Matrix mit Einträgen aus dem Körper K. Dann heißt die Zahl

$$\det A := |A| := \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} := \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\cdot a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)}$$

Determinante von A. Die Summe wird dabei über sämtliche Permutationen σ der Zahlen 1,...,n erstreckt. Der Faktor sign(σ) bezeichne das Vorzeichen der Permutation σ.

Bemerkung
Das Produkt

$$a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)}$$

unter dem Summenzeichen enthält aus jeder Zeile und jeder Spalte der Matrix genau einen Faktor. Die Determinante ist gerade die vorzeichenbehaftete Summen aller Produkte, die man auf diese Weise bilden kann.

4.3.2 Die Determinante der transponierten Matrix

Sei

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in K^{n\times n}$$

eine quadratische n-reihige Matrix mit Einträgen aus dem Körper K. Dann gilt

$$\det A = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\cdot a_{\sigma(1)1}\cdot a_{\sigma(2)2}\cdot\ldots\cdot a_{\sigma(n)n}.$$

Mit anderen Worten, die Determinante von A stimmt mit der der Transponierten überein,

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

Beweis von (ii).
Die Reihenfolge der Faktoren unter der Summe der Determinantenformel ist unwesentlich für den Wert der Determinante. Sind also $i_1,\ldots,i_n$ die Zahlen 1,...,n in irgendeiner Reihenfolge, so gilt,

$$a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)} = a_{i_1\sigma(i_1)}\cdot a_{i_2\sigma(i_2)}\cdot\ldots\cdot a_{i_n\sigma(i_n)}$$

Mit anderen Worten, für jede fest vorgegebene Permutation $\tau\in S_n$ gilt

$$a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)} = a_{\tau(1)\sigma(\tau(1))}\cdot a_{\tau(2)\sigma(\tau(2))}\cdot\ldots\cdot a_{\tau(n)\sigma(\tau(n))}.$$

Setzt man speziell $\tau = \sigma^{-1}$, so erhält man

$$a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)} = a_{\sigma^{-1}(1)1}\cdot a_{\sigma^{-1}(2)2}\cdot\ldots\cdot a_{\sigma^{-1}(n)n}$$

Die Determinantenformel bekommt damit die Gestalt

$$\det A = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma)\cdot a_{\sigma^{-1}(1)1}\cdot a_{\sigma^{-1}(2)2}\cdot\ldots\cdot a_{\sigma^{-1}(n)n}.$$

Nun durchläuft mit σ auch σ^{-1} die Gruppe S_n. Wir können in der letzten Identität σ durch σ^{-1} ersetzen und erhalten

$$\det A = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma^{-1})\cdot a_{\sigma(1)1}\cdot a_{\sigma(2)2}\cdot\ldots\cdot a_{\sigma(n)n}.$$

Weiter gilt

$$\mathrm{sign}(\sigma)\bullet\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma\circ\sigma^{-1}) = \mathrm{sign}(\mathrm{Id}) = 1.$$

Durch Multiplikation beider Seiten mit $\mathrm{sign}(\sigma)$ erhalten wir wegen $\mathrm{sign}(\sigma)^2 = 1$ schließlich

$$\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma).$$

Durch einsetzen ergibt sich die behauptete Identität.
QED.

4.3.3 Die Determinante einer 2×2-Matrix

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

Beweis. Da S_2 aus zwei Elementen besteht, kommen in der Determinantenformen zwei Summand vor. Diese Summanden sind bis aufs Vorzeichen[20]

ad und bc.

In S_2 gibt es eine gerade und eine ungerade Permutation (da es von jeder Sorte Permutationen gleichviele gibt)[21], d.h. genau ein Summand hat ein negatives Vorzeichen. Da ad zur identischen Permutation gehört, ist das Vorzeichen zu ad positiv. Daraus ergibt sich die behauptete Formel..
QED.
Alternativer Beweis. Es ist

$$S_n = \{\mathrm{Id}, (1,2)\},$$

also

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} = \mathrm{sign}\ \mathrm{Id}\bullet a_{1,\mathrm{Id}(1)}a_{2,\mathrm{Id}(2)} + \mathrm{sign}\ (1,2)\bullet a_{1,(1,2)(1)}a_{2,(1,2)(2)}$$
$$= a_{11}a_{22} - a_{12}a_{21}$$

[20] Wenn a in einem Produkt vorkommt, so muß der andere Faktor aus der zweiten Zeile und zweiten Spalte kommen, also gleich d sein. Analog sieht man, daß auch bc (bis aufs Vorzeichen) ein möglicher Summand ist. Da nur zwei Summanden vorkommen, sind damit alle Möglichkeit erfaßt.

[21] Die Abbildung

$$S_n \longrightarrow S_n,\ \sigma \mapsto \sigma\circ(1,2),$$

ist eine zu sich selbst inverse Abbildung und als solche bijektiv. Sie überführt gerade in ungerade und ungerade in gerade Permutationen.

QED.

4.3.4 Die Determinante einer 3×3-Matrix (Sarrussche Regel)

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - gec - hfa - idb. \qquad (1)$$

Um sich diese Formel zu merken, schreibe man ein zweites Exemplar der Matrix neben die Ausgangsmatrix.

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

In der entstehenden 3×6-Matrix gibt es drei zur Hauptdiagonale parallele Diagonalen und entsprechend drei zur Nebendiagonale parallele. Jeder dieser Diagonalen entspricht ein Produkt von drei Einträgen der Ausgangsmatrix. Man versehe die zur Hauptdiagonalen gehörigen Produkte mit dem positiven und die übrigen mit dem negativen Vorzeichen und bilde die Summe. Das Ergebnis ist die Determinante der Ausgangsmatrix.
Formel (1) heißt <u>Sarrussche Regel.</u>
Beweis. Die symmetrische Gruppe S_3 besitzt $3! = 1\bullet 2\bullet 3 = 6$ Elemente. Zu diesen Elementen gehören die identische Abbildung, welche im Kontext von Permutationen auch mit

$$(1)$$

bezeichnet wird, die Transpositionen

$$(1\ 2), (1\ 3), (2\ 3)$$

und die Dreierzyklen

$$(1\ 2\ 3), (3\ 2\ 1).$$

Dies sind bereits sechs Elemente, d.h. es gilt

$$S_3 = \{(1), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (3\ 2\ 1)\}$$

Das Vorzeichen der drei Transpositionen ist -1. Da die Anzahl der geraden Permutationen gleich der Anzahl der ungeraden ist, haben alle anderen Permutationen von S_3 das Vorzeichen -1:

$$\begin{aligned} \operatorname{sign}(1) &= 1, \\ \operatorname{sign}(1\ 2\ 3) &= 1 \\ \operatorname{sign}(3\ 2\ 1) &= 1 \\ \operatorname{sign}(1\ 2) &= -1 \\ \operatorname{sign}(1\ 3) &= -1 \\ \operatorname{sign}(2\ 3) &= -1 \end{aligned}$$

Die Determinanten-Formel von Leibniz bekommt damit die oben angegebene Gestalt.

QED.
Bemerkungen
(i) Die Formel für die Determinante einer mehr als dreireihigen Matrix läßt sich nicht in ähnlicher Weise vereinfachen wie in den Fällen n=2 und n=3.
(ii) Unser nächstes Ziel besteht im Beweis von Umformungsregeln für Determinanten, die es uns ermöglichen, diese zu berechnen.

4.4 Eigenschaften der Determinante

4.4.1 Linearität in jeder Zeile und Spalte

(i) Seien $a_1,...,a_{i-1},x,a_{i+1},...,a_n \in K^{n\times 1}$ Spaltenvektoren,

$$A = (a_1,...,a_{i-1},x,a_{i+1},...,a_n)$$

die Matrix, deren Spalten gerade diese Vektoren sind und

$$f(x) := \det (a_1,...,a_{i-1},x,a_{i+1},...,a_n)$$

die Determinante dieser Matrix. Dann gilt

$$f(c'x' + c''x'') = c'\cdot f(x') + c''\cdot f(x'')$$

für beliebige $c',c''\in K$ und $x',x''\in K^{n\times 1}$. Mit anderen Worten, die Determinate det A einer Matrix ist eine lineare Funktion der i-ten Spalte von A (für i = 1,...,n).

(ii) Seien $b_1,...,b_{i-1},y,b_{i+1},...,b_n \in K^{1\times n}$ Zeilenvektoren,

$$B = \begin{pmatrix} b_1 \\ \dots \\ b_{i-1} \\ y \\ b_{i+1} \\ \dots \\ b_n \end{pmatrix}$$

die Matrix, deren Zeilen gerade diese Vektoren sind und

$$g(y) := \det \begin{pmatrix} b_1 \\ \dots \\ b_{i-1} \\ y \\ b_{i+1} \\ \dots \\ b_n \end{pmatrix}$$

die Determinante dieser Matrix. Dann gilt

$$g(c'y' + c''y'') = c'\cdot g(y') + c''\cdot g(y'')$$

für beliebige $c',c''\in K$ und $y',y''\in K^{n\times 1}$. Mit anderen Worten, die Determinate det A einer Matrix ist eine lineare Funktion der i-ten Zeile von A (für i = 1,...,n).

Beweis. Zu (i). Bezeichne $a_{\alpha\beta}$ den Eintrag von A in der Position (α,β) mit $\alpha\neq i$ und x_j die j-te Koordinate des Vektors x. Auf Grund der Determinanten-Definitoin gilt dann

$f(x) = \det A$

$$= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\cdot a_{\sigma(1)1}\cdot a_{\sigma(2)2}\cdot\ldots\cdot a_{\sigma(i-1)i-1}\cdot x_{\sigma(i)}\cdot a_{\sigma(i+1)i+1}\cdot\ldots\cdot a_{\sigma(n)n}$$

$$= \sum_{\sigma\in S_n} c_{\sigma,i}\cdot x_{\sigma(i)}$$

mit festen von x unabhängigen $c_{\sigma,i}\in K$. Der letzte Ausdruck ist offensichtlich linear in den Koordinaten des Vektors x (die ihrerseits lineare Funktionen von x sind).

Zu (ii). Wenn A sämtliche Matrizen von $K^{n\times n}$ durchläuft, so gilt dasselbe auch für A^T. Wir könne also annehmen $B = A^T$. Dann gilt $b_i = a_i^T$ für jedes i und $y=x^T$, also

$$g(y) = \det B = \det A = f(x) = f(y^T).$$

Mit anderen Worten, g ist die Zusammensetzung der linearen Abbildungen

$$y \mapsto y^T \text{ und } x \mapsto f(x)$$

und als solche linear.
QED.

4.4.2 Verhalten beim Vertauschen von Zeilen bzw. Spalten

(i) Sei A eine quadratische Matrix mit den Spalten $a_1,...,a_n\in K^{n\times 1}$. Dann besteht für jede Permutation $\sigma\in S_n$ die Relation

$$\det (a_{\sigma(1)},...,a_{\sigma(n)}) = \operatorname{sign}(\sigma)\cdot\det(a_1,...,a_n).$$

Insbesondere wechselt die Determinante beim Vertauschen zweier Spalten das Vorzeichen.

(ii) Sei A eine quadratische Matrix mit den Zeilen $a^1,...,an\in K^{1\times n}$. Dann besteht für jede Permutation $\sigma\in S_n$ die Relation

$$\det\begin{pmatrix} a^{\sigma(1)} \\ \dots \\ a^{\sigma(n)} \end{pmatrix} = \operatorname{sign}(\sigma)\cdot\begin{pmatrix} a^1 \\ \dots \\ a^n \end{pmatrix}.$$

Insbesondere wechselt die Determinante beim Vertauschen zweier Zeilen das Vorzeichen.

Beweis. Zu (i). 1. Schritt: Reduktion auf den Fall, daß σ ein Nachbartausch ist.

Jede Permutation ist ein Produkt von Transpositionen. Wir können die veränderte Reihenfolge der Spalten also dadurch erreichen, daß wir nacheinander Vertauschungen von Spalten ausführen. Bei solchen wiederholten Vertauschungen multiplizieren sich deren Vorzeichen, d.h. man erhält insgesamt das Vorzeichen der gegebenen Permutation. Wir können also annehmen,, σ ist eine Transposition, σ =(u,v). Weiter wissen wir , jede Transposition ist ein Produkt von Nachbartauschen. Wir können also sogar annehmen, daß σ von der Gestalt

$$\sigma = (u,u+1)$$

ist.

2. Schritt: der Fall daß σ ein Nachbartausch ist.

Wir haben zu zeigen, beim Vertauschen zweier benachbarter Spalten ändert die Determinante nur das Vorzeichen. Bezeichne

$$A' = (a'_{ij})$$

die Matrix, die aus A durch Vertauschen von u-ter und (u+1)-ter Spalte entsteht, d.h.

$$a'_{ij} := \begin{cases} a_{i,u+1} & \text{falls } j = u \\ a_{i,u} & \text{falls } j=u+1 \\ a_{ij} & \text{sonst} \end{cases}$$

Wir haben zu zeigen,

$$\det(A') = \det(A).$$

Um die Abhängigkeit der Einträge vom Spaltenindex besser zu erkennen, benutzen wir die Formel für die transponierte Matrix. Es gilt

$$\begin{aligned} \det(A') &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\cdot a'_{\sigma(1)1}\cdot\ldots\cdot a'_{\sigma(u)u}\cdot a'_{\sigma(u+1)u+1}\cdot\ldots\cdot a'_{\sigma(n)n} \\ &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(u)u+1}\cdot a_{\sigma(u+1)u}\cdot\ldots\cdot a_{\sigma(n)n} \end{aligned}$$

Den Ausdruck unter der Summe kann man auch in der folgenden Gestalt schreiben.

$$\operatorname{sign}(\sigma)\cdot a_{\tau(1)1}\cdot\ldots\cdot a_{\tau(u)u}\cdot a_{\tau(u+1)u+1}\cdot\ldots\cdot a_{\tau(n)n}$$

mit

$$\tau(i) := \begin{cases} \sigma(u+1) & \text{für } i=u \\ \sigma(u) & \text{für } i=u+1 \\ \sigma(i) & \text{sonst} \end{cases}$$

Mit anderen Worten, τ ist die Zusammensetzung von σ mit der Transposition (u,u+1),

$$\tau = \sigma\circ(u,u+1).$$

Wenn σ die gesamte Gruppe S_n durchläuft, so gilt dasselbe auch für τ.[22] Wir können die Summe also auch schreiben als

$$\begin{aligned} \det(A') &= \sum_{\tau\in S_n} \operatorname{sign}(\tau\circ(u,u+1)^{-1})\cdot a_{\tau(1)1}\cdot\ldots\cdot a_{\tau(u)u}\cdot a_{\tau(u+1)u+1}\cdot\ldots\cdot a_{\tau(n)n} \\ &= \operatorname{sign}((u,u+1)^{-1})\cdot \sum_{\tau\in S_n} \operatorname{sign}(\tau)\cdot a_{\tau(1)1}\cdot\ldots\cdot a_{\tau(u)u}\cdot a_{\tau(u+1)u+1}\cdot\ldots\cdot a_{\tau(n)n} \\ &= \operatorname{sign}((u,u+1)^{-1})\cdot\det(A). \end{aligned}$$

Schließlich ist das Inverse der Transposition (u,u+1) gleich der Transposition (u,u+1) selbst und hat insbesondere das Vorzeichen -1.

Zu (ii). Da beim Transponieren Zeilen in Spalten und Spalten in Zeilen übergehen, die Determinante sich jedoch nicht ändert, folgt die Behauptung von (ii) aus (i). Explizit erhalten wir

---

[22] Sie Abbildung

$$S_n \longrightarrow S_n,\ f \mapsto f\circ(u,\ u+1),$$

ist wohldefiniert und zu sich selbst invers, also insbesondere bijektiv.

$$\det \begin{pmatrix} a^{\sigma(1)} \\ \dots \\ a^{\sigma(n)} \end{pmatrix} = \det \begin{pmatrix} a^{\sigma(1)} \\ \dots \\ a^{\sigma(n)} \end{pmatrix}^T$$

$$= \det (a^{\sigma(1)T},...,a^{\sigma(n)T})$$

$$= \mathrm{sign}(\sigma)\cdot\det (a^{1T},...,a^{nT})$$

$$= \mathrm{sign}(\sigma)\cdot\det (a^{1T},...,a^{nT})^T$$

$$= \mathrm{sign}(\sigma)\cdot \begin{pmatrix} a^1 \\ \dots \\ a^n \end{pmatrix}$$

Man kann die Aussage aber auch in derselben Weise wie (i) beweisen.
QED.

Bemerkungen

(i) Matrizen mit zwei gleichen Zeilen oder Spalten. Sei $A \in K^{n\times n}$ eine Matrix mit zwei gleichen Zeilen oder Spalten. Beim Vertauschen dieser Zeilen oder Spalten ändert sich das Vorzeichen der Determinante. Gleichzeitig bleibt die Matrix dabei unverändert. Es gilt also in diesem Fall

$$\det(A) = - \det (A),$$

d.h.

$$2\bullet\det(A) = 0.$$

Ist K ein Körper, in welchem

$$2 = 1 + 1$$

von Null verschieden ist, so können wir mit dem Inversen von 2 multiplizieren, und erhalten

$$\det (A) = 0.$$

(ii) Im Körper $K = \mathbb{F}_2$ mit zwei Elementen gilt

$$2 = 1 + 1 = 0,$$

d.h. die obige Argumentation funktioniert nicht für diesen Körper. Ein Körper K mit der Eigenschaft

$$2\bullet x = 0 \text{ für jedes } x \in K$$

heißt Körper der Charakteristik 2.

(iii) Charakteristik eines Körpers. Die Charakteristik eines Körpers K ist die kleinste natürlichen Zahl n (≥2) mit der Eigenschaft

$$n\bullet x = 0 \text{ für jedes } x \in K$$

oder sie ist gleich 0, falls es keine solche natürliche Zahl n gibt. Die Charatkeristik von K wird mit

$$\mathrm{char}(K)$$

bezeichnt.

(iv) Die Charakteristik des Körpers mit p Elementen ist p,

$$\mathrm{char}(\mathbb{F}_p) = p.$$

Die Charakteristik der Körper der rationalen, reellen und komplexen Zahlen ist Null,

$$\mathrm{char}(\mathbb{Q}) = \mathrm{char}(\mathbb{R}) = \mathrm{char}(\mathbb{C}) = 0.$$

(v) Ist die Charakteristik des Körpers K von Null veschieden,

$$\mathrm{char}(K) = m \neq 0,$$

so ist diese Charakteristik stets ein Primzahl.

char(K) ist eine Primzahl.

Ist nämlich

$$m = p_1 \bullet \ldots \bullet p_r$$

die Zerlegung der natürlichen Zahl in ein Produkt von Primfaktoren, so gilt in K

$$0 = m \bullet 1_K = (p_1 \bullet 1_K) \bullet \ldots \bullet (p \bullet 1_K),$$

wobei 1_K das Einselement von K bezeichne.Weil K ein Körper ist, muß einer der Faktoren rechts gleich Null sein, sagen wir

$$p_i \bullet 1_K = 0.$$

Dann ist aber $p_i \bullet K = 0$ und p_i ist die Charakteristik von K.

(vi) Die Aussage von (i) ist auch für den Fall richtig, daß K ein Körper der Charakteristik 2 ist.

Beweisskitze* für Aussage (vi).
Bezeichnen wir mit $X = (X_{ij})$ eine Matrix vom selben Typ wie A, deren Einträge Unbestimmte sind, wobei die Zeilen bzw. Spalten, die bei A übereinstimmen, auch bei X übereinstimmen sollen. Die nach der Leibniz-Formel definierte Determinante

$$\det(X)$$

ist dann ein Polynom in den Unbestimmten X_{ij} mit ganzzahligen Koeffizienten, d.h. det(X) ist ein Element des Rings

$$R := \mathbb{Z}[X_{ij} \mid i, j = 1, \ldots , n]$$

der Polynome in den X_{ij} mit Koeffizienten aus $\mathbb{Z}$. Dieser Ring ist enthalten im Körper

$$L := \mathbb{Q}(X_{ij} \mid i, j = 1, \ldots , n)$$

der rationalen Funktionen in den X_{ij} mit Koeffizienten aus $\mathbb{Q}$. Der Körper L hat die Charakteristik 0,

$$\mathrm{char}(L) = 0$$

und X ist eine Matrix mit Einträgen aus L. Weil X zwei gleiche Zeilen bzw. Spalten hat, ist die Determinante von X nach (i) gleich 0,

$$\det X = 0 \text{ in } L.$$

Nun liegt aber det X auch in R, d.h. es gilt

$$\det X = 0 \text{ in } R.$$

Betrachten wir jetzt die folgende Abbildung

$$h: R \longrightarrow K,$$

Zur Beschreibung des Bildes eines Polynom $p \in R$ schreiben wir dieses in in der Gestalt

$$p = \sum_{\nu} g_\nu \mu_\nu(\ldots, X_{ij}, \ldots)$$

mit ganzen Zahlen g_ν und Produkten μ_ν der Unbestimmten X_{ij}. Das Bild von p bei,

$$h(p) := \sum_\nu (g_\nu \cdot 1_K)\mu_\nu(\ldots,a_{ij},\ldots),$$

entstehe dann aus p, indem man jeden der Koeffizienten g_ν durch des Element $g_\nu \cdot 1_K$ des Körpers K ersetzt und jeden Eintrag X_{ij} von X durch den entsprechenden Eintrag $a_{ij} \in K$ der Matrix A. Die Abbildung h ist dann ein Homomorphismus von Ringen mit 1. Deshalt gilt

$$0 = h(0) = h(\det(X_{ij})) = \det(h(X_{ij})) = \det(a_{ij}) = \det A.$$

QED.

4.4.3 Verhalten bei elementaren Operationen

Die Determinanten der quadratischen Matrix A ändert sich nicht, wenn man ein Vielfaches einer Spalte zu einer anderen Spalte addiert.

(1) $\det(\ldots,a_i,\ldots,a_j,\ldots) = \det(\ldots,a_i+c\cdot a_j,\ldots,a_j,\ldots)$

Die anaolge Aussage gilt auch für die Zeilen,

(2) $\det(\ldots,a_i,\ldots,a_j,\ldots)^T = \det(\ldots,a_i+c\cdot a_j,\ldots,a_j,\ldots)^T$

Beweis. Da die Determinante eine lineare Funktion der i-ten Spalte ist, gilt

$$\det(\ldots,a_i+c\cdot a_j,\ldots,a_j,\ldots) = \det(\ldots,a_i,\ldots,a_j,\ldots) + c\cdot\det(\ldots,a_j,\ldots,a_j,\ldots).$$

Die zweite Determinanten rechts hat zwei gleiche Zeilen, ist also Null, d.h. es gilt (1). Formel (2) gilt damit natürlich auch, da sich die Determinante beim Transponieren nicht ändert.
QED.

4.4.4 Die Determinante einer Diagonalmatrix

(i) Die Determinante einer oberen Dreiecksmatrix ist gleich dem Produkt der Elemente auf der Hauptdiagonalen,

$$\det\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ 0 & \ldots & a_{nn} \end{pmatrix} = a_{11}\cdot\ldots\cdot a_{nn}$$

(ii) Allgemeiner, zerfällt die Matrix A wir folgt in Blöcke,

$$A = \begin{pmatrix} U_1 & \ldots & * \\ \ldots & \ldots & \ldots \\ 0 & \ldots & U_n \end{pmatrix}$$

mit quadratischen Matrizen U_i auf der Hauptdiagonalen, so gilt

$$\det A = \det U_1 \cdot\ldots\cdot \det U_n$$

Beweis. Zu (i). Ein Eintrag a_{ij} der Matrix

$$A := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ 0 & .. & a_{nn} \end{pmatrix}$$

ist höchstens dann ungleich Null, wenn

$$i \le j$$

gilt. Der Summand

$$\mathrm{sign}(\sigma)\bullet a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)}$$

in der Leibniz-Formel für det A ist höchstens dann ungleich Null, wenn gilt

$$1 \le \sigma(1),\ 2 \le \sigma(2),\ \ldots\ ,\ n\text{-}2 \le \sigma(n\text{-}2),\ n - 1 \le \sigma(n\text{-}1),\ n \le \sigma(n).$$

Aus der letzten Ungleichung folgt

$$\sigma(n) = n \tag{1}$$

da für $\sigma(n)$ nur die Werte 1, ... , n in Frage kommen. Wegen (1) kommen für $\sigma(n\text{-}1)$ nur die Werte 1, ... , n-1 in Frage. Aus der vorletzten Ungleichung ergibt sich damit

$$\sigma(n\text{-}1) = n\text{-}1.$$

Durch Wiederholen dieses Schlussen folgt

$$\sigma(i) = i$$

für alle i. In der Leibniz-Formel ist höchstens ein Summand ungleich Null, nämlich der zur identischen Permutation. Da letztere das Vorzeichen 1 hat, folgt die Behauptung.

<u>Zu (ii)</u> Durch elementare Zeilenumformungen der Matrix A vom Typ $Q_{ij}(c)$ und P_{ij} überführen wir jedes der U_i in obere Dreieckgestalt. Die Zahl der Vertauschungen, die wir dabei ausführen sei u_i. Diese Operationen führen wir für i = 1, ... , n durch. Die Block gestalt der Matrix bleibt davon unberührt, d.h. die neue Matrix hat die Gestalt

$$A' = \begin{pmatrix} U'_1 & \dots & * \\ \dots & \dots & \dots \\ 0 & .. & U'_n \end{pmatrix} \tag{2}$$

mit quadratischen Matrizen U'_i , wobei U'_i denselben Typ hat wie U_i und aus U_i durch elementare Zeilen-Operationen entsteht, von denen u_i Operationen Vertauschungen sind. Insbesondere ist

$$\det U'_i = (-1)^{u_i} \det U_i. \tag{3}$$

Dieselbe Argumentation läßt sich auf A' und A anwenden und zeigt,

$$\det A' = (-1)^u\bullet\det A \text{ mit } u = u_1+\ldots+u_n. \tag{4}$$

Die Matrizen A' , U_1, ... , U_n haben Diagonalgestalt. Ihre Determinanten sind nach (i) gleich dem Produkt ihrer Einträge auf der Hauptdiagonalen. Wegen (1) folgt

$$\det A' = \det U'_1\bullet\ldots\bullet U'_n$$

Wir multiplizieren beide Seiten dieser Identität mit

$$(-1)^u = (-1)^{u_1}\bullet\ldots\bullet(-1)^{u_n}$$

und erhalten wegen (4) und (3) die Behauptung.
QED.

4.4.5 Charakterisierung der Umkehrbarkeit einer Matrix

Für quadratische Matrizen $A \in K^{n\times n}$ sind folgende Aussagen äquivalent.
(i) A ist umkehrbar.
(ii) rk A = n.
(iii) det A ≠ 0.

Beweis. Die Äquivalenz von (i) und (ii) wurde bereits bewiesen. Es reicht also, die von (ii) und (iii) nachzuweisen.
Sei A eine beliebige quadratische Matrix mit den Spalten

$$a_1,\dots,a_n.$$

Der von den Spalten erzeugte Vektorraum bleibt unverändert, wenn man die Vektoren permutiert, einzelne Vektoren mit einem von Null verschiedenen Faktor multipliziert oder zu einem Vektor das Vielfache eines anderen Addiert. Insbesondere bleibt dabei auch die Dimension des Vektorraums unverändert. Mit anderen Worten, der Rang einer Matrix ändern sich nicht bei elementaren Spaltenoperationen. Die Determinante kann sich dabei wohl ändern. Aber die Eigenschaft, eine von Null verschiedene Determinante zu sein, bleibt dabei unberührt.
Nun kann man durch elementare Spaltenoperationen, A auf obere Dreiecksgestalt bringen. Mit anderen Worten, es gibt eine Matrix B mit folgenden Eigenschaften.

1. rk B = rk A
2. det B ≠ 0 ⇔ det A ≠ 0
3. B hat obere Dreieckgestalt und in jeder Zeile höchstens einen von Null verschiedenen Eintrag.

$$\begin{pmatrix} * & & \\ & \dots & \\ 0 & & * \end{pmatrix}$$

Es genügt also, die Behauptung für obere Dreieckmatrizen zu beweisen. O.B.d.A. sei also A eine obere Dreiecksmatrix wie in 3. beschrieben. Es reicht zu zeigen:

rk A = n ⇔ die Hauptdiagonalelemente von A sind sämtlich ≠ 0,

denn die Aussage auf der rechten Seite ist nach 4.4.4 (i) äquivalent zu

$$\det A \neq 0.$$

<u>Beweis von</u> '⇒'.
Nach Voraussetzung sind die Spalten der Matrix linear unabhängig. Bezeichne

$$V_i = \{\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \mid x_{i+1} = x_{i+2} = \dots = x_n = 0\} \subseteq K^n$$

den K-lineaern Unterraum der Spalten von K^n, deren Einträge ab der (i+1)-ten Zeile Null sind. Dann gilt $V_i \cong K^i$, also

$$\dim V_i = i.$$

Die ersten i-Spalten von B liegen in V_i und sind linear unabhängig, d.h. sie erzeugen sie V_i,

V_i wird von den ersten i-Spalten von B erzeugt (für i = 1,...,n).

Wäre der i-te Eintrag auf der Hauptdiagonalen von B gleich Null, so läge die i-te Spalte von B in V_{i-1}, wäre also eine Linearekombination der ersten (i-1)-Spalten. Das steht im

Widerspruch zur linearen Unabhängigkeit der Spalten von B. Auf der Hauptdiagonalen stehen also lauter von Null verschiedene Einträge.

Beweis von '$\Leftarrow$'.
Weil der erste Eintrag auf der Hauptdiagonalen ungleich Null ist, ist die erste Spalte linaer unabhängig.
Angenommen, wir haben bereits gezeigt, die ersten i Spalten sind unabhängig. Dann ist keine eine Linearkombination der übrigen. Es reicht also zu zeigen, die (i+1)-Spalte ist keine Linearkombination der vorhergehenden.
Das ist aber der Fall, denn die ersten i Spalten liegen in V_i und die (i+1)-te liegt nicht in diesem Unterraum.
QED.

4.4.6 Produktsatz für quadratische Matrizen

Für je zwei Matrizen $A,B\in K^{n\times n}$ gilt

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

Beweis. 1. Schritt. Der Fall det B = 0.

Nach 4.4.5 reicht es zu zeigen,

$$\text{rk } AB < n.$$

Es gilt

$$\text{rk } AB = \dim f_{AB}(K^n) = \dim f_A(f_B(K^n)) \le^{23} \dim f_B(K^n) = \text{rk } B.$$

Er reicht also zu zeigen, rk B < n. Das ist aber der Fall wegen det B = 0 und 4.4.5.

2. Schritt. Der Fall, det B $\ne$ 0.
Wir beginnen mit dem Fall, daß B eine Elementarmatrix ist,

$$B = M_i(c),\ \ Q_{ij}(c) \text{ oder } P_{ij}.$$

Für $B = M_i(c)$ gilt

$$\det M_i(c) = c$$

da $M_i(c)$ eine Diagonalmatrix ist. Weiter gilt

$$\det A\cdot M_i(c) = \det A \cdot c = \det A \cdot \det M_i(c)$$

da det A linear in der i-ten Spalte von A ist.

Für $B = Q_{ij}(c)$ gilt

$$\det Q_{ij}(c) = 1$$

da $Q_{ij}(c)$ eine obere (oder untere) Dreiecksmatrix ist. Weiter gilt

$$\det A\cdot Q_{ij}(c) = \det A = \det A \cdot \det Q_{ij}(c)$$

da sich die Determinante von A nicht ändert, wenn man das c-fache der i-ten Spalte zur j-ten addiert.

Für $B = P_{ij}$ gilt schließlich

$$\det A\cdot P_{ij} = - \det A$$

---

[23] Für jede lineare Abbildung f: V $\longrightarrow$ W gilt dim Im(f) $\le$ dim V, denn dim V ist die Summe aus der Dimensin der Bildes und der Dimension des Kerns.

da sich das Vorzeichen einer Determinante ändert, wenn man zwei Spalten der Matrix vertauscht. Ist speziell A = Id die Einheitsmatrix, so erhalten wir

$$\det P_{ij} = -1$$

Zusammen ergibt sich also auch in diesem Fall

$$\det A\cdot P_{ij} = \det A \cdot \det P_{ij}$$

Wir haben gezeigt

$$\det AB = \det A \cdot \det B,$$

falls B eine Elementarmatrix ist.

Sei jetzt B beliebig. Auf Grund unserer Voraussetzung ist B eine umkehrbare Matrix und als solche ein Produkt von Elementarmatrizen,

$$B = B_1\cdot\ldots\cdot B_s$$

Damit erhalten wir durch wiederholtes Anwenden des eben Bewiesenen

$$\begin{aligned}\det AB &= \det\ (A\cdot B_1\cdot\ldots\cdot B_s)\\ &= \det(A)\cdot\det\ (B_1)\cdot\ldots\cdot\det(B_s)\\ &= \det(A)\cdot\det(B_1\cdot\ldots\cdot B_s)\\ &= \det(A)\cdot\det(B).\end{aligned}$$

QED.

4.4.7 Axiomatische Charakterisierung der Determinante

Sei eine Funktion

$$f\colon K^n\times\ldots\times K^n \to K,\ (v_1,\ldots,v_n) \mapsto f(v_1,\ldots,v_n),$$

in n Variablen gegeben, die folgende Eigenschaften hat.

(i) f ist linear in der der i-ten Variablen v_i für $i = 1,\ldots,n$.

(ii) Sind in $f(v_1,\ldots,v_n)$ zwei der Spaltenvektoren v_i gleich, so gilt $f(v_1,\ldots,v_n) = 0$

(iii) $f(e_1,\ldots,e_n) = 1$ (e_i := i-ter Einheitsvektor).

Dann ist $f(v_1,\ldots,v_n)$ gerade die Determinante der Matrix mit den Spalten $v_1,\ldots,v_n$,

$$f(v_1,\ldots,v_n) = \det\ (v_1,\ldots,v_n).$$

Bemerkungen

(i) Wir werden im folgenden f bei Bedarf als Funktion der Matrix A mit den Spalten $v_1,\ldots,v_n$ betracthen,

$$f(A) = f(v_1,\ldots,v_n) \text{ für } A = (v_1,\ldots,v_n).$$

Als solche erfüllt f Bedingungen, die für die Determinante erfüllt sind.

(ii) Aus der Bedingung (ii) des Satzes ergibt sich, daß auch die folgende Bedingung erfüllt ist.

(ii)' Vertaucht man in $f(v_1,\ldots,v_n)$ zwei beliebige Variablen, so ändert sich das Vorzeichen.[24]

---

[24] Der Wert von f ist Null wenn das i-te und das j-te Argument gleich v_i+v_j sind. Auf Grund der Linearität von f in jedem Argument folgt

$$0 \quad = f(\ldots v_i+v_j\ \ldots\ v_i+v_j\ldots)$$

(iii) Für Körper einer Charakteristik $\neq 2$ ist (ii)' sogar äquivalent zu (ii).
(iv) Aus (i) und (ii) ergibt sich, daß sich f(A) nicht ändert, wenn man ein Vielfaches einer Spalte von A zu einer anderen Spalte addiert.

Beweis. Auf Grund der obigen Bemerkungen ändert sich f bei elementaren Operationen in derselben Weise wie die Determinante. Wir können also eine obere Dreiecksmatrix B finden mit folgenden Eigenschaften.

1. $f(A) = (-1)^r f(B)$
2. $\det(A) = (-1)^r \det(B)$.
3. B ist obere Dreiecksmatrix, die in jeder Zeile höchstens einen von Null verschiedenen Eintrag besitzt.

Wir können also ohne Beschränkung der Allgemeinheit annehmen, A ist eine obere Dreickmatrix der in Bedingung 3 beschriebenen Gestalt.

Der Fall det A = 0. Ist det A = 0, so ist das Produkt der Einträge auf der Hauptdiagonalen gleich Null, d.h. einer dieser Einträge ist Null. Dann ist aber die erste Spalte, für die das zutrifft gleich dem Nullvektor. Weil f eine lineare Funtion dieser Spalte ist, folgt

$$f(A) = 0 = \det A.$$

Der Fall det A $\neq$ 0. Ist det A $\neq$ 0, so sind alle Eintrage von A auf der Hauptdiagonalen ungleich Null. Alle anderen Einträge müssen dann aber gleich Null sein, d.h. A ist eine Diagonalmatrix:

$$A = (a_{11}\cdot e_1,\dots,a_{nn}\cdot e_n).$$

Auf Grund der Linearität von f bezüglich jeder Spalte folgt

$$f(A) = a_{11}\cdot\ldots\cdot a_{nn} f(e_1,\dots, e_n) = a_{11}\cdot\ldots\cdot a_{nn}\cdot 1 = \det(A).$$

QED.

4.4.8 Teilmatrizen, Minoren, Unterdeterminanten

Sei A eine m×n-Matrix mit Einträgen aus dem Körper K,

$$A \in K^{m\times n}.$$

Eine Teilmatrix von A ist eine Matrix, die aus A durch Streichen von Zeilen und Spalten entsteht. Dabei kann die Zahl der gestrichenen Zeilen oder Spalten auch Null sein. Insbesondere gehört auch A selbst zu den Teilmatrizen von A. Eine von A verschieden Teilmatrix heißt echte Teilmatrix von A.
Ein r-reihiger Minor oder auch r-Minor von A ist die Determinante einer quadratischen Teilmatrix von A des Typs (r,r).

Die Matrix

$$= f(\dots v_i \dots v_i+v_j \dots) + f(\dots v_j \dots v_i+v_j \dots)$$
$$= f(\dots v_i \dots v_i \dots) + f(\dots v_i \dots v_j \dots)$$
$$+ f(\dots v_j \dots v_{ij} \dots) + f(\dots v_j \dots v_j \dots)$$

Im ersten und letzten Summanden kommen zwei gleiche Argumente vor, die beiden Summanden sind also Null. Die beiden anderen Summanden unterscheiden sich also nur im Vorzeichen.

$$U_{ij}(A) := \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{mn} \end{pmatrix}.$$

welche aus A durch Streichen von i-ter Zeile und j-ter Spalte entsteht heißt Teilmatrix von A zur Position (i,j) oder oach $i\times j$-Teilmatrix.

Betrachten wir jetzt denn Fall $m = n$, d.h. sei A die quadratische Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdot\cdot & a_{nn} \end{pmatrix}.$$

Die Determinante der Teilmatrix $U_{ij}(A)$ heißt dann Unterdeterminante zur Position (i,j) oder auch $i\times j$-Unterdeterminante von A. Die mit dem Faktor $(-1)^{i+j}$ multiplizierte Unterdetermiante

$$A_{ij} := (-1)^{i+j} \det U_{ij}(A)$$

heißt auch Adjunkte zur Position (i, j) von A oder auch adjungierte Unterdeterminante.

Weiter sei

$$U'_{ij}(A) = \begin{array}{c} \begin{array}{ccc} & j & \end{array} \\ \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix} \end{array} \; i$$

die Matrix, welche aus A entsteht, indem man die i-te Zeile durch Nullen und anschließend die j-te Spalte durch den i-ten Standard-Einheitsvektor e_i ersetzt. Diese Matrix heißt die zu A komplementäre Matrix bezüglich der Position (i,j) oder auch die zu A $i\times j$-komplementäre Matrix.

Bemerkungen

(i) Nach Definition sind die oben definierten Unterdeterminanten (n-1)-reihige Minoren. Ihre Bedeutung besteht darin, daß man die Berechnung von det A auf die Berechnung dieser Minoren zurückführen kann.

(ii) Die Determinante der komplementären Matrix $U'_{ij}(A)$ ändert sich nicht, wenn man Vielfache der j-ten Spalte zu anderen Spalten addiert. Insbesondere gilt damit

$$\det U'_{ij}(A) = \det(a_1,...,a_{j-1},e_j,a_{j+1},...,a_n)$$

wenn a_v die v-te Spalte von A bezeichnet,

$$A = (a_1,..., a_n)$$

4.4.9 Die Berechnung der Adjunkten

Sei $A\in K^{n\times n}$ eine quadratische Matrix mit Einträgen aus dem Körper K. Dann ist die Adjunkte zur Position (i,j) gerade die Determinante der komplementären Matrix zu dieser Position.

$$A_{ij} = (-1)^{i+j} \det U_{ij}(A) = \det U'_{ij}(A). \tag{1}$$

Bemerkung zur Bestimmung des Vorzeichens (Schachmuster-Regel)

Das Vorzeichen $(-1)^{i+j}$ der Unterdeterminante zur Position (i,j) kann man dadurch bestimmen, daß man die beiden möglichen Vorzeichen ‘+’ und ‘-’ auf die Positionen der Matrix alternierend verteilt, wie etwa auf einem Schachbrett die schwarzen und weißen Felder. Man hat dabei mit dem Pluszeichen in der linken oberen Ecke zu beginnen. Zur Position (i,j) gehört dann das Vorzeichen $(-1)^{i+j}$.

Beweis von (1). Nach Bemerkung 4.4.8(ii) gilt

$$\det U'_{ij}(A) = \det (a_1,...,a_{j-1},e_i,a_{j+1},...,a_j).$$

Durch Vertauschen von Nachbarspalten erreichen wir, daß der i-te Standardeinheitsvektor e_i in die erste Spalte verschoben wird. Die Zahl der benötigten Vertauschungen ist j-1, d.h. es gilt

$$\det U'_{ij}(A) = (-1)^{j-1}\det(e_i,a_1,...,a_{j-1},a_{j+1},...,a_j),$$

Wir vertauschen jetzt solange benachbarte Zeilen auf der rechten Seite, bis die i-te Zeile in die erste gelangt. Die Anzahl der benötigten Vertauschungsoperationen ist i-1, d.h. es gilt

$$\det U'_{ij}(A) = (-1)^{i+j-2}\det(e_1,a'_1,...,a'_{j-1},a'_{j+1},...,a'_j),$$

Dabei entsteht a'_v aus der v-ten Spalte von a_v von A, indem man die i-te Koordinate streicht und als erste Koordinate in den Vektor einfügt.

Als nächstes benutzen wir die 1 in der Position (1,1) um die übrigen Einträge der ersten Zeile durch Null zu ersetzen. Dabei ändert sich die Determinante nicht, d.h. es gilt

$$\det U'_{ij}(A) = (-1)^{i+j-2} \det\begin{pmatrix} 1 & 0 \\ 0 & U_{ij}(A)\end{pmatrix},$$

wobei wie oben angegeben $U_{ij}(A)$ aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Die Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & U_{ij}(A)\end{pmatrix}$$

hat Blockgestalt wie in 4.4.4 (ii), d.h. ihre Determinante ist gleich

$$\det\begin{pmatrix} 1 & 0 \\ 0 & U_{ij}(A)\end{pmatrix} = 1\cdot\det U_{ij}(A) = \det U_{ij}(A)$$

Damit gilt

$$\det U'_{ij}(A) = (-1)^{i+j-2} \det U_{ij}(A) = A_{ij},$$

d.h. es gilt die Behauptung.
QED.

4.4.10 Entwicklungssatz von Laplace

Für $A\in K^{n\times n}$ gilt

(i) $\quad \det A = \sum_{i=1}^{n} a_{ij}\cdot A_{ij}$ für $j=1,...,n$.

(ii) $\quad \det A = \sum_{j=1}^{n} a_{ij}\cdot A_{ij}$ für $i=1,...,n$.

Dabei bezeichne A_{ij} die Adjunkte von A zur Position (i,j).

Beweis. Bezeichne a_ν die ν-te Spalte von A und $a_{\nu\mu}$ den Eintrag in der Position (ν,μ).
Dann gilt

$$
\begin{aligned}
\det A &= \det (a_1,...,a_{j-1},a_j,a_{j+1},...,a_j) \\
&= \det (a_1,...,a_{j-1}, \sum_{i=1}^{n} a_{ij}e_i, a_{j+1},...,a_j) \\
&= \sum_{i=1}^{n} a_{ij}\cdot\det (a_1,...,a_{j-1},e_i, a_{j+1},...,a_j) \\
&= \sum_{i=1}^{n} a_{ij}\cdot\det U'_{ij}(A)
\end{aligned}
$$

Zusammen mit 4.3.9 folgt Aussage (i). Aussage (ii) ergibt aus (i) durch Übergang zur transponierten Matrix (oder durch eine analoge Rechnung mit den Zeilen von A).
QED.

4.4.11 Rechnen mit Determinanten

Die bisher bewiesenen Eigenschaften der Determinante, versetzen uns in die Lage, Determinanten auch von Matrizen mit mehr als drei Reihen auszurechnen. Bevor man den Entwicklungssatz von Laplace anwendet, sollte man die Matrix so umformen, daß sie oder Spalten mit möglichst vielen Nullen enthält.

Wir illustrieren dies an einem Beispiel.

$$d := \det \begin{pmatrix} 3 & 4 & 2 & 5 \\ 2 & 3 & 1 & 4 \\ 4 & 4 & 4 & 5 \\ 3 & 4 & 3 & 4 \end{pmatrix}$$

Wir ziehen von der 1., 3. und 4. Zeile das 2-fache, 4-fache bzw. 3-fache der zweiten ab:

$$d = \det \begin{pmatrix} -1 & -2 & 0 & -3 \\ 2 & 3 & 1 & 4 \\ -4 & -8 & 0 & -11 \\ -3 & -5 & 0 & -8 \end{pmatrix}$$

$$= - \det \begin{pmatrix} -1 & -2 & -3 \\ -4 & -8 & -11 \\ -3 & -5 & -8 \end{pmatrix} \qquad \text{(Entwicklung nach der dritten Spalte)}$$

$$= \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 11 \\ 3 & 5 & 8 \end{pmatrix} \qquad \text{(Linearität in den Spalten)}$$

Von der 2. und 3. Spalte der letzten Matrix ziehen wir das 2- bzw. 3-fache der ersten ab.

$$d = \det \begin{pmatrix} 1 & 0 & 0 \\ 4 & 0 & -1 \\ 3 & -1 & -1 \end{pmatrix}$$

$$= \det \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix} \qquad \text{(Entwicklung nach der ersten Zeile)}$$

$$= 0\cdot(-1) - (-1)(-1) \qquad \text{(Determinantenformel für 2×2-Matrizen)}$$
$$= -1.$$

4.4.12 Die Cramersche Regel

Seien $A\in K^{n\times n}$ eine quadratische Matrix mit von Null verschiedener Determinante,

$$\det A\neq 0.$$

und $b\in K^{n\times 1}$ ein Spaltenvektor. Dann hat das lineare Gleichungssystem

$$Ax = b$$

genau eine Lösung. Die i-te Koordinate des Lösungsvektors ist dabei gleich

$$x_i = \frac{\det A_i}{\det A},$$

wobei A_i Matrix ist, welche man aus A erhält indem man die i-te Spalte durch b ersetzt,

$$A_i = (a_1,...,a_{i-1},b,a_{i+1},...,a_n)$$

Dabei bezeichne a_j die j-te Spalte von A.

Beweis. Wegen $\det A\neq 0$ ist die Matrix A umkehrbar. Es gibt also genau eine Lösung des Systems, nämlich

$$x = A^{-1}b.$$

Sei jetzt $A = (a_{ij})$ und b_i bezeichne die i-te Koordinate von b und x_i die i-te Koordinate des Vektors x. Nach Voraussetzung gilt

$$\sum_{v=1}^{n} a_{uv}x_v = b_u \text{ für } u=1,...,n. \qquad | \cdot A_{uj},\ \sum_{u=1}^{n}$$

Wir multiplizieren die u-te Gleichung dieses Systems mit der Adjunkten A_{uj} von A und bilden die Summe der entstehenden Gleichungen:

$$\sum_{u=1}^{n} \left(\sum_{v=1}^{n} a_{uv}x_v\right)\bullet A_{uj} = \sum_{u=1}^{n} b_u\cdot A_{uj} = \det A_j.$$

Das letzte Gleichheitszeichen rechts gilt dabei nach dem Laplaceschen Einwicklungssatz.
Also ist

$$\det A_j = \sum_{v=1}^{n} \left(\sum_{u=1}^{n} a_{uv}\cdot A_{uj}\right)\cdot x_v. \tag{1}$$

Versuchen wir die innere Summe des letzten Ausdrucks zu verstehen. Im Fall v=j ist

$$\sum_{u=1}^{n} a_{uv} \cdot A_{uj} = \det A \text{ im Fall } v=j.$$

denn den Ausdruck links erhält man gerade, indem man det A nach der j-ten Spalte entwickelt. Im Fall $v\neq j$ ist der Ausdruck links ebenfalls das Ergebnis einer Entwicklung nach der j-Spalte einer Matrix,

$$\sum_{u=1}^{n} a_{uv} \cdot A_{uj} = \det B \text{ im Fall } v\neq j$$

Als Koeffizienten a_{uv} treten in diesem Fall aber nicht die Einträge der j-ten Spalte von A auf, sondern die Einträge von deren v-te Spalte. Die Matrix B entsteht also aus A durch Ersetzung der j-ten Spalte durch die v-te. Mit anderen Worten, B hat zwei gleiche Spalten und damit die Determinante Null,

$$\det B = 0.$$

Zusammenfasssend ergibt sich, die innere Summe ist nur im Fall v=j ungleich Null, nämlich gleich der Determinante von A,

$$\sum_{u=1}^{n} a_{uv} \cdot A_{uj} = \det A\cdot\delta_{jv}.$$

Durch Einsetzen in (1) erhalten wir

$$\det A_j = \sum_{v=1}^{n} (\det A\cdot\delta_{jv})\cdot x_v = \det A\cdot x_j.$$

Da j beliebig gewählt war, ergibt sich die Behauptung.
QED.

4.4.13 Die inverse Matrix

Sei $A\in K^{n\times n}$ eine umkehrbare Matrix. Dann gilt

$$\det A^{-1} = \frac{1}{\det A}\cdot(A_{ij})^T_{i,j=1,\dots,n}$$

Beweis. Beim Beweis von 4.4.12 haben wir gezeigt

$$\sum_{u=1}^{n} a_{uv} \cdot A_{uw} = \begin{cases}\det A & \text{falls } v=w\\ 0 & \text{sonst}\end{cases}$$

oder, anders ausgedrückt,

$$\sum_{u=1}^{n} a_{uv} \cdot A_{uw} = \det A\cdot\delta_{uw}.$$

Mit $A^T_{ij} := A_{ji}$ kann man diese Identitäten auch in der Gestalt

$$\sum_{u=1}^{n} A^T_{wu}\cdot a_{uv} = \det A\cdot\delta_{uw}.$$

schreiben. Bezeichnet $A' = (A^T_{ij})$ die Matrix mit den Einträgen A^T_{ij}, so bedeutet letzteres

$$A'\cdot A = \det A\cdot \mathrm{Id},$$

Wir multiplizieren diese Identität von rechts mit A^{-1} und teilen durch die Determinante. Wir erhalten

$$A^{-1} = \frac{1}{\det A}\cdot A'.$$

Nun ist A' aber gerade transponiert zur Matrix der Adjunkten A_{ij}, d.h. es gilt die Behauptung.

QED.

4.4.14 Die Determinante eines Endomorphismus

Seien V ein endlich-dimensionaler K-Vektorraum, $f:V\to V$ eine K-lineare Abbildung und $v=(v_1,...,v_n)$ eine Basis von V. Dann ist der Wert der Determinante

$$\det M_v^v(f)$$

der zu f gehörigen Matrix unabhängig von der speziellen Wahl der Basis v. Er wird mit

$$\det(f) = \det M_v^v(f)$$

bezeichnet und heißt <u>Determinante</u> des Endomorphismus f.

Beweis der Unabhängigkeit der Determinante. Sei v' eine zweite Basis von V. Dann gilt

$$\begin{aligned} M_{v'}^{v'}(f) &= M_v^{v'}(\mathrm{Id})\, M_v^v(f)\, M_{v'}^v(\mathrm{Id}) \\ &= T^{-1}\cdot M_{v'}^{v'}(f)\cdot T \end{aligned}$$

mit $T := M_{v'}^v(\mathrm{Id})$. Es folgt

$$\begin{aligned} \det M_{v'}^{v'}(f) &= \det T^{-1}\cdot\det M_v^v(f)\cdot\det T \\ &= \det T^{-1}\cdot\det T\cdot\det M_v^v(f) \\ &= \det (T^{-1}\cdot T)\cdot\det M_v^v(f) \\ &= \det (\mathrm{Id})\cdot\det M_v^v(f) \\ &= \det M_v^v(f) \end{aligned}$$

QED.

4.5 Determinanten-Kriterium für den Rang

4.5.1 Das Rangkriterium I

Seien $A\in K^{m\times n}$ eine Matrix mit Einträgen aus K und r eine nicht-negative ganze Zahl. Dann sind folgende Aussagen äquivalent.
(i) rk A < r.
(ii) Für jede r-reihige quadratische Teilmatrix B von A gilt det B = 0.

Beweis. <u>(i) ⇒ (ii)</u>. Seien $a_1,...,a_n$ die Spalten von A und $a^1,...,a^m$ die Zeilen von A. Die Matrix B entstehe aus A indem zunächst alle Spalten von A gestrichen werden mit Ausnahme der Spalten $a_{i_1},...,a_{i_r}$ und dann anschließend in der entstehenden Matrix

$$A' := (a_{i_1}...a_{i_r})$$

m-r Zeilen.
Nach Voraussetzung hat A einen Rang <r, d.h. je r Spalten von A sind linear abhängig. Insbesondere besteht also eine lineare Abhängigkeit zwischen den Spalten von A',

$$\text{rk } A' < r.$$

Dann sind aber auch jeweils r Zeilen von A' linear abhängig. Insbesondere sind also die Zeilen von B linear abhängig, d.h. es gilt

$$\mathrm{rk}\ B < r.$$

Da B eine r-reihige quadratische Matrix ist, folgt

$$\det B = 0.$$

(ii) ⇒ (i). Nach Voraussetzung hat jede r-reihige Untermatrix von A die Determinante Null. Angenommen, es wäre

$$\mathrm{rk}\ A \geq r.$$

Dann gibt es in A mindestens r linear unabhängige Spalten, sagen wir $a_{i_1},...,a_{i_r}$. Die Matrix

$$A' := (a_{i_1} ... a_{i_r})$$

hat deshalb den Rang r. In A' gibt es damit aber auch r linear unabhängige Zeilen. Sei B eine von solchen r linear unabhängigen Zeilen gebildete Teilmatrix. Dann gilt

$$\mathrm{rk}\ B = r,$$

also $\det B \neq 0$. Das steht aber im Widerspruch zu unserer Annahme. Also muß rk A < r gelten.

QED.

4.5.2 Das Rangkriterium II

Seien $A \in K^{m\times n}$ eine Matrix mit Einträgen aus K und r eine nicht-negative ganze Zahl. Dann sind folgende Aussagen äquivalent.

(i) rk A = r.

(ii) Für jede (r+1)-reihige quadratische Teilmatrix B von A gilt det B = 0 und für mindestens eine r-reihige Teilmatrix B ist det B ≠ 0.

Beweis. Folgt direkt aus 4.5.1.

QED.

Beispiel

Sei $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ -1 & 4 & -1 \end{pmatrix}$. Dann gilt

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

und

$$\det A^{2,3}_{2,3} = \det \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} = -2 - 8 = -10.$$

Also hat A den Rang 2.

4.5.3 Das Rangkriterium III

Seien $A \in K^{m\times n}$ eine Matrix mit Einträgen aus K und r eine nicht-negative ganze Zahl. Dann sind folgende Aussagen äquivalent.

(i) rk A = r.

(ii) Es gibt eine r-reihige quadratische Teilmatrix B von A mit

$$\det B \neq 0$$

und für jede (r+1)-reihige quadratische Teilmatrix B' von A, welche B als Teilmatrix besitzt, gilt

$$\det B' = 0.$$

Beweis. (i) ⇒ (ii). Gilt auf Grund von 4.5.2.

(ii) ⇒ (i). Sei

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

Da sich der Rang von A beim Permutieren von Zeilen und Spalten nicht ändert, können wir annehmen, die Matrix B befindet sich im linken oberen Teil der Matrix A, d.h. A zerfällt wie folgt in Blöcke.

$$A = \begin{pmatrix} B & B' \\ B'' & B''' \end{pmatrix}$$

Wenden wir das Rangkriterium von 4.5.2 auf die Matrizen

$$A' := (B\ B') \text{ und } A'' := \begin{pmatrix} A \\ B'' \end{pmatrix}$$

an. Wegen $\det B \ne 0$ ist der Rang dieser Matrizen gleich r, d.h. die ersten r Zeilen und die ersten r Spalten von A sind linear unabhängig.

Insbesondere gilt

$$\text{rk } A \ge r.$$

Zum Beweis der Behauptung reicht es zu zeigen, jede der letzten n-r Spalten von A ist Linearkombination der ersten r. Betrachten wir die Matrix

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

welche entstehe, indem man zur Matrix A" noch eine weitere Spalte von A hinzufügt.[25] Wir haben zu zeigen,

$$\text{rk}\begin{pmatrix} B & * \\ * & * \\ \cdots & \cdots \\ * & * \end{pmatrix} < r + 1.$$

Dieselbe Argumentation wie im Fall der Matrix A zeigt, die erstren r Zeilen der Matrix (1) sind linear unabhängig. Es reicht also zu zeigen, die übrigen Zeilen sind Linearkombinationen der ersten r Zeilen. Betrachten wir die Teilmatrix

$$\begin{pmatrix} B & * \\ * & * \end{pmatrix} \tag{2}$$

von (1), wobei die Sterne in der unteren Zeile eine der letzten n - r Zeilen von (1) bezeichnen sollen. Zum Beweis der Behauptung reicht es zu zeigen

$$\text{rk}\begin{pmatrix} B & * \\ * & * \end{pmatrix} < r+1.$$

Weil es sich um den Rang einer (r+1)-reihigen quadratischen Matrix handelt, reicht es zu zeigen,

$$\det\begin{pmatrix} B & * \\ * & * \end{pmatrix} = 0.$$

Das ist aber der Fall auf Grund der Voraussetzung (ii).

QED.

---

[25] d.h. die Matrix bestehe aus den ersten r Spalten von A und einer weiteren Spalte.

Beispiel
Sei

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

Es gilt

$$\det A^{2,3}_{2,3} = \det \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} \neq 0.$$

Jede Teilmatirx von A, welche die Matrix $A^{2,3}_{2,3}$ als Teilmatrix enthält, hat zusätzlich Einträge aus der ersten Zeile und aus der ersten oder letzen Spalte. Es gilt also zwei solche Matrizen. Für deren Determinanten erhalten wir

$$\det A^{1,2,3}_{1,2,3} = \det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ -1 & 4 & -1 \end{pmatrix} = 0$$

nach dem Beispiel von 4.5.3, undWeiter gilt

$$\det A^{1,2,3}_{2,3,4} = \det \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 4 & -1 & 1 \end{pmatrix} = \det \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -2 & 1 \end{pmatrix} = \det \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} = 0$$

Also hat A den Rang 2

4.6 Allgemeiner Entwicklungssatz von Laplace

4.6.1 Komplementäre Folgen von Indizes und komplementäre Matrizen

Seien $A=(a_{ij})\in K^{m\times n}$ eine Matrix mit den Einträgen $a_{ij}\in K$ und

$$1 \le u_1 < \dots < u_k \le m$$
$$1 \le v_1 < \dots < v_\ell \le n$$

zwei endliche Folgen von paarweise verschiedenen ganzen Zahlen. Dann setzen wir

$$A^{u_1\dots u_k}_{v_1\dots v_\ell} := (a_{u_iv_j}) := \begin{pmatrix} a_{u_1v_1} & \dots & a_{u_1v_\ell} \\ \dots & \dots & \dots \\ a_{u_kv_1} & \dots & a_{u_kv_\ell} \end{pmatrix}$$

Seien jetzt weiter ganze Zahlen

$$1 \le u_{k+1}, \dots , u_{m\ k} \le m$$

gegeben mit

$$\{u_1, \dots , u_k\} \cup \{u_{k+1}, \dots , u_m\} = \{1,\dots,m\}.$$

Wir sagen dann die beiden Folgen

$$\{u_\mu\}_{\mu=1,\dots,k} \text{ und } \{u_\mu\}_{\mu=k+1,\dots,m}$$

sind zwei <u>Folgen komplementärer Zeilen-Indizes</u>. Analog sei eine Folge

$$1 \le v_{\ell+1},\dots,v_n \le n$$

gegeben mit

$$\{v_1, \ldots, v_\ell\} \cup \{v_{\ell+1}, \ldots, v_n\} = \{1,\ldots,n\}.$$

Dann sagen wir, die Folgen

$$\{v_\nu\}_{\nu=1,\ldots,k} \text{ und } \{v_\nu\}_{\nu=k+1,\ldots,m}$$

sind zwei Folgen von komplementären Spalten-Indizes.Die Marrix

$$A^{u_{k+1}\ldots u_m}_{v_{\ell+1}\ldots v_n}$$

heißt in dieser Situation die zu $A^{u_1\ldots u_k}_{v_1\ldots v_\ell}$ komplementäre Teilmatrix von A.

4.6.2 Der Entwicklungssatz

Seien $A \in K^{n\times n}$ eine quadratische Matrix mit Einträgen aus K und

$$u_1 < \ldots < u_k \text{ und } u_{k+1} < \ldots < u_n$$

zwei Folgen komplementärer Zeilen-Indizes.
Entwicklung nach den Zeilen $u_1, \ldots, u_k$: Es gilt

$$\det A = (-1)^{\sum_{\mu=1}^{k} u_\mu} \sum_{v_1<\ldots<v_k, v_{k+1}<\ldots<v_n} (-1)^{\sum_{\nu=1}^{k} v_\nu} \det A^{u_1\ldots u_k}_{v_1\ldots v_k} \det A^{u_{k+1}\ldots u_n}_{v_{k+1}\ldots v_n}$$

wobei die Summe zu erstrecken ist über alle Zerlegungen der Menge {1,...,n} in zwei Folgen komplementärer Spalten-Indizes (wobei die erste Folge k Glieder hat).
Die analoge Formel, die man erhält, indem man die Rolle von Zeilen und Spalten vertauscht, gilt ebenfalls.

Beispiel

Durch Entwicklung nach den ersten zwei Zeilen ergibt sich

$$(-1)^{1+2}\cdot\det\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = (-1)^{1+2}\cdot\det\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\det\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

$$+ (-1)^{1+3}\cdot\det\begin{pmatrix} a & -c \\ b & -d \end{pmatrix}\det\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$+ (-1)^{1+4}\cdot\det\begin{pmatrix} a & -d \\ b & c \end{pmatrix}\det\begin{pmatrix} d & a \\ -c & b \end{pmatrix}$$

$$+ (-1)^{2+3}\cdot\det\begin{pmatrix} -b & -c \\ a & -d \end{pmatrix}\det\begin{pmatrix} c & -b \\ d & a \end{pmatrix}$$

$$+ (-1)^{2+4}\cdot\det\begin{pmatrix} -b & -d \\ a & c \end{pmatrix}\det\begin{pmatrix} c & a \\ d & b \end{pmatrix}$$

$$+ (-1)^{3+4}\cdot\det\begin{pmatrix} -c & -d \\ -d & c \end{pmatrix}\det\begin{pmatrix} c & d \\ d & -c \end{pmatrix}$$

$$= -(a^2+b^2)^2 - (-ad+bc)^2 - (ac+bd)^2 - (bd+ac)^2 - (ad-bc)^2 - (c^2+d^2)^2$$

$$= -(a^2+b^2)^2 - (c^2+d^2)^2 - 2(ad-bc)^2 - 2(ac+bd)^2$$

$$= - a^4 - b^4 - 2a^2b^2$$
$$- c^4 - d^4 - 2c^2d^2$$
$$- 2a^2d^2 - 2b^2c^2 + 4abcd$$
$$- 2a^2c^2 - 2b^2d^2 - 4abcd$$
$$= - (a^2+b^2+c^2+d^2)^2$$

Also ist

$$\det\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = (a^2+b^2+c^2+d^2)^2$$

Berechnung ohne Verwendung des verallgemeinerten Entwicklungssatzes:

Mit $A := \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ und $B := \begin{pmatrix} c & d \\ d & -c \end{pmatrix}$ kann man auch schreiben

$$\det\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} = \det\begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \det\begin{pmatrix} A+iB & i(A+iB) \\ B & A \end{pmatrix} = \det\begin{pmatrix} A+iB & 0 \\ B & A-iB \end{pmatrix}$$

$$= \det(A+iB)\det(A-iB) = |\det(A+iB)|^2$$

$$= \left|\det\begin{pmatrix} a+ic & -(b-id) \\ b+id & a-ic \end{pmatrix}\right|^2$$

$$= \left|\det\begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix}\right|^2 \text{ mit } z = a + ic \text{ und } w = b + id$$

$$= (|z|^2 + |w|^2)^2$$

$$= (a^2+b^2+c^2+d^2)^2$$

Beispiel

Durch Entwicklung nach den ersten zwei Zeilen ergibt sich

$$(-1)^{1+2}\cdot\det\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

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

$$= - 3\cdot 3 + 1\cdot 1 - 3\cdot(-1) - (-1)\cdot 1 + 0\cdot(-1) - 1\cdot 0$$
$$= -9 + 1 + 3 +1 = -4$$

d.h.

$$\det\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} = 4.$$

Beispiel

Sei

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

Durch Entwicklung nach den ersten drei Zeilen erhält man

$(-1)^{1+2+3}\det A$
$= (-1)^{1+2+3}\cdot 0 + (-1)^{1+2+4}\cdot\alpha + (-1)^{1+2+5}\cdot 0 + (-1)^{1+2+6}\cdot 0$
$+ (-1)^{1+3+4}\cdot 0 + (-1)^{1+3+5}\cdot 0 + (-1)^{1+3+6}\cdot 0$
$+ (-1)^{1+4+5}\cdot 0 + (-1)^{1+4+6}\cdot\beta$
$+ (-1)^{1+5+6}\cdot 0$
$+ (-1)^{2+3+4}\cdot\gamma + (-1)^{2+3+5}\cdot 0 + (-1)^{2+3+6}\cdot 0$
$+ (-1)^{2+4+5}\cdot\delta + (-1)^{2+4+6}\cdot 0$
$+ (-1)^{2+5+6}\cdot 0$
$+ (-1)^{3+4+5}\cdot 0 + (-1)^{3+4+6}\cdot\varepsilon$
$+ (-1)^{3+5+6}\cdot 0$
$+ (-1)^{4+5+6}\cdot\delta$

mit

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

$$\beta = \det\begin{pmatrix} 0 & 1 & -2 \\ 2 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}\cdot\det\begin{pmatrix} 0 & 4 & 3 \\ -2 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} = (-2\bullet 4)\bullet(2\bullet 1) = -2^4$$

$$\gamma = \det\begin{pmatrix} 3 & 0 & 1 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}\cdot\det\begin{pmatrix} -4 & 3 & 0 \\ 0 & 0 & 4 \\ 3 & 1 & 0 \end{pmatrix} = (3+3)\bullet(-4\bullet -13) = 2^3\bullet 3\bullet 13$$

$$\delta = \det\begin{pmatrix} 3 & 1 & 0 \\ 0 & 0 & 4 \\ -3 & 1 & 0 \end{pmatrix}\cdot\det\begin{pmatrix} -4 & 4 & 0 \\ 0 & 0 & 4 \\ 3 & 1 & 0 \end{pmatrix} = (-4\bullet 6)\bullet(-4\bullet -16) = -2^9\bullet 3$$

$$\varepsilon = \det\begin{pmatrix} 0 & 1 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}\cdot\det\begin{pmatrix} -4 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix} = (-4)\bullet(-2\bullet -13) = -2^3\bullet 13$$

$$\delta = \det\begin{pmatrix} 1 & 0 & -2 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix}\cdot\det\begin{pmatrix} -4 & 0 & 4 \\ 0 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix} = (4\bullet 4)\bullet(-2\bullet -16) = 2^9$$

$$\det\begin{pmatrix} 0 & 3 & 0 & 1 & 0 & -2 \\ 2 & 0 & 1 & 0 & 4 & 0 \\ 0 & -3 & 0 & 1 & 0 & 2 \\ -4 & 0 & 4 & 0 & 3 & 0 \\ 0 & -2 & 0 & 3 & 0 & 4 \\ 3 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} = -2^6\bullet 3 + 2^4 - 2^3\bullet 3\bullet 13 + 2^9\bullet 3 + 2^3\bullet 13 - 2^9$$

$= -2^6\bullet 3 + 2^4 - 2^3\bullet 2\bullet 13 + 2^9\bullet 3$
$= -2^6\bullet 3 - 2^3\bullet 24 + 2^9\bullet 3$
$= -2^6\bullet 3 - 2^3\bullet 24 + 2^9\bullet 3$
$= -2^6\bullet 3 - 2^3\bullet 2^3\bullet 3 + 2^9\bullet 3$
$= -2^6\bullet 6 + 2^9\bullet 3$
$= -2^7\bullet 3 + 2^9\bullet 3$
$= 3\bullet(512 - 128)$
$= 3\bullet 128\bullet 3$
$= 2^7\bullet 3^2$
$= 768$

<u>Alternative Berechnungsweise</u>:

$$\det A = -\det\begin{pmatrix} 0 & 3 & 0 & 1 & 0 & -2 \\ 2 & 0 & 1 & 0 & 4 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 \\ -4 & 0 & 4 & 0 & 3 & 0 \\ 0 & -2 & 0 & 3 & 0 & 4 \\ 0 & -3 & 0 & 1 & 0 & 2 \end{pmatrix} \quad \text{(Vertauschen von 3.und 6.Zeile)}$$

$$= \det\begin{pmatrix} -4 & 0 & 4 & 0 & 3 & 0 \\ 2 & 0 & 1 & 0 & 4 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 & 0 & -2 \\ 0 & -2 & 0 & 3 & 0 & 4 \\ 0 & -3 & 0 & 1 & 0 & 2 \end{pmatrix} \quad \text{(Vertauschen von 1. und 4. Zeile)}$$

$$= -\det\begin{pmatrix} -4 & 3 & 4 & 0 & 0 & 0 \\ 2 & 4 & 1 & 0 & 0 & 0 \\ 3 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 3 & -2 & 4 \\ 0 & 0 & 0 & 1 & -3 & 2 \end{pmatrix} \quad \text{(Vertauschen von 2. und 5. Spalte)}$$

$$= -\det\begin{pmatrix} -4 & 3 & 4 \\ 2 & 4 & 1 \\ 3 & 1 & 1 \end{pmatrix}\det\begin{pmatrix} 1 & 3 & -2 \\ 3 & -2 & 4 \\ 1 & -3 & 2 \end{pmatrix}$$

$$=- \det\begin{pmatrix} -16 & -1 & 0 \\ -1 & 3 & 0 \\ 3 & 1 & 1 \end{pmatrix}\det\begin{pmatrix} 1 & 3 & -2 \\ 0 & -11 & 10 \\ 0 & -6 & 4 \end{pmatrix}$$

$= -(-16\bullet 3 - 1)\bullet(-11\bullet 4 + 6\bullet 10)$
$= 47\bullet 16$

4.6.3 Ein Lemma zum Vorzeichen von Permutationen

Seien

$$u_1 < \dots < u_k \text{ und } u_{k+1} < \dots < u_m$$

zwei Folgen komplementärer Zeilen-Indizes und sei σ die Permutation

$$\sigma = \begin{pmatrix} 1 & \dots & k & k+1 & \dots & m \\ u_1 & \dots & u_k & u_{k+1} & \dots & u_m \end{pmatrix}$$

Dann gilt

$$\operatorname{sign}(\sigma) = (-1)^{\sum_{\mu=1}^{k} (u_\mu - \mu)} = (-1)^{\sum_{\mu=1}^{k} u_\mu - k(k+1)/2}$$

Beweis. Wir führen den Beweis durch Induktion nach k. Im Fall k=0 ist σ die identische Permutation und die Aussage des Satzes ist trivial.
Sei jetzt der Satz bereits für Folgen bewiesen, deren erste Familie aus k-1 Elementen besteht. Betrachten wir die Folgen komplementärer Zeilen-Indizes

(1) $$u_1 < \dots < u_k \text{ und } u_{k+1} < \dots < u_m .$$

Dann ist u_k der größte Index der ersten Folge. Alle größeren Indizes gehören zur zweiten Folge und stimmen mit ihren Platznummern überein[26], d.h.

$$u_\mu = \mu$$

für alle u_μ, die größer sind als u_k. Verlegen wir jetzt u_k in die zweite Folge und betrachten die beiden Folgen komplementärer Indizes

(2) $$u_1 < \dots < u_{k-1} \text{ und } u_{k+1} < \dots\ u_k \dots < u_m .$$

Dann ist u_k größer als der größte Index der linken Folge, hat also die Platz-Nummer u_k. Die zweite Folge hat also die Gestalt

$$u_{k+1} < \dots\ u_k < u_k+1 < u_k+2 < \dots < u_k+x = m$$

Bezeichne σ' die zu (2) gehörige Permutation. Nach Induktionsvoraussetzung gilt dann

$$\operatorname{sign}(\sigma') = (-1)^{\sum_{\mu=1}^{k-1} (u_\mu - \mu)}$$

Die Permutation σ' entsteht aus σ indem man den Index u_k von der Position k in die Position u_k bringt, d.h. durch Ausführen von u_k-k Nachbartauschen. Also gilt

$$\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma') \cdot (-1)^{u_k - k}$$

Daraus ergibt sich die behauptete Formel.

---

[26] Im Fall $u_k < u_m$ ist u_m das größte Element von {1, ... , m}. Also gilt

$$u_m = m.$$

Im Fall $u_k < u_{m-1}$ ist u_{m-1} außerdem das größte Element von {1, ... , m-1}. Also gilt

$$u_{m-1} = m-1.$$

u.s.w.

QED.

4.6.4 Ein Spezialfall

Seien $A\in K^{n\times n}$ eine quadratische Matrix und k eine natürliche Zahl mit $1\le k\le n$.
Wir zerlegen A in Blöcke, sagen wir

$$A = \begin{pmatrix} A' & * \\ * & A'' \end{pmatrix} \text{ mit } A' := A^{1,\dots,k}_{1,\dots,k} \text{ und } A'' := A^{k+1,\dots,n}_{k+1,\dots,n}.$$

Dann ist die Summe s der Glieder der Determinante

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

in denen nur Faktoren aus A' und A" vorkommen gerade

$$s = \det A'\cdot\det A''$$

Beispiel
Das Produkt

$$a_{11}a_{22}\cdot\ldots\cdot a_{nn}$$

ist ein Glied der beschriebenen Art, nicht aber das Produkt

$$\pm a_{1n}a_{2n-1}\cdots a_{n1}$$

Beweis. Die Summe s ändert sich nicht, wenn man alle Einträge der Matrix A, deren Positionen weder in A' noch in A" liegen, durch Nullen ersetzt. Deshalb gilt

$$s = \det\begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix} = \det A' \bullet \det A''.$$

QED.

4.6.5 Beweis des Satzes

<u>1. Schritt</u>: Die Formel für die Entwicklung nach den ersten k Spalten.
Seien

$$v_1 < \dots < v_k \text{ und } v_{k+1} < \dots < v_n$$

zwei Folgen komplementärer Spalten-Indizes. Durch geeignete Nachbartausche kann man die Spalten $v_1, \dots, v_k$ in die ersten k Spalten bringen. Die Zahl der Nachbartausche, die man dabei benötigt ist

$$(v_1-1) + (v_2-2) + \dots + (v_k-k).$$

Das Vorzeichen ändert sich also um den Faktor

$$(-1)^{\sum_{\nu=1}^{k} v_\nu - \nu}$$

Macht man alle dieser Vertauschungen wieder rückgängig, so sieht man, daß die Summe aller Glieder von det A, deren Faktoren entweder in einer der ersten k Zeilen und einer der Spalten $v_1, \dots, v_k$ stehen oder in einer der letzten n-k Zeilen und einer der Spalten $v_{k+1},\dots,v_n$ stehen gerade gleich

$$(-1)^{k(k+1)/2}(-1)^{\sum_{\nu=1}^{k} v_\nu} \det A^{1,\dots,k}_{v_1\dots v_k} \det A^{k+1\dots n}_{v_{k+1}\dots v_n}$$

ist. Nun liegen aber die Faktoren eines Gliedes der ersten k Zeilen stets in irgendwelchen wohlbestimmten Spalten, d.h. die Determinante schreibt sich als Summe der Ausdrücke der obigen Gestalt:

$$\det A = (-1)^{k(k+1)/2} \sum_{v_1<\dots<v_k, v_{k+1}<\dots<v_n} (-1)^{\sum_{\nu=1}^{k} v_\nu} \det A^{1,\dots,k}_{v_1\dots v_k} \det A^{k+1\dots n}_{v_{k+1}\dots v_n}$$

Dies ist gerade die behauptete Formel für den Fall einer Entwicklung nach den ersten k Zeilen.

<u>2. Schritt</u>: der allgemein Fall.

Die Formel für die Entwicklung nach den Zeilen

$$u_1, \dots, u_k$$

können wir aus dem gerade behandelten Spezialfall gewinnen, indem wir eine Folge von Nachbartauschen von Zeilen ausführen, welche die gegebenen Zeilen in die ersten k Zeilen überführen. Die Determinante wird dabei mit dem Faktor

$$(-1)^{k(k+1)/2}(-1)^{\sum_{\mu=1}^{k} u_\mu}$$

multipliziert. Daraus ergibt sich die behauptete Formel.

QED.

5. Eigenwerte und Eigenvektoren

5.1 Eigenwerte und Eigenvektoren

5.1.0 Vorbemerkungen: konjugierte Matrizen und die Klassifikation linearer Abbildungen

(i) <u>Matrizen zu unterschiedlichen Basen sind konjugiert</u>.In 3.4.4 haben wir gesehen, daß die Matrix

$$A = M^v_v(f),\ v = (v_1,\dots,v_n)$$

einer linearen Abbildung

$$f: V \to V$$

bezüglich einer Basis $v_1,\dots,v_n$ von V beim Ersetzen der Basis v durch eine neue Basis v' in eine Matrix $A' := M^{v'}_{v'}(f)$ übergeht, die sich aus der alten Matrix A nach der Formel

(1) $$A' := SAS^{-1}$$

berechnen läßt, wobei $S = M^{v'}_{v}(Id)$ die sogenannte Basiswechsel-Matrix ist.

Matrizen A, $A' \in K^{n\times n}$, die in der Relation (1) zueinander stehen mit einer umkehrbaren Matrix S heißen <u>konjugiert</u>.

(ii) Konjugierte Matrizen gehören zur selben Abbildung. Umgekehrt kann man zwei Matrizen A und A', die in einer Beziehung der Gestalt (1) zueinander stehen, als Matrizen von ein und derselben Abbildung ansehen (bezüglich verschiedener Basen). Mit anderen Worten, Matrizen dieser Art sind in einem gewissen Sinne äquivalent.

(iii) Gegenstand des Kapitels.In diesem Kapitel wollen wir der Fragen nachgehen, wann zwei gegebene Matrizen zu ein und derselben Abbildung gehören, d.h. wann sie im oben beschriebenen Sinne konjugiert sind.

(iv) Wir werden dabei so vorgehen, daß wir in jeder Menge äquivalenter Matrizen eine Matrix auszeichen, d.h. wir konstruieren für diese Matrizen eine Normalform, so daß zwei Matrizen genau dann äquivalent sind, wenn sie dieselbe Normalform besitzen.

(v) Der erste Schritt bei der Verfolgung dieses Ziels ist die Konstruktion von Invaranten einer Matrix, d.h. von Zahlen, die zu der Matrix gehören und die sich nicht ändern, wenn man zu einer äquivalenten Matrix übergeht.

(vi) Problem. Die Äquivalenz zweier Matrizen A und A' kann man zeigen, indem man eine Matrix S angibt, so daß (1) gilt. Wenn man von zwei Matrizen zeigen will, sie sind nicht äquivalent, so muß man nachweisen, daß es keine solche Matrix S gibt, was zunächst ungleich schwerer ist. Die wichtigste Methode bei der Lösung dieses schwierigeren Problems besteht in der Angabe einer Invarianten, die für die betrachteten Matrizen unterschiedliche Werte annimmt.

(vii) Diagonalisierung. Die wichtigsten Konstruktionen unserer Theorie treten bereits für den Fall auf, daß die gesuchte Normalform eine besonders einfache Gestalt hat, nämlich eine Diagonalmatrix ist. Wir beschäftigen uns deshalb zunächst mit diesem Spezialfall, d.h. mit der Frage, ob eine gegebene quadratische Matrix diagonalisierbar ist.

5.1.1 Eigenvektoren, Eigenwerte, Eigenbasen, Eigenräume

Sei $f: V\to V$ eine lineare Abbildung eines endlich-dimensionalen Vektorraums V in sich.

Ein von Null verschiedener Vektor $v\in V-\{0\}$ heißt Eigenvektor von f, wenn es ein $c\in K$ gibt mit

$$f(v) = c\cdot v.$$

In dieser Situation heißt c Eigenwert von f zum Eigenvektor v. Mit anderen Worten, $c\in K$ heißt Eigenwert von f, wenn es einen von Null verschiedenen Vektor $v\in V-\{0\}$ gibt, mit

$$f(v) = c\cdot v.$$

Eine Basis $v_1,\dots,v_n$ von V heißt Eigenbasis von f, wenn sämtliche Vektoren v_i Eigenvektoren von f sind,

$$f(v_i) = c_i v_i \text{ für } i = 1,\dots,n.$$

Zu vorgegebenen $c\in K$ heißt

$$V_c := V_c(f) := \{v\in V \mid f(v) = c\cdot v\}$$

Eigenraum zum Eigenwert c.

Sei $A\in K^{n\times n}$ eine quadratische Matrix. Dann versteht man unter einem Eigenvektor, einem Eigenwert, bzw. einer Eigenbasis der Matrix A einen Eigenvektor, einen Eigenwert bzw. eine Eigenbasis der zugehörigen linearen Abbildung

$$f_A: K^n \to K^n,\ x \mapsto Ax.$$

Entsprechend setzt man

$$V_c(A) := V_c(f_A) := \{v\in V \mid Av = c\cdot v\}$$

und nennt $V_c(A)$ <u>Eigenraum</u> von A zum Eigenwert c.

Bemerkungen

(i) Die Mengen $V_c(f)$ bzw. $V_c(A)$ sind lineare Unterräume von V:

$$V_c(f) = \ker(f - c\cdot Id),\ V_c(A) = \ker(f_{A-c\cdot Id})$$

(ii) Die von Null verschiedenen Elemente von $V_c(f)$ bzw. $V_c(A)$ sind Eigenvektoren von f bzw. A zum Eigenwert c.

(iii) Wie wir demnächst sehen werden, gibt es höchstens endlich viele $c\in K$, für welche die Räume $V_c(f)$ bzw. $V_c(A)$ vom Nullraum verschieden sind. Diese c sind gerade die Eigenwerte von f bzw. c.

(iv) Mit anderen Worten, die Räume $V_c(f)$ bzw. $V_c(A)$ kann man für beliebiges c bilden. Sie sind jedoch nur für die (endlich vielen) Eigenwerte ungleich Null.

5.1.2 Ein Beispiel: Eigenbasen und Diagonalmatrizen

Sei

$$A = \begin{pmatrix} c_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & c_n \end{pmatrix}$$

eine Diagonalmatrix mit den paarweise verschiedenen Einträgen $c_1,...,c_n$ in der Hauptdiagonalen. Dann gilt

$$A\cdot e_i = c_i e_i,$$

d.h. der i-te Standardeinheitsvektor ist ein Eigenvektor e_i zum Eigenwert c_i und die Standardeinheitsvektoren

$$e_1,...,e_n$$

bilden eine Eigenbasis. Es ist nicht schwer einzusehen, daß es keine weiteren Eigenwerte und (bis auf Vielfache) keine weiteren Eigenvektoren gibt. Wir kennen damit sämtliche Eigenwerte, Eigenvektoren und Eigenräume.

Sei jetzt umgekehrt $f:V\to V$ eine lineare Abbildung eines endlich-dimensionalen Vektorraums V in sich und sei

$$v = (v_1,...,v_n)$$

eine Eigenbasis von f. Dann gilt

(1) $$f(v_i) = c_i v_i \text{ für } i=1,...,n$$

und gewisse $c_i\in K$. Die Identitäten (1) besagen gerade, daß die Matrix von f bezüglich der Basis v die folgende ist

$$M_v^v(f) = \begin{pmatrix} c_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & c_n \end{pmatrix}$$

Mit anderen Worten, die Eigenbasen eines Endomorphismus $f:V\to V$ sind gerade diejenigen Basen von V, bezüglich der die Matrix von f Diagonalgestalt hat.

Probleme

(i) Besitzt jeder lineare Endomorphismus eine Eigenbasis, bzw. wie entscheidet man, ob ein gegebener Endomorphismus eine solche besitzt?

(ii) Wie kann man die Eigenvektoren bzw. Eigenwerte eines Endomorphismus bestimmen?

5.1.3 Das charakteristische Polynom eines Endomorphismus

Sei $f: V \to V$ eine lineare Abbildung des endlich-dimensionalen Vektorraums V in sich. Dann heißt

$$\chi_f(T) := \det(f - T\cdot Id)$$

charakteristisches Polynom von f. Sei $v = (v_1,...,v_n)$ eine Basis von V und

$$A := M_v^v(f)$$

die Matrix von f bezüglich v. Dann gilt

$$\chi_f(T) := \det(f - T\cdot Id) := \det(A-T\cdot Id),$$

d.h. $\chi_f(T)$ ist tatsächlich ein Polynom (n-ten Grades) in T. Für beliebige quadratische Matrizen A schreibt man auch

$$\chi_A(T) := \det(A-T\cdot Id)$$

und spricht vom charakteristischen Polynom der Matrix.

Beispiel

Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Dann gilt

$$\begin{aligned} \chi_A(T) &= \det \begin{pmatrix} 1-T & 2 \\ 3 & 4-T \end{pmatrix} \\ &= (T-1)(T-4) - 6 \\ &= T^2 - 5T - 2 \\ &= (T - \tfrac{1}{2}(5+\sqrt{33}))\,(T - \tfrac{1}{2}(5-\sqrt{33})) \end{aligned}$$

5.1.4 Die Nullstellen des charakteristischen Polynoms

Seien $f: V \to V$ eine lineare Abbildung des endlich-dimensionalen Vektorraums V und $c\in K$ ein Element. Dann sind folgende Aussagen äquivalent.

(i) c ist ein Eigenwert von f.

(ii) $\chi_f(c) = 0$.

Beweis. Zum Beweis können wir eine Basis von V fixieren und anstelle von f die Matrix von f bezüglich dieser Basis betrachten. Sei A diese Matrix. Wir haben dann zu zeigen, daß die folgenden Aussagen äquivalent sind.

(i)' c ist Eigenwert von A.

(ii)' $\chi_A(c) = 0$.

<u>(i)' ⇒ (ii)'</u>. Sei c ein Eigenwert von A. Es gibt also einen Vektor $v \in K^n - \{0\}$ mit

$$Av = c \cdot v.$$

Diese Identität kann man auch wie folgt schreiben.

$$0 = Av - cv = A \cdot v - c \cdot Id \cdot v = (A - c \cdot Id)v .$$

Mit $B := A - c \cdot Id$ gilt folglich

$$Bv = 0,$$

d.h. das Gleichungssystem $Bx = 0$ hat eine nicht-triviale Lösung, d.h. neben der Lösung $x = 0$ gibt es noch mindestens eine weitere Lösung (nämlich $x = v$). Dann mußt aber B die Determinante Null haben,

$$0 = \det (B) = \det (A - c \cdot Id) = \chi_A(c).$$

Es gilt also (ii)'.

<u>(ii)' ⇒ (i)'</u>. Sei c eine Nullstelle von χ_A. Mit $B := A - c \cdot Id$ gilt dann

$$\det B = 0,$$

also

$$\operatorname{rk} B < n.$$

Die Spalten $b_1,\dots,b_n$ von B sind also linear abhängig, d.h. es gibt Elemente $v_1,\dots,v_n \in K$ mit

(1) $$\sum_{i=1}^{n} v_i b_i = 0 \text{ und } v_i \neq 0 \text{ für ein i.}$$

Wir setzen $v = \begin{pmatrix} v_i \\ \dots \\ v_n \end{pmatrix}$. Dann kann man (1) auch wir folgt ausdrücken.

$$B \cdot v = 0,\ v \neq 0.$$

Wegen $B = A - c \cdot Id$ bedeutet dies,

$$0 = Bv = (A - c \cdot Id)v = Av - cv,$$

d.h.

$$Av = cv.$$

Wir haben gezeigt, c ist Eigenwert von A.
QED.

5.1.5 Ein Beispiel: Eigenwerte für verschiedene Grundkörper

In 5.4.3 hatten wir die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

mit dem charakteristischen Polynom

$$\chi_A(T) = T^2 - 5T - 2 = (T - \frac{1}{2}(5+\sqrt{33}))\ (T - \frac{1}{2}(5-\sqrt{33}))$$

betrachtet. Im Fall $K = \mathbb{Q}$ hat die Matrix also keinerlei Eigenwerte oder Eigenvektoren. Über den reellen Zahlen dagegen sind die Eigenwerte gleich

$$c_{1'2} = \frac{1}{2}(5\pm\sqrt{33})$$

Die Matrix $A - c\cdot Id$ hat, falls c ein Eigenwert ist, einen Rang <2. Die zugehörigen Eigenvektoren sind also durch die einzige Gleichung

$$\frac{1}{2}(3+\sqrt{33})\cdot x - 2\cdot y = 0 \text{ (im Fall } c = c_1)$$

bzw.

$$\frac{1}{2}(3-\sqrt{33})\cdot x - 2\cdot y = 0 \text{ (im Fall } c = c_1)$$

gegeben. Insbesondere sind

$$\binom{x}{y} = \binom{4}{3+\sqrt{33}} \text{ und } \binom{x}{y} = \binom{4}{3-\sqrt{33}}$$

Eigenvektoren. Diese Vektoren sind nicht proportional, also linear unabhängig. Sie bilden also eine Eigenbasis von K^2.

Bemerkung

Die Situation des obigen Beispiels ist typisch für Betrachtungen im Kontext von Eigenwerten: Eigenwerte lassen sich stets finden, solange der Grundkörper nicht zu klein ist. In der Algebra-Vorlesung des zweiten Studienjahres zeigt man:

1. Man kann den Körper K stets soweit vergrößern, daß das charakteristische Polynom in Linearfaktoren zerfällt (also n nicht notwendig verschiedene Eigenwerte existieren).
3. Allgemeiner: Jeder Körper k liegt in einem Körper K derart, daß jedes nichtkonstante Polynom mit Koeffizienten aus K in K eine Nullstelle besitzt. Solche Körper K heißen algebraisch abgeschlossen.
2. Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen (Fundamentalsatz der Algebra).

5.1.6 Vereinbarung, algebraische und geometrische Vielfachheiten

Wir werden im folgenden oft annehmen, unser Körper K ist so groß, daß die Eigenwerte der betrachteten Matrizen in K liegen. Wir werden in dieser Situation sagen, der Körper K sei hinreichend groß.

Bemerkung

Auf Grund des nachfolgenden Ergebnisses bedeutet dies gerade, daß das Charakterische Polynon in Linearfaktoren zerfällt,

$$\chi(T) = \pm(T-c_1)^{\nu_1}(T-c_2)^{\nu_2}...(T-c_r)^{\nu_r}$$

wobei $c_1,...,c_r \in K$ die Eigenwerte sind. Die natürlichen Zahlen $\nu_1,...,\nu_r$ heißen algebraische Vielfachheiten der Eigenwerte im Gegensatz zu deren geometrischen Vielfachheiten

$\mu_i := \dim V_{c_i}$

5.1.7 Zerlegung eines Polynoms in Linearfaktoren

Seien f(T) ein nicht-konstantes Polynom mit Koeffizienten aus K und $\overline{K}$ ein algebraisch abgeschlossener Körper, welcher K enthält. Dann gibt es einen Körper L zwischen K und $\overline{K}$ derart, daß f über L in Linearfaktoren zerfällt, d.h. es gibt Elmente c, $c_1,...,c_r$ $\in$L und natürliche Zahlen $n_1,...,n_r$ mit

$$f(T) = c\cdot(T-c_1)^{n_1}(T-c_2)^{n_2}...(T-c_r)^{n_r}$$

Bemerkung
Der kleinste solche (hinreichend große) Körper wird mit $K(c_1,...,c_1)$ bezeichnet und heißt der von $c_1,...,c_r$ über K erzeugte Körper. Er besteht (wie leicht zu sehen ist) aus allen Quotienten

$$u(c_1,...,c_r)/v(c_1,...,c_r)$$

von Polynome in $c_1,...,c_r$ mit Koeffizienten aus K.

Beweis. Es reicht zu zeigen, f zerfällt über L := $\overline{K}$ in Linearfaktoren. Wir führen den Beweis durch Induktion nach dem Grad d = deg f von f. Im Fall d = 1 ist f linear und die Behauptung ist trivial. Sei also d>1. Dann besitzt f in $\overline{K}$ eine Nullstelle, sagen wir c_1. Division mit Rest durch T - c_1 liefert

$$f(T) = q(T)\cdot(T-c_1) + r$$

mit Polynomen q und r, wobei der Grad von r kleiner ist als deg $T-c_1$ = 1, d.h. r ist eine Konstante aus $\overline{K}$. Wir setzen T = c_1 in diese Gleichung ein und erhalten

$$0 = f(c) = q(c_1)\cdot 0 + r,$$

d.h. es ist r = 0 und

$$f(T) = q(T)\cdot(T-c_1)$$

Das Polynom q(T) hat einen Grad < d, zerfällt also nach Induktionsvoraussetzung in Linearfaktoren. Dann gilt dasselbe aber auch für f.
QED.

5.1.8 Existenz von Eigenbasen und algebraische Vielfachheiten

Sei f:V$\to$V eine K-lineare Abbildung endlich-dimensionaler Vektorräume, deren (paarweise verschiedenen) Eigenwerte $c_r,...,c_r$ sämtlich in K liegen. Die geometrischen und algebraischen Vielfachheiten der Eigenwerte seien $\mu_1,...,\mu_r$ bzw. $\nu_1,...,\nu_r$. Dann gilt

$$\mu_i \le \nu_i \text{ für } i=1,...,r.$$

Weiter sind die beiden folgenden Aussagen äquivalent.
(i) f besitzt eine Eigenbasis.

(ii) $V =$[27] $V_{c_1} \oplus ... \oplus V_{c_r}$

(iii) $\mu_i = \nu_i$ für i=1,...,r.

Beweis. Beweis der behaupteten Ungleichungen. Sei c einer der Eigenwerte von f, μ die zugehörige geometrische und ν die zugehörige algebraische Vielfachheit. Nach Definition gilt

$$\mu := \dim V_c = \dim \ker(f - c\cdot Id).$$

Wir wählen eine Basis $v_1,...,v_\mu$ von V_c und ergänzen diese durch zu einer Basis

$$v_1,...,v_\mu,...,v_d \text{ , } d = \dim V,$$

von V. Wegen $f(v_i) = c\cdot v_i$ für $i = 1,...,\mu$ hat f bezüglich dieser Basis die Matrix

$$M_v^v(f) = \begin{pmatrix} c & \dots & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & c & * & \dots & * \\ 0 & \dots & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} = \begin{pmatrix} c\cdot Id & A \\ 0 & B \end{pmatrix},$$

d.h. die Matrix zerfällt in Blöcke mit einer $\mu\times\mu$-Matrix in der linken oberen Ecke. Es folgt

$$\begin{aligned} \chi_f(T) &= \det(M_v^v(f) - T\cdot Id) \\ &= \det\begin{pmatrix} (c-T)\cdot Id & A \\ 0 & B-T\cdot Id \end{pmatrix} \\ &= \det((c-T)\cdot Id)\cdot\det(B-T\cdot Id) \\ &= (c-T)^\mu\cdot\det(B-T\cdot Id) \end{aligned}$$

Mit anderen Worten, das charakteristische Polynom $\chi_f(T)$ hat c als Nullstelle mit einer Vielfachheit $\geq \mu$, d.h. es ist

$$\nu \geq \mu.$$

Beweis der Äquivalenz der Bedingungen (i)-(iii).

(i) $\Rightarrow$ (iii). Besitze f eine Eigenbasis $v_1, ..., v_d$. Die Anzahl μ'_j der Eigenvektoren v_i zum Eigenwert c_j ist höchstens so groß wie die Dimension des Eigenraums V_{c_i}, d.h. die geometrische Vielfachheit μ_j von c_j:

$$\mu'_j \leq \mu_j = \dim V_{c_j} \leq \nu_j \text{ für } j = 1, ..., r.$$

Es gilt

[27] d.h. der Raum V läßt sich in natürlicher Weise mit der direkten Summe der Eigenräume V_{c_i} identifizieren. Genauer, die lineare Abbildung $V_{c_1} \oplus ... \oplus V_{c_r} \to V$, $(v_1,...,v_r) \mapsto v_1+...+v_r$, ist bijektiv.

$$\dim V = \sum_{j=1}^{r} \mu'_j \le \sum_{j=1}^{r} \mu_j \le \sum_{j=1}^{r} \nu_j = \deg \chi_f = \dim V.$$

In der Abschätzung muß überall das Gleichheitszeichen stehen. Das ist aber nur möglich, wenn

$$\mu'_j = \mu_j = \nu_j$$

gilt für jedes j. Insbesondere gilt (iii). Wir haben außerdem bewiesen:

Folgerung
<u>In jeder Eigenbasis kommen alle Eigenvektoren mit ihren geometrischen Vielfachheiten vor</u>.

<u>(iii) ⇒ (ii)</u>. Nach Voraussetzung ist die Dimension

$$\mu_i := \dim V_{c_i}$$

gleich der algebraischen Vielfachheit ν_i. Insbesondere gilt

$$\dim V_{c_1} \oplus \ldots \oplus V_{c_r} = \sum_{i=1}^{r} \mu_i = \sum_{i=1}^{r} \nu_i = \deg \chi_f(T) = \dim V.$$

Die Dimension der Räume, deren Isomorphie wir zeigen müssen, ist somit gleich. Es reicht also, wenn wir zeigen, die lineare Abbildung

$$\varphi: V_{c_1} \oplus \ldots \oplus V_{c_r} \to V \text{ , } (v_1,\ldots,v_r) \mapsto v_1 + \ldots + v_r,$$

ist injektiv. Angenommen die Abbildung wäre nicht injektiv. Dann gibt es Vektoren

$$v_i \in V_{c_i},$$

welche nicht sämtlich gleich Null sind, mit

(1) $$v_1 + \ldots + v_r = 0.$$

Sei s die Anzahl der von Null verschiedenen v_i . Diese Anzahl ist mindestens 1,

$$s \ge 1.$$

Die Idee des nachfolgenden Beweises besteht darin, danach zu fragen, ob es vielleicht ein Tupel $(v_1,\ldots,v_r)$ derselben Art gibt mit weniger als s von Null verschiedenen v_i. Falls ja, so können wir das betrachtete Tupel durch dasjenige mit dem kleinerem s ersetzen. Wir können also annehmen, die Zahl s ist für das hier von uns betrachtete Tupel $(v_1,\ldots,v_r)$ minimal, d.h. es gibt kein Tupel derselben Art, mit weniger als s von Null verschiedenen Vektorren v_i.

Durch geeignets Abändern der Bezeichnungen können wir erreichen, daß gerade die ersten s Vektoren in der Summe (1) von Null verschieden sind. Bedingung (1) bekommt dann die Gestalt,

(1') $$v_1 + \ldots + v_s = 0,$$

wobei jetzt sämtliche Summanden ungleich Null sind. Dann muß s > 1 gelten. Es gibt also mindestens zwei verschiedene Eigenwerte c_i und mindestens einer davon muß ungleich Null sein. O.B.d.A. sei

$$c_1 \ne 0.$$

Wir gewinnen jetzt zwei verschiedene Relationen aus (1'), einmal indem wir f auf (1') anwenden und einmal indem wir (1') mit c_1 multiplizieren. Es ergibt sich

(2) $$c_1v_1 + \ldots + c_sv_s = 0$$

und

(3) $$c_1v_1 + ... + c_1v_s = 0.$$

Wir bilden die Differenz dieser beiden Relationen und erhalten

(4) $$(c_2-c_1)v_2 + ... + (c_s-c_1)v_s = 0.$$

Die Koeffizienten in dieser neuen Relation sind sämtlich von Null verschieden (da die Eigenwerte c_i nach Voraussetzung paarweisen verschieden sein sollen). Die Zahl der Summanden ist kleiner als s. Das steht aber im Widerspruch zur Minimalität der Zahl s. Dieser Widerspruch beweist die die behauptete Implikation.

(ii) ⇒ (i). Wir benutzen den Isomorphismus von (ii) um V mit der direkten Summe der V_{c_i} zu identifizieren. Wir wählen in jedem der Eigenräume V_{c_i} eine Basis und vereinigen alle diese Basen zu einer Basis v von V. Jeder Vektor dieser Basis liegt in einem der V_{c_i}, ist also ein Eigenvektor. Die so erhaltene Basis ist somit eine Eigenbasis.

QED.

Bemerkung

Die Frage nach der Existenz von Eigenbasen (bzw. nach der Diagonalisierbarkeit von Matrizen) wird im gesamten nachfolgenden Verlauf der Vorlesung eine Rolle spielen. Um die Frage abschließend[28] zu beantworten, müssen wir jedoch erst einige einfachere Probleme lösen:

1. Besitzt jede Abbilung wenigsten einen Eigenvektor?
2. Kann man jede Matrix wenigstens in eine obere Dreieckgestalt überführen.

Beispiel

Bei einer Drehung im $\mathbb{R}^2$ um einen kleinen Winkel wird kein Vektor in ein (reelles) Vielfaches von sich selbst überführt. Eine solche Drehung besitzt also keinen Eigenvektor.

Wie wir sehen werden, liegt das einfach daran, daß der Körper $\mathbb{R}$ dafür zu klein ist. Über den komplexen Zahlen besitzt die entsprechende Matrix sehr wohl einen Eigenwert.

Bemerkungen

(i) Die nachfolgenden Ergebnisse gelten also insbesondere auf Grund unserer Annahme, daß unser Grundkörper so groß sein soll, daß er alle Nullstellen des charakteristischen Polynoms enthält.

(ii) Allgemein gilt: Die über 'kleinen' Körpern wie $\mathbb{R}$ auftretenden Phänomene sind sehr viel komplizierter und reichhaltiger als die Phänomene über $\mathbb{C}$. Es ist deshalb typisch für die Vorgehensweise in der Mathematik, zuerst solche Körper wie $\mathbb{C}$ zu behandeln.

5.1.9 Existenz von Eigenwerten

Sei f: V→V eine K-lineare Abbildung. Dann besitzt f (falls K hinreichend groß ist) mindestens einen Eigenvektor.

Analog besitzt eine beliebige Matrix über K (falls K hinreichend groß ist) mindestens einen Eigenvektor.

Beweis. Das charakteristische Polynom von f besitzt in einen algebraisch abgeschlossenen Oberkörper von K eine Nullstelle. Da K hinreichend groß sein soll, liegt diese Nullstelle in K. Diese Nullstelle ist dann aber ein Eigenwert von f, d.h. f besitzt einen Eigenvektor.

QED.

---

[28] über algebraisch abgeschlossenen Körpern

5.1.11 Fahnen von Vektorräumen

Sei V ein K-Vektorraum. Eine <u>Fahne</u> der Länge r von V ist eine echt aufsteigende Folge

$$V_0 \subset V_2 \subset \ldots \subset V_r$$

von Unterräumen von V. Die Fahne heißt <u>vollständig</u>, falls folgende Bedingungen erfüllt sind.

(i) $V_0 = \{0\}$

(ii) $V_r = V$

(iii) $\dim V_{i+1} = \dim V_i + 1$ für $i=0,\ldots,r-1$.

Seien $f: V \to V$ ein linearer Endomorphismus und $W \subseteq V$ ein linearer Unterraum. Dann heißt W auch f-<u>invariant</u>, falls gilt $f(W) \subseteq W$.

5.1.12 Existenz von Fahnen invarianter Unterräume

Sei $f: V \to V$ ein linearer Endomorphismus eines endlich-dimensionalen K-Vektorraums. Falls K hinreichend groß ist, so existiert eine eine vollständige Fahne

$$0 = V_0 \subset V_1 \subset \ldots \subset V_d = V$$

von V, deren Unterräume V_i sämtlich f-invariant sind.

Beweis. Wir führen den Beweis durch Induktion nach $d = \dim V$. Im Fall $d = 1$ ist nicht zu beweisen. Sei also $d > 1$. Weil K hinreichend groß ist, gibt es einen Eigenvektor $v_1 \in V$ von f. Wir setzen

$$W = K \cdot v_1$$

Da v_1 ein Eigenvektor ist, gilt $f(W) \subseteq W$, d.h. W ist f-invariant. Wir setzen

$$V' := V/W$$

und bezeichnen mit

$$\rho: V \to V'$$

die natürliche Abbildung. Weiter sei f' die Abbildung

$$f': V' \to V',\ v + W \mapsto f(v) + W.$$

Diese Abbildung ist wohldefiniert, denn mit $v + W = w + W$ gilt $v - w \in W$, also

$$f(v-w) \in f(W) \subseteq W,$$

also $f(v) - f(w) \in W$, also $f(v) + W = f(w) + W$. Die Abbildung f' ist offensichtlich linear. Sie ist gerade so definiert worden, daß das folgende Diagramm kommutativ ist.

$$\begin{array}{ccccccc}
V & \xrightarrow{f} & V & \qquad & v & \qquad & f(v) \\
\downarrow\rho & & \downarrow\rho & & & & \\
V' & \xrightarrow{f'} & V' & & v+W & & f(v)+W
\end{array}$$

Es gilt $\dim V' = \dim V - \dim W = \dim V - 1$. Nach Induktionsvoraussetzung gibt es eine Fahne aus f'-invarianten Unterräumen in V', sagen wir

$$0 = V'_1 \subset V'_2 \subset \ldots \subset V_d = V'$$

Wir setzen

$$V_i := f^{-1}(V'_i) \quad \text{für } i = 1,\ldots,d$$

und $V_0 := \{0\}$. Dann gilt

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_d = V.$$

Zum Beweis von 5.1.10 reicht es, wenn wir zeigen

1. Die Räume V_i sind f-invariant.
2. $\dim V_{i+1} = \dim V_i$ für $i = 1,\ldots, r-1$.

Zu 1. Sei $v \in V_i$. Wir haben zu zeigen $f(v) \in V_i$. Nach Voraussetzung gilt $\rho(v) \in V'_i$. Da der Raum V'_i invariant bezüglich f' ist, folgt

$$(f' \circ \rho)(v) = f'(\rho(v)) \in f'(V'_i) \subseteq V'_i.$$

Wegen der Kommutativität des obigen Vierecks folgt

$$\rho(f(v)) = (\rho \circ f)(v) = (f' \circ \rho)(v) \in V'_i,$$

also $f(v) \in V_i$.

Zu 2. Wir betrachten die Einschränkung der natürlichen Abbildung ρ auf V_i,

$$\rho_i : V_i \to V'_i \qquad \begin{array}{ccc} V & \xrightarrow{\rho} & V' \\ \cup & & \cup \\ V_i & \xrightarrow{\rho'} & V'_i \end{array}$$

Da ρ surjektiv ist, ist auch die Einschränkung ρ_i surjektiv und ihr Kern ist gerade

$$\ker(\rho_i) = \{v \in V_i \mid \rho(v) = 0\} = V_i \cap \ker(\rho) = V_i \cap W = W \text{ (falls } i>0).$$

Damit gilt für $i > 0$:

$$\dim V_i = \dim \operatorname{im}(\rho_i) + \dim \ker(\rho_i) = \dim V'_i + \dim W = \dim V'_i + 1.$$

also ist

$$\dim V_{i+1} - \dim V_i = \dim V'_{i+1} - \dim V'_i = 1,$$

d.h. Aussage 2 gilt zumindest für $i>0$. Für $i = 0$ erhalten wir

$$\dim V_1 = \dim W = 1 = 0 + 1 = \dim V_0 + 1.$$

QED.

5.1.12 Überführung von Matrizen in obere Dreiecksgestalt

Sei $A \in K^{n\times n}$ eine Matrix über einem hinreichend großen Körper K. Dann gibt es eine umkehrbare Matrix $T \in K^{n\times n}$ derart, daß

$$TAT^{-1} = \begin{pmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & * \end{pmatrix}$$

obere Dreieckgestalt besitzt.
Beweis. Sei f die Abbildung

$$f = f_A: K^n \to K^n, x \mapsto Ax.$$

Wir haben zu zeigen, bezüglich einer geeigneten Basis v von K^n besitzt die Matrix

$$M_v^v(f)$$

obere Dreiecksgestalt.Da K hinreichend groß ist, gibt es eine vollständige Fahne von f-invarianten Unterräumen

$$0 = V_0 \subseteq V_1 \subseteq \dots \subseteq V_n = K^n.$$

Jede Basis von V_i läßt sich durch Hinzufügen eines Vektors zu einer Basis von V_{i+1} ergänzen. Es gibt also eine Basis $v_1,\dots,v_n$ von K^n derart, daß $v_1,\dots,v_i$ für jedes i eine Basis von V_i ist. Wegen $v_i \in V_i$ und $f(V_i) \subseteq V_i$ gilt $f(v_i) \in V_i$, d.h.

$$f(v_i) = \text{Linearkombination von } v_1,\dots,v_i.$$

Das bedeutet, $M_v^v(f)$ besitzt obere Dreiecksgestalt.

QED.
Bemerkung
Wir haben gezeigt, durch eine geeignete Wahl der Basis, bekommt die Matrix eines linearen Endomorphismus die Gestalt

$$M = D + N$$

mit einer Diagonalmatrix D und einer Matrix in oberer Dreiecksgestalt N, auf deren Hauptdiagonalen lauter Nullen stehen. Wenn wir die Matrix M diagonalisieren wollen, müssen wir uns also noch um den 'störenden Rest' N kümmern. Dies ist der Inhalt des nächst Abschnitts.

5.2. Nilpotente Endomorphismen

5.3 Die Jordansche Normalform (beliebiger Matrizen)

5.4 Satz von Cayley-Hamilton

6. Bilineare Abbildungen

7. Ergänzungen

Index

Inhalt