Lineare Algebra

B. Herzog, Universität Leipzig, Institut für Mathematik und Informatik,
Vorlesung des ersten Studienjahrs im Herbstsemester 2007

Hinweise

Aufgaben

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 2/3 der Gesamtpunktzahl für die Lösung der Aufgaben erhalten haben.

Sie sollten versuchen alle Aufgaben zu lösen, auch wenn ihre aktuelle Punktzahl oberhalb von 2/3 des Maximums liegt, denn die Aufgaben werden nicht leichter werden.

Vorlesungsmanuskript

Diese Vorlesung wird einer Vorlesung, die ich im Wintersemester 2003/2004 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
$\mathbb{C}$	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. 3.2.7
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. 3.2.8
$\mathbb{F}_2$	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 $n\times n$-Matrizen über den reellen Zahlen $\mathbb{R}$, vgl. 2.6.1
$GL(n, R)$	Gruppe der umkehrbaren $n\times n$-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
$J_d(c)$	$d\times d$-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 $m\times n$-Matrizen mit Einträgen aus K, vgl. 3.2.2
$L_K(V', V'')$	Vektorraum der K-linearen Abbildungen $V'\to V''$, vgl. 3.1
$\bigwedge(V)$	die äußere Algebra des K-Vektorraums V, vgl. 6.6.20
$M_n(\mathbb{R})$	Ring der $n\times n$-Matrizen mit Einträgen aus dem Körper $\mathbb{R}$ der reellen Zahlen, vgl. 2.6.2
$M_{m,n}(K)$	K-Vektoraum der $m\times n$-Matrizen mit Einträgen aus K, vgl 3.2.2
$M_w^v(f)$	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
$\mathbb{P}_K^n$	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
$\mathbb{Z}$	Ring der ganzen Zahlen, vgl. 2.6.2
$\|A\|$	Determinante der quadratischen Matrix A, vgl. 4.3.1
$\|z\|$	Betrag der komplexen Zahl z, vgl. 2.8.1
$\|q\|$	Betrag des Quaternions q, vgl. 2.8.2
$\{v_i^*\}$	die zur Basis $\{v_i\}$ duale Basis, vgl. 3.3.17.
$\bar{z}$	die zur komplexen Zahl z konjugierte komplexe Zahl, vgl. 2.8.1
$\bar{q}$	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
$<M>$	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 $n\times n$-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. 3.2.8
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

1.1 Eine Lösungsformel

Wie löst man

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

oder allgemeiner

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

Eine einfache Rechnung liefert:

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

mit

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

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 <u>Gauß-Algorithmus</u> (obwohl Gauß wohl kaum der erste war, der ihn verwendet hat).

Das Problem:

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

Grundidee:

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

Was ist eine Lösung von (1)?

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

$$(c_1,...,c_n)$$

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:

$$\begin{aligned} a_1 \cdot x_1 \qquad\qquad &= b_1 \\ a_2 \cdot x_2 \qquad &= b_2 \\ \dots \quad & \\ a_n x_n &= b_n \end{aligned}$$

Bemerkungen

(i) Diese System besitzt die einzige Lösung

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

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 alle zugehörigen b_i ebenfalls Null, sind 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

$$\begin{aligned} A \qquad &\equiv 1x + 2y - 3 &&= 0 \\ A + 17B \qquad &\equiv (4x + 5y - 3)+17(1x + 2y - 6) &&= 0 \end{aligned}$$

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:

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

Allgemeiner Fall

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

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

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 $x = (x_1, \ldots, x_n)$. Das Gleichungssystem bekommt dann die Gestalt

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

Grundlegener Fakt: 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

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

schreibt man

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

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.

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:

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

Lösung: (-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:

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

Äquivalent dazu ist das System:

$$x = z - 2$$

$$y \qquad = -2z + 3$$

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.
Menge der Lösungen: (x,y,z) = (t-2,-2t+3,t) mit t beliebig. Insbesondere bekommt man 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 ist gerade

$$0 \cdot z = 1.$$

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

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].

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 Null werden. Die erweiterte Koeffizientenmatrix bekommt dadurch die Gestalt

[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.

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

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

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

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

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

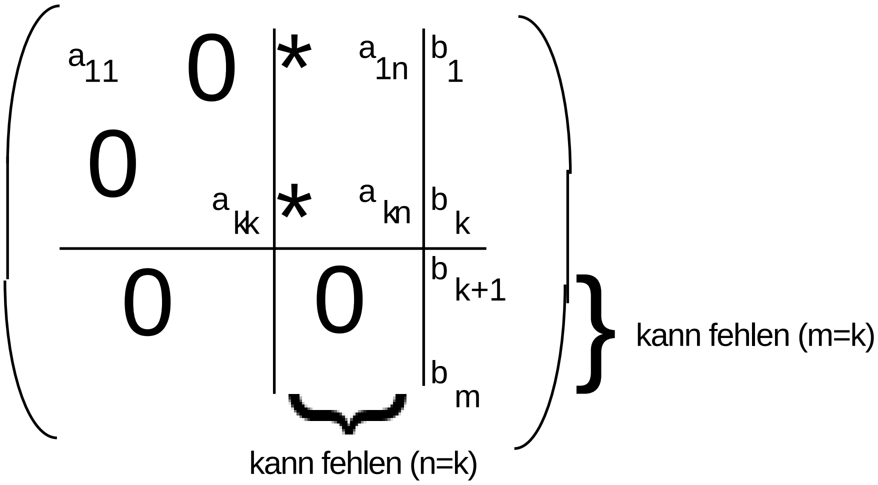
3. Wir wiederholen jetzt die eben ausgeführten Opertionen in analoger Weise, indem wir erst dafür sorgen, daß der Eintrag a_{22} ungleich Null wird und danach dafür, daß alle anderen Einträge der zweiten Spalte Null werden. Danach gehen wir analog mit der dritten Spalte vor. Wir erreichen nach endlich vielen Schritten,daß die Matrix die folgende Gestalt bekommt:

(*) In den Spalten 1 bis k ist genau der Eintrag a_{ii} (i=1,...,k) 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 (*) 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:

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

mit $a_{11} \neq 0, \dots, a_{kk} \neq 0$. 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:



1.4.3 Das Lösungsverhalten

Fall 1: Falls einer der Einträge $b_{k+1},\dots,b_m$ ungleich Null ist, so hat das System keine Lösung.

Fall 2: Falls $b_{k+1}=\dots=b_m=0$ gilt (oder $m\leq k$, d.h. die b_i mit $k < i$ garnicht vorkommen) so gibt es die beiden folgenden Möglichkeiten.

Fall 2.1: k ist gleich n (d.h. gleich der Anzahl der Unbestimmten). Dann gibt es genau eine Lösung des Gleichungssystems. Der Wert der Unbekannten zur i-ten Spalte ist gleich

$$\frac{b_i}{a_{ii}}$$

Fall 2.2: k ist kleiner als n. Dann gibt es unendlich viele Lösungen. Die Unbestimmten zu den Spalten k+1 bis n können beliebig gewählt werden. Zu jeder solchen Wahl gibt es eindeutig bestimmte Werte für die übrigen Unbestimmten derart, das insgesamt eine Lösung entsteht:

$$x_i = \frac{b_i}{a_{ii}} - \frac{a_{ik+1}}{a_{ii}} x_{k+1} - \frac{a_{ik+2}}{a_{ii}} x_{k+2} - \dots - \frac{a_{in}}{a_{ii}} x_n$$

wenn x_i die Unbekannte zur i-ten Spalte ist.

1.4.4 Zusammenfassung

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

Bemerkung

Im dritten Fall ist die Beschreibung ungenau: nicht jede unendliche Menge von n-Tupeln kommt als Lösungsmenge in Frage. Später werden wir in der Lage sein, die Beschreibung zu präzisieren.

1.5 Matrizenmultiplikation

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.

$$1x + 2y + 3z = u$$

$$2x + 3y + 4z = v$$
$$3x + 4y + 6z = w$$

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.

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

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 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.

Verallgemeinerung

Seien die beiden folgenden Systeme gegeben.

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= u_1 \\ &\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= u_m \end{aligned}$$

und

$$\begin{aligned} b_{11}u_1 + \dots + b_{1m}u_m &= v_1 \\ &\dots \\ b_{p1}u_1 + \dots + b_{pm}u_m &= v_p \end{aligned}$$

Die zugehörigen Matrizen sind die folgenden

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

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

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

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

und für das zweite

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

Durch einsetzen erhalten wir

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

d.h.

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

Anders ausgedrückt,

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

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

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

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×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 <u>Summe</u> 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.

<u>Kurzschreibweise</u>:

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

Zwei Matrizen werden Addiert, indem man alle Einträge in denselben Positionen addiert. Eine <u>Nullmatrix</u> ist eine Matrix, deren sämtliche Einträge Null sind.

Ist a eine Zahl, so heißt die Matrix

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

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

$$M - N$$

und nennt diese Matrix <u>Differenz</u> von M und N.

2.2 Eigenschaften der Matrizenaddition

(i) $M+N = N + M$

(ii) $(M+N) + P = M + (N+P)$

(iii) $M + 0 = 0 + M = M$ falls 0 die Nullmatrix ist.

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

(Eine oder zwei Eigenschaften beweisen).

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

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

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 verkettet 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 Einheitsmatrix.

2.4 Eigenschaften der Matrizenmultiplikation

(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) EM = M und ME' = M für Einheitsmatrizen E, E' geeigneten Typs.

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 Anwendung auf lineare Gleichungssysteme

2.5.1 Lineare Gleichungssysteme in Matrizenschreibweise

Sei ein lineares Gleichungssystem gegeben.

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$

(1) $\quad\dots$

$$a_{m1}x_1 + ... + a_{mn}x_n = b_m$$

Dann heißt die Matrix der a_{ij},

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

Koeffizientenmatrix 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.

2.5.2 Vereinbarung

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

2.5.3 Homogene und inhomogen Gleichungssysteme

Ein lineare Gleichungssystem

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

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

(2) $$Ax = b$$

ein lineares Gleichungssystem in Matrizenschreibweise und

(3) $$Ax = 0$$

das zugehörige homogene Gleichungssystem. Dann gilt:

(i) Sind u und v zwei Lösungen des homogenen Systems (2) 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

$$2x + 3y = 6$$

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 wirklich exakt 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.

2.6 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.

2.6.1 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 $ae = ea = a$ 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 bezügtlich 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

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

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

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

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

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

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

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

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

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

ein Gruppen-Homomorphismus.

Beweis. Beweis von (1). Sei

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

Dann hat α die Gestalt

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

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

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

Sie umgekehrt

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

Dann hat α die Gestalt

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

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

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

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

$$\bar{i} + \bar{0} = \overline{i+0} = \bar{i} = \overline{0+i} = \bar{0} + \bar{i}.$$

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

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

Kommutativgesetz: für beliebige i, $j \in\mathbb{Z}$ gilt

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

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

QED.

2.6.2 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.

Ein Ringhomomorphismus 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 $x, y \in R$.
Dieser Ringhomomorphismus heißt Homomorphismus von Ring 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 Ring mit 1. Eine R-Algebra (mit 1) ist ein Ring mit 1 zusammen mit einem Homomorphismus

$$R \to S$$

von Ringen mit 1.

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 $n\times n$-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 $n\times n$-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 $n\times n$-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: direkte 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 direktes Produkt 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

$$\overline{i} \cdot \overline{j} = \overline{i\cdot j} \tag{2}$$

Insbesondere ist die Gruppe der Restklassen modulo n,

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

ein kommutativer Ring mit 1 bezüglich der Multiplikation (2) und die Abbildung

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

ein Homomorphismus von Ringen mit 1.

Beweis. Beweis von (2). 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

$$\alpha = kn + ij = (k\cdot n + i)(0\cdot n + j) + kn - knj = (k\cdot n + i)(0\cdot n + j) + (k\text{-}kj)n,$$

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 $i, j, k\in\mathbb{Z}$:

$$(\overline{i} \cdot \overline{j})\cdot\overline{k} = \overline{i\cdot j} \cdot\overline{k} = \overline{ijk} = \overline{i}\cdot\overline{jk} = \overline{i} \cdot (\overline{j}\cdot\overline{k}),$$

d.h. es gilt das Assoziativgesetz. Die Kommutativität von $\mathbb{Z}/(n)$ ergibt sich aus der von $\mathbb{Z}$:

$$\overline{i} \cdot \overline{j} = \overline{i\cdot j} = \overline{j\cdot i} = \overline{j}\cdot\overline{i}$$

Der Ring $\mathbb{Z}/(n)$ hat ein Einselement, denn es gilt

$$\overline{1}\cdot\overline{j} = \overline{1\cdot j} = \overline{j}.$$

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 $\overline{1}$ das Einselement des Restklassenrings ist.

QED.

2.6.3 Begriff des Körpers

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

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) mit den durch die folgenden Tabellen gegebenen Operationen ist ein Körper.

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

·	**0**	**1**
0	0	0
1	0	1

Beispiel 4

Die Menge

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

(der Restklassen modulo 6) mit den durch die folgenden Tabellen gegeben Operationen

+	**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

ist ein kommutativer Ring mit 1, jedoch 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

$$(a_0 \ge)\ a_1 > a_2 > ... > a_r ,$$

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

$$a \ge b \ge 0$$

(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 $0 \le r_i < a_i$ $\qquad (q_i , r_i \in \mathbb{Z})$

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

<u>Folgerung 1</u>

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.

<u>Folgerung 2</u>

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:

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

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

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

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

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

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

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

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 $\bar{i} \in \mathbb{Z}/(n)$ von Null verschieden, d.h. die ganze Zahl ist kein Vielfaches von n. Weil n eine Primzahl ist, gilt d.h. es gibt ganze Zahlen u, v mit

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

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

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

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

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

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

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

2.7 Eine weitere Matrizenoperation

2.7.1 Transponierte Matrizen

Seien K ein Körper und $A = (a_{ij}) \in K^{m\times n}$ eine Matrix mit Einträgen aus K. Dann heißt die Matrix

$$A^T = (a'_{ij}) \in K^{n\times m}$$

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

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

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

Insbesondere ist der Übergang zur transponierten Matrix,

$$K^{m\times n} \to K^{n\times m},\ A \mapsto A^T,$$

eine K-lineare Abbildung.
Beweis. Seien $A = (a_{ij})$ und $B = (b_{ij})$. Dann gilt

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

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

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

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

Zu (iv). 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

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

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 $v := a'b''+b'a''$. Die Menge $\mathbb{C}$ ist mit der Addition von Matrizen und der Multiplikation von Matrizen ein Ring mit 1.[2]

Bemerkungen

(i) Es gilt das Kommutativgesetz.[3]

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

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

[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 $n\times n$-Matrizen gelten. Also ist $\mathbb{C}$ ein Ring. . Es ist sogar ein Ring mit 1, die Einheitsmatrix spielt die Rolle des Einselements.

[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.

Man beachte, die Abbildung ist injektiv. Es ist sogar ein Homomorphismus von Ringen mit 1.

(iii) 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

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

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

(iv) 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

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

(v) 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:

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

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

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

Nach Definition gilt[5]

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

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

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

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

---

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

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

(2) $\mathrm{Re}(z) \leq |z|$ und $\mathrm{Im}(z) \leq |z|$.

Weiter gilt

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

Wegen (2) erhält man damit

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

also gilt die Behauptung.

2.8.2 Der Schiefkörper der Quaternionen

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

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

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.

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

Mit der Addition und der Multiplikation von Matrizen ist $\mathbb{H}$ ein Ring mit 1. Man zeigt das auf dieselbe Art und Weise wie bei $\mathbb{C}$.

Bemerkungen

(i) Der Ring der Quaternionen ist nicht kommutativ. Jedes Quaternion kommutiert allerding mit jeder reellen Zahl,

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

(ii) 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.

(iii) 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.

$$i^2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = (-1) + i\cdot 0 = -1.$$

$$j^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = (-1) + i\cdot 0 = -1.$$

$$k^2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = (-1) + i\cdot 0 = -1.$$

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

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

Weiter ist damit

$ik = iij = -j$

$ki = iji = -jii = j$

$jk = jij = -\ jji = i$

$kj = ijj = -i$

(iv) 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.

(v) 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 überführt Summen und Produkte in Summen und Produkte, wobei sich bei den Produkten die Reihenfolge der Faktoren umkehrt.

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

Diese Abbildung ist ein $\mathbb{R}$-Algebra-Isomorphismus[6]

$$\mathbb{H} \to \mathbb{H}^{op},\ q \ \mathrm{a}\ \bar{q}.$$

Sie heißt <u>Konjugation</u>. Das Quaternion $\overline{a+ib+jc+kd}$ heißt das zu a+ib+jc+kd konjugierte Quaternion.

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

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

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

$$q\cdot\bar{q} =^{7} a^2+b^2+c^2+d^2.$$

Nach Definition ist

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

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

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

---

[6] 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.

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

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

Die Konjugationsabbildung bekommt dann die Gestalt

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

d.h. es ist gerade der Übergang zur transponierten Matrix. Von der Transposition wisser wir aber, daß es sich um einen Algebrahomorphismus handelt, der zu sich selbst invers ist.

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

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

(vii) 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:

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

und

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

3. Vektorräume

In 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) ______m für Isomorphie. Jede 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 folgendne 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
$$\begin{aligned}(g\circ f)(\lambda v+\lambda' v') &= g(f(\lambda v+\lambda' v'))\\ &= g(\lambda\, f(v)+\lambda'\, f(v')) && \text{(da f linear ist)}\\ &= \lambda g(f(v))+\lambda'\, g(f(v')) && \text{(da g linear ist)}\\ &= \lambda(g\circ f)(v)+\lambda'(g\circ f)(v')\end{aligned}$$
Also ist auch $g\circ f$ linear.

Zu (iii). Nach (ii) ist die Komposition von linearen Automorphismen eine wohldefinierte Abbildung
$$\mathrm{Aut}_K(V)\times \mathrm{Aut}_K(V) \to \mathrm{Aut}_K(V).$$
Diese Komposition ist als Komposition von Abbildungen bekanntermaßen assoziativ. 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 $\cdot$ dieselben sind wie die von V (genauer : die Einschränkungen der entsprechenden Operationen von V).

Zu (v). Wir haben zu zeigen, die Vektorraumaxiome sind für die Menge W erfüllt. 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 folgenden Operationen.

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

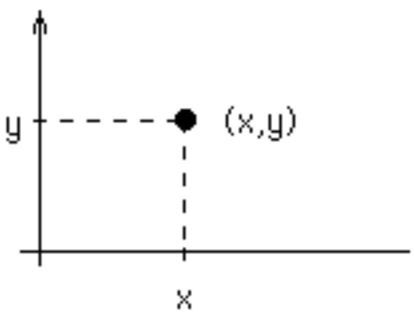
$$\lambda \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \dots \\ \lambda a_n \end{pmatrix}$$

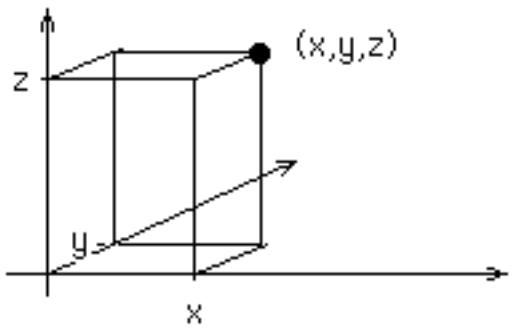
für a_i, b_i, $\lambda\in K$. Mit diesen Operationen ist K^n ein K-Vektorraum.

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.

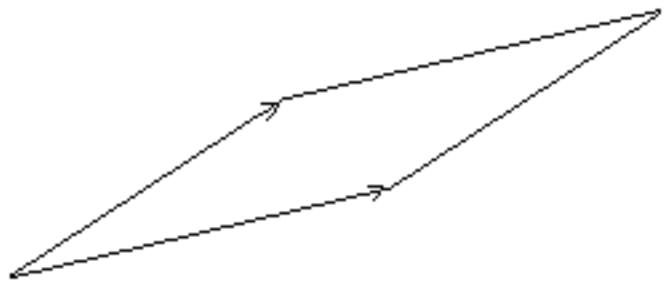




(iii) Im Fall n=3 betrachtet man die Elemente von K^n auch als Punkte eines Raumes, die durch drei Koordinaten gegeben sind. Mit der Menge K^3 selbst verbindet man die Vorstellung eines dreidimensionalen Raumes.

(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 $(0,...,0)^T$ und der Spitze in $(a_1,...,a_n)^T$ bezeichnet man dabei einfach mit $(a_1,...,a_n)^T$.

(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.

3.2.3 Abbildungen mit Werten in einem Vektorraum

Seien M eine Menge, K ein Körper und V ein K-Vektorraum. Dann ist die Menge

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

mit den folgenden Operationen ein K-Vektoraum.

$$(f+g)(m) := f(m) + g(m)$$
$$(a\cdot f)(m) := a\cdot f(m).$$

3.2.3 Lineare Abbildungen

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

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

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

$$\mathrm{Abb}(V', V").$$

3.2.4 Direktes Produkt

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 direktes Produkt oder auch direkte Summe 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.

$$\{v_i\} + \{v'_i\} := \{v_i + v'_i\}$$
$$a\cdot\{v_i\} := \{a\cdot v_i\}$$

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

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

(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 ihr j-tes Glied abbilded, eine K-lineare Abbildung. Sie heißt Projektion auf den j-ten Faktor.

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

3.2.6 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

$$\langle M\rangle := \{c_1 v_1 + ... + c_k v_k \mid c_1,...,c_k \in K,\ v_1,...,v_k \in M,\ k=1,2,3,...\}$$

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

$$\langle M\rangle = V$$

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,

$$M \subseteq W,\ W \text{ linearer Unterraum} \Rightarrow \langle M\rangle \subseteq W.$$

(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 $\langle M\rangle$ ist selbt ein linearer Unterraum, welcher die Menge M enthält. Der Durchschnitt aller linearen Unterräume mit dieser Eigenschaft, ist also gleich $\langle M\rangle$,

$$\langle M\rangle = \bigcap_{\substack{M\subseteq W \\ W\subseteq V \text{ linear}}} W$$

(iv) Die Eigenschaft von $\langle M\rangle$, 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 $\langle M\rangle$ der kleinste Unterraum ist, welcher die Menge M enthält.

$\langle M\rangle$:= kleinster linearer Unterraum, welcher die Menge M enthält.

3.2.7 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.

3.2.8 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 Abbildung, welche gleich 1 an der Stelle m und sonst gleich Null ist,

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

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 frei erzeugte K-Vektorraum. Er wird mit

$$F_K(M)$$

bezeichnet.

3.2.9 Faktorräume

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

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

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 (*).

2. Schritt: 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.

...

Zu (iii). Folgt unmittelbar aus der Definition der Vektorraumoperationen auf V/W.
QED.

3.2.10 Bild und Kern einer linearen Abbildung

Seien K ein Körper und f: V $\to$ V' eine K-lineare Abbildung. Dann gilt:

(i) Das Bild von f,

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

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 Kern der Abbildung f).

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

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

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

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

(1) $v+w\in im(f)$
(2) $c\cdot v\in im(f)$
Nach Voraussetzung gibt es Vektoren $v, w\in V$ mit

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

Also gilt

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

d.h. es gelten (1) und (2).
2. Schritt. 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 $0\in V'$ in im(f) liegt: $0 = f(0)\in im(f)$.
Jedes Element $v'\in im(f) \subseteq V'$ hat die Gestalt $v' = f(v)$ für ein $v\in V$. 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 $f(-v)\in im(f)$ 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.
3. Schritt. 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

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

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

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

Andernfalls heißen die Vektoren $v_1,...,v_k$ (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

$$c_1v_1 + ... + c_kv_k$$

mit $v_i \in V$ und $c_i \in K$ für alle i heißt Linearkombination der Vektoren $v_1,...,v_k$ 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

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

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

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

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

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

im Widerspruch zu unserer Annahme.

(iii) ⇒ (ii). 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

$$
\begin{aligned}
0 &= c_1v_1 + ... + c_{i-1}v_{i-1} + (-1)v_i + c_{i+1}v_{i+1} + ... + c_kv_k \\
&= \varphi(c_1, ... ,c_{i-1},-1,c_{i+1}, ...,c_k)
\end{aligned}
$$

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

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

$$c_1v_1 + ... + c_kv_k = 0.$$

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

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

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

$$\varphi: \oplus_{i\in I} K \to V,\ \{c_i\}_{i\in I} \mapsto \sum_{i\in I} c_iv_i,$$

injektiv.

Beweis. Übungsaufgabe.

QED.

3.3.4 Basen eines Vektorraumes

Seien K ein Körper und V ein K-Vektorraum. Eine <u>Basis</u> von V (über K) ist eine Menge

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

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

$$c_{i_1} v_{i_1} + ... + c_{i_k} v_{i_k} = 0 \text{ mit } c_{i_1},...,c_{i_k} \in K$$

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). <u>1. Schritt</u>. 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.
<u>2. Schritt</u>. 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

$$\varphi:K^n \to V, (c_1, ... , c_n) \mapsto c_1v_1 + ... + c_nv_n,$$

ist ein Isomorphismus.

Sind diese Bedingungen erfüllt, so nennt man die Abbilung φ von Bedingung (iv) auch Koordinatenabbildung zur Basis $v_1,...,v_n$.

Beweis. (i) $\Rightarrow$ (iv). Sei $v_1,...,v_n$ eine Basis. Dann ist jedes Element von V Linearkombination der $v_1,...,v_n$, d.h. die Abbildung ist surjektiv. Auf Grund unserer Charakterisierung der linearen Unabhängigkeit (vgl. 4.3.3) ist die Abbildung auch injektiv.

(iv) $\Rightarrow$ (iii). Auf Grund der Surjektivität der Abbildung φ bilden die Vektoren $v_1,...,v_n$ ein Erzeugendensystem von V. Angenommen, der Raum V wird auch von den Vektoren $v_1,...,v_{i-1},v_{i+1},...,v_n$ erzeugt. Dann ist v_i eine Linearkombination dieser Vektoren, sagen wir

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

Dann ist aber $(c_1,...,c_{i-1}, -1, c_{i+1},...,c_n)\in K^n$ ein von Null verschiedener Vektor mit demselben Bild wir der Nullvektor:

$$\begin{aligned}\varphi(c_1,...,c_{i-1}, -1, c_{i+1},...,c_n) &= c_1v_1+...+c_{i-1}v_{i-1}+(-1)v_i+c_{i+1}v_{i+1}+...+c_nv_n\\ &= v_i - v_i\\ &= 0\\ &= \varphi(0,...,0).\end{aligned}$$

Dies steht aber im Widerspruch zur Injektivität von φ, d.h. die Vektoren $v_1,...,v_{i-1},v_{i+1},...,v_n$ erzeugen nicht den ganzen Raum V.

(iii) $\Rightarrow$ (ii). Sei $v_1,...,v_n$ 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) $$v_1 = c_2v_2 + ... + c_nv_n.$$

Sei jetzt $v\in V$ ein beliebiger Vektor. Da die $v_1,...,v_n$ den Raum V erzeugen, gilt

$$v = c'_1v_1+...+c'_nv_n$$

mit gewissen $c'_1,...,c'_n\in K$. Durch Einsetzen von (1) erhalten wir

$$\begin{aligned}v &= c'_1(c_2v_2 + ... + c_nv_n)+c'_2v_2...+c'_nv_n\\ &= (c'_1c_2+c'_2)v_2 + (c'_1c_3+c'_3)v_3+ ... + (c'_1c_n+c'_n)v_n.\end{aligned}$$

Wir haben gezeigt, jeder Vektor $v\in V$ läßt sich als Linearkombination der Vektoren $v_2, ... , v_n$ schreiben. Das stehe aber im Widerspruch zu unserer Voraussetzung, daß die Vektoren $v_1,...,v_n$ ein minimales Erzeugendensystem bilden sollen. Also ist die Annahme, daß sie linear abhängig sind, falsch. Wir haben damit gezeigt,

(2) $v_1,...,v_n$ 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 $v_1,...,v_n$ ein Erzeugendensystem bilden, läßt sich der Vektor v als Linearkombination der $v_1,...,v_n$ schreiben. Dann sind die Vektoren

$$v_1,...,v_n, v$$

aber linear abhängig.

(ii) ⇒ (i). Sei jetzt $v_1,...,v_n$ eine maximale linear unabhängige Familie. Wir haben zu zeigen, die Vektoren $v_1,...,v_n$ erzeugen den ganzen Raum V. Sei $v \in V$ beliebig. Dann sind die Vektoren

$$v_1,...,v_n, v$$

nicht mehr linear unabhängig, d.h. eine nicht-triviale Linearkombination dieser Vektoren ist Null,

(3) $$c_1v_1+...+c_nv_n+ cv = 0.$$

Dabei kann der Koeffizienten c nicht gleich Null sein, denn andernfalls wären die Vektoren $v_1,...,v_n$ 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

$$v =(- c_1)v_1+...+(-c_n)v_n.$$

Wir haben gezeigt, jeder Vektor $v \in V$ läßt sich als Linearkombination der Vektoren v_1 ,...,v_n schreiben. Diese Vektoren $v_1,...,v_n$ 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 $\mathfrak{M} := \{M_i \mid i\in I\}$ eine Familie von nicht-leeren Mengen. Dann kann man aus jeder der Mengen von $\mathfrak{M}$ ein Element auswählen, d.h. es gibt eine Abbildung

$$\varphi: \mathfrak{M} \to \cup\mathfrak{M} := \cup_{i\in I} M_i$$

mit der Eigenschaft, daß für jedes $M\in\mathfrak{M}$ das Bild $\varphi(M)$ ein Element der Menge M ist,

$$\varphi(M)\in M \text{ für jedes } M\in\mathfrak{M}.$$

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 Beweise 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 Gleichheit 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

$$R_2 := \{(x,y)\in \mathbb{R}\times\mathbb{R} \mid x-y\in \mathbb{R}_+\}$$

heißt Größerrelation 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

$$R_3 := \{(x,y)\in \mathbb{R}\times\mathbb{R} \mid y-x\in \mathbb{R}_+\cup\{0\}\}$$

heißt Kleinergleichrelation auf $\mathbb{R}$ und wird mit ≤ bezeichnet, d.h.

$$x\le y$$

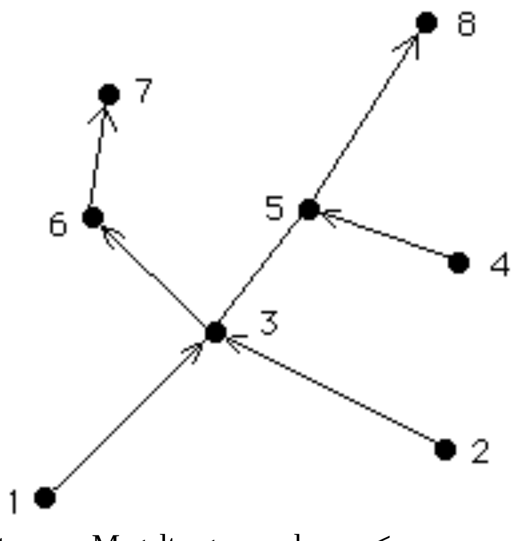
bedeutet, die reelle Zahl x ist kleiner oder gleich der reellen Zahl y. Diese Relation ist reflexiv, anti-symmetrisch, nicht symmetrisch, aber sie ist transitiv. Es handelt sich also um keine Äquivalenzrelation aber um eine Halbordnung, und zwar sogar um eine lineare Halbordnung.

4. Die Relation

$$R_4 := \{(x,y)\in \mathbb{R}\times\mathbb{R} \mid |y-x| = 1\}$$

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

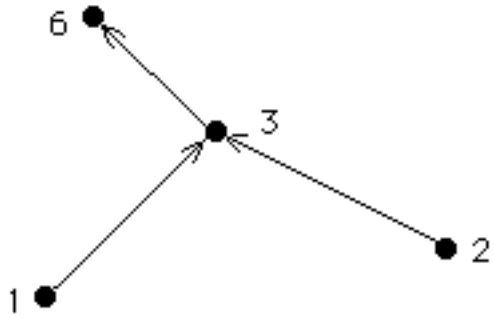
Seien M eine Menge, ≤ eine Halbordnung auf M und $S\subseteq M$ eine Teilmenge von M. Ein Element $m\in M$ heißt obere Schranke von S in M, wenn

$$x \le m$$

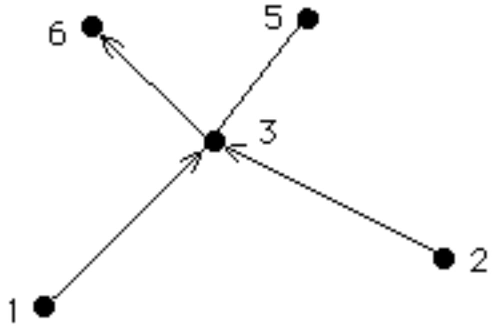
gilt für jedes $x\in S$. Ein Element $m\in M$ heißt maximal, wenn für jedes $x\in M$ mit $m\le x$ automatisch sogar $m = x$ gilt.

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 M 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.

Zum Beweis siehe A. G. Kuroš: Allgemeine Algebra, §6, Satz von Kuratowski-Zorn.

3.3.7.5 Die Exisitenz Vektorraumbasen

Seien K ein Körper und V eine K-Vektorraum. Dann gibt es eine Familie $\{v_i\}_{i\in I}$ von Elementen von 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

$$\mathfrak{M} := \{S \mid S \text{ ist K-linear unabhängige Teilmenge von V}\}$$

die Menge aller linear unabhängigen Teilmengen des Vektorraumes V. Für je zwei Elemente

$$S',\ S'' \in \mathfrak{M}$$

schreiben wir

$$S' \le S''$$

wenn $S' \subseteq S''$ gilt. Auf diese Weise ist auf der Menge $\mathfrak{M}$ eine Halbordnung $\le$ definiert. Ein maximales Element bezüglich dieser Halbordnung ist nach unserer Charakterisierung der Basen (vgl.4.3.6) gerade eine Basis von V. Es reicht also, wenn wir zeigen, die Menge $\mathfrak{M}$ 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

$$\mathfrak{S} := \{S_i \mid i\in I\}$$

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 $\mathfrak{M}$ ist, so ist S eine obere Schranke von $\mathcal{S}$ und der Beweis der Behauptung ist abgeschlossen.

Es genügt also zu zeigen, $S\in\mathfrak{M}$, d.h. S besteht aus linear unabhängigen Vektoren. Seien

$$v_1,...,v_n \in S$$

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

$$v_j \in S_{i_j} \text{ (j=1,...,n).}$$

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

$$S_{i_1} \subseteq S_{i_2} \subseteq ... \subseteq S_{i_n}.$$

Dann gilt aber

$$v_1,...,v_n \in S_{i_n} \; (\in \mathcal{S} \subseteq \mathfrak{M}).$$

Als Element von $\mathfrak{M}$ besteht die Menge $\mathcal{S}$ aus linear unabhängigen Vektoren. Insbesondere sind die Vektoren $v_1,...,v_n$ 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, $v_1,...,v_n$ ein Erzeugendensystem von V und $w_1,...,w_m$ beliebige linear unabhängige Vektoren von V. Dann gibt es n-m Vektoren

$$v_{i_1},...,v_{i_{n-m}}$$

des gegebenen Erzeugendensystems derart, daß

$$v_{i_1},...,v_{i_{n-m}}, w_1,...,w_m$$

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 $v_1,...,v_n$ ein Erzeugendensystem ist, läßt sich der Vektor $w = w_1$ als Linearkombination der v_i schreiben,

$$w = c_1v_1 +...+ c_nv_n.$$

Nach Voraussetzung ist $w \neq 0$, also muß auch einer der Koeffizienten c_i von Null verschieden sein,

$$c_i \neq 0.$$

Dann gilt

$$v_i = (-\frac{c_1}{c_i})v_1 + ... + (-\frac{c_{i-1}}{c_i})v_{i-1} + \frac{1}{c_i}w + (-\frac{c_{i+1}}{c_i})v_{i+1} + ... + (-\frac{c_n}{c_i})v_n.$$

Mit anderen Worten, v_i liegt in dem von w und den $v_1, ..., v_{i-1}, v_{i+1}, ..., v_n$ erzeugten Unterraum. Es gilt damit

$$v_1, v_2, ..., v_n \in \langle w, v_1, ..., v_{i-1}, v_{i+1}, ..., v_n\rangle,$$

also

$$(V =)\langle v_1, v_2, ..., v_n\rangle \subseteq \langle w, v_1, ..., v_{i-1}, v_{i+1}, ..., v_n\rangle \ (\subseteq V),$$

also

$$V = \langle w, v_1, ..., v_{i-1}, v_{i+1}, ..., v_n\rangle$$

Mit anderen Worten, die Vektoren $w, v_1, ..., v_{i-1}, v_{i+1}, ..., v_n$ 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

$$w = c_1v_1 +...+ c_nv_n.$$

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) $$v_{i_1},...,v_{i_{n-m+1}}$$

derart, daß

(2) $$v_{i_1},...,v_{i_{n-m+1}}, w_1,...,w_{m-1}$$

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:

$$w_m = c_1 v_{i_1} +...+c_{n-m+1} v_{i_{n-m+1}} + c'_1 w_1 +...+c'_{m-1} w_{m-1}.$$

Dabei können die Koeffizienten c_1, ..., c_{n-m+1} nicht sämtlich Null sein, denn andernfalls wären die Vektoren $w_1,...,w_m$ linear abhängig. Nach der Bemerkung am Ende des ersten Schritts läßt sich also w_m gegen einen der Vektoren $v_{i_1},...,v_{i_{n-m+1}}$ 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) $$\{v_i \mid i\in I\} = \{v_{i_1},...,v_{i_n}\},$$

wäre endlich und die andere unendlich. Dann können wir aus der unendlichen Menge n+1 Vektoren auswählen.

(2) $$w_{j_1}, ..., w_{j_{n+1}}.$$

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.
2. Schritt. Gleichheit der Elementzahl im endliche Fall.
Seien also beide Basen endlich,

$$\{v_i \mid i\in I\} = \{v_{i_1},...,v_{i_n}\},$$
$$\{\{w_j \mid j\in J\} = \{w_{j_1},...,w_{j_m}\}.$$

Da beide Basen insbesondere Erzeugendensysteme sind und aus linear unabhängigen Vektoren bestehen, gilt nach dem Satz von Steinitz sowohl n≤m also auch m≤n.
QED.
Bemerkung
Bijektive lineare Abbildungen f:V→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 $f: V \to W$ ist somit linear.

QED.

3.3.12 Die Dimension von Kern und Bild einer linearen Abbildung

Sei $f: V \to V'$ 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. Wir haben zu zeigen, die Vektoren v_i und w_j sind linear unabhängig und erzeugen den Raum V.

1. Schritt. Die v_i und w_j erzeugen den Raum V.

Sei $v\in V$ 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

$$v = c_1 \cdot v_{i_1} + ... + c_r \cdot v_{i_r} + d_1 w_{j_1} + ... + d_s w_{j_s}.$$

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.

2. Schritt. 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) $$c_1 \cdot v_{i_1} + ... + c_r \cdot v_{i_r} + d_1 w_{j_1} + ... + d_s w_{j_s} = 0$$

mit $c_1, ... + c_r, d_1, ... , d_s \in K$. Wir haben zu zeigen, sämtliche Koeffizienten sind Null. Wir wenden f aus die Identität (3) an und erhalten

$$\begin{aligned} 0 &= c_1 \cdot f(v_{i_1}) + ... + c_r \cdot f(v_{i_r}) + d_1 f(w_{j_1}) + ... + d_s f(w_{j_s}) \\ &= c_1 \cdot f(v_{i_1}) + ... + c_r \cdot f(v_{i_r}). \end{aligned}$$

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) $$c_1 = ... = c_r = 0.$$

Wir setzen (4) in die Ausgangsidentität (3) ein und erhalten

(5) $d_1 w_{j_1} + ... + d_s w_{j_s} = 0.$

Nach Voraussetzung bilden die w_j eine Basis (von ker(f)) sind also linear unabhängig. Mit (5) gilt also

(6) $d_1 = ... = d_s = 0.$

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.

Folgerung: 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

Beweis: ergibt sich aus den Beweisen der letzten beiden Sätze.

3.3.14 Exakte Sequenzen

Eine Folge von K-linearen Abbildungen

$$... \to V_i \xrightarrow{f_i} V_{i+1} \xrightarrow{f_{i+1}} V_{i+2} \to ...$$

heißt Komplex, wenn die Zusammensetzung von je zwei benachbarten Abbildungen Null ist,

$$f_{i+1} \circ f_i = 0 \text{ für jedes } i \in \mathbb{Z}.$$

Diese Bedingungen sind äquivalent zu den Inklusionen

$$\operatorname{im}(f_i) \subseteq \ker(f_{i+1}) \text{ für jedes } i \in \mathbb{Z}.$$

Der Komplex heißt exakt an der Stelle V_{i+1}, wenn anstelle der Inklusion sogar das Gleichheitszeichen gilt,

$$\operatorname{im}(f_i) = \ker(f_{i+1}).$$

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 von 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).

3.3.15 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) $$0 \to 0 \xrightarrow{\alpha} \operatorname{im}(f_2) \xrightarrow{i} V_3 \xrightarrow{f_3} \ldots$$

Dabei sei i(v) = v für jedes $v\in \operatorname{im}(f_2)$. Dieser Komplex ist an allen Stellen mit eventueller Ausnahmen der Stellen $\operatorname{im}(f_2)$ 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.
<u>Exaktheit an der Stelle</u> $\operatorname{im}(f_2)$: Nach Definition von i ist
$$\ker(i) = 0 = \operatorname{im}(\alpha).$$
<u>Exaktheit an der Stelle</u> V_3: Es gilt
$$\operatorname{im}(i) = \{i(v) \mid v\in \operatorname{im}(f_2)\} = \{v \mid v \in \operatorname{im}(f_2)\} = \operatorname{im}(f_2) = \ker(f_3).$$
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) $$\dim \operatorname{im}(f_2) + \sum_{i=3}^{\infty} (-1)^i \dim V_i = 0.$$

Weiter ist
$$\dim \operatorname{im}(f_2) = \dim V_2 - \dim \ker(f_2) \qquad \text{(nach 3.3.11)}$$

$= \dim V_2 - \dim im(f_1)$ (weil (1) exakt ist)
$= \dim V_2 - (\dim V_1 - \dim ker(f_1))$ (nach 3.3.11)
$= \dim V_2 - \dim V_1$,

denn $ker(f_1) = im(0 \to V_1) = 0$. Wir setzen das Ergebnis dieser Berechnung in (3) und erhalten die Behauptung:

$$\sum_{i=1}^{\infty} (-1)^i \dim V_i = 0.$$

QED.

3.3.16 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

$$ker(f) + im(f) = \dim V'\oplus V".$$

Es reicht also zu zeigen,

(1) $ker(f) = V"$
(2) $im(f) = V'$

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

$$im(f) = \{f(v',v") \mid v'\in V' \text{ und } v"\in V"\} = \{v' \mid v'\in V'\} = V'.$$

QED.

3.3.17 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'∩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) $0 \to W'\cap W" \xrightarrow{f} W'\otimes W" \xrightarrow{g} W'+W" \to 0$

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'∩W" exakt. Weiter gilt

$$\begin{aligned} ker(g) &= \{(w', w") \mid w'\in W', w"\in W", g(w', w") = 0\} \\ &= \{(w', w") \mid w'\in W', w"\in W", w'+ w" = 0\} \end{aligned}$$

$$= \{(w, -w) \mid w\in W', w\in W''\}$$
$$= \{(w, -w) \mid w\in W'\cap W''\}$$
$$= \{ f(w) \mid w\in W'\cap W''\}$$
$$= \mathrm{im}(f),$$

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.

3.3.18 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 $V \to K$ dieselbe Dimension

$$\dim \mathrm{Hom}(V, K) = \dim V$$

wie 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,

$$(1) \qquad v = \sum_{i\in I} f_i v_i = \sum_{i\in I} f_i(v)\cdot v_i$$

mit eindeutig bestimmten Koeffizienten

$$f_i = f_i(v)\in K.$$

Betrachten wir die Abbildungen

$$(2) \qquad f_i: V \to K,\ v \mapsto f_i(v).$$

<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,

$$f_i\in \mathrm{Hom}_K(V,K).$$

<u>2.Schritt</u>: Die $f_i\in \mathrm{Hom}_K(V,K)$ sind linear unabhängig. Insbesondere gilt im Fall $\dim V = \infty$ auch $\dim \mathrm{Hom}_K(V,K) = \infty$.

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

(3) $$0 = \sum_{i\in I} c_i \cdot f_i(v_j).$$

Wegen (1) ist nun

(4) $$f_i(v_j) = \begin{cases} 1 \text{ für } i=j \\ 0 \text{ sonst} \end{cases}$$

Auf der rechten Seite von (3) 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.

3. Schritt: Abschluß des Beweises.

Wir haben noch zu zeigen, daß im Fall $\dim V < \infty$ gilt

$$\dim V = \dim \mathrm{Hom}_K(V,K).$$

Dazu reicht es zu zeigen, daß die im zweiten Schritt konstruierte Familie von Linearformen

$$f_i \in \mathrm{Hom}_K(V,K)$$

eine Basis von $\mathrm{Hom}_K(V,K)$ bilden. Im 2. Schritt haben wir bereits gezeigt, die f_i sind linear unabhängig. Es reicht also zu zeigen, sie bilden ein Erzeugendensystem von $\mathrm{Hom}_K(V,K)$.

Sei also $f \in \mathrm{Hom}_K(V,K)$ ein beliebiges Element. Wir setzen

$$c_i := f(v_i)$$

und

(5) $$g := f - \sum_{i\in I} c_i f_i$$

Es reicht zu zeigen, daß g die Nullabbildung ist, denn dann ist jede Linearform f auf V eine Linearkombination der f_i, d.h. die f_i erzeugen den Raum der Linearformen.

Wegen (4) gilt zunächst für jedes $j \in I$,

$$g(v_j) = f(v_j) - \sum_{i\in I} c_i f_i(v_j) = c_j - c_j \cdot f_j(v_j) = 0.$$

Sei jetzt $v \in V$ beliebig. Dann ist v Linearkombination der v_i,

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

Also ist

$$g(v) = g(\sum_{i\in I} c_i v_i) = \sum_{i\in I} c_i g(v_i) = \sum_{i\in I} c_i\cdot 0 = 0,$$

d.h. g ist tatsächlich die Nullabbildung

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

$$f_j(v_i) = \delta_{ij} = \begin{cases} 1 \text{ für } i = j \\ 0 \text{ für } i \ne j \end{cases}$$

Diese Basis heißt die zur Basis der v_i <u>duale Basis</u>. Ihre Vektoren werden wir oft mit

$$v_j^* := f_j$$

bezeichnen.

3.3.19 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) $$\varphi: \oplus_{i\in I}\mathrm{Hom}(V,K) \to \mathrm{Hom}_K(V,V'),\ \{f_i\}_{i\in I} \mapsto (v \mapsto \sum_{i\in I} f_i(v)v'_i),$$

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

$$\varphi(c'\cdot\{f'_i\}_{i\in I} + c''\cdot\{f''_i\}_{i\in I})(v) \quad = \varphi(\{c'\cdot f'_i + c''\cdot f''_i\}_{i\in I})(v)$$

$$= \sum_{i\in I} (c'\cdot f_i + c''\cdot f''_i)(v)v'_i$$

$$= \sum_{i\in I} (c'\cdot f_i(v) + c''\cdot f''_i(v))v'_i$$

$$= c' \sum_{i\in I} f_i(v)v'_i + c''\cdot \sum_{i\in I} f''_i(v)v'_i$$

$$= c'\cdot\varphi(\{f'_i\}_{i\in I}) + c''\cdot\varphi(\{f''_i\}_{i\in I}).$$

2. Schritt. $\ker(\varphi) = 0$.
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).

3. Schritt. 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 $v' = v''$.

4. Schritt. Der Fall $\dim V' = \infty$
Auf Grund des 2. und 3. Schritts ist die Abbildung (1) injektiv, d.h. die direkte Summe

$$\oplus_{i\in I}\mathrm{Hom}(V,K) \qquad (2)$$

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.

5. Schritt. 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

$$f(v) = \sum_{i\in I} f_i(v)\cdot v'_i \tag{3}$$

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 $f: V\to W$ eine K-lineare Abbildung von endlich-dimensionalen Vektorräumen. Weiter seien

$$v_1,\dots,v_m \text{ Basis von } V$$
$$w_1,\dots,w_n \text{ Basis von } W$$

Dann läßt sich jeder der Vektoren $f(v_i)$ auf genau eine Weise als Linearkombination der Basisvektoren $w_1,\dots,w_n$ schreiben,

$$f(v_i) = c_{1i}w_1+\dots+c_{ni}w_n \qquad (i=1,\dots,m)$$

mit eindeutig bestimmten Koeffizienten $c_{ij}\in K$, $i=1,\dots,m, j=1,\dots,n$. 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

$$M(f) := M^{v_1,\dots,v_m}_{w_1,\dots,w_n}(f) := \begin{pmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nm} \end{pmatrix}$$

Beispiel 1

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

$$f(a) = \ \ r + 2s + 3t + 4u$$
$$f(b) = 5r + 6s + 7t + 8u$$
$$f(c) = 9r + 8s + 7t + 6u$$

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.

Beispiel 2

Seien K ein Körper,

$$A = (a_{ij}) \in K^{m\times n}$$

eine Matrix und

$$f = f_A: K^n \to K^m,\ x \mapsto Ax,$$

die Abbildung, welche jedem Vektor $x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ das Matrizenprodukt Ax zuordnet. Diese Abbildung ist K-linear. Wir wählen als Basen der Vektorräume K^n und K^m die <u>Standard-Einheits-Vektoren</u>

$$e_1,\dots,e_n \in K^n \text{ bzw. } e_1,\dots,e_m \in K^m$$

des K^n bzw. K^m[8], d.h. e_i bezeichne den Vektor des K^n bzw. K^m, dessen einzige von Null verschiedene Koordinaten die i-te Koordinate ist, welche gleich 1 sein soll. Dann gilt für i = 1,...,n:

$$f_A(e_i) = Ae_i = \text{i-te Spalte von } A = \sum_{j=1}^{m} a_{ji}e_j .$$

Also ist

$$M(f) = A$$

die Matrix von f bezüglich der gegegeben <u>Standard-Einheits-Basen</u>.

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

$$A = (a_{ij}) := M(f) = M_e^e(f)$$

die Matrix der Abbildung f bezüglich der gegebenen Basen.

Dann ist das folgende Diagramm kommutativ,

$$\begin{array}{ccc} K^m & \xrightarrow{\varphi_V} & V \\ f_A\downarrow & & \downarrow f \\ K^n & \xrightarrow{\varphi_W} & W \end{array}$$

d.h. es gilt

---

[8] Insbesondere kann e_1 einen Vektor des K^n bzw. einen des K^m bezeichnen. Aus dem Kontext wird stets hervorgehen, welcher Vektor gemeint ist.

$$f \circ \varphi_v = \varphi_w \circ f_A$$

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 aus dem Beispiel 2 von 3.4.1, d.h.

$$f_A(x) = Ax.$$

Beweis. Für i = 1, ... , m gilt

$$\begin{aligned} f \circ \varphi_v(e_i) &= f(v_i) \\ &= \sum_{j=1}^{n} a_{ji} w_j \\ &= \sum_{j=1}^{n} a_{ji} \varphi_w(e_j) \\ &= \varphi_w(\sum_{j=1}^{n} a_{ji} e_j) \\ &= \varphi_w(\text{ i- te Spalte vn A}) \\ &= \varphi_w(Ae_i) \\ &= \varphi_w(f_A(e_i)) \end{aligned}$$

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) Jede K-lineare Abbildung

$$K^n \to K^m$$

ist von der Gestalt f_A wie im Beispiel von 3.4.1 mit einer Matrix $A \in K^{m\times n}$. Das ergibt aus 3.4.2 für den Spezialfall

$$V := K^n, W := K^m, v_i := e_i \in K^m, w_j := e_j \in K^n.$$

Die horizontalen Abbildungden des kommutativen Vierecks sind dann die identischen Abildungen.

(ii) Im allgemeinen Fall sehen wir, daß wir bei fixierten Basen zu jeder Matrix eine lineare Abbildung und zu jeder linearen Abbildung eine Basis erhalten. Genauer wir erhalten eine bijektive Abildung

$$K^{m\times n} \to \mathrm{Hom}_K(V, W), A \mapsto \varphi_w \circ f_A \circ \varphi_v^{-1}$$

mit der Umkehrung

$$\mathrm{Hom}_K(V, W) \to K^{m\times n},\ f \mapsto M^e_e(\varphi_w^{-1}\circ f\circ\varphi_v).$$

Diese Abbildungen gestatten es uns, die beiden Mengen zu identifizieren (wobei dei Art der Identifzikation von der Wahl der Basen abhängt.

(iii) Die Aussage kann man auch wie folgt interpretieren: benutzt man die zu den gegebenen Basen gehörigen Koordinaten-Abbildungen φ_v und φ_w um V mit dem K^m und W mit dem K^n zu identifizieren,

$$K^m \overset{\varphi_v}{\cong} V$$
$$K^n \overset{\varphi_w}{\cong} W$$

so identifizieren sich die linearen Abbildungen $V \to W$ mit den Matrize(-Multiplikationen der Matrizen) von $K^{m\times n}$.

(iv) Die Abbildungen φ_v, φ_w, φ_v^{-1}, φ_w^{-1} sind K-lineare Abildungen. Deshalb sind die bijektiven Abbildungen von (ii) ebenfalls K-lineare Abbildungen, also Isomorphismen von K-Vektorräumen.

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

$$u_1,\ldots,u_\ell \in U$$
$$v_1,\ldots,v_m \in V$$
$$w_1,\ldots,w_n \in W$$

die Vektoren der mit u, v bzw. w bezeichneten Basen.

. Nach 3.4.2 haben wir kommutive Diagramme

$$\begin{array}{ccc} K^\ell & \xrightarrow{\varphi_u} & U \\ f_A\downarrow & & \downarrow f \\ K^m & \xrightarrow{\varphi_v} & V \end{array} \quad \text{mit } A := M^u_v(f)$$

und

$$\begin{array}{ccc} K^m & \xrightarrow{\varphi_v} & V \\ f_B\downarrow & & \downarrow g \\ K^n & \xrightarrow{\varphi_w} & W \end{array} \quad \text{mit } B := M^v_w(g)$$

und

$$\begin{array}{ccc} K^\ell & \xrightarrow{\varphi_u} & U \\ f_C\downarrow & & \downarrow g\circ f \\ K^n & \xrightarrow{\varphi_w} & W \end{array} \quad \text{mit } C := M^u_w(g\circ f)$$

Durch Zusammensetzen der ersten beiden Diagramme erhalten wir ein kommuatativen 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

$$f_A\circ f_B = \varphi_w^{-1}\circ(g\circ f)\circ\varphi_v = f_C,$$

für jedes $x\in K^{\ell}$ gilt

$$(BA)x = B(Ax) = f_A(f_B(x)) = f_C(x) = Cx.$$

Speziell für $x = e_i$ sehen wir, daß die i-te Spalte der Matrix BA mit der i-ten Spalte der Matrix C übereinstimmt. Da dies für jedes i gilt, folgt

$$BA = C.$$

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,

$$M_v^v(\mathrm{Id}) = \mathrm{Id}$$

für beliebige Basen v von V.

(ii) Sind $f: V\to W$ und $g:W\to V$ zueinander inverse lineare Abbildungen, d.h. gilt $f\circ g = \mathrm{Id}$ und $g\circ f = \mathrm{Id}$, so folgt aus der eben bewiesenen Formel und Bemerkung (i), daß $M(f)M(g) = \mathrm{Id}$ und $M(g)M(f) = \mathrm{Id}$ 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→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 ↦ v' und v' ↦ v gehören zu zueinander inversen Matrizen,

$$M^{v'}_v(Id) = M^v_{v'}(Id)^{-1}.$$

(vi) Seien f:V→V ein lineare Endomorphismus des (endlich-dimensionalen Vektorraums) und v,v' zwei Basen von V. Dann gilt

$$M^{v'}_{v'}(f) = M^{v'}_v(Id)^{-1}M^v_v(f)M^{v'}_v(Id)$$

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

$$f_A \circ f_B = f_{AB} = f_{Id} = Id.$$

Insbesondere ist die Abbildung f_A surjektiv, denn für jedes $x \in K^n$ gilt

$$f_A(f_B(x)) = x,$$

d.h. x liegt im Bild von f_A. Damit ist

$$\begin{aligned} n &= \dim K^n \\ &= \dim \ker(f_A) + \dim \operatorname{im}(f_A) \\ &= \dim \ker(f_A) + \dim K^n \\ &= \dim \ker(f_A) + n, \end{aligned}$$

also $\dim \ker(f_A) = 0$, also $\ker(f_A) = 0$. Wir haben gezeigt, f_A ist nicht nur surjektiv, sondern auch injektiv, also bijektiv. Mit anderen Worten, f_A ist eine umkehrbare Abbildung (und die Umkehrung ist linear). Insbesondere gilt

$$f_A^{-1} \circ f_A = Id.$$

Wir jehen von den Abbildungen zu den zugehörigen Matrizen (bezüglich der Standardbasis des K^n) über und erhalten nach 4.4.2,

$$Id = M(Id) = M(f_A^{-1} \circ f_A) = M(f_A^{-1})M(f_A) = M(f_A^{-1}) \cdot A.$$

Mit anderen Worten, die Matrix A besitzt außer der Rechtsinversen B auch eine linksinverse Matrix (nämlich $M(f_A^{-1})$). Also ist A umkehrbar.
Wir multiplizieren die Identität (1) mit dem Inversen von A und erhalten

$$A^{-1} = A^{-1}AB = B,$$

d.h. B ist gerade die zu A inverse Matrix (und also solche ebenfalls umkehrbar). Insbesondere gilt

$$BA = A^{-1}A = Id.$$

QED.

3.4.6 Fortsetzbarkeit von linearen Abbbildungen auf Unterräumen

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. Seien $\{v'_i\}_{i\in I'}$ eine Basis von V' und $\{\bar{v}_j\}_{j\in J}$ eine Basis des Faktorraumes V/V'. Da die natürliche Abbildung

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

surjektiv ist, gibt es für jedes $j \in J$ ein $v_j \in V$ mit $\rho(v_j) = \bar{v}_j$. Nach 3.3.11 bilden die beiden Familien

$$\{v'_i\}_{i\in I'} \text{ und } \{v_j\}_{j\in J}$$

zusammen eine Basis von V. Wir können deshalb die gesuchte lineare Abbildung f dadurch definieren, daß wir die Bilder der Basiselemente v'_i und v_j angeben, wobei wir diese Bilder beliebig wählen können. Wir setzen

$$f(v'_i) := f'(v'_i) \text{ für alle } i\in I'.$$
$$f(v_j) := 0 \text{ für alle } j\in J$$

Auf diese Weise ist eine lineare Abbildung f: V→W definiert. Sie stimmt nach Definition auf den Elementen v'_i der Basis $\{v'_i\}_{i\in I'}$ von V' mit f' überein. Deshalb muß

$$f(v) = f'(v) \text{ für alle } v\in V'$$

gelten, d.h. f stimmt mit f' auf V' überein.
QED.

3.4.7 Die duale Abbildung

3.4.7.1 Definition

Seien f:V→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→V und g:V→W gilt

$$(g\circ f)^* = f^*\circ g^*$$

(iii) Die zur identischen Abbildung Id:V→V duale Abbildung ist die identische Abbildung von V*,

$$\mathrm{Id}_V^* = \mathrm{Id}_{V^*}.$$

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 $\ell', \ell''\in W^*$, 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 ℓ: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 $g\circ f$ = Id, so gilt nach (ii) und (iii) auch $f^*\circ g^*$ = Id* = 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.

$$\begin{tikzcd}
 & K & \\
V \arrow[ur, "\ell"] \arrow[r, "g"', "\cong"] & f(V) \arrow[u, "\ell''"] \arrow[r, phantom, "\subseteq"] & W \arrow[ul, "\ell'"']
\end{tikzcd}$$

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 ℓ':W$\to$K der Linearform ℓ'':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.

3.4.7.3 Die duale Basis

Sei V ein endlich-dimensionaler K-Vektoraum und sei

(1) $$v_1,\ldots,v_n\in V$$

eine Basis von V. Wir erinnern daran (vgl. die Bemerkung von 3.3.18), die zu dieser Basis duale Basis besteht dann aus Vektoren

$$v_1^*,\ldots,v_n^*\in V^*$$

des dualen Raums V* mit

$$v_i^*(v_j) = \delta_{ij}$$

für alle i,j= 1,...,n. Diese bilden eine Basis von V*.

$$v_1^*,...,v_n^* \in V^*$$

Bemerkung
Unser nächstes Ziel besteht darin, zu zeigen, daß zwischen der Matrix einer Abbildung und der Matrix der dualen Abbildung ein Zusammenhang besteht.

3.4.7.4 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,

$$v_1,...,v_m \in V \text{ und } w_1,...,w_n \in W$$

Basen von V bzw. W und

$$v_1^*,...,v_m^* \in V \text{ und } w_1^*,...,w_n^* \in W$$

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

$$f(v_i) = \sum_{\alpha=1}^{n} a_{\alpha i} w_\alpha \text{ für } i = 1,...,m.$$

Nach Definition der dualen Abbildung $f^*:W^*\to V^*$ ist

$$\begin{aligned}
f^*(w_j^*)(v_i) &= (w_j^* \circ f)(v_i) \\
&= w_j^*(f(v_i)) \\
&= w_j^*(\sum_{\alpha=1}^{n} a_{\alpha i} w_\alpha) \\
&= \sum_{\alpha=1}^{n} a_{\alpha i}\, w_j^*(w_\alpha) \\
&= a_{ij} \\
&= \sum_{\alpha=1}^{n} a_{i\alpha} v_\alpha^*(v_i).
\end{aligned}$$

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[9]

$$\rho := \rho_V: V \to V^{**},\ v \mapsto (\ell \mapsto \ell(v)),$$

(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: $V \to V$ eine K-lineare Abbildung. Dann ist das folgende Diagramm kommutativ.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow\rho_V & & \rho_W\downarrow \\ V^{**} & \xrightarrow{f^{**}} & W^{**} \end{array},$$

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

$$g: K \to Kv,\ c \mapsto cv,$$

ein Isomorphismus. Die Abbildung

$$g^{-1}: Kv \to K,\ cv \mapsto c,$$

[9] $\rho(v)$ ist die Auswertungsabbildung ℓ a $\ell(v)$ an der Stelle v.

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 $im(\rho)$ ein echter Unterraum von V^{**}, also

$$\dim im(\rho) < \dim V^{**} = \dim V.$$

Da aber ρ, wie wir gesehen haben, injektiv ist, muß das Bild von ρ isomorph zum Definitionsbereich sein,

$$\rho: V \to im(\rho)$$

ist ein Isomorphismus, d.h. es gilt $\dim im(\rho) = \dim V$.

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

$$f^{**}(\rho_V(v))(\ell) = (\rho_V(v))(f^*(\ell)) = (\rho_V(v))(\ell\circ f) =^{10} (\ell\circ f)(v)$$

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

Sei $A \in K^{m\times n}$ eine Matrix mit Einträgen aus K. Die Zeilen von A sind somit Vektoren von $K^{1\times n}$ und erzeugen damit einen Unterraum von $K^{1\times n}$. Die Dimension dieses Unterraums heißt Zeilenrang von A und wird mit

$$rk' A$$

bezeichnet. Mit anderen Worten, rk' A ist die Anzahl linear unabhängiger Zeilen von A.

Analog sind die Spalten von A Vektoren von $K^{m\times 1}$ und erzeugen einen Unterraum von $K^{m\times 1}$. Die Dimension dieses Unterraums heißt Spaltenrang von A und wird mit

$$rk A$$

bezeichnet. Mit anderen Worten, rk A ist die Anzahl linear unabhängiger Spalten von A.

Der Rang einer linearen Abbildung $f: V\to W$ ist definiert als Dimension des Bildes und wird mit

$$rk f := \dim im(f).$$

bezeichnet.
Beispiel

---

[10] $\rho_V(v)$ ist die Auswertung an der Stelle v.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 9 \\ 6 & 7 & 13 \end{pmatrix}$$

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

$$\mathrm{rk}\ A = 2.$$

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,

$$\mathrm{rk}'\ A = 2.$$

Diese Übereinstimmung ist, wie wir später sehen werden, kein Zufall.

Problem

Gilt stets rk A = rk' A?

Die Lösung des Problems besteht in seiner Übersetzung in die Sprache der dualen Räume und dualen Abbildungen. Wir beginnen diese Übersetzung mit einigen Bemerkungen.

Bemerkungen

(i) Sind $a_1,...,a_n$ die Spalten von A,

$$A = (a_1...a_n),$$

so ist nach Definition rk A die Dimension des Raums

$$<a_1,...,a_n> = \{ \sum_{i=1}^m c_i a_i \mid c_1,...,c_n \in K\}$$

Mit $c := \begin{pmatrix} c_1 \\ ... \\ c_n \end{pmatrix}$ gilt $Ac = \sum_{i=1}^m c_i a_i$, d.h. wir können diesen Raum auch in der folgenden Gestalt schreiben.

$$<a_1,...,a_n> = \{Ac \mid c\in K^n\} = \mathrm{im}(f_A),$$

d.h. dies ist das Bild der durch A definierten linearen Abbbildung

$$f_A: K^n \to K^m.$$

Wir haben gezeigt,

$$\mathrm{rk}(A) = \dim \mathrm{im}(f_A) = \mathrm{rk}\ f_A.$$

(ii) Die Spalten von A sind gerade die Zeilen der transponierten Matrix A^T. Es gilt deshalb

$$\mathrm{rk}'\ A = \mathrm{rk}\ A^T = \dim \mathrm{im}(f_{A^T}).$$

(iii) Wie wir in 4.4.6.6 gesehen haben, kann man die Abbildung f_{A^T} (bezüglich geeigneter Basen der beteiligten Räume) mit der zu f_A dualen Abbildung identifizieren,

$$f_{A^T} = (f_A)^*.$$

Wir haben damit die folgenden Formeln zur Verfügung.

$$\mathrm{rk}\ A = \dim \mathrm{im}(f_A)$$

$$rk'A = \dim im(f_A)^*$$

Um das angekündigte Ergebnis über die Gleichheit von Zeilen- und Spaltenrang einer Matrix zu beweisen, genügt es also, wenn wir zeigen, für jede lineare Abbildung f endlich-dimensionaler Vektorräume gilt

$$\dim im(f) = \dim im(f^*).$$

Dies ist die Aussage des nächsten Abschnittes.

3.4.10 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

$$\dim im(f) = \dim im(f^*).$$

Insbesondere ist für jede Matrix A der Zeilenrang gleich dem Spaltenrang von A,

$$rk\ A = rk'\ A.$$

Beweis. Seien $U := ker(f)$ und

$$u: U \to V, u \mapsto u,$$

die natürliche Einbettung von U in V, d.h. die lineare Abbildung auf U, die jedes Element in sich abbilddet. Dann gilt

(1) $\dim im(f) = \dim V - \dim U.$

(2) $f\circ u = 0$ (Nullabbildung).

Wegen (2) gilt

$$u^*\circ f^* = (f\circ u)^* = 0^* = 0 \text{ (= Nullabbildung).}$$

Die Zusammensetzung von $f^*:W^*\to V^*$ und $u^*: V^*\to U^*$ ist also Null. Das Bild von f^* liegt somit ganz im Kern von u^*,

$$\dim im(f^*) \le \dim ker(u^*) = \dim V^* - \dim im(u^*).$$

Man beachte, nach Definition ist u injektiv, also u^* surjektiv. Anstelle von $im(u^*)$ kann man also auch U^* schreiben. Also gilt

$$\dim im(f^*) \le \dim V - \dim U = \dim V - \dim ker(f) = \dim im(f).$$

Wir haben gezeigt,

(1) $rk(f^*) \le rk(f).$

Dies gilt für alle linearen Abbildungen f, also auch für $f^*: W^* \to V^*$, d.h. es gilt damit auch

$$rk(f^{**}) \le rk(f^*) \le rk(f).$$

Zum Beweis der Behauptung genügt es somit, wenn wir zeigen,

$$rk(f^{**}) = rk(f).$$

Auf Grund des kommutativen Diagramms von 4.4.7 erhält man f^{**} aus f durch Zusammensetzung von f mit zwei Isomorphismen,

$$f^{**} = \rho_W\circ f\circ\rho_V^{-1}.$$

Also gilt

$$\begin{aligned}
rk(f^{**}) &= \dim f^{**}(V^{**}) \\
&= \dim \rho_W(f(\rho_V^{-1}(V^{**}))) \\
&= \dim f(\rho_V^{-1}(V^{**}) && \text{(weil } \rho_W \text{ ein Isomorphismus ist)} \\
&= \dim f(V) && \text{(weil } \rho_V^{-1} \text{ ein Isomorphismus ist)} \\
&= rk(f).
\end{aligned}$$

QED.

3.4.11 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) $\Rightarrow$ (ii). Nach Voraussetzung existiert A^{-1}. Wegen $A^{-1}A = Id$ gilt für die zugehörige Abbildung $f_A:K^n \to K^n$ die Relation

$$f_{A^{-1}} \circ f_A = f_{Id} = Id.$$

Damit ist

$$n = \dim K^n = \dim Id(K^n) = \dim f_{A^{-1}} \circ f_A(K^n) \le^{11} \dim f_A(K^n) \le^{12} \dim K^n = n.$$

In dieser Abschätzung gilt also überall das Gleichheitszeichen, d.h. es ist

$$rk\ A = \dim f_A(K^n) = n.$$

(ii) $\Rightarrow$ (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

$$\dim \ker(f) = \dim K^n - \dim f_A(K^n) = n - n = 0,$$

d.h. der Kern von f ist trivial, d.h. f ist injektiv. Wir haben gezeigt, f ist bijektiv, also ein Isomorphismus.
QED.

4. Determinanten

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)	:= $\{f \in Abb(M) \mid f \text{ bijektiv}\}$
	(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
Wir haben zu zeigen, die Komposition von zwei bijektiven Abbildungen ist wieder bijektiv. Seien also $f:M \to N$ und $g:M \to M$ bijektive Abbildungen. Wir haben zu zeigen, $f \circ g$ ist dann ebenfalls bijektiv, d.h. die Fasern

$$(f \circ g)^{-1}(m)$$

von $f \circ g$ bestehen für jedes $m \in M$ aus genau einem Element. Sei $m \in M$ ein fest vorgegebenes Element. Für ein beliebiges Element $x \in M$ gilt dann

---

[11] Wir benutzen die Ungleichung

$$\dim f(V) = \dim V - \dim Ker(f) \le \dim V$$

für beliebige lineare Abbildungen f.

[12] Wegen $f_A(K^n) \subseteq K^n$.

$$x \in (f \circ g)^{-1}(m) \Leftrightarrow m = (f \circ g)(x) = f(g(x)) \Leftrightarrow g(x) \in f^{-1}(m).$$

Nach Voraussetzung ist f bijektiv, d.h. die Menge $f^{-1}(m)$ besteht aus genau einem Element, sagen wir $m' \in M$. Die Bedingung

$$g(x) \in f^{-1}(m) = \{m'\}$$

ist daher gleichbedeutend mit

$$g(x) = m'$$

Damit gilt

$$x \in (f \circ g)^{-1}(m) \Leftrightarrow g(x) = m' \Leftrightarrow x \in g^{-1}(m')$$

Wir haben gezeigt,

$$(f \circ g)^{-1}(m) = g^{-1}(m').$$

Da g nach Voraussetzung bijektiv ist, ist die Menge rechts einelementig, d.h. auch die Menge

$$(f \circ g)^{-1}(m)$$

besteht aus genau einem Element. Da dies für beliebige $m \in M$ gilt, ist $f \circ g$ bijektiv.

2. Die Komposition von Abbildungen ist assozialtiv

Seien f, g, $h:M \to M$ drei bijektive Abbildungen. Wir haben zu zeigen,

(1) $$(f \circ g) \circ h = f \circ (g \circ h),$$

d.h. zu zeigen ist,

(2) $$((f \circ g) \circ h)(m) = (f \circ (g \circ h))(m) \text{ für beliebiges } m \in M$$

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

Bezeichne Id: $M \to M$ die identische Abbildung, d.h. es sei

$$Id(m) = m \text{ für jedes } m \in M.$$

Diese Abbildung ist bijektiv, denn das einzige $x \in M$ mit $Id(x) = m$ ist $x=m$, d.h.

$$(Id)^{-1}(m) = \{m\} \text{ für jedes } m \in M.$$

Zeigen wir, Id hat die Eigenschaften des Einselements, d.h. es gilt

$$f \circ Id = Id \circ f = f \text{ für jedes } f \in G.$$

Zu zeigen ist,

(3) $$(f \circ Id)(m) = (Id \circ f)(m) = f(m) \text{ für jedes } m \in M \text{ und jedes } f \in G.$$

Es gilt

LHS von (3) $= f(Id(m)) = f(m)$
MHS von (3) $= Id(f(m)) = f(m)$
RHS von (3) $= f(m)$

Es gilt also tatsächlich (3).

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. 3.2.3).

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 $e_1, \dots, e_r \in M$ gibt mit

$$f(e_1) = e_2,\ f(e_2) = e_3,\ \dots,\ f(e_{r-1}) = e_r,\ f(e_r) = e_1$$

und

$$f(e) = e \text{ für } e \in M - \{e_1, \dots, e_r\}.$$

Für einen solchen Zyklus verwendet man die Bezeichnung

$$f = (e_1, \dots, e_r) = (e_2, \dots, e_r, e_1) = \dots$$

Zyklen, deren zugehörige Mengen $\{e_1, \dots, e_r\}$ disjunkt sind, heißen elementfremd. Zyklen, deren zugehörige Menge $\{e_1, \dots, e_r\}$ aus zwei Elementen bestehen, heißen Transpositionen.

Bemerkungen

(i) Nicht jede Permutation ist ein Zyklus.

(ii) Es ist leicht einzusehen, daß jede Permutation das Produkt von elementfremden Zyklen ist.

(iii) Jeder Zyklus ist ein Produkt von Transpositionen,

$$(e_1, \dots, e_r) = (e_1,e_2)\circ(e_2,e_3)\circ\dots\circ(e_{r-2},e_{r-1})\circ(e_{r-1},e_r)$$

(iv) 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

$$\begin{pmatrix} 1 & \dots & i & \dots & i & \dots & n \\ f(1) & \dots & f(i) & \dots & f(j) & \dots & f(n) \end{pmatrix} \circ (i,j) = \begin{pmatrix} 1 & \dots & i & \dots & i & \dots & n \\ f(1) & \dots & f(j) & \dots & f(i) & \dots & f(n) \end{pmatrix}$$

Konjugation eines Zyklus mit einer Permutation f

$$f\circ(e_1, \dots , e_r)\circ f^{-1} = (f(e_1), \dots , f(e_r))$$

$$\begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 4\,3\,7\,1\,6\,8\,2\,5 \end{pmatrix}\circ(1,3,5)(2,4,6)\circ\begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 4\,3\,7\,1\,6\,8\,2\,5 \end{pmatrix}^{-1} = (4,7,6)(3,1,8)$$

4.1.2.3 Inversionen und Vorzeichen einer Permutation

Seien M die Menge der ersten n natürlichen Zahlen

$$M = \{1,...,n\}$$

und f: M → M eine Permutation von M,

$$f = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ f(1) & f(2) & f(3) & \dots & f(n) \end{pmatrix}.$$

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

$$\mathrm{sign}(f) = (-1)^{13} = -1.$$

4.1.2.4 Das Vorzeichen bei Multiplikation mit einer Transposition

$$\mathrm{sign}(f\circ(i,j)) = - \mathrm{sign}(f)$$

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.
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) $(i,j) =^{13} (i,i+1)\circ(i+1,i+2)\circ...\circ(j-2,j-1)\circ(j,j-1)\circ(j-1,j-2)\circ...\circ(i+2,i+1)\circ(i+1,i)$

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 ist 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
$$\text{sign}\,(1,2)\circ(2,5)\circ(3,4) = -\,\text{sign}\,(1,2)\circ(2,5) = +\,\text{sign}\,(1,2) = -1.$$

4.1.2.5 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,
$$f = (a_1,a_2)\circ(a_3,a_4)\circ...\circ(a_r,a_{r+1})$$
$$g = (b_1,b_2)\circ(b_3,b_4)\circ...\circ(b_s,b_{s+1})$$
Dann gilt

[13] Wendet man die rechte Seite auf i an, so wird i nacheinander durch 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).

$$f\circ g = (a_1,a_2)\circ(a_3,a_4)\circ\ldots\circ(a_r,a_{r+1})\circ(b_1,b_2)\circ(b_3,b_4)\circ\ldots\circ(b_s,b_{s+1}).$$

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,

$$H\times H \to H,\ (a,b) \mapsto a\cdot b,$$

und die Menge H ist zusammen mit dieser Abbildung eine Gruppe ist.
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 $H \to G$, $h \mapsto h$, die jedes Element auf sich selbst abbildet, ist ein Homomorphismus.

4.1.3.2 Beispiel: Abbildungen und lineare Abbildungen

Für jede $n\times n$-Matrix $A\in \mathbb{R}^{n\times n}$ und jeden n-zeiligen Spaltenvektor $x\in \mathbb{R}^n := \mathbb{R}^{n\times 1}$ setzen wir

$$f_A(x) \quad := Ax \text{ (Matrizenprodukt)}$$

Dann ist

$$f_A\colon \mathbb{R}^n\to\mathbb{R}^n,\ x \mapsto f_A(x) = Ax,$$

eine wohldefinierte lineare Abbildung. Wir setzen

M := $\mathbb{R}^n = \mathbb{R}^{n\times 1}$ die Menge der n-zeiligen Spaltenvektoren
G := Abb(M) die Gruppe der bijektiven Abbildungen $M\to M$
H := $\{f_A\colon M\to M \mid A\in GL(n, \mathbb{R})\}$

Dann ist H eine Untergruppe von G.

Beweis. 1. Schritt. Die Abbildung $GL(n, \mathbb{R}) \to H$, $A \mapsto f_A$, ist bijektiv.

Auf Grund der Definition der Menge H ist die Abbildung surjektiv. Wir haben noch zu zeigen, daß sie injektiv ist. Seien A, A' $\in$ GL(n, $\mathbb{R}$) zwei Matrizen mit

(1) $$f_A = f_{A'}\ .$$

Wir haben zu zeigen, daß beiden Matrizen gleich sind. Zum Beweis führen wir folgende Bezeichnung ein.

e_i := i-ter Einheits-Spalten-Vektor,

d.h. $e_i \in \mathbb{R}^{n\times 1}$ sei der Spaltenvektor mit n-Koordinaten, dessen einzige von Null verschiedene Koordinate - die i-te Koordinate - gleich Eins ist.

Auf Grund von Voraussetzung (1) gilt jedes i:

$$\text{i-te Spalte von } A = Ae_i = f_A(e_i) = f_{A'}(e_i) = A' e_i = \text{i-te Spalte von } A',$$

d.h. A und A' haben dieselbe i-te Spalte. Da dies für jedes i gilt, folgt A = A'. Damit ist die Aussage des ersten Schrittes bewiesen.

Folgerung aus dem ersten Schritt:

Wir können die Menge $GL(n, \mathbb{R})$ mit der Bildmenge H identifizieren, indem wir keinen Unterschied mehr machen zwischen einer Matrix $A\in GL(n,\mathbb{R})$ und der zugehörigen Abbildung f_A. Insbesondere wird auf diese Weise H zu einer Gruppe (indem wir die Gruppenoperation von $GL(n, \mathbb{R})$ auf H übertragen.

2. Schritt. Die Abbilung $i:GL(n, \mathbb{R}) \to G=Abb(M)$, $A \mapsto f_A$, ist ein Homomorphismus.

Seien $A, B\in GL(n, \mathbb{R})$. Wir haben zu zeigen,

(3) $$i(AB) = i(A)\circ i(B).$$

Auf beiden Seiten der zu beweisenden Identität stehen Abbildungen mit dem Definitionsbereich $M = \mathbb{R}^n$. Wir haben also zu zeigen,

(4) $$i(AB)(x) = (i(A)\circ i(B))(x) \text{ für jedes } x\in M.$$

Es gilt

$$i(AB)(x) = f_{AB}(x) = (AB)x = A(Bx) = f_A(f_B(x)) = (f_A\circ f_B)(x) = (i(A)\circ i(B))(x).$$

Zusammenfassung

Wir wissen aus dem ersten Schritt, die Teilmenge $H\subseteq G$ ist eine Gruppe, welche man mit der Gruppe $GL(n, \mathbb{R})$ identifizieren kann. Identifiziert man H mit der Gruppe $GL(n, \mathbb{R})$, so wird die natürliche Einbettung

(5) $$H \to G,\ f \mapsto f,$$

zur Abbildung

$$H \to G,\ A \mapsto f_A.$$

Von dieser Abbildung haben wir im zweiten Schritt gezeigt, daß es sich um einen Homomorphismus handelt. Damit ist aber auch die Einbettung (5) ein Homomorphismus. Mit anderen Worten, H ist eine Untergruppe von G.

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. (i) ⇒ (iii). trivial.

(iii) ⇒ (ii). 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>. Wir 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 alternierden 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.

$$E_{ij} := \begin{array}{cc} \begin{array}{ccccc} & & j & & \end{array} & \\ \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} & \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} \end{array}$$

Bezeichne E_{ij} die Matrix, deren Eintrag in der Position (i,j) gleich 1 ist, und deren sämtliche anderen Einträge Null sind.

$$M_i(c) := Id + (c-1)E_{ii} = \begin{array}{cc} \begin{array}{ccccc} & & i & & \end{array} & \\ \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & c & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} & \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} \end{array}$$

Die Multiplikationsmatrix $M_i(c)$ sei die Diagonalmatrix[14], deren Eintrag in der Position (i,i) gleich c ist, und deren übrige Hauptdiagonal-Einträge gleich Eins sind.

$$Q_{ij}(c) = Id + c\cdot E_{ij} = \begin{array}{cc} \begin{array}{ccccc} & & j & & \end{array} & \\ \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & c & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} & \begin{array}{c} \\ i \\ \\ \\ \\ \\ \\ \end{array} \end{array} \quad (i\neq j)$$

sei die Matrix, deren Hauptdiagonaleinträge sämtlich 1 sind, deren Eintrag in der Position (i,j) gleich c ist, und deren übrige Einträge Null sind.

$$P_{ij} = Id - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$

Die Permutationsmatrix 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 Elementarmatrizen.

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.

Multiplikation der i-ten Spalte mit c:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto A\cdot M_i(c).$$

Addition des c-fachen der i-ten Spalte zur j-ten Spalte:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto A\cdot Q_{ij}(c).$$

Vertauschen von i-ter und j-ter Spalte:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto A\cdot P_{ij}.$$

[14] d.h. die einzigen von Null verschiedenen Einträge befinden sich auf auf der Hauptdiagonalen.

Durch Multiplikation von links erhält man die entsprechenden Zeilenoperationen.

Multiplikation der i-ten Zeile mit c:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto M_i(c)\cdot A.$$

Addition des c-fachen der j-ten Zeile zur i-ten Zeile:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto Q_{ij}(c)\cdot A.$$

Vertauschen von i-ter und j-ter Zeile:

$$K^{n\times n} \to K^{n\times n},\ A \mapsto P_{ij}\cdot A.$$

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:

$$\begin{aligned} Q_{ij}(c) &= M_j(\tfrac{1}{c})Q_{ij}(1)M_j(c) \\ P_{ij} &= Q_{ji}(1)Q_{ij}(-1)Q_{ji}(1)M_j(-1). \end{aligned}$$

(ii) Die Elementarmatrizen sind umkehrbar und ihre Inversen sind wieder Elementarmatrizen. Genauer gilt:

$$\begin{aligned} M_i(c)^{-1} &= M_i(\tfrac{1}{c}) \\ Q_{ij}(c)^{-1} &= Q_{ij}(-c) \\ P_{ij}^{-1} &= P_{ij} \end{aligned}$$

(iii) Jede umkehrbare Matrix ist das Produkt von endlich vielen Elementarmatrizen.

Beweis. 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

$$A = (...,a_i,...,a_j,...)$$

mit den Spaltenvektoren $...,a_i,...,a_j,...$ die folgende Situation:

$$\begin{aligned} AQ_{ji}(1) &= (...,a_i+a_j,...,a_j,...) \\ AQ_{ji}(1)Q_{ij}(-1) &= (...,a_i+a_j,...,- a_i,...) \\ AQ_{ji}(1)Q_{ij}(-1)Q_{ji}(1) &= (...,a_j,...,- a_i,...) \end{aligned}$$

Beachtet man noch, wie die Multiplikation von rechts mit $)M_j(-1)$ wirkt, so ergibt sich die zweiter der zu beweisenden Identitäten.

Weiter erhält man

$$\begin{aligned} A\,M_j(\tfrac{1}{c}) &= (...,a_i,...,\tfrac{1}{c}\cdot a_j,...) \\ A\,M_j(\tfrac{1}{c})Q_{ij}(1) &= (...,a_i,...,\tfrac{1}{c}\cdot a_j + a_i,...) \\ A\,M_j(\tfrac{1}{c})Q_{ij}(1)\,M_j(c) &= (...,a_i,...,a_j + c\cdot a_i,...) \end{aligned}$$

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). Sei eine umkehrbare Matrix. Durch elementare 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 $B_1,\dots,B_s$ mit

$$A' = A\cdot B_1\cdot\ldots\cdot B_s.$$

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 $B_1\cdot\ldots\cdot B_s$ mit

$$\mathrm{Id} = A\cdot B_1\cdot\ldots\cdot B_s .$$

Wir mutiplizieren von rechts nacheinander mit den Inversen von B_s, B_{s-1}, ..., B_1 und erhalten

$$B_s^{-1}\cdot B_{s-1}^{-1}\cdot\ldots\cdot B_1^{-1} = A.$$

Da die Inversen von Elementarmatrizen wieder Elementarmatrizen sind, haben wir damit A als Produkt von Elementarmatrizen dargestellt.
QED.

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

$$\sum_{\sigma\in S_n} \mathrm{sign}(\sigma)\cdot a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdot\ldots\cdot a_{n\sigma(n)}$$

Determinante von A und wird mit

det(A) oder auch |A|

bezeichnet. 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

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \in K^{n\times n}$$

eine quadratische n-reihige Matrix mit Einträgen aus dem Körper K. Dann gilt

$$\det A = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma)\cdot a_{\sigma(1)1}\cdot a_{\sigma(2)2}\cdot\ldots\cdot a_{\sigma(n)n}.$$

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 $i_1,...,i_n$ 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

$$\det A = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma^{-1})\cdot a_{\sigma(1)1}\cdot a_{\sigma(2)2}\cdot\ldots\cdot a_{\sigma(n)n}.$$

Schließtlich gilt $\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma)$, d.h. es gilt 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[15]

ad und bc.

In S_2 gibt es eine gerade und eine ungerade Permutation (da es von jeder Sorte Permutationen gleichviele gibt), 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.

4.3.4 Die Determinante einer 3×3-Matrix (Sarrussche Regel)

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - gec - hfa - idb.$$

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.

Beweis. Weil S_3 sechs Elemente hat, müssen auf der rechten Seite der gesuchten Formel sechs Summanden stehen. Die angegebenen Summanden gehören zu den folgenden Permutationen von {1,2,3}.

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

Durch Bestimmung der Vorzeichen dieser Permutationen sieht man das die behauptete Vorzeichenverteilung tatsächlich auftritt. Es gilt

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1,2,3), \quad \text{d.h. } \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = +1$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1,3,2), \quad \text{d.h. } \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = +1$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (1,3), \quad \text{d.h. } \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = -1$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (2,3), \quad \text{d.h. } \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = -1$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (1,2), \quad \text{d.h. } \operatorname{sign}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = -1$$

QED.

Bemerkungen

[15] 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.

(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,

$$B = \begin{pmatrix} b_1 \\ ... \\ b_{i-1} \\ y \\ b_{i+1} \\ ... \\ b_n \end{pmatrix}$$

die Matrix, deren Zeilen gerade diese Vektoren sind und

$$g(y) := \det \begin{pmatrix} b_1 \\ ... \\ b_{i-1} \\ y \\ b_{i+1} \\ ... \\ b_n \end{pmatrix}$$

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

$$\begin{aligned} f(x) &= \det A \\ &= \sum_{\sigma\in S_n} \text{sign}(\sigma)\cdot a_{\sigma(1)1}\cdot a_{\sigma(2)2}\cdot...\cdot a_{\sigma(i-1)i-1}\cdot x_{\sigma(i)}\cdot a_{\sigma(i+1)i+1}\cdot...\cdot a_{\sigma(n)n} \end{aligned}$$

$$= \sum_{\sigma\in S_n} c_{\sigma,i}\cdot x_{\sigma(i)}$$

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 $b = a_i^T$ 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

$y \mapsto y^T$ und $x \mapsto f(x)$

und als solche linear.

QED.

4.4.2 Verhalten beim Vertauschen von Zeilen bzw. Spalten

(i) Sei A eine 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 Matrix mit den Zeilen $a_1,...,a_n\in K^{1\times n}$. Dann besteht für jede Permutation $\sigma\in S_n$ die Relation

$$\det\begin{pmatrix} a_{\sigma(1)} \\ ... \\ a_{\sigma(n)}\end{pmatrix} = \operatorname{sign}(\sigma)\cdot\begin{pmatrix} a_1 \\ ... \\ a_n\end{pmatrix}.$$

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, $\sigma =(u,v)$. 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

$$\det(A') = \sum_{\sigma\in S_n} \text{sign}(\sigma)\cdot a'_{\sigma(1)1}\cdot\ldots\cdot a'_{\sigma(u)u}\cdot a'_{\sigma(u+1)u+1}\cdot\ldots\cdot a'_{\sigma(n)n}$$

$$= \sum_{\sigma\in S_n} \text{sign}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(u)u+1}\cdot a_{\sigma(u+1)u}\cdot\ldots\cdot a_{\sigma(n)n}$$

Den Ausdruck unter der Summe kann man auch in der folgenden Gestalt schreiben.

$$\text{sign}(\sigma)\cdot a_{\tau(1)1}\cdot\ldots\cdot a_{\tau(u)u}\cdot a_{\tau(u+1)u+1}\cdot\ldots\cdot a_{\tau(n)n}$$

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. Wir könne die letzte Summe also auch schreiben als

$$\det(A') = \sum_{\tau\in S_n} \text{sign}(\tau\circ(u,u+1)^{-1})\cdot a_{\tau(1)1}\cdot\ldots\cdot a_{\tau(u)u}\cdot a_{\tau(u+1)u+1}\cdot\ldots\cdot a_{\tau(n)n}$$

$$= \text{sign}((u,u+1)^{-1})\cdot \sum_{\tau\in S_n} \text{sign}(\tau)\cdot a_{\tau(1)1}\cdot\ldots\cdot a_{\tau(u)u}\cdot a_{\tau(u+1)u+1}\cdot\ldots\cdot a_{\tau(n)n}$$

$$= \text{sign}((u,u+1)^{-1})\cdot\det(A).$$

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). Man kann sie aber auch durch explizite Rechnung beweisen. Nach (i) gilt

$$\det\begin{pmatrix} a_{\sigma(1)} \\ \ldots \\ a_{\sigma(n)} \end{pmatrix} = \det(a^T_{\sigma(1)},\ldots,a^T_{\sigma(n)})$$

$$= \text{sign}(\sigma)\cdot\det(a^T_1,\ldots,a^T_n)$$

$$= \text{sign}(\sigma)\cdot\begin{pmatrix} a_1 \\ \ldots \\ a_n \end{pmatrix}$$

QED.

Bemerkung: Körper der Charakteristik 2

Stimmen in einer Matrix zwei Zeilen oder zwei Spalten überein, so ergibt sich sofort aus dem eben Bewiesenen, daß deren Determinante Null sein mußt,

$$\det(A) = 0$$

falls A zwei gleiche Zeilen oder zwei gleiche Spalten hat. Vertauscht man nämlich diese Zeilen bzw. Spalten, so bleibt det(A) unverändert und wechselt andererseits das Vorzeichen.

Das eben gegebene Argument versagt für Körper, in denen

$$2 = 1 + 1$$

gleich Null ist (zum Beispiel für den Körper aus 2 Elementen). Das Argument zeigt nämlich nur,

$$\det(A) = -\det(A),$$

d.h.

$$2\cdot\det(A) = 0.$$

Nur im Fall, daß der Koeffizient 2 in K ungleich Null ist, existiert das Inverse von 2 in K und wir können die obige Gleichung mit diesem Inversen multiplizieren und so die Behauptung

$$\det(A) = 0$$

erhalten.

Man sagt von Körpern K, in denen 2 = 1 + 1 Null ist, es seien <u>Körper der Charakteristik 2</u> und man schreibt in dieser Situation

$$\mathrm{char}(K) = 2.$$

Unsere Aussage ist auch in diesem Fall richtig und folgt aus der Betrachtung der Determinante als ganzzahliges Polynom

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) $$\det(\dots,a_i,\dots,a_j,\dots) = \det(\dots,a_i+c\cdot a_j,\dots,a_j,\dots)$$

Die anaolge Aussage gilt auch für die Zeilen,

(2) $$\det(\dots,a_i,\dots,a_j,\dots)^T = \det(\dots,a_i+c\cdot a_j,\dots,a_j,\dots)^T$$

Beweis. Da die Determinante eine lineare Funktion der i-ten Spalte ist, gilt

$$\det(\dots,a_i+c\cdot a_j,\dots,a_j,\dots) = \det(\dots,a_i,\dots,a_j,\dots) + c\cdot\det(\dots,a_j,\dots,a_j,\dots).$$

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,

$$\det\begin{pmatrix} c_{11} & \dots & c_{1n} \\ \dots & \dots & \dots \\ 0 & .. & c_{nn} \end{pmatrix} = c_{11}\cdot\dots\cdot c_{nn}$$

(ii) Allgemeiner, zerfällt die Matrix A wir folgt in Blöcke,

$$A = \begin{pmatrix} U & V \\ 0 & W \end{pmatrix}$$

mit quadratischen Matrizen U und V, so gilt

$$\det A = \det U \cdot \det V.$$

Beweis. Zu (i). In der Leibnizschen Determinantenformel ist nur ein Summand ungleich Null, nämlich der zur identischen Permutation[16]. Dieser Summand liefert den Beitrag

$$(+1)\cdot c_1 \cdot \ldots \cdot c_n.$$

Zu (ii). Durch elementare Zeilenumformungen der Matrix A vom Typ $Q_{ij}(c)$ und P_{ij} überführen wir U in obere Dreieckgestalt. Die Zahl der Vertauschungen, die wir dabei ausführen sei u. Anschließend überführen wir durch dieselbe Art von Zeilenümformungen die Matrix W in obere Dreiecksgestalt. Die Zahl der Vertauschungen sei dabei v. Wir erhalten eine Matrix

$$A' = \begin{pmatrix} U' & V' \\ 0 & W' \end{pmatrix}$$

mit

$$\begin{aligned} \det U' &= (-1)^u \cdot \det(U) \\ \det V' &= (-1)^v \cdot \det(V) \\ \det A' &= (-1)^{u+v} \cdot \det(A). \end{aligned}$$

Dabei sind U', V' (und damit auch A') obere Dreiecksmatrizen. Auf Grund von (i) gilt also

$$\det(A') = \det(U')\cdot\det(V').$$

Durch Einsetzen erhalten wir 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) $\det A \neq 0$.

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

$$a_1,\ldots,a_n.$$

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 Wort, 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 haben, 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. $\det B = 0 \Leftrightarrow \det A = 0$
3. B hat obere Dreieckgestalt und in jeder Zeile höchstens einen von Null verschiedenen Eintrag.

$$\begin{pmatrix} * & & \\ & \ldots & \\ 0 & & * \end{pmatrix}$$

Es genügt also, die Behauptung für obere Dreieckmatrizen zu beweisen. O.B.d.A. sei also A eine obere Dreiecksmatrix wie in 3. beschrieben. Dann gilt

[16] In jedem anderen Summanden gibt es einen Faktor $a_{i\sigma(i)}$ mit $i\neq\sigma(i)$ und damit auch einen solchen mit $i>\sigma(i)$. Letzterer ist aber Null, d.h. der gesamte Summand ist Null.

$\text{rk } A = n \quad \Leftrightarrow$ die Hauptdiagonalelemente von A sind sämtlich $\neq 0$
$\Leftrightarrow \det A \neq 0$.

Die einzige nicht-triviale Implikation ist dabei die obere Implikation '$\Rightarrow$'. Angenommen ein Hauptdiagonaleintrag a_{ii} ist gleich Null. Wir fixieren das kleinste i mit

$$a_{ii} = 0.$$

Die einzigen Einträge $\neq 0$ der i-ten Spalte befinden sich dann über der Hauptdiagonalen. Dann ist die die i-te Spalte aber Linearkombination früherer Spalten im Widerspruch zur Voraussetzung, daß die Spalten von A lineare unabhängig sein sollen. Man beachte, die Matrix A muß damit eine Diagonalmatrix sein.
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 $\det B = 0$.
Es reicht zu zeigen, $\det(AB) = 0$, oder, was dasselbe ist, $\text{rk}(AB) < n$.
Wegen $\det(B) = 0$ gilt $\text{rk}(B) < n$, also, wenn f_A bzw. f_B die zugehörigen linearen Abbildungen bezeichnen,

$$\text{rk}(AB) = \dim f_{AB}(K^n) = \dim f_A(f_B(K^n)) \leq \dim f_B(K^n) = \text{rk } B < n.$$

2. Schritt. Der Fall, $\det B \neq 0$.
Wir beginnen mit dem Fall, daß B eine Elementarmatrix ist.
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

$$\det A\cdot P_{ij} = - \det A$$

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

$$f: K^n\times\ldots\times K^n \to K, (v_1,\ldots,v_n) \mapsto f(v_1,\ldots,v_n),$$

in n Variablen gegeben, die folgende Eigenschaften hat.

(i) f ist linear in der der i-ten Variablen v_i für i = 1,...,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.

(iii) Für Körper einer Charakteristik $\neq 0$ 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 Matrix 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:

$$A = (a_{11}\cdot e_1,\ldots,a_{nn}\cdot e_n).$$

Auf Grund der Linearität von f bezüglich jeder Spalte folgt

$$f(A) = a_{11}\cdot\ldots\cdot a_{nn} f(e_1,\ldots, e_n) = a_{11}\cdot\ldots\cdot a_{nn}\cdot 1 = \det(A).$$

QED.

4.4.8 Minoren

Für $A\in K^{n\times n}$ bezeichne

$$A'_{ij} = \begin{array}{cc} \begin{array}{c} j \end{array} & \\ \begin{pmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{pmatrix} & \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} \end{array}$$

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,j)-komplementäre Matrix. Ihre Determinante

$$\det A'_{ij}$$

heißt (n-1)-reihiger Minor von A bezüglich der Position (i,j) oder auch (i,j)-Minor. Weiter bezeichne

$$A_{ij}$$

die Matrix, die man aus A durch Streichen der i-ten Zeile und j-ten Spalte erhält.

Bemerkungen

(i) Die Bedeutung der (n-1)-reihigen Minoren besteht darin, daß man die Berechnung von det A auf die Berechnung der Minoren zurückführen kann.

(ii) Die Determinante von A'_{ij} ändert sich nicht, wenn man Vielfache der j-ten Spalte zu anderen Spalten addiert. Insbesondere gilt damit

$$\det A'_{ij} = \det(a_1,\ldots,a_{j-1},e_j,a_{j+1},\ldots,a_n)$$

wenn a_v die v-te Spalte von A bezeichnet,

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

4.4.9 Die Berechnung eines Minors

Seien $A\in K^{n\times n}$ eine quadratische Matrix, A'_{ij} die (i,j)-komplementäre Matrix und

$$A_{ij}$$

die Matrix, welche aus A durch Streichen von i-ter Zeile und j-ter Spalte entsteht. Dann gilt

$$\det A'_{ij} = (-1)^{i+j} \det A_{ij} \tag{1}$$

Bemerkung zur Bestimmung des Vorzeichens (Schachmuster-Regel)

Das Vorzeichen $(-1)^{i+j}$ des Minors zur Position (i,j) kann man dadurch bestimmen, indem man die beiden mögliche Vorzeichen '+' und '-' auf die Positionen der Matrix alternierend verteilt, wie etwa auf einem Schachbrett die schwarzen und weißen Felder verteilt sind. 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

$$\det A'_{ij} = \det(a_1,...,a_{j-1},e_i,a_{j+1},...,a_j).$$

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

$$\det A'_{ij} = (-1)^{j-1}\det(e_i,a_1,...,a_{j-1},a_{j+1},...,a_j),$$

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 j-1, d.h. es gilt

$$\det A'_{ij} = (-1)^{j-1}\,(-1)^{j-1}\det(e_1,a'_1,...,a'_{j-1},a'_{j+1},...,a'_j),$$

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

$$\det A'_{ij} = (-1)^{i+j-2}\det\begin{pmatrix}1 & 0\\ 0 & A_{ij}\end{pmatrix},$$

wobei wie oben angegeben A_{ij} aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Die Matrix

$$\begin{pmatrix}1 & 0\\ 0 & A_{ij}\end{pmatrix}$$

hat Blockgestalt wie in 4.4.4 (ii), d.h. ihre Determinante ist gleich

$$\det\begin{pmatrix}1 & 0\\ 0 & A_{ij}\end{pmatrix} = 1\cdot\det A_{ij} = \det A_{ij}$$

Damit gilt

$$\det A'_{ij} = (-1)^{i+j-2}\det A_{ij},$$

d.h. es gilt die Behauptung.
QED.

4.4.10 Entwicklungssatz von Laplace

Für $A\in K^{n\times n}$ gilt

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

(ii) $\det A = \sum_{j=1}^{n}(-1)^{i+j}\cdot a_{ij}\cdot\det A_{ij}$ für i=1,...,n.

Dabei bezeichne A_{ij} die Matrix, welche durch Streichen der i-ten Zeilen und j-Spalte aus A entsteht.

Beweis. Bezeichne a_ν die ν-te Spalte von A und $a_{\nu\mu}$ den Eintrag in der Position (ν,μ). Dann gilt

$$\det A = \det(a_1,...,a_{j-1},a_j,a_{j+1},...,a_j)$$

$$= \det(a_1,\dots,a_{j-1}, \sum_{i=1}^{n} a_{ij}e_i, a_{j+1},\dots,a_j)$$

$$= \sum_{i=1}^{n} a_{ij}\cdot\det(a_1,\dots,a_{j-1},e_i, a_{j+1},\dots,a_j)$$

$$= \sum_{i=1}^{n} a_{ij}\cdot\det A'_{ij}$$

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. 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-fach bzw. 3-fache der zweite 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.

$$d = \det\begin{pmatrix} 1 & 0 & 0 \\ 4 & 0 & -1 \\ 3 & -1 & -1 \end{pmatrix}$$

$$= \det\begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix} \qquad \text{(Entwicklung nach der ersten Zeile)}$$

$$= 0\cdot(-1) - (-1)(-1) \qquad \text{(Determinantenformel für } 2\times 2\text{-Matrizen)}$$

$$= -1.$$

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,

$$A_i = (a_1,\ldots,a_{i-1},b,a_{i+1},\ldots,a_n)$$

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

$$\sum_{v=1}^{n} a_{uv}x_v = b_u \text{ für } u=1,\ldots,n. \qquad | \cdot \det A'_{uj}\,,\ \sum_{u=1}^{n}$$

Wir multiplizieren die u-te Gleichung dieses Systems mit $\det A'_{uj}$ von A und bilden die Summe der entstehenden Gleichungen und erhalten nach dem Laplaceschen Entwicklungssatz

$$\sum_{u=1}^{n} \det A'_{uj}\cdot \sum_{v=1}^{n} a_{uv}x_v = \sum_{u=1}^{n} b_u\cdot\det A'_{uj} = \det A_j.$$

Also ist

(1) $$\det A_j = \sum_{v=1}^{n} \Big(\sum_{u=1}^{n} a_{uv}\cdot \det A'_{uj}\Big)\cdot x_v.$$

Versuchen wir die innere Summe des letzten Ausdrucks zu verstehen. Im Fall v=j ist

$$\sum_{u=1}^{n} a_{uv}\cdot \det A'_{uj} = \det A \text{ im Fall } v=j.$$

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,

$$\sum_{u=1}^{n} a_{uv}\cdot \det A'_{uj} = \det B \text{ im Fall } v\neq j$$

Als Koeffizienten a_{uv} treten in diesem Fall aber nicht die Elemente der j-Spalte von A auf, sondern die Elemente von deren v-ter Spalte auf. Die Matrix B kann man sich also aus A entstanden denken durch Ersetzung der Einträge der j-ten Spalte durch die der v-ten Spalte. 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,

$$\sum_{u=1}^{n} a_{uv}\cdot \det A'_{uj} = \det A\cdot\delta_{jv}.$$

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

$$\det A^{-1} = \frac{1}{\det A}\cdot(\det A'_{ij})^T_{i,j=1,...,n}$$

Beweis. Beim Beweis von 4.4.12 haben wir gezeigt

$$\sum_{u=1}^{n} a_{uv}\cdot \det A'_{uw} = \begin{cases}\det A & \text{falls } v=w\\ 0 & \text{sonst}\end{cases}$$

oder, anders ausgedrückt,

$$\sum_{u=1}^{n} a_{uv}\cdot \det A'_{uw} \qquad = \det A\cdot\delta_{uw}.$$

Mit $a'_{ij} := \det A'_{ji}$ kann man diese Identitäten auch in der Gestalt

$$\sum_{u=1}^{n} a'_{wu}\cdot a_{uv} \quad = \det A\cdot\delta_{uw}.$$

schreiben. Bezeichnet $A' = (a'_{ij})$ die Matrix mit den Einträgen a'_{ij}, so bedeutet letzteres

$$A'\cdot A = \det A\cdot \mathrm{Id},$$

d.h.

$$A^{-1} = \frac{1}{\det A}\cdot A'.$$

Dies ist aber gerade 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 Determinante 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 Untermatrizen

Sei $A \in K^{m\times n}$ eine Matrix mit Einträgen aus K. Eine Teilmatrix von A ist eine Matrix B welche durch streichen gewisser Zeilen und/oder Spalten von A entsteht. Wir wollen dabei zulassen, daß die Anzahl der gestrichenen Zeilen bzw. Spalten auch 0 sein kann. Insbesondere wollen wir A selbst zu den Teilmatrizen von A zählen.

Beispiel

Zur Matrix $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ gibt es $3\cdot 3 = 9$ zweireihige quadratische Untermatrizen

$$A_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \dots , A_{33} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$$

4.5.2 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

$$\text{rk } B < r.$$

Da B eine r-reihige quadratische Matrix ist, folgt

$$\det B = 0.$$

<u>(ii) ⇒ (i)</u>. Nach Voraussetzung hat jede r-reihige Untermatrix von A die Determinante Null. Angenommen, es wäre

$$\text{rk } A \geq r.$$

Dann gibt es in A mindestens 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

$$\text{rk } B = r,$$

also det $B \neq 0$. Das steht aber im Widerspruch zu unserer Annahme. Also muß rk $A < r$ gelten.

QED.

4.5.3 Das Rangkriterium II (Folgerung)

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 $\neq$ 0.

Beweis. Folgt direkt aus 4.5.2
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.

4.5.4 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) 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) $\Rightarrow$ (ii). Gilt auf Grund von 4.5.3.
(ii) $\Rightarrow$ (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 in der linken oberen Teil der Matrix A, d.h.

$$B = (a_{ij})_{1\le i\le r, 1\le j\le r}.$$

Weiter seien $a_1,...,a_n$ die Spalten von A,

$$A = (a_1,...,a_n).$$

Dann ist B auch eine Teilmatrix von

$$A' := (a_1,...,a_r) = \begin{pmatrix} B \\ * \\ ... \\ * \end{pmatrix}$$

Nach 4.5.3 ist der Rang der Matrix A' gleich r,

(1) $$\operatorname{rk} A' = r,$$

d.h. die Spalten von A' sind linear unabhängig.

Betrachten wir jetzt die Matrix

$$A'' = (a_1,...,a_r, a_j) = \begin{pmatrix} B & * \\ * & * \\ ... & ... \\ * & * \end{pmatrix} \text{mit } r < j \le n.$$

Die ersten r Spalten dieser Matrix sind linear unabhängig, d.h.

$$\text{rk } A'' \ge r.$$

Zum Beweis der Behauptung reicht es (für beliebiges j) zu zeigen,

(*) $$\text{rk } A'' = r,$$

denn dann ist jede Spalte a_j von A eine Linearkombination der ersten r Spalten.

Untersuchen wir also den Rang der Matrix A”. Es gilt

(2) Die ersten r Zeilen von A” sind linear unabhängig,

denn der Spaltenrang der Matrix aus den ersten r Zeilen ist $\ge$ rk B = r. Zum Beweis von (*) reicht es also zu zeigen, die letzten n-r Zeilen von A” sind Linearkombinationen der ersten r Zeilen.

Streichen wir in A” mit einer Ausnahme die letzten n-r Zeilen und bezeichnen die entstehende $(r+1)\times(r+1)$-Matrix mit

$$A^{(3)} = \begin{pmatrix} B & * \\ * & * \end{pmatrix}.$$

Wir haben zu zeigen, die letzte Zeile dieser Matrix ist eine Linearkombination der übrigen r Zeilen.

Nach Konstruktion enthält $A^{(3)}$ die Matrix B als Teilmatrix. Nach Voraussetzung (ii) gilt also

$$\det A^{(3)} = 0,$$

d,h. rk $A^{(3)} < r+1$. Die ersten r Zeilen sind linear unabhängig, die Gesamtheit aller r+1 Zeilen ist es nicht. Die letzte Zeile ist somit eine Linearkombination der übrigen r Zeilen.
QED.
Beispiel
Sei $A = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ -1 & 4 & -1 & 1 \end{pmatrix}$. Im Beispiel von 4.5.3 haben wir gesehen

$$\det A^{1,2,3}_{1,2,3} = \det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ -1 & 4 & -1 \end{pmatrix} = 0$$

und

$$\det A^{2,3}_{2,3} = \det \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} \ne 0.$$

Weiter 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(weggelassen)

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

$$1 \le u_1 < \ldots < u_k \le m$$
$$1 \le v_1 < \ldots < v_\ell \le n$$
zwei endliche Folgen von paarweise verschiedenen ganzen Zahlen. Dann setzen wir

$$A^{u_1 \ldots u_k}_{v_1 \ldots v_\ell} := (a_{u_i v_j}) := \begin{pmatrix} a_{u_1 v_1} & \ldots & a_{u_1 v_\ell} \\ \ldots & \ldots & \ldots \\ a_{u_k v_1} & \ldots & a_{u_k v_\ell} \end{pmatrix}$$

Seien jetzt weiter ganze Zahlen
$$1 \le u_{k+1}, \ldots , u_{m\ k} \le m$$
gegeben mit
$$\{u_1, \ldots , u_k\} \cup \{u_{k+1}, \ldots , u_m\} = \{1,\ldots,m\}.$$
Wir sagen dann die beiden Folgen
$$\{u_\mu\}_{\mu=1,\ldots,k} \text{ und } \{u_\mu\}_{\mu=k+1,\ldots,m}$$
sind zwei Folgen komplementärer Zeilen-Indizes. Analog sei eine Folge

$$1 \le v_{\ell+1},\ldots,v_n \le n$$
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

$$(-1)^{1+2}\cdot\det\begin{pmatrix} a & -b & -c & -d\\ b & a & -d & c\\ c & d & a & -b\\ d & -c & b & a\end{pmatrix} = (-1)^{1+2}\cdot\det\begin{pmatrix}a & -b\\ b & a\end{pmatrix}\det\begin{pmatrix}a & -b\\ b & a\end{pmatrix}$$

$$+ (-1)^{1+3}\cdot \det\begin{pmatrix}a & -c\\ b & -d\end{pmatrix}\det\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$$

$$+ (-1)^{1+4}\cdot \det\begin{pmatrix}a & -d\\ b & c\end{pmatrix}\det\begin{pmatrix}d & a\\ -c & b\end{pmatrix}$$

$$+ (-1)^{2+3}\cdot \det\begin{pmatrix}-b & -c\\ a & -d\end{pmatrix}\det\begin{pmatrix}c & -b\\ d & a\end{pmatrix}$$

$$+ (-1)^{2+4}\cdot \det\begin{pmatrix}-b & -d\\ a & c\end{pmatrix}\det\begin{pmatrix}c & a\\ d & b\end{pmatrix}$$

$$+ (-1)^{3+4}\cdot \det\begin{pmatrix}-c & -d\\ -d & c\end{pmatrix}\det\begin{pmatrix}c & d\\ d & -c\end{pmatrix}$$

$$= -(a^2+b^2)^2 - (-ad+bc)^2 - (ac+bd)^2 - (bd+ac)^2 - (ad-bc)^2 - (c^2+d^2)^2$$

$$= - (a^2+b^2)^2 - (c^2+d^2)^2 - 2(ad-bc)^2 - 2(ac+bd)^2$$

$$= - a^4 - b^4 - c^4 - d^4 - 2a^2b^2 - 2c^2d^2$$

$$- 2a^2d^2 - 2b^2c^2 + 4abcd$$

$$- 2a^2c^2 - 2b^2d^2 - 4abcd$$

$$= - (a^2+b^2+c^2+d^2)^2$$

Also ist

$$\det\begin{pmatrix} a & -b & -c & -d\\ b & a & -d & c\\ c & d & a & -b\\ d & -c & b & a\end{pmatrix} = (a^2+b^2+c^2+d^2)^2$$

<u>Berechnung ohne Verwendung des verallgemeinerten Entwicklungssatzes:</u>

Mit $A := \begin{pmatrix}a & -b\\ b & a\end{pmatrix}$ und $B := \begin{pmatrix}c & d\\ d & -c\end{pmatrix}$ kann man auch schreiben

$$\det\begin{pmatrix} a & -b & -c & -d\\ b & a & -d & c\\ c & d & a & -b\\ d & -c & b & a\end{pmatrix} = \det\begin{pmatrix}A & -B\\ B & A\end{pmatrix} = \det\begin{pmatrix}A+iB & i(A+iB)\\ B & A\end{pmatrix} = \det\begin{pmatrix}A+iB & 0\\ B & A-iB\end{pmatrix}$$

$$= \det(A+iB)\det(A-iB) = |\det(A+iB)|^2$$

$$= \left|\det\begin{pmatrix}a+ic & -(b-id)\\ b+id & a-ic\end{pmatrix}\right|^2$$

$$= \left|\det\begin{pmatrix}z & -\overline{w}\\ w & \overline{z}\end{pmatrix}\right|^2$$

$$= (|z|^2 + |w|^2)^2$$

$$= (a^2+b^2+c^2+d^2)^2$$

mit $z := a+ic$ und $w := b+id$.

Beispiel

Durch Entwicklung nach den ersten zwei Zeilen ergibt sich

$$(-1)^{1+2}\cdot\det\begin{pmatrix}2&1&1&1\\1&2&1&2\\1&1&2&1\\1&1&1&2\end{pmatrix}$$

$$\begin{aligned}
&= (-1)^{1+2}\cdot\det\begin{pmatrix}2&1\\1&2\end{pmatrix}\cdot\det\begin{pmatrix}2&1\\1&2\end{pmatrix}\\
&\quad + (-1)^{1+3}\det\begin{pmatrix}2&1\\1&1\end{pmatrix}\cdot\det\begin{pmatrix}1&1\\1&2\end{pmatrix}\\
&\quad + (-1)^{1+4}\det\begin{pmatrix}2&1\\1&2\end{pmatrix}\cdot\det\begin{pmatrix}1&2\\1&1\end{pmatrix}\\
&\quad + (-1)^{2+3}\det\begin{pmatrix}1&1\\2&1\end{pmatrix}\cdot\det\begin{pmatrix}1&1\\1&2\end{pmatrix}\\
&\quad + (-1)^{2+4}\det\begin{pmatrix}1&1\\2&2\end{pmatrix}\cdot\det\begin{pmatrix}1&2\\1&1\end{pmatrix}\\
&\quad + (-1)^{3+4}\det\begin{pmatrix}1&1\\1&2\end{pmatrix}\cdot\det\begin{pmatrix}1&1\\1&1\end{pmatrix}\\
&= -\ 3\cdot 3 + 1\cdot 1 - 3\cdot(-1) - (-1)\cdot 1 + 0\cdot(-1) - 1\cdot 0\\
&= -9 + 1 + 3 + 1 = -4
\end{aligned}$$

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

Durch Entwicklung nach den ersten drei Zeilen erhält man

$$(-1)^{1+2+3}\det\begin{pmatrix}0&3&0&1&0&-2\\2&0&1&0&4&0\\0&-3&0&1&0&2\\-4&0&4&0&3&0\\0&-2&0&3&0&4\\3&0&1&0&1&0\end{pmatrix}$$

$$\begin{aligned}
&= (-1)^{1+2+3}\cdot 0 + (-1)^{1+2+4}\cdot x + (-1)^{1+2+5}\cdot 0 + (-1)^{1+2+6}\cdot 0\\
&+ (-1)^{1+3+4}\cdot 0 + (-1)^{1+3+5}\cdot 0 + (-1)^{1+3+6}\cdot 0\\
&+ (-1)^{1+4+5}\cdot 0 + (-1)^{1+4+6}\cdot y\\
&+ (-1)^{1+5+6}\cdot 0\\
&+ \ldots
\end{aligned}$$

mit $x = \det\begin{pmatrix}0&3&1\\2&0&0\\0&-3&1\end{pmatrix}\cdot\det\begin{pmatrix}4&3&0\\0&0&4\\1&1&0\end{pmatrix}$ und $y = \det\begin{pmatrix}0&1&-2\\2&0&0\\0&1&2\end{pmatrix}\cdot\det\begin{pmatrix}0&4&3\\-2&0&0\\0&1&1\end{pmatrix}$

4.6.3 Ein Lemma über das Vorzeichen von Permutationen

Seien

$$u_1 < \ldots < u_k \text{ und } u_{k+1} < \ldots < u_m$$

zwei Folgen komplementärer Zeilen-Indizes und sei σ die Permutation

$$\sigma = \begin{pmatrix}1 & \ldots & k & k+1 & \ldots & m\\ u_1 & \ldots & u_k & u_{k+1} & \ldots & u_m\end{pmatrix}$$

Dann gilt

$$\mathrm{sign}(\sigma) = (-1)^{\sum_{\mu=1}^{k} (u_\mu - \mu)} = (-1)^{\sum_{\mu=1}^{k} u_\mu - k(k+1)/2}$$

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) $u_1 < ... < u_k$ und $u_{k+1} < ... < u_m$.

Dann ist u_k der größte Index der ersten Folge. Alle größeren Indizes gehören zur zweiten Folge und stimmen mit ihren Platznummern überein, 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) $u_1 < ... < u_{k-1}$ und $u_{k+1} < ... u_k ... < u_m$.

Der Index u_k steht dann in der zweiten Folge an der u_k-ten Stelle[17] und diese zweite Folge hat genauer die Gestalt[18]

$$u_{k+1} < ... u_k, u_k+1, u_k+2, ... < u_k+x = m$$

Bezeichne σ' die zu (2) gehörige Permutation. Nach Induktionsvoraussetzung gilt dann

$$\mathrm{sign}(\sigma') = (-1)^{\sum_{\mu=1}^{k-1} (u_\mu - \mu)}$$

Die Permutation σ' entsteht aus σ indem man den Index u_k von der k-ten Position in die Position u_k bringt, d.h. durch Ausführen von u_k-k Nachbartauschen. Also gilt

$$\mathrm{sign}(\sigma) = \mathrm{sign}(\sigma') \cdot (-1)^{u_k - k}$$

Daraus ergibt sich die behauptete Formel.
QED.

4.6.4 Ein Spezialfall

Seien $A \in K^{n\times n}$ eine quadratische Matrix und k eine natürliche Zahl mit $1 \le k \le n$. Dann ist die Summe der Glieder von

$$\det A = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma)\, a_{1\sigma(1)} ... a_{n\sigma(n)}$$

in denen nur Faktoren der ersten k Zeilen und Spalten und der letzten n-k Zeilen und Spalten[19] vorkommen gleich

---

[17] wenn man der erste Glied u_{k+1} der zweiten Folge als an der (k+1)-ten Stelle befindlich ansieht.

[18] Kein Index größer als u_k kommt in der ersten Folge vor, d.h. die Indizes m, m-1,m-2,...,u_k sind gerade die x+1 letzten Glieder der zweiten Folge (in umgekehrter Reihenfolge).

[19] d.h. jeder Faktor, der aus einer der ersten k Zeilen kommt, soll auch in einer der ersten k Spalten liegen und analog, jeder Faktor, der aus einer der letzten n-k Zeilen kommt, soll auch in einer der letzten n-k Spalten liegen.

$$\det A^{1,\dots,k}_{1,\dots,k} \cdot \det A^{k+1,\dots,n}_{k+1,\dots,n}$$

Beispiel

Das Produkt

$$a_{11}a_{22}\cdot\ldots\cdot a_{nn}$$

ist ein Glied der beschriebenen Art, nicht aber das Produkt

$$\pm a_{1n}a_{2n-1}\ldots a_{n1}$$

Beweis. Wir fragen zunächst nach den Gliedern in der Determinanten-Formel, welche das Produkt

$$a_{11}\cdot\ldots\cdot a_{kk}$$

enthalten. In diesen Gliedern kommt kein anderer Eintrag aus den ersten k Zeilen oder Spalten vor. Die Summe x dieser Glieder ändert sich also nicht, wenn wir alle Einträge der ersten k Zeilen und Spalten bis auf $a_{11},\dots,a_{kk}$ durch Null ersetzen. Deshalb gilt

$$x = \det\begin{pmatrix} a_{11} & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{kk} & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & * & \dots & * \\ \dots & \dots & & & & & \end{pmatrix} = a_{11}\cdot\ldots\cdot a_{kk}\cdot \det A^{k+1,\dots,n}_{k+1,\dots,n}$$

Wir können also einige der Glieder in der Determinantenformel so herausgreifen, daß sich die Determinante

$$\det A^{k+1,\dots,n}_{k+1,\dots,n} \tag{1}$$

als Faktor ausklammern läßt. Fragen wir jetzt nach der Gesamtheit aller dieser Glieder. Die Einträge a_{ij} der letzten n-k Zeilen, welche in diesen Gliedern vorkommen, stehen gleichzeitig in den letzten n-k Spalten: es sind gerade diejenigen, die in der Determinantenformel für (1) vorkommen. Diese Glieder bleiben unverändert, wenn wir die übrigen Einträge der letzten n-k Zeilen bzw. letzen n-k Spalten durch Null ersetzen. Die Summe dieser Glieder ist also gleich

$$y = \det\begin{pmatrix} A^{1,\dots,k}_{1,\dots,k} & 0 \\ 0 & A^{k+1,\dots,n}_{k+1,\dots,n} \end{pmatrix} = \det A^{1,\dots,k}_{1,\dots,k} \cdot \det A^{k+1,\dots,n}_{k+1,\dots,n}$$

QED.

4.6.5 Beweis des Satzes

Seien jetzt

$$v_1 < \dots < v_k \text{ und } v_{k+1} < \dots < v_n$$

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

$$(-1)^{k(k+1)/2}(-1)^{\sum_{\nu=1}^{k} v_\nu}$$

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 Ausdrücken 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 beinahe schon die allgemeine Formel, die wir beweisen wollen. Der Unterschied zur allgemeinen Formel besteht darin, daß wir hier nach den ersten k Zeilen entwickelt haben, aber nach beliebigen Zeilen entwickeln wollen. Letzteres können wir aber dadurch tun, indem wir die uns interessierenden Zeilen $u_1, \dots, u_k$ durch eine Permutation in die ersten k Zeilen überführen und dann nach den ersten k Zeilen entwickeln. Die Determinante multipliziert sich dabei mit dem Faktor

$$(-1)^{k(k+1)/2}\,(-1)^{\sum_{\mu=1}^{k} u_\mu}$$

Wir bekommen damit die gesuchte Formel.
QED.

5. Eigenwerte und Eigenvektoren

Vorbemerkungen

(i) Matrizen zu unterschiedlichen Basen sind konjugiert. 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 $S = M^{v'}_{v}(\mathrm{Id})$ 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 konjugiert.

(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.

5.1 Eigenwerte und Eigenvektoren

5.1.1 Definition

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 $v_1,...,v_n$ von V heißt Eigenbasis von f, wenn sämtliche Vektoren v_i Eigenvektoren von f sind,

$$f(v_i) = c_i v_i \text{ für } i = 1,...,n.$$

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 $V_c(A)$ Eigenraum 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 $c_1,\dots,c_n$ 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

$$e_1,\dots,e_n$$

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

$$v = (v_1,\dots,v_n)$$

eine Eigenbasis von f. Dann gilt

(1) $$f(v_i) = c_i v_i \text{ für } i=1,\dots,n$$

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

$$\chi_f(T) := \det(f - T \cdot \mathrm{Id})$$

<u>charakteristisches Polynom</u> von f. Sei $v = (v_1,\dots,v_n)$ 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
$$\begin{aligned}\chi_A(T) &= \det \begin{pmatrix} 1-T & 2 \\ 3 & 4-T \end{pmatrix} \\ &= (T-1)(T-4) - 6 \\ &= T^2 - 5T - 2 \\ &= (T - \frac{1}{2}(5+\sqrt{33}))\ (T - \frac{1}{2}(5-\sqrt{33}))\end{aligned}$$

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)’ $\chi_A(c) = 0$.

(i)’ $\Rightarrow$ (ii)’. 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)’.

(ii)’ $\Rightarrow$ (i)’. Sei c eine Nullstelle von χ_A. Mit $B := A - c\cdot Id$ gilt dann
$$\det B = 0,$$
also
$$\mathrm{rk}\, B < n.$$
Die Spalten $b_1,...,b_n$ von B sind also linear abhängig, d.h. es gibt Elemente $v_1,...,v_n\in K$ mit

(1) $$\sum_{i=1}^{n} v_i b_i = 0 \text{ und } v_i \neq 0 \text{ für ein i.}$$

Wir setzen $v = \begin{pmatrix} v_i \\ \ldots \\ v_n \end{pmatrix}$. 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

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 3+\sqrt{33} \end{pmatrix} \text{ und } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 3-\sqrt{33} \end{pmatrix}$$

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 <u>algebraisch abgeschlossen</u>.
2. Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen (Fundamentalsatz der Algebra).

Vereinbarung

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 <u>hinreichend groß</u>.

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

$$\mu_i := \dim V_{c_i}$$

5.1.6 Zerlegung eines Polynoms in Linearfaktoren

Seien f(T) ein nicht-konstantes Polynom mit Koeffizienten aus K und $\bar{K}$ ein algebraisch abgeschlossener Körper, welcher K enthält. Dann gibt es einen Körper L zwischen K und $\bar{K}$ 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 $L := \bar{K}$ 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 $\bar{K}$ 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 $\bar{K}$. 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.

5.1.7 Existenz von Eigenbasen und algebraische Vielfachheiten

Sei $f:V\rightarrow V$ 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) $V =$[20] $V_{c_1} \oplus ... \oplus V_{c_r}$
(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

$$M_v^v(f) = \begin{pmatrix} c & ... & 0 & * & ... & * \\ ... & ... & ... & ... & ... & ... \\ 0 & ... & c & * & ... & * \\ 0 & ... & 0 & * & ... & * \\ ... & ... & ... & ... & ... & ... \end{pmatrix} = \begin{pmatrix} c\cdot Id & A \\ 0 & B \end{pmatrix},$$

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

$$\dim V = \sum_{j=1}^{r} \mu'_j \leq \sum_{j=1}^{r} \mu_j \leq \sum_{j=1}^{r} \nu_j = \deg \chi_f = \dim V.$$

In der Abschätzung muß überall das Gleichheitszeichen stehen. Das ist aber nur möglich, wenn

$$\mu'_j = \mu_j = \nu_j$$

[20] 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.

gilt für jedes j. Insbesondere gilt (iii). Wir haben außerdem bewiesen:

Folgerung
In jeder Eigenbasis kommen alle Eigenvektoren mit ihren geometrischen Vielfachheiten vor.

(iii) $\Rightarrow$ (ii). Nach Voraussetzung ist die Dimension

$$\mu_i := \dim V_{c_i}$$

gleich der algebraischen Vielfachheit ν_i. Insbesondere gilt

$$\dim V_{c_1} \oplus ... \oplus V_{c_r} = \sum_{i=1}^{r} \mu_i = \sum_{i=1}^{r} \nu_i = \deg \chi_f(T) = \dim V.$$

Die Dimension der Räume, deren Isomorphie wir zeigen müssen, ist somit gleich. Es reicht also, wenn wir zeigen, die lineare Abbildung

$$\varphi: V_{c_1} \oplus ... \oplus V_{c_r} \to V\ ,\ (v_1,...,v_r) \mapsto v_1 + ... + v_r,$$

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) $$v_1 + ... + v_r = 0.$$

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 $(v_1,...,v_r)$ 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 $(v_1,...,v_r)$ 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’) $$v_1 + ... + v_s = 0,$$

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) $$c_1v_1 + ... + c_sv_s = 0$$

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) $\Rightarrow$ (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[21] 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.

5.1.8 Existenz von Eigenwerten

Sei f: V$\to$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.

5.1.9 Fahnen von Vektorräumen

Sei V ein K-Vektorraum. Eine Fahne der Länge r von V ist eine echt aufsteigende Folge
$$V_0 \subset V_2 \subset ... \subset V_r$$
von Unterräumen von V. Die Fahne heißt vollständig, 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,...,r-1.

[21] über algebraisch abgeschlossenen Körpern

Seien $f:V\to V$ ein linearer Endomorphismus und $W \subseteq V$ ein linearer Unterraum. Dann heißt W auch f-invariant, falls gilt $f(W) \subseteq W$.

5.1.10 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

$$0 = V_0 \subset V_1 \subset ... \subset V_d = V$$

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.

$$\begin{array}{ccccccc} V & \xrightarrow{f} & V & \qquad & v & \mapsto & f(v) \\ \downarrow\rho & & \downarrow\rho & & \downarrow & & \downarrow \\ V' & \xrightarrow{f'} & V' & & v+W & \mapsto & f(v)+W \end{array}$$

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

$$0 = V'_1 \subset V'_2 \subset ... \subset V_d = V'$$

Wir setzen

$$V_i := f^{-1}(V'_i) \quad \text{für } i = 1,...,d$$

und $V_0 := \{0\}$. Dann gilt

$$0 = V_0 \subseteq V_1 \subseteq ... \subseteq V_d = V.$$

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,..., 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.

5.1.11 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

5.2.1 Definition

Ein linearer Endomorphismus

$$f: V \to V$$

heißt nilpotent, wenn es eine natürliche Zahl g gibt mit

$$f^g = f\circ ... \circ f \text{ (g-fache Komposition)} = 0.$$

Die kleinste solche natürliche Zahl g heißt Ordnung von f. Eine Matrix $N\in K^{n\times n}$ heißt nilpotent, falls die zugehörige Abbildung f_N nilpotent ist, d.h. falls $N^g = 0$ gilt für ein natürliche g. Die Ordnung von f_N heißt dann auch Ordnung von N.

Die wichtigste Eigenschaft nilpotenter Endomorphimen wird im nachfolgenden Lemma beschrieben, auf welches wir uns im folgenden nur als auf das Lemma beziehen wollen.

Lemma über nilpotente Endomorphismen

Seien $f:V\to V$ ein nilpotenter linearer Endomorphismus und $W\subseteq V$ ein nicht-trivialer linearer f-invarianter Unterraum. Dann gilt

$$f(W) \subset W \text{ (echtes Enthaltensein).}$$

Beweis. Angenommen es gilt f(W) = W. Dann gilt $f^j(W) = W$ für jedes j. Weil f nilpotent ist, können wir j so wählen, daß gilt

$$0 = f^j(V) \supseteq f^j(W) = W.$$

Also muß W = 0 sein im Widerspruch zur Wahl von W.
QED.

5.2.2 Nilpotenz und Fahnen invarianter Unterräume

Sei $f:V\to V$ ein linearer Endomorphismus mit $d:=\dim V < \infty$. Dann sind folgende Aussage äquivalent.

(i) f ist nilpotent.

(ii) Es gibt eine vollständige Fahne von f-invarianten Unterräumen

$$0 = V_0 \subseteq V_1 \subseteq ... \subseteq V_d = V$$

mit $f(V_i) \subseteq V_{i-1}$ für i = 1,...,d.

(iii) $f^d = 0$ für d = dim V.

Beweis. (i) ⇒ (ii). Wir setzen $V_d = V$. Seien jetzt $V_d, V_{d-1}, ..., V_i$ bereits konstruiert. Wir haben die Konstruktion von V_{i-1} zu beschreiben. Falls $V_i = 0$ ist, gibt es nichts zu beschreiben: die Konstruktion ist abgeschlossen. Sei also $V_i\neq 0$. Weil V_i invariant bezüglich f ist, gilt

$$f(V_i) \subseteq V_i.$$

Wegen des Lemmas über nilpotente Endomorphismen ist diese Enthaltenseinsrelation echt. Es gibt also einen Unterraum V_{i-1} mit

$$f(V_i) \subseteq V_{i-1} \subseteq V_i \text{ und } \dim V_{i-1} = \dim V_i - 1.$$

Dieser Raum V_{i-1} ist invariant, denn es gilt

$$f(V_{i-1}) \subseteq f(V_i) \subseteq V_{i-1}.$$

Nach Konstruktion gilt $f(V_i) \subseteq V_{i-1}$.

(ii) ⇒ (iii). Nach Voraussetzung gilt $f^j(V_i) \subseteq V_{i-j}$. Insbesondere ist

$$f^d(V) \subseteq V_{d-d} = V_0 = 0.$$

(iii) ⇒ (i). trivial.
QED.
Bemerkung
Man beachte, im Beweis wurde nicht verwendet, daß K hinreichend groß sein soll.

5.2.3 Die Matrix eines nilpotenten Endomorphismus

Sei $f:V\to V$ ein linearer Endomorphismus eines Vektorraums V endlicher Dimension. Dann sind folgende Aussagen äquivalent.

(i) f ist nilpotent.

(ii) Es gibt eine Basis v von V derart, daß $M_v^v(f)$ obere Dreiecksgestalt hat, wobei in der Hauptdiagonalen lauter Null stehen.

Beweis. (i) ⇒ (ii). Wir wählen eine vollständige Fahne

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_d = V$$

mit $f(V_i) \subseteq V_{i-1}$ für alle i. Dann gibt es eine Basis v von V derart, daß für jedes i die ersten i Vektoren dieser Basis gerade V_i erzeugen. Wie im Beweis von 5.1.11 ergibt sich, daß die Matrix $M_v^v(f)$ gerade die behauptete Gestalt hat.

(ii) ⇒ (i). Sei v eine Basis von V, für welche $M_v^v(f)$ obere Dreiecksgestalt hat, wobei auf der Hauptdiagonalen lauter Nullen stehen. Das i-te Basiselement wird dann in eine Linearkombination der ersten i-1 Basiselemente abgebildet. Also gilt

$$f(V_i) \subseteq V_{i-1},$$

wenn V_i den von den ersten i Basiselementen erzeugten Unterraum bezeichnet. Die Räume V_i bilden also eine Fahne wie in 5.2.2(ii). Also ist f nilpotent.

QED.

5.2.4 Beispiel: Jordanblöcke

Eine d×d-Matrix der Gestalt

$$J_d(c) = \begin{pmatrix} c & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & c & 1 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & c & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & c \end{pmatrix}$$

in deren Hauptdiagonalen der feste Wert $c\in K$ steht, in deren Diagonalen unmittelbar über der Hauptdiagonalen lauter Einsen stehen und deren sonstigge Einträge sämtlich Null sind, heißt Jordanblock zum Eigenwert c.

Wie wir eben gesehen haben sind Jordanblöcke zum Eigenwert Null nilpotent. Alle übrigen Jordanblöcke sind nicht nilpotent[22].

5.2.5 Beispiel: direkte Summen von Matrizen

Seien $A_1,\ldots,A_r$ quadratische Matrizen mit Einträgen aus K. Dann heißt die Matrix

[22] In der Hauptdiagonalen der j-ten Potenz von $J_d(c)$ steht c^j.

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_r \end{pmatrix},$$

welche sich aus den auf der Hauptdiagonalen angeordneten Matrizen A_i zusammensetzt und welche außerhalb dieser so angeordneten Blöcke als Einträge nur Nullen besitzt, direkte Summe von $A_1,\dots,A_r$ und wird auch wie folgt bezeichnet,

$$A = A_1 \oplus \dots \oplus A_r.$$

Für die n-te Potenz der direkten Summe A gilt

$$A^n = A_1^n \oplus \dots \oplus A_r^n$$

Insbesondere ist die direkte Summe von Matrizen genau dann nilpotent, wenn alle direkten Summanden nilpotent sind.

Bemerkung

Das Ziel dieses Abschnitts besteht darin, zu zeigen, jede nilpotente Matrix ist konjugiert zu einer direkten Summe von Jordan-Blöcken zum Eigenwert Null.

5.2.6 Zyklische Basen

Sei $f: V \to V$ ein K-linearer Endomorphismus mit $d := \dim V < \infty$. Eine f-zyklische Basis von V ist eine Basis von V der Gestalt

(1) $$f^{d-1}(v), f^{d-2}(v), \dots, f^2(v), f(v), v$$

mit einem Vektor $v \in V$, welcher Hauptvektor der zyklischen Basis heißt, wobei zusätzlich gefordert wird, daß

$$f^d(v) = 0$$

gilt. Der Vektorraum V heißt f-zyklisch, falls er eine f-zyklische Basis besitzt. Weiter wollen wir in dieser Situation sagen, der Endomorphismus f ist zyklisch.

Ein linearer Unterraum W von V heißt f-zyklisch, wenn er f-invariant ist und die Einschränkung von f auf W zyklisch ist.

Die f-Ordnung eines Vektors $v' \in V$ ist definiert als die kleinste nicht-negative ganze Zahl i mit $f^i(v') = 0$. Zum Beispiel hat $0 \in V$ die Ordnung 0 und die Vektoren der zyklischen Basis haben die Ordnungen 1, 2, 3, ...d-1, d.

Bemerkungen

(i) Die Bedingung $f^d(v) = 0$ ist notwendig und hinreichend dafür, daß f nilpotent ist[23].

(ii) Die Matrix von f bezüglich der zyklischen Basis $f^{d-1}(v), f^{d-2}(v), \dots, v$ ist gerade ein Jordanblock zum Eigenwert 0,

$$M(f) = \begin{pmatrix} 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix} = J_d(0)$$

---

[23] Ist die Bedingung erfüllt, so ist $f^d(\sum_{i=0}^{d-1} a_i f^i(v)) = \sum_{i=0}^{d-1} a_i f^{i+d}(v) = 0$ für beliebige $a_i \in K$, d.h. f ist nilpotent. Ist umgekehrt f nilpotent (und V von der Dimension d), so gilt sogar $f^d(v') = 0$ für beliebiges $v' \in V$ (siehe 5.2.2).

5.2.7 Beispiel

Seien $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ und $f = f_A$. Wir führen Bezeichnungen für die von Null verschiedenen Spalten von A ein:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Es gilt $f(v_1) = 0$, $f(v_2) = v_1$. Schließlich kann man hier durch Raten noch einen Vektor v_3 finden mit $f(v_3) = v_2$, nämlich

$$v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die Vektoren v_1, v_2, v_3 sind linear unabhängig, denn $\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 \neq 0$. Also ist v_3 Hauptvektor einer zyklischen Basis. Die Matrix von f bezüglich der Basis $v = (v_1, v_2, v_3)$ ist:

$$M_v^v(f) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

5.2.8 Der Kern eines zyklischen Endomorphismus

Seien $f: V \to V$ ein linearer Endomorphismus und $v \in V$ ein Hauptvektor einer f-zyklischen Basis. Dann gilt

$$\ker(f) = K \cdot f^{d-1}(v)$$

mit $d = \dim V$, also

$$\dim \ker(f) = 1.$$

Beweis. Nach Definition des Begriffs der zyklischen Basis gilt

$$0 \neq f^{d-1}(v) \in \ker(f).$$

Es reicht also zu zeigen,

$$\dim \ker(f) = 1.$$

Nun gilt,

$$\operatorname{rk} f = \dim \operatorname{im}(f) = \dim V - \dim \ker(f) = d - \dim \ker(f),$$

d.h. es reicht zu zeigen $\operatorname{rk}(f) = d-1$. Sei $M = M(f)$ die Matrix von f bezüglich der vorliegenden zyklischen Basis. Wie wir wissen, stehen in M über der Hauptdiagonalen lauter Einsen und alle anderen Einträge von M sind Null,

$$M = (0, e_1, e_2, \ldots, e_{d-1}).$$

Daher gilt

$$\operatorname{rk}(f) = \operatorname{rk}(M) = d-1.$$

QED.

Bemerkungen

(i) Seien $f_i: V_i \to V_i$ für $i=1,\ldots,r$ zyklische Endomorphismen und sei

$$f = f_1 \oplus \ldots \oplus f_r : V_1 \oplus \ldots \oplus V_r \to V_1 \oplus \ldots \oplus V_r,$$

d.h. es sei $f(v_1,\ldots,v_r) = (f_1(v_1),\ldots,f_r(v_r))$. Dann gilt

$$\ker(f) = \ker(f_1) \oplus \ldots \oplus \ker(f_r)$$

also dim ker(f) = r. Insbesondere besitzt $V = V_1 \oplus ... \oplus V_r$ im Fall r>1 keine zyklische Basis. Es gibt also nicht-zyklische lineare Endomorphismen.

(ii) Wir wollen im folgenden zeigen, die eben beschriebenen Situation tritt immer ein, d.h. jeder nilpotente Endomorphismus f ist bis auf Isomorphie eine direkte Summe von zyklischen Endomorphismen. Die Anzahl der direkten Summanden ist dabei leicht zu bestimmen: es ist gerade die Dimension des Kerns, denn dieser wird von den Basiselementen der Ordnung 1 erzeugt.

5.2.9 Zerlegung nilpotenter Endomorphismen in zyklische

Sei f: $V \to V$ ein nilpotenter Endomorphismus mit

$$d := \dim V < \infty.$$

Dann gibt es lineare f-invariante Unterräume $V_1,...,V_r$ mit folgenden Eigenschaften.

(i) $V = V_1 \oplus ... \oplus V_r$. Genauer, die lineare Abbildung

$$V_1 \oplus ... \oplus V_r \to V, (v_1,...,v_r) \mapsto v_1 + ... + v_r,$$

ist bijektiv.

(ii) Für jedes i ist V_i zyklisch bezüglich $f|_{V_i}$.

Bezeichne

$$\rho_k(f) := \# \{ V_i \mid \dim V_i = k \}$$

die Anzahl der direkten Summanden V_i der Dimension k. Dann gilt

$$(1) \qquad \sum_{k=i+1}^{d} (k-i)\cdot\rho_k(f) = \mathrm{rk}(f^i) \text{ für } i = 0, 1, 2, ...,d-1$$

Dabei bezeichne f^0 die identische Abbildung, d.h. es ist

$$\dim \mathrm{rk}(f^0) = \dim \mathrm{im}(f^0) = \dim V.$$

Bemerkungen

(i) Von der Zahl r wissen wir, daß sie unabhängig von der speziellen Wahl der V_i ist, denn, wie wir gesehen haben, gilt

$$r = \dim \ker(f).$$

(ii) Durch die Bedingungen (1) sind die Zahlen $\rho_k(f)$ eindeutig festgelegt.[24] Sie hängen nur von den Rängen der Potenzen von f ab und nicht von der speziellen Wahl der V_i.

Beweis. Wir führen den Beweis durch Induktion nach der Ordnung

$$m := \mathrm{ord}\,(f) = \min \{n \in \mathbb{Z} \mid n \ge 0, f^n = 0 \}$$

des Endomorphismus f: $V \to V$:

$$f^m = 0, f^{m-1} \ne 0.$$

Der Fall m = 1. f ist die Nullabbildung. Wir wählen eine Basis

$$v_1, ..., v_d \in V$$

von V und betrachten die zugehörige Zerlegung in eine direkte Summe 1-dimensionaler Unterräume,

---

[24] Wir können (1) als Gleichungssystem in den Unbestimmten $\rho_k(f)$ auffassen. Dieses Gleichungssystem hat ein Koeffizientenmatrix von Dreiecksgestalt, wobei auf der Hauptdiagonalen lauter Einsen stehen.

$$V = Kv_1 \oplus \ldots \oplus Kv_d$$

Jeder der Unterräume Kv_i ist zyklisch (und v_i ist Hauptvektor einer zyklischen Basis). Weiter gilt

$$\rho_k(f) = \begin{cases} \dim V & \text{für } k = 1 \\ 0 & \text{sonst} \end{cases}$$

und

$$rk(f^i) = \begin{cases} \dim V & \text{für } i = 0 \\ 0 & \text{sonst} \end{cases}$$

Die Identitäten (1) sind damit trivialerweise erfüllt.

Der Fall $m > 1$. Seien

$$W := \ker(f)$$

und

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

die natürliche Abbildung. Weiter betrachten wir den linearen Endomorphismus

$$f': V' \to V',\ v + W \mapsto f(v) + W.$$

Zwischenbemerkung. f' ist wohldefiniert.

Aus $v + W = v' + W$ folgt $v-v' \in W$, also $f(v) - f(v') \in f(W) = 0$, also $f(v) = f(v')$, also $f(v) + W = f(v') + W$.

Nach Konstruktion ist f' die eindeutig bestimmte Abbildung, für welche das folgende Diagramm kommutativ ist.

(2)
$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \rho\downarrow & & \downarrow\rho \\ V' & \xrightarrow{f'} & V' \end{array}$$

Weil f nilpotent ist, gilt $\ker f \neq 0$, also

$$\dim V' = \dim V - \dim W < \dim V = d.$$

und wegen des kommutativen Diagramms ist mit f auch f' nilpotent:

$$f'^k \circ \rho = \rho \circ f^k = 0 \text{ für } k \text{ groß, d.h. } f'^k = 0 \text{ für } k \text{ groß.}$$

Wir wollen jetzt die Induktionsvoraussetzung auf den Endomorphismus f' anwenden. Dazu müssen wir noch zeigen,

$$\mathrm{ord}\ f' < m.$$

Es reicht zu zeigen, $f'^{m-1} = 0$. Nach Definition von m gilt $f^m = 0$, d.h. für jedes $v \in V$ gilt

$$0 = f^m(v) = f(f^{m-1}(v)),$$

d.h. $f^{m-1}(v) \in \ker(f) = W = \ker(\rho)$. Es ist also

$$0 = \rho(f^{m-1}(v)) = f'^{m-1}(\rho(v)) \text{ für jedes } v \in V.$$

Da ρ surjektiv ist, folgt $f'^{m-1} = 0$. Die Induktionsvoraussetzung kann also auf f' angewandt werden. Wir erhalten eine Zerlegung

$$V' = V'_1 \oplus \ldots \oplus V'_{r'},$$

in f'-zyklische invarianten Unterräume mit

(3)
$$\sum_{k=i+1}^{\infty} \rho_k(f')(k-i) = rk(f'^i) \text{ für } i = 0, 1, 2, \ldots$$

Wir schreiben hier ∞ für die obere Summationsgrenze um keine Bezeichnung für die Dimension von V' einführen zu müssen. Die zusätzlichen Summanden sind sämtlich gleich Null, da $\rho_k(f') = 0$ sein muß für $k > \dim V'$. Außerdem können wir auch i bis ins Unendliche laufen lassen, da die zusätzlichen Gleichungen von der Gestalt $0 = 0$ sind.

Bezeichne

$$v'_i \in V'_i$$

den Hauptvektor einer zyklischen Basis von V'_i. Für jedes i wählen wir einen Vektor

$$v_i \in V \text{ mit } \rho(v_i) = v'_i\ .$$

Wir führen den weiteren Beweis des Satzes in mehreren Schritten.

1. Schritt. Für jedes i ist v_i der Hauptvektor einer zyklischen Basis eines linearen Unterraums $V_i \subseteq V$ der Dimension

$$\dim V_i = \dim V'_i + 1.$$

Die Summe der V_i ist direkt, d.h. die lineare Abbildung

$$V_1 \oplus ... \oplus V_{r'} \to V_1 + ... + V_{r'}\ (\subseteq V),\ (x_1,...,x_{r'}) \mapsto x_1 + ... + x_{r'},$$

ist ein Isomorphismus.

Sei

(4) $\qquad d_i := \dim V'_i + 1$

und V_i der von den d_i Vektoren v_i , $f(v_i)$, $f^2(v_i)$, ... , $f^{d_i-1}(v_i)$ erzeugte lineare Unterraum von V,

(5) $\qquad V_i := K\, v_i + Kf(v_i) + Kf^2(v_i) + ... + Kf^{d_i-1}(v_i)$

Zum Beweis der Aussage des ersten Schritts reicht es zu zeigen:

1. $f^{d_i}(v_i) = 0$ für i = 1, ... , r'.
2. Die Vektoren $f^k(v_i)$, i = 1, ... , r', k = 0, ... , d_i - 1 sind linear unabhängig.

Zu 1. Es reicht zu zeigen $f(f^{d_i-1}(v_i)) = 0$, d.h.

$$f^{d_i-1}(v_i) \in \ker(f)\ (= W = \ker(\rho)),$$

d.h.

$$\rho(f^{d_i-1}(v_i)) = 0.$$

Es gilt

$$\begin{aligned} \rho(f^{d_i-1}(v_i)) &= f'^{d_i-1}(\rho(v_i)) && \text{(wegen der Kommutativität von (2))} \\ &= f'^{d_i-1}(v'_i) && \text{(nach Definition von } v_i\text{)} \end{aligned}$$

Nun ist V'_i ein Unterraum der Dimension

$$\dim V'_i = d_i - 1$$

(vgl. (4)) und $f'|_{V'_i}$ ist nilpotent. Deshalb gilt $f'^{d_i-1} = 0$ auf V'_i also ist auch

$$f'^{d_i-1}(v'_i) = 0.$$

Zu 2. Angenommen es besteht eine lineare Relation

$$0 = \sum_{i=1}^{r'} \sum_{k=0}^{d_i-1} c_{k,i}\, f^k(v_i) \text{ mit } c_{k,i} \in K.$$

Wir bringen die Summanden der Gestalt $c\cdot f^{d_i-1}(v_i)$ auf die andere Seite und erhalten

$$\sum_{i=1}^{r'}\sum_{k=0}^{d_i-2} c_{k,i}\, f^k(v_i) = -\sum_{i=1}^{r'} c_{d_i-1,i}\, f^{d_i-1}(v_i) \in W$$

Man beachte, wegen 1. liegen die Vektoren $f^{d_i-1}(v_i)$ im linearen Unterraum

$$\ker(f) = W = \ker(\rho).$$

Wir wenden die natürliche Abbilung ρ an und erhalten wegen $\rho\circ f^j = f'^j\circ\rho$ die Identität

$$0 = \sum_{i=1}^{r'}\sum_{k=0}^{d_i-2} c_{k,i}\, f'^k(\rho(v_i)) = \sum_{i=1}^{r'}\sum_{k=0}^{d_i-2} c_{k,i}\, f'^k(v'_i)$$

Nun ist

(6) $$f'^{d_i-2}(v'_i), f'^{d_i-3}(v'_i), \dots , f'(v'_i), v'_i$$

nach Wahl von v'_i eine zyklische Basis von V'_i (vgl. (4)). Weil V' eine direkte Summe der V'_i ist, bilden die Basen (6) zusammen ($i = 1, \dots , r'$) eine Basis von V' und sind insbesondere linear unabhängig. Es gilt also

$$c_{k,i} = 0 \text{ für } i = 1, \dots , r' \text{ und } k = 0, \dots , d_i - 2.$$

Wir haben noch zu zeigen, die $c_{k,i}$ sind auch für $k = d_i-1$ gleich Null. Die Ausgangsrelation bekommt, da die $c_{k,i}$ für $k < d_i-1$ gleich Null sind, die Gestalt

$$0 = \sum_{i=1}^{r'} c_{d_i-1,i}\, f^{d_i-1}(v_i) = f\Big(\sum_{i=1}^{r'} c_{d_i-1,i}\, f^{d_i-2}(v_i)\Big)$$

Man beachte, es gilt $d_i = \dim V'_i + 1 \ge 2$ für jedes i, da keiner der zyklischen direkten Summanden V'_i von V' Null ist. Wir erhalten

$$\sum_{i=1}^{r'} c_{d_i-1,i}\, f^{d_i-2}(v_i) \in \ker(f) = W = \ker(\rho).$$

Wir wenden ρ auf diese Summe an und erhalten wegen der Kommutativität des Diagramms (2)

$$0 = \sum_{i=1}^{r'} c_{d_i-1,i}\, f'^{d_i-2}(\rho(v_i)) = \sum_{i=1}^{r'} c_{d_i-1,i}\, f'^{d_i-2}(v'_i)$$

Wie wir oben angemerkt haben, sind die Vektoren

$$f'^{d_i-2}(v'_i)\ ,\ i = 1, \dots , r'$$

linear unabhängig, d.h. es gilt

$$c_{d_i-1,i} = 0 \text{ für } i = 1, \dots , r'.$$

<u>2. Schritt</u>. Es gibt einen linearen Unterraum $W' \subseteq \ker(f)$ mit

$$V = V_1\oplus\dots\oplus V_{r'}\oplus W'.$$

Insbesondere ist V eine direkte Summe f-zyklischer Unterräume.

Nach dem ersten Schritt ist die Summe der V_i direkt. Es reicht zu zeigen, eine Basis von

$$V_1+\dots+V_{r'}$$

läßt sich zu einer Basis von V ergänzen, wobei man die zusätzlichen Basisvektoren aus W nehmen kann.

Betrachten wir die natürliche Abbildung

$$\rho':V \to V/(V_1+...+V_{r'})$$

Zur Konstruktion einer Basis von V genügt es, irgendeine Basis des Faktorraums rechts zu wählen. Die Urbilder der Basiselemente in V bilden dann zusammen mit der gegebenen Basis von $V_1+...+V_{r'}$ eine solche von V.

Es reicht also zu zeigen, daß wir die Urbilder der Basiselemente so abändern können, daß sie in W liegen. Mit anderen Worten, es reicht zu zeigen:

(7) Für jedes $v\in V$ gibt es ein $w\in W$ mit $\rho'(v) = \rho'(w)$.

Sei $v\in V$. Dann liegt $\rho(v)$ in V' und ist somit eine Linearkombination von Vektoren der Gestalt

$$f'^k(v'_i) \text{ im } i = 1, ... , r', k = 0, ... , d_i - 2,$$

sagen wir

$$\rho(v) = \sum_{i=1}^{r'} \sum_{k=0}^{d_i-2} c_{k,i} f'^k(v'_i) = \sum_{i=1}^{r'} \sum_{k=0}^{d_i-2} c_{k,i} f'^k(\rho(v_i))$$

$$= \rho(\tilde{v}) \text{ mit } \tilde{v} = \sum_{i=1}^{r'} \sum_{k=0}^{d_i-2} c_{k,i} f^k(v_i) \in V_1+...+V_{r'}.$$

Es folgt

$$\rho(v - \tilde{v}) = 0$$

also

$$w := v - \tilde{v} \in \ker(\rho) = W.$$

Nun liegt $\tilde{v}$ in $V_1+...+V_{r'}$ und es gilt somit $\rho'(\tilde{v}) = 0$. Dann ist aber

$$\rho'(v) = \rho'(w+\tilde{v}) = \rho'(w).$$

Wir haben gezeigt

$$V = V_1\oplus...\oplus V_{r'}\oplus W'.$$

mit einem linearen Unterraum $W' \subseteq W = \ker(f)$. Da f auf W die Ordnung 1 hat, ist W' (auf Grund des bereits behandelten Falls m = 0) eine direkte Summe von zyklischen Unterräumen (der Dimension 1), die wir mit $V_{r'+1}$, ... , V_r bezeichnen wollen. Damit bekommt V die behauptete Gestalt

$$V = \oplus_{r}^{i=1} V_i$$

Wir haben noch die Gültigkeit der Formeln (1) zu beweisen. Die direkte Summe über alle zyklischen Unterräume der Dimension k von V hat die Dimension

$$k\times \text{ Anzahl der direkten Summanden} = k\cdot\rho_k(f)$$

Da V die direkte Summe über alle zyklischen Unterräume V_i gleich V ist, folgt

$$\sum_{i=1}^{r} k\cdot\rho_k(f) = \dim V = \mathrm{rk}(f^0).$$

Dies ist die erste der Identitäten (1) (mit i = 0). Beweisen wir die übrigen. Nach Induktionsvoraussetzung gilt

$$\sum_{k=i+1}^{\infty} \rho_k(f')(k-i) = \mathrm{rk}(f'^i) \text{ für } i = 0, 1, 2, ...$$

(vgl. Formel (3)). Wegen (4) ist die Anzahl $\rho_k(f')$ der k-dimensionalen zyklischen direkten Summanden von V' gleich der Anzahl der (k+1)-dimensionalen zyklischen direkten Summanden von V,

$$\rho_k(f') = \rho_{k+1}(f)$$

Durch Einsetzen in die obigen Identitäten erhalten wir

$$\sum_{k=i+1}^{\infty} \rho_{k+1}(f)(k-i) = rk(f'^i) \text{ für } i = 0, 1, 2, \ldots$$

und durch Index-Verschiebung

$$\sum_{k=i+2}^{\infty} \rho_k(f)(k-i-1) = rk(f'^i) \text{ für } i = 0, 1, 2, \ldots$$

Mit anderen Worten,

$$\sum_{k=i+1}^{\infty} \rho_k(f)(k-i) = rk(f'^{i-1}) \text{ für } i =1, 2, \ldots$$

Zum Beweis der verbleibenden Identitäten (1) reicht es somit, wenn wir zeigen

<u>3. Schritt</u>: $rk(f'^{i-1}) = rk(f^i)$ für i = 1, 2, ...

Es gilt

$$rk(f'^{i-1}) = \dim V' - \dim \ker(f'^{i-1}) = \dim V - \dim W - \dim \ker(f'^{i-1})$$

und

$$rk(f^i) \quad = \dim V - \dim \ker(f^i).$$

Es reicht also zu zeigen,

$$\dim \ker(f'^{i-1}) = \dim \ker(f^i) - \dim W.$$

Dazu wiederum reicht es zu zeigen,

$$\ker(f'^{i-1}) = \ker(f^i)/W.$$

Es gilt

$$\begin{aligned} v\in \ker(f^i) &\Leftrightarrow f^i(v) = 0 \Leftrightarrow f(f^{i-1}(v)) = 0 \Leftrightarrow f^{i-1}(v)\in \ker(f) = W = \ker(\rho) \\ &\Leftrightarrow 0 = \rho(f^{i-1}(v)) = f'^{i-1}(\rho(v)) \\ &\Leftrightarrow \rho(v) \in \ker(f'^{i-1}) \\ &\Leftrightarrow v \in \rho^{-1}(\ker(f'^{i-1})) \end{aligned}$$

Damit ist (wegen der Surjektivität von ρ)

$$\begin{aligned} \ker(f'^{i-1}) &= \rho(\rho^{-1}(\ker(f'^{i-1})) = \rho(\ker(f^i)) \\ &= \{ v + W \mid v\in \ker(f^i) \} \\ &= \ker(f^i)/W. \end{aligned}$$

<u>4. Schritt</u>: Durch die Identitäten (1) sind die $\rho_k(f)$ eindeutig festgelegt.

Wir betrachten die Identitäten als lineares Gleichungssystem zur Bestimmung der $\rho_k(f)$.

Die Zahl der Gleichungen ist dann gleich der Zahl der Unbestimmten und die Koeffizientenmatrix ist quadratisch von der Gestalt

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

Das Gleichungssystem hat somit eine eindeutig bestimmte Lösung.

QED.

Bemerkungen

(i) Eine Zerlegung wie in 5.2.9 (i) heißt auch Jordan-Zerlegung von V bezüglich f oder auch Zerlegung in f-zyklische Unterräume.

(ii) Die wichtigste Frage im Zusammenhang mit dieser Zerlegung ist die nach der Anzahl $\rho_k(f)$ der direkten Summanden einer vorgegebenen Dimension k. Die Formeln (1) gestatten es die Zahlen $\rho_k(f)$ zu berechnen.

(ii) Wie wir bereits in 5.2.8 gesehen haben, gilt außerdem

$$\sum_k \rho_k(f) = r = \dim \ker(f).$$

Diese Formel gestattet es oft, die Berechnung der $\rho_k(f)$ zu vereinfachen, bzw. sie bietet eine Möglichkeit einer Probe.

5.2.10 Die Jordansche Normalform eines nilpotenten Endomorphismus

Sei $f:V\to V$ ein nilpotenter Endomorphismus mit $\dim V < \infty$. Dann gibt es eine Basis v von V derart, daß die zugehörige Matrix $M_v^v(f)$ eine direkte Summe von Jordanblöcken zum Eigenwert Null ist,

$$M_v^v(f) = J_{k_1}(0)\oplus\ldots\oplus J_{k_r}(0).$$

Wählt man noch die Reihenfolge der Basiselemente von v derart, daß

$$k_1 \ge k_2 \ge \ldots \ge k_r$$

gilt, so ist die Folge der k_i unabhängig von der speziellen Wahl der Basis.

Beweis. Folgt aus 5.2.9.
QED.

5.2.11 Beispiel

Sei $f = f_A:K^3\to K^3$ mit $A = \begin{pmatrix} 5 & 2 & -1 \\ -10 & -4 & 2 \\ 5 & 2 & -1 \end{pmatrix}$. Wir wollen die Jordansche Normalform dieser Abbildung berechnen. Wir müssen zunächst überprüfen, ob A auch wirklich nilpotent ist, denn nur in diesem Fall können wir bisher die Jordansche Normalform berechnen. Da die Spalten von A proportional sind, hat f den Rang 1 und das Bild

$$\mathrm{im}(f) = K\cdot\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Damit ist

$$\mathrm{im}(f^2) = f(\mathrm{im}(f)) = K\cdot f\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = K\cdot\begin{pmatrix} 5 & 2 & -1 \\ -10 & -4 & 2 \\ 5 & 2 & -1 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = 0,$$

Wir erhaltwen:

$\mathrm{rk}\, f = 1,$

$\dim \ker(f) = \dim K^3 - 1 = 2.$

Damit ist

$$\begin{aligned} \rho_1(f) + \rho_2(f) \quad + \rho_3(f) &= \dim \ker(f) = 2 \\ \rho_1(f) + 2\cdot\rho_2(f) + 3\cdot\rho_3(f) &= r(f^0) = 3 \\ \rho_2(f) + 2\cdot\rho_3(f) &= r(f) = 1 \end{aligned}$$

Die Dimension des Kerns von f ist gleich der Anzahl der zyklischen direkten Summanden. Da diese größer als 1 ist, kann es keinen direkten Summanden der Dimension 3 geben, d.h. es gilt

$$\rho_3(f) \quad = 0$$

(das folgt auch aus der dritten Bestimmungsgleichung). Dann ist aber

$$\rho_2(f) \quad = 1$$

$$\rho_1(f) \quad = 1.$$

Als Jordansche Normalform erhalten wir damit

$$M(f) = J_2(0) \oplus J_1(0).$$

Die Bestimmung einer Basis, bezüglich welcher die Matrix von f die Jordansche Normalform annimmt, ist schwieriger. Die systematische Behandlung dieses Problems verschieben wir auf später (bis wir geeignete Begriffe zur Verfügung haben). Im hier vorliegenden Fall ist dies vergleichsweise einfach, da bereits das Quadrat von f Null ist.

Wir berechnen dazu zunächst den Kern von f. Da f den Rang 1 hat, besteht der Kern aus den Lösungen der einzelnen Gleichung

$$5x + 2y - z = 0.$$

Wir erhalten z = 5x + 2y, d.h. $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 5x+2y \end{pmatrix} = x\cdot\begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} + y\cdot\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, d.h. es ist

$$\ker(f) = K\cdot\begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} + K\cdot\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Der Raum $K^3/\ker(f)$ ist 1-dimensional, also zyklisch bezüglich der durch f induzierten Abbildung. Jeder von Null verschiedene Vektor ist Hauptvektor. Jeder Repräsentant eines solchen Vektors in K^3 ist Hauptvektor eines f-zyklischen Unterraums der Dimension 2 von K^3. Es reicht also, einen beliebigen Vektor zu wählen, der nicht in ker(f) liegt. Zum Beispiel können wir

$$v_1 := \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}$$

wählen, denn es ist

$$v'_1 = f(v_1) = \begin{pmatrix} 5 & 2 & -1 \\ -10 & -4 & 2 \\ 5 & 2 & -1 \end{pmatrix}\begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 30 \\ -60 \\ 30 \end{pmatrix} \neq 0,$$

d.h. v_1 ist nicht im Kern von f, d.h. v_1 ist Hauptvektor eines zyklischen Unterraums

$$V_1 = K\, v_1 + Kv'_1$$

der Dimension 2.

Die Frage, wie man den Vektor v_1 in der allgemeinen Situation findet, behandeln wir später. Wir haben noch einen zyklischen Unterraum der Dimension 1 zu finden, dessen Summe mit V_1 direkt ist. Dazu müssen wir die Basis v_1, v'_1 von V_1 so zu einer Basis

$$v_1, v'_1, v_2 \text{ von } V$$

ergänzen, daß v_2 im Kern von f liegt, d.h. gesucht ist ein Vektor von ker(f), der nicht in V_1 liegt. Aus theoretischen Gründen wissen wir, daß es einen solchen Vektor gibt. Dann hat aber einer der Basisvektoren

$$\cdot\begin{pmatrix}1\\0\\5\end{pmatrix},\begin{pmatrix}0\\1\\2\end{pmatrix}$$

von ker(f) diese Eigenschaft. Ein solcher Vektor ist

$$v_2 = \begin{pmatrix}0\\1\\2\end{pmatrix}$$

Damit hat die Matrix von f die Gestalt $M(f) = J_2(0)\oplus J_1(0)$ bezüglich der Basis v_1, v'_1, v_2 von V.

Bemerkung

Bei der Konstruktion einer Basis, bezüglich welcher ein nilpotenter Endomorphismus eine Matrix in Jordanscher Normalform besitzt, ist es oft nützlich, die Anzahl der zyklischen direkten Summanden der Dimension 1 zu kennen, d.h. die Anzahl der in jedem Induktionsschritt des obigen Beweises hinzukommenden Summanden.

5.2.12 Die Anzahl der zyklischen direkten Summanden der Dimension 1

Seien $f: V\to V$ ein nilpotenter Endomorphismus mit $\dim V < \infty$ und

$$V = \oplus_{i\in I} V_i$$

eine Zerlegung in f-zyklische Unterräume (d.h. eine Jordanzerlegung bezüglich f). Dann ist die Anzahl der direkten Summanden der Dimension 1 gerade gleich

$$\rho_1(f) = \dim \ker(f) - \dim \ker(f)\cap \mathrm{im}(f).$$

Insbesondere im Fall $f = f_A$ mit $A\in K^{n\times n}$ ist dies gerade die Anzahl der 1×1-Jordanblöcke von A.

Beweis. Wir setzen

$$V' := \bigoplus_{\dim V_i=1} V_i$$

$$V'' := \bigoplus_{\dim V_i>1} V_i$$

Dann ist $V = V'\oplus V''$ und

$$\rho_1(f) = \dim V'.$$

Wir haben also zu zeigen, V' hat die Dimension $\dim \ker(f) - \dim \ker(f)\cap \mathrm{im}(f)$. Dazu wiederum genügt es zu zeigen,

$$V'\oplus \ker(f)\cap \mathrm{im}(f) = \ker(f),$$

genauer, die lineare Abbildung

$$\varphi: V'\oplus \ker(f)\cap \mathrm{im}(f) \to \ker(f),\ (a,b) \mapsto a+b,$$

ist bijektiv.

<u>φ ist wohldefiniert</u>.Ein zyklischer Unterraum V_i der Dimension 1 wird von einem Basisvektor der Ordung 1 erzeugt und liegt damit im Kern von f. Also gilt $V'\subseteq \ker(f)$, d.h. die Abbildung φ ist wohldefiniert.

<u>Injektivität von φ</u>. Für jedes $i\in I$ wählen wir einen Hauptvektor v_i einer f-zyklischen Basis von V_i. Sei

$$k_i := \mathrm{ord}\, v_i = \dim V_i$$

die Ordnung von v_i, d.h. die Vektoren

$$f^j(v_i) \text{ mit } j = 0,\dots,k_i -1$$

bilden eine Basis von V_i (und es ist $f^j(v_i) = 0$ für $j = k_i$). Sei jetzt

$$(a,b) \in \ker(\varphi),$$

d.h. es gelte $a + b = 0$. Wir schreiben $a\in V'$ in der Gestalt

(1) $$a = \sum_{k_i=1} c_i \cdot v_i \text{ mit } c_i\in K$$

Wegen $b\in \ker(f)\cap \mathrm{im}(f)$ gibt es ein $b'\in V$ mit $b = f(b')$. Wir schreiben b' in der Gestalt

$$b' = \sum_{j<k_i} c_{ij}\cdot f^j(v_i) \text{ mit } c_{ij} \in K.$$

Dabei können wir b' um Elemente aus $\ker(f)$ abändern, denn dadurch ändert sich $b = f(b')$ nicht. Mit anderen Worten, wir können alle Summanden mit $j = k_i - 1$ in der Summe weglassen und annehmen

$$b' = \sum_{j<k_i-1} c_{ij}\cdot f^j(v_i) \text{ mit } c_{ij} \in K.$$

Insbesondere fallen damit alle Summanden mit $k_i = 1$ weg, d.h. gerade diejenigen i, über welche die Summe von (1) erstreckt wird. Wir wenden f an und erhalten

(2) $$b = \sum_{j<k_i-1} c_{ij}\cdot f^{j+1}(v_i)$$

Wir sehen, a und b sind Linearkombinationen von disjunkten Mengen der Basisvektoren $f^j(v_i)$ (in (1) treten nur Basisvektoren mit $j=0$ auf und in (2) nur solche mit $j\ge 1$). Damit folgt aber aus $a + b = 0$ (und der linearen Unabhängigkeit der $f^j(v_i)$ mit $j<k_i$ und $i\in I$), daß alle c_i und alle c_{ij} Null sein müssen. Mit anderen Worten, es ist $a = b = 0$.

Surjektivität von φ. Sei $v\in \ker(f)$ $(\subseteq V)$. Wir schreiben v mit Hilfe der oben gewählten Basis der $f^j(v_i)$ in der Gestalt

$$v = \sum_{j<k_i} c_{ij} f^j(v_i) \text{ mit } c_{ij}\in K$$

und setzen

$$a := \sum_{k_i=1} c_{i0} v_i \text{ und } b := \sum_{j<k_i, k_i\neq 1} c_{ij} f^j(v_i)$$

Dann gilt $a + b = v$ und $a\in \ker(f)$ (weil die v_i in der ersten Summe die Ordnung $k_i=1$ haben). Wir haben also nur noch zu zeigen,

(4) $$b \in \ker(f)\cap \mathrm{im}(f).$$

Wegen $v\in \ker(f)$ und $a\in \ker(f)$ gilt auch

(3) $$b = v - a \in \ker(f),$$

also

$$0 = f(b) = \sum_{j<k_i, k_i\neq 1} c_{ij} f^{j+1}(v_i),$$

d.h. alle c_{ij} mit $f^{j+1}(v_i) \neq 0$ sind Null. Genauer gilt,

$$c_{ij} = 0 \text{ für } j = 0,1,...,k_i\text{-}2.$$

In der Summe

$$b = \sum_{j<k_i, k_i\neq 1} c_{ij} f^j(v_i)$$

kann man also insbesondere alle Summanden mit j = 0 weglassen. Das bedeutet aber, b liegt im Bild von f,

$$b \in \text{im}(f).$$

Dies zusammen mit (3) ist aber gerade die zu beweisende Aussage (4).
QED.

5.3 Die Jordansche Normalform (beliebiger Matrizen)

5.3.1 Vorbemerkung (Wdhlg)

Sei $f:V\to V$ ein K-linearer Endomorphismus mit $d := \dim V < \infty$ und r verschiedenen Eigenwerten $c_1, ... , c_r$. Wie wir bereits wissen, ist in "günstigen" Fällen der Vektorraum V eine direkte Summe der Eigenräume

$$V_{c_i} := \ker(f\text{-}c_i\cdot\text{Id}),$$

d.h. f besitzt eine Eigenbasis, d.h. eine Basis bezüglich welcher die Matrix von f Diagonalgestalt besitzt. "Günstig" bedeutet hier, die Dimensionen der Eigenräume sind groß genug, und zwar so groß, daß die Summe der Dimensionen gerade d ist,

$$\sum_{i=1}^{r} \dim V_{c_i} = \dim V.$$

Unser Ziel in diesem Abschnitt besteht darin, die Eigenräume durch etwas größere Räume zu ersetzen, für die diese Dimensionsbedingung immer erfüllt ist, und eine explizite Beschreibung der Matrix von f auf diesen Räumen anzugeben. Diese Räume werden die Gestalt

$$\ker(f\text{-}c_i\cdot\text{Id})^{n_i}$$

besitzen mit n_i hinreichend groß. Zunächst müssen wir deshalb die Kerne für sämtliche Werte von n_i = 1,2,3,... betrachten. Anstelle von $f\text{-}c_i\cdot\text{Id}$ werden wir vorerst einen beliebigen Endomorphismus $g:V\to V$ betrachten. Später wird uns nur der Fall

$$g = f\text{-}c_i\cdot\text{Id}$$

interessieren.
Wir beginnen mit einigen allgemeinen Betrachtungen zur Menge aller Endomorphismen eines Vektorraums.

5.3.2 Der Endomorphismenring eines Vektoraums

Sei V ein K-Vektorraum. Wie wir bereits wissen ist ein K-Endomorphismus von V eine K-lineare Abbildung

$$f: V\to V.$$

Die Menge aller K-Endomorphismen von V wird mit

$$\text{End}(V) = \text{End}_K(V) := \text{Hom}_K(V,V)$$

bezeichnet und ist offensichtlich ein K-Vektorraum. Außerdem definiert die Komposition von Abbildungen auf End(V) eine weitere Operation,

$$f\cdot g := f\circ g,$$

durch welche End(V) zu einem nicht-notwendig kommutativen Ring mit Eins wird, d.h. "$\circ$" spielt die Rolle der Multiplikation und es gelten die üblichen Rechengesetze (außer eventuell dem Kommutativgesetz).

Für jedes Polynom

$$f(x) = a_n x^n + \ldots + a_0$$

mit Koeffizienten $a_i \in K$ und jeden Endomorphismus $\varphi \in \mathrm{End}(V)$ ist somit der Ausdruck

(1) $$f(\varphi) = a_n \varphi^n + \ldots + a_0$$

wohldefiniert und beschreibt einen Endomorphismus von V, d.h.

$$f(\varphi) \in \mathrm{End}(V).$$

Die Teilmenge aller Endomorphismen der Gestalt (1) mit festem φ, wobei f alle Polynome mit Koeffizienten aus K durch läuft, wird mit

$$K[\varphi]$$

bezeichnet. Diese Teilmenge von End(V) ist sogar ein Teilring[25]. In diesem Teilring gilt stets das Kommutativgesetzt.[26]

5.3.3 Das Minimalpolynom eines Endomorphismus

Sei $f: V \to V$ linearer Endomorphismus mit $\dim V < \infty$. Dann gibt es genau ein Polynom kleinsten Grades $m_f(T) \in K[T]$ mit dem höchsten Koeffizienten 1 und der Nullstelle f,

$$m_f(f) = 0.$$

Dieses Polynom teilt jedes Polynom aus K[T] mit der Nullstelle f. Es heißt Minimalpolynom von f.

Beweis. Wir zeigen zunächst, es gibt ein Polynom $p(T) \in K[T]$ mit

$$p(f) = 0.$$

Dazu fixieren wir eine Basis und betrachten die zugehörige Matrix A von f. Es reicht zu zeigen, es gibt ein Polynom p mit $p(A) = 0$.

Ist $d := \dim V$, so ist A ein $d\times d$-Matrix. Der Raum $K^{d\times d}$ aller $d\times d$-Matrizen hat die Dimension d^2. Deshalb sind die d^2+1 Elemente

$$A^0, A^1, A^2, \ldots, A^{d^2} \in K^{d\times d}$$

linear abhängig. Das bedeutet aber gerade, es gibt ein Polynom p der behaupteten Art.

Aus der Existenz eines Polynoms p mit $p(A) = 0$ folgt natürlich auch die Existenz eines Polynoms m minimalen Grades mit $m(A) = 0$, d.h. eines Minimalpolynoms.

Sei jetzt p irgendein Polynom mit $p(A) = 0$. Polynomiale Division mit Rest liefert dann

$$p = q \cdot m + r$$

mit Polynomen q und r, wobei außerdem noch $\deg r < \deg m$ gilt. Wir setzen A ein und erhalten

$$0 = p(A) = q(A)\cdot m(A) + r(A) = q(A)\cdot 0 + r(A),$$

also $r(A) = 0$. Wäre $r \neq$ würde dies der Minimalität des Grades von m widersprechen. Es muß also $r = 0$ gelten, d.h. es ist

$$p = q\cdot m.$$

Wir haben gezeigt, jedes Polynom mit der Nullstelle A ist Vielfaches von m. Insbesondere sind zwei Minimalpolynome Vielfache voneinander. Sie haben also denselben Grad. Weil der höchst Koeffizienten eines Minimalpolynoms 1 sein soll, kann es somit nur ein Minimalpolynom von A geben.

[25] Addition und Multiplikation von Elementen dieser Teilmenge liefert wieder Elemente von dieser Teilmenge.

[26] Wegen $f^i \circ f^j = f^{i+j} = f^j \circ f^i$.

QED.

Bemerkung

Aus dem Beweis ergibt sich, daß $\deg m_f \le d^2$ gilt. Wir werden später sehen, es ist sogar $\deg m_f \le d$ (:= dim V).

Beispiel

Seien $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ und $f(T) = T^2 - 5T - 2$. Es gilt

$$A^2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

also

$$\begin{aligned} f(A) &= A^2 - 5A - 2\cdot Id \\ &= \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - 5\cdot\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\cdot\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - 2\cdot\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Überzeugen wir uns nun davon, daß f sogar das Minimalpolynom von A ist. Dazu müssen wir zeigen, daß kein Polynom eines Grades < 2 (außer dem Nullpolynom) die Nullstelle f hat. Sei g ein solches Polynom. Wir können annehmen, g ist nicht konstant, d.h.

$$\deg g = 1.$$

Außerdem können wir durch den höchsten Koeffizienten teilen und so erreichen, daß g die Gestalt

$$g(T) = T - a$$

hat. Dann gilt

$$g(A) = A - a\cdot\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = \begin{pmatrix} 1-a & 2 \\ 3 & 4-a \end{pmatrix}.$$

Diese Matrix ist niemals Null, ganz gleich wie man a wählt.

5.3.4 Potenzen eines Endomorphismus, Stabilisieren von Kern und Bild

Sei $g: V\to V$ ein K-linearer Endomorphismus mit $\dim V < \infty$. Dann gelten die folgenden Aussagen.

(i) Die Kerne bzw. die Bilder der Potenzen g^i, i=0,1,2, 3, ... bilden eine aufssteigende bzw. absteigende Kette von K-linearen Unterräumen von V,

$$0 \subseteq \ker(g)\subseteq\ker(g^2)\subseteq...\subseteq\ker(g^i)\subseteq...\subseteq V$$

$$V\supseteq im(g)\supseteq im(g^2)\supseteq...\supseteq im(g^i)\supseteq...\supseteq 0$$

(ii) Es gibt nicht-negative ganze Zahlen u bzw. v mit

$$\ker(g^i) = \ker(g^{i+1}) \text{ für } i=u, u+1, u+2, ...$$

bzw.

$$im(g^i) = im(g^{i+1}) \text{ für } i=v, v+1, v+2, ...$$

Wählt man u und v minimal, so gilt u=v. Wir wollen dieses u = v den ersten stabilen Exponenten von g nennen und mit

$$u = stab(g)$$

bezeichnen. Entsprechend heißt g^u dann erste stabile Potenz von g.

(iii) Ist i kleiner als der erste stabile Exponent von g, so bestehen sogar die echten Inklusionen,

$$\ker(f^i) \subset \ker(f^{i+1}) \text{ und } im(f^i) \supset im(f^{i+1}).$$

Beweis. Zu (i): Wir haben zu zeigen, für jedes i gilt

1. $\ker(g^i) \subseteq \ker(g^{i+1})$
2. $\mathrm{im}(g^i) \supseteq \mathrm{im}(g^{i+1})$.

Zu 1. Für $x \in \ker(g^i)$ gilt $g^i(x) = 0$, also $g^{i+1}(x) = 0$, also $x \in \ker(g^{i+1})$.

Zu 2. Für $y \in \mathrm{im}(g^{i+1})$ gibt es ein $x \in V$ mit $y = g^{i+1}(x) = g^i(g(x))$, d.h. es gilt $y \in \mathrm{im}(g^i)$.

Zu (ii). Die Existenz von u bzw. v folgt aus (i) wegen der Endlichkeit der Dimension von V. Die Gleichheit der beiden Zahlen ergibt sich aus der Identität

$$\dim \ker(g^i) + \dim \mathrm{im}(g^i) = \dim V, \tag{1}$$

welche für alle i gilt: wenn bei Vergrößerung von i die Dimension des Kerns zunimmt, so muß dabei die Dimension des Bildes abnehmen.

Zu (iii). Es reicht die folgende Implikation zu beweisen.

$$\mathrm{im}(g^i) = \mathrm{im}(g^{i+1}) \Rightarrow \mathrm{im}(g^{i+1}) = \mathrm{im}(g^{i+2}). \tag{2}$$

Die analoge Implikation für die Kerne der Potenzen von g ergibt sich dann aus der Identität (1). Sei also

$$\mathrm{im}(g^i) = \mathrm{im}(g^{i+1}).$$

Es reicht zu zeigen, es gilt

$$\mathrm{im}(g^{i+1}) \subseteq \mathrm{im}(g^{i+2}),$$

denn die umgekehrte Inklusion wurde bereits in (i) bewiesen. Sei also $y \in \mathrm{im}(g^{i+1})$. Dann gibt es ein $x \in V$ mit $y = g^{i+1}(x)$, d.h. es gilt

$$y = g(z) \text{ mit } z = g^i(x).$$

Der Vektor z liegt in $\mathrm{im}(g^i) = \mathrm{im}(g^{i+1})$, d.h. es gibt ein $w \in V$ mit $z = g^{i+1}(w)$. Dann ist aber

$$y = g(z) = g(g^{i+1}(w)) \in \mathrm{im}(g^{i+2}).$$

QED.

5.3.5 Haupträume

Seien $f:V \to V$ ein K-linearer Endomorphismus mit $d := \dim V < \infty$ und $c \in K$ ein Eigenwert von f. Dann heißt der Kern

$$V'_c(f) := V'_c := \ker(f - c\cdot\mathrm{Id})^u \text{ mit } u := \mathrm{stab}(f - c\cdot\mathrm{Id})$$

der ersten stabilen Potenz von $f - c\cdot\mathrm{Id}$ Hauptraum von f zum Eigenwert c. Mit anderen Worten, V'_c ist der größte K-lineare Unterraum von V, der von einer Potenz der Abbildung

$$f - c\cdot\mathrm{Id}$$

in die Null abgebildet wird. Nach Konstruktion ist der Eigenraum zum Eigenwert c ganz im Hauptraum enthalten,

$$V_c \subseteq V'_c .$$

Bemerkungen

(i) Der Hauptraum V'_c ist f-invariant: mit $v \in \ker(f - c\cdot\mathrm{Id})^u$ gilt

$$(f - c\cdot\mathrm{Id})^u (f(v)) = (f - c\cdot\mathrm{Id})^u \circ f\,(v) = f \circ (f - c\cdot\mathrm{Id})^u(v) = f(0) = 0.$$

(ii) Insbesondere ist das charakteristische Polynom von f auf V'_c ein Teiler des charakteristischen Polynoms von f auf V. Genauer, es gilt

$$\chi_f(T) = \chi_{f|V'_c}(T) \cdot \chi_{f'}(T),$$

wenn $f':V'\to V'$ die durch f auf $V':=V/V'_c$ induzierte Abbilung bezeichnet.

(iii) Nach Definition ist die Einschränkung von $f - c\cdot Id$ auf V'_c nilpotent, d.h. diese Einschränkung besitzt eine Matrix, die in der Hauptdiagonalen und darunter lauter Nullen hat. Insbesondere gilt

$$\det(f-c\cdot Id - T\cdot Id) = (-T)^k \text{ mit } k = \dim V'_c$$

also ist

$$\chi_{f|V'_c}(T) = \det(f - T\cdot Id)|_{V'_c} = (c-T)^k$$

(iv) Weil $\chi_{f|V'_c}$ ein Teiler von χ_f ist, kann die Dimension des Hauptraum höchstens so groß sein wie die algebraische Vielfachheit des Eigenwertes c von f,

$$\dim V'_c \le \nu_f(c).$$

Unser nächstes Ziel ist es, zu zeigen, daß sogar das Gleichheitszeichen gilt.

5.3.6 Die Dimension der Haupträume

Seien $f:V\to V$ ein K-linearer Endomorphismus mit $\dim V<\infty$ und $c\in K$ ein Eigenwert von f. Falls K hinreichend groß ist, ist die Dimension des Hauptraums zum Eigenwert c gleich der algebraischen Vielfachheit von c,

$$\dim V'_c(f) = \nu_f(c).$$

Beweis. Angenommen, es wäre

(1) $$k := \dim V'_c < \nu_f(c).$$

Wir setzen

$$W := V/V'_c$$

und betrachten den durch f auf W induzierten Endomorphismus

$$f': W \to W,\ w+ V'_c \mapsto f(w)+ V'_c.$$

Es gilt

$$\chi_f(T) = \chi_{f|V'_c}(T) \cdot \chi_{f'}(T) = (c-T)^k\cdot \chi_{f'}(T)$$

Wegen (1) kommt in der Faktorzerlegung des Polynoms $\chi_{f'}(T)$ noch mindestens ein Faktor der Gestalt T-c vor, d.h. es ist

$$\chi_{f'}(c) = 0,$$

d.h. c ist Eigenwert von f'. Deshalb hat f' einen (von Null verschiedenen) Hauptraum zum Eigenwert c, d.h. es gibt einen Unterraum

$$\tilde{U} \subseteq W,\ \tilde{U} \ne 0,$$

und eine natürliche Zahl

$$\ell \in \mathbb{N}$$

mit

(2) $$(f' - c\cdot Id)^{\ell}(\tilde{U}) = 0.$$

Seien

$$\rho:V\to W = V/V'_c$$

die natürliche Abbildung und $U := \rho^{-1}(\tilde{U})$. Dann ist U ein Unterraum von V mit

(3) $$V'_c \subset U \subseteq V,$$

wobei die erste Inklusion eine echte Inklusion ist. Für jedes $u\in U$ gilt

$$\rho((f - c\cdot Id)^{\ell}(u)) = (f - c\cdot Id)^{\ell}(u) + V'_c = (f' - c\cdot Id)^{\ell}(u + V'_c) = (f' - c\cdot Id)^{\ell}(\rho(u)) = 0$$

(wegen $\rho(u) \in \tilde{U}$ und (2)). Mit anderen Worten, es ist $(f - c\cdot Id)^{\ell}(u) \in \ker(\rho) = V'_c$. Nach Definition des Hauptraums ist dann aber

$$(f - c\cdot Id)^{\ell+k}(u) = 0 \text{ für jedes } u\in U \text{ und } k \text{ hinreichend groß.}$$

Insbesondere ist

$$\ker(f - c\cdot Id)^{\ell+k}$$

echt größer als der Hauptraum V'_c. Das steht aber im Widerspruch zur Definition des Hauptraums V'_c (als der größte Unterraum von V, der von einer Potenz von $f - c\cdot Id$ in die Null abgebildtet wird). Dieser Wiederspruch zeigt, daß unsere Annahme (1) falsch sein muß.

QED.

5.3.7 Abschätzung der stabilen Potenzen von f

Seien $f:V\to V$ ein K-linearer Endomorphismus von f mit $\dim V < \infty$ und $c\in K$ ein Eigenwert. Dann ist der erste stabile Exponent von $f - c\cdot Id$ höchstens so groß wie die algebraische Vielfachheit von c, genauer,

$$\mathrm{stab}(f - c\cdot Id) \le \nu_f(c) - \mu_f(c) + 1.$$

Insbesondere ist der Hauptraum zum Eigenwert c gleich

$$V'_c(f) = \ker(f - c\cdot Id)^{\nu_f(c)}.$$

Beweis. Angenommen es ist

$$\nu_f(c) - \mu_f(c) + 1 < \mathrm{stab}(f - c\cdot Id).$$

Dann ist die Folge der Unterräume

(1) $\qquad \ker(f - c\cdot Id)^i$, $i = 1, ..., \nu_f(c) - \mu_f(c) + 2$.

echt aufsteigend. Insbesondere ist

(2) $\qquad \dim \ker(f - c\cdot Id)^i \ge i - j + \dim (f - c\cdot Id)^j$ für $j \le i \le \nu_f(c) - \mu_f(c) + 2$

(der Unterschied in den Dimensionen ist mindestens so groß wie der in den Indizes). Speziell für j = 1 ist

$$\dim (f - c\cdot Id)^j = \mu_f(c)$$

die Dimension des Eigenraums. Für $j = 1$ und $i = \nu_f(c) - \mu_f(c) + 2$ erhalten wir damit aus (2) die Ungleichung

(2) $\qquad \dim \ker(f - c\cdot Id)^i \ge \nu_f(c) + 1$ mit $i = \nu_f(c) - \mu_f(c) + 2$

Da der Kern (2) im Hauptraum $V'_c(f)$ liegt, folgt

$$\dim V'_c(f) \ge \nu_f(c)+1.$$

Dies steht im Widerspruch zu 5.3.6.

QED.

Bemerkung

Insbesondere gilt

$$V'_c(f) = \ker(f - c\cdot Id)^{\nu_f(c)-\mu_f(c)+1},$$

5.3.8 Hauptraumzerlegung

Sei $f:V\to V$ ein K-linearer Endomorphismus mit $\dim V < \infty$. Die Eigenwerte von f seien mit $c_1,...,c_r$ bezeichnet. Falls K hinreichend groß ist, ist die K-lineare Abbildung

$$\varphi: V'_{c_1}\oplus...\oplus V'_{c_r} \to V, (v_1,...,v_r) \mapsto v_1 + ... + v_r,$$

bijektiv.

Beweis. Nach 5.3.6 sind die Dimensionen der Haupträume gleich den algebraischen Vielfachheiten der Eigenwerte. Die Summe ihrer Dimensionen ist somit gleich dim V. Mit anderen Worten, die Abbildung bildet Räume gleicher Dimension ineinander ab. Es genügt also, ihre Injektivität zu beweisen.

Sei also

$$\varphi(v_1,...,v_r) = 0,$$

d.h.

(1) $$v_1+...+v_r = 0.$$

Wir haben zu zeigen, jeder einzelne Summand ist Null. Sei s die Anzahl der von Null verschiedenen Summanden.

$$s = \#\{i \mid v_i \ne 0\}$$

Wir führen den Beweis durch Induktion nach s. Im Fall s = 1 ist nichts zu beweisen: aus dem Verschwinden der Summe folgt das Verschwinden des einzigen Summanden. Sei jetzt s > 1 und sei $v_i \ne 0$. Wir betrachten die Abbilung

$$g := (f - c_i\cdot Id)^{\nu_i} \text{ mit } \nu_i := \nu_f(c_i).$$

Aus (1) folgt

$$g(v_1)+...+g(v_r) = 0.$$

Da alle Haupträume f-invariant sind, sind sie auch invariant bezüglich der Abbildung[27] g. Also ist

(2) $$(g(v_1),...,g(v_r)) \in V'_{c_1}\oplus...\oplus V'_{c_r}$$

Auf V'_{c_i} ist g die Nullabbildung, d.h. die Anzahl der von Null verschiedenen Koordinaten des Elements (2) ist kleiner als s. Nach Induktionsvoraussetzung folgt,

$$g(v_1) = ... = g(v_r) = 0.$$

Wegen (1) reicht es zu zeigen, die Vektoren $v_1,...,v_r$ sind Null mit eventueller Ausnahme von einem. Dazu reicht es zu zeigen,

$$g|_{V'_{c_j}} \text{ ist injektiv für alle } j \ne i.$$

Dazu wiederum reicht es, wenn wir zeigen,

(3) $$f - c_i\cdot Id\,|_{V'_{c_j}} \text{ ist injektiv für } j \ne i.$$

Sei $v\in V'_{c_j}$ ein Element aus dem Kern der Abbilung (3). Dann gilt

$$f(v) = c_i\cdot v,$$

also

$$(f - c_j\cdot Id)(v) = (c_i-c_j)\cdot v,$$

also

[27] Sogar bezüglich aller Polynome in f.

$$(f - c_j\cdot Id)^k(v) = (c_i - c_j)^k\cdot v \text{ für } k=1,2,3,...$$

Wegen $v\in V'_{c_j}$ ist die linke Seite dieser Identität Null für große k. Es gilt also

$$(c_i - c_j)^k\cdot v = 0 \text{ für ein } k\in \mathbb{N}.$$

Im Fall $i\neq j$ ist aber $c_i - c_j$ von Null verschieden und es folgt $v = 0$. Damit ist (3) bewiesen, und damit wiederum der Satz.

QED.

5.3.9 Jordansche Normalform eines Endomorphismus

Sei $f: V\to V$ ein K-linearer Endomorphismus mit $\dim V < \infty$. Falls K hinreichend groß ist, so gibt es eine Jordan-Basis von V bezüglich f, d.h. eine Basis, bezüglich welcher die Matrix von V eine direkte Summe von Jordan-Blöcken ist. Die Anzahl

$$\rho_k(f, c)$$

der direkten Summanden mit vorgegebenen Format $k\times k$ und Eigenwert c ist dabei für jede Wahl von k und c unabhängig von der Wahl der speziellen Basis.

Beweis. Seien $c_1,...,c_r$ die Eigenwerte von f. Weil K hinreichend groß ist, zerfällt V in eine direkte Summe der Haupträumen,

$$V = V_{c_1}\oplus...\oplus V_{c_r}.$$

Da die Haupträume V_{c_i} invariant bezüglich f sind, reicht es die Behauptung für die Einschränkungen von f auf die Haupträume zu beweisen, denn für eine Basis, die gerade Vereinigung von Basen der direkten Summanden ist, gilt

$$M(f) = M(f_1)\oplus...\oplus M(f_r)$$

Dabei bezeichne $f_i: V_{c_i}\to V_{c_i}$ die Einschränkung von f auf V_{c_i}. Beweisen wir also die Behauptung für jedes der f_i. Mit anderen Worten, beweisen wir die Behauptung für den Fall, daß f nur einen einzigen Eigenwert, sagen wir $c\in K$, besitzt. Dann gibt es nur einen Hauptraum und V fällt mit diesem Hauptraum (zum Eigenwert c) zusammen. Das bedeutet,

$$g = f - c\cdot Id$$

ist nilpotent. Dann gibt es aber eine Basis von V, bezüglich welcher die Matrix von g eine direkte Summe von Jordan-Blöcken zum Eigenwert 0 ist,

$$M(f) - M(c\cdot Id) = M(f - c\cdot Id) = M(g) = J_{a_1}(0) \oplus ... \oplus J_{a_s}(0).$$

Es folgt

$$M(f) = M(c\cdot Id) + M(g) = c\cdot Id + J_{a_1}(0) \oplus ... \oplus J_{a_s}(0) = J_{a_1}(c) \oplus ... \oplus J_{a_s}(c),$$

d.h. M(f) ist eine direkte Summe von Jordan-Blöcken zum Eigenwert c.

QED.

5.3.10 Jordansche Normalform einer Matrix

Sein $A\in K^{n\times n}$ eine n-reihige quadratische Matrix. Falls K groß genug ist, so gibt es eine umkehrbare n-reihige Matrix $B\in K^{n\times n}$ derart, daß

$$BAB^{-1}$$

eine direkte Summe von Jordan-Blöcken ist. Die Anzahl

$$\rho_k(A,c)$$

der direkten Summanden mit vorgegebenen Format $k\times k$ und Eigenwert c ist dabei für jede Wahl von k und c unabhängig von der Wahl der speziellen Basis.

Beweis. Das folgt unmittelbar aus 5.3.9.
QED.

5.4 Satz von Cayley-Hamilton

Sei $f:V\to V$ ein K-linearer Endomorphismus mit $\dim V < \infty$. Dann ist f Nullstelle seines charakteristischen Polynoms,

(1) $$\chi_f(f) = 0.$$

Analog ist jede quadratische Matrix $A\in K^{n\times n}$ Nullstelle ihres charakteristischen Polynoms,

(2) $$\chi_A(A) = 0.$$

Beweis. Offensichtlich sind die beiden Aussagen (1) und (2) äquivalent. Beim Beweis von (2) können wir bei Bedarf den Körper A vergrößern und zum Beispiel durch einen algebraisch abgeschlossenen Körper ersetzen. Wir können also annehmen, der Körper K ist groß genug. Der Raum V zerfällt dann in eine direkte Summe von Haupträumen,

$$V = V'_{c_1}\oplus...\oplus V'_{c_r},$$

und das charakteristische Polynom von f zerfällt in Linearfaktoren,

$$\chi_f(T) = (c_1\text{-}T)^{\nu_1}\cdot...\cdot(c_r\text{-}T)^{\nu_r}$$

Wir wollen zeigen, der Endomorphismus

(3) $$\chi_f(f) = (c_1\cdot Id -f)^{\nu_1}\cdot...\cdot(c_r\cdot Id -f)^{\nu_r}$$

ist identisch Null auf V. Dazu reicht es zu zeigen, er ist identisch Null auf jedem der Haupträume V'_{c_i}. Nun sind diese Haupträume invariant bei f also auch invariant bezüglich der Endomorphismen $c_i\cdot Id - f$ und aller ihrer Potenzen. Wenn wir zeigen wollen,

$$\chi_f(f)|_{V'_{c_i}} = 0,$$

so reicht es zu zeigen, einer der Faktoren

$$(c_j\cdot Id -f)^{\nu_j}$$

auf der rechten Seite von (3) ist identisch Null auf V'_{c_i}. Das ist aber der Fall für den Faktor mit j = i, denn es gilt

$$V'_{c_i} = \ker(f - c_i\cdot Id)^{\nu_i} = \ker (c_i\cdot Id - f)^{\nu_i}.$$

QED.
Aufgabe
Man finde den Fehler im folgenden “Beweis” des Satzes von Cayley-Hamilton. Wegen

$$\chi_A(T) = \det (A - T\cdot Id)$$

gilt

$$\chi_A(A) = \det (A - A\cdot Id) = \det (A - A) = \det (0) = 0.$$

6. Bilineare Abbildungen

6.1 Räume mit Skalarprodukt

6.1.1 Nicht-entartete symmetrische Bilinearformen

Sei V ein endlich-dimensionaler K-Vektorraum. Eine Bilinearform auf V ist eine Abbildung

$$b: V\times V \to K$$

mit folgenden Eigenschaften.

(i) Für jedes $v\in V$ ist die Abbildung

$$V \to K, x \mapsto b(x,v),$$

K-linear, d.h. b ist linear bezüglich der zweiten Variablen.

(ii) Für jedes $v\in V$ ist die Abbildung

$$V \to K, x \mapsto b(v,x),$$

K-linear, d.h. b ist linear bezüglich der zweiten Variablen.

Die Bilinearform b heißt symmetrisch, wenn gilt

$$b(v',v'') = b(v'',v')$$

für alle $v',v''\in V$. Die Bilinearform b heißt nicht-entartet, wenn es für jedes $v'\in V-\{0\}$ ein $v''\in V-\{0\}$ gibt mit

$$b(v',v'') \neq 0.$$

Eine Bilinearform heißt anisotrop , wenn für jeden von Null verschiedenen Vektor $v\in V$ gilt

$$b(v,v) \neq 0.$$

Im Fall $K = \mathbb{R}$ heißt eine Bilinearform b positiv definit (bzw. negativ definit) wenn für jeden von Null verschiedenen Vektor $v\in V$ gilt

$$b(v,v) > 0 \text{ (bzw. } b(v,v) < 0.$$

Eine nicht-entartete symmetrische Bilinearform heißt auch Skalarprodukt von V. Im Fall $K = \mathbb{R}$ heißt ein Skalarprodukt euklidisch, wenn es positiv definit ist.

Eine Paarung der K-Vektorräume V und W ist eine Abbildung

$$b: V \times W \to K, (v,w) \mapsto b(v,w),$$

welche für jedes feste v linear in w und für jedes fest w linear in v ist. Eine Paarung heißt nicht-entartet, wenn es für beliebige $v\in V-\{0\}$ und $w\in W-\{0\}$ Vektoren $v'\in V$ und $w'\in W$ gibt mit

$$b(v,w') \neq 0 \text{ und } b(v', w) \neq 0.$$

Bemerkungen

(i) Positiv bzw. negativ definite Bilinearforms sind offensichtlich insbesondere anisotrop. Definite Bilinearformen sind offensichtlich nicht-entartet. Eine Bilinearform über $\mathbb{R}$ ist genau dann anisotrop, wenn sie positiv oder negativ definiert ist (nach dem Zwischenwertsatz der Analysis).

(ii) Das Standard-Skalarprodukt des $\mathbb{R}^n$,

$$\left\langle \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1} x_i y_i ,$$

ist euklisch, denn <x,x> ist eine Summe von Quadraten reeller Zahlen, also stets >0 sobald $x \neq 0$ ist.

(iii) Wir werden sehen, die Bilinearform von (ii) ist über jedem Körper K ein Skalarprodukt (obwohl es für manche K nicht anisotrop ist).

(iv) Für jeden K-Vektorraum V ist die folgende Abbildung ein Beispiel für eine Paarung.

$$V\times V^* \to K,\ (v,\ell)\mapsto \ell(v).$$

Die Abbildung ist offensichtlich bilinear und nach dem Fortsetzungssatz (3.4.6) für lineare Abbildungen nicht entartet. Sie heißt natürliche Paarung des Vektorraums V.

(v) Sei V ein K-Vektorraum mit Skalarprodukt $<\ ,\ > : V \times V \to K$. Ein Vektor

$$v \in V - \{0\}$$

heißt isotrop, falls

$$< v, v > = 0$$

gilt. Der Vektorraum heißt isotrop, falls er einen isotropen Vektor enthält. Andernfalls heißt er anisotrop . Der Vektorraum heißt also genau dann anisotrop, wenn dessen Bilinearform anisotrop ist. Ein Unterraum $U \subseteq V$ heißt total isotrop, falls alle Vektoren von $U - \{0\}$ isotrop sind.

6.1.2 Die Bilinearform zu einer Matrix

Die bilineare Abbildung zu einer quadratischen Matrix

Sei $A = (a_{ij}) \in K^{n\times n}$ eine quadratische Matrix mit Einträgen aus K. Für je zwei Spaltenvektoren

$$x,y\in K^n = K^{n\times 1},\ x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$$

setzen wir

$$b(x,y) := x^TAy = \sum_{i,j=1}^{n} a_{ij}x_iy_j.$$

Die Abbildung

$$b: K^n\times K^n \to K,\ (x,y) \mapsto b(x,y) = \sum_{i,j=1}^{n} a_{ij}x_iy_j,$$

ist dann bilinear. Das folgt unmittelbar aus den Eigenschaften der Matrizenmultiplikation. Zum Beispiel ist

$$\begin{aligned} b(x,\ c'y'+c''y'') &= (x^TA)(c'y'+c''y'') \\ &= c'(x^TA)y'+c''(x^TA)y'' \\ &= c'b(x,\ y')+c''b(x,\ y''), \end{aligned}$$

d.h. b(x,y) ist linear in der zweiten Variablen y.

Symmetrie von b im Fall symmetrischer Matrizen A

Sei jetzt A eine symmetrische Matrix, d.h. es gelte

$$A = A^T.$$

Die Matrix

$$x^TAy$$

besitzt genau eine Zeile und genau eine Spalte. Sie verändert sich also nicht, wenn wir sie transponieren,

$$\begin{aligned} b(x,y) &= x^TAy \\ &= (x^TAy)^T \\ &= y^TA^Tx^{TT} \\ &= y^TAx \qquad \text{(A ist symmetrisch)} \\ &= b(y,x). \end{aligned}$$

Wir haben gezeigt, die Bilinearform b ist im Fall symmetrischer Matrizen symmetrisch.

Der Fall von Matrizen maximalen Rangs

Sei jetzt rk A = n. Wir wollen zeigen, daß dann die Bilinearform b nicht-entartet ist.

Beweis. Angenommen, b wäre entartet. Dann gibt es einen Spaltenvektor

$$y\in K^n - \{0\}$$

derart, daß für jeden Spaltenvektor $x\in K^n$

$$0 = b(y,x) = y^T Ax$$

gilt. Insbesondere gilt dies für $x=e_i$ (i=1,...,n),

$$(y^T A)e_i = 0$$

Nun ist $(y^T A)e_i$ gerade die i-te Spalte von $y^T A$. Mit anderen Worten, für jedes i ist die i-te Spalte von $y^T A$ gleich Null, d.h. $y^T A$ ist die Nullmatrix,

$$y^T A = 0.$$

Seien $a_1,...,a_n$ die Zeilen von A und $y_1,...,y_n$ die Koordinaten von y. Dann kann man die letzte Identität auch in der folgenden Gestalt schreiben,

$$y_1 a_1 + ... + y_n a_n = 0.$$

Da y vom Nullvektor verschieden sein soll, bedeutet letzteres, die Zeilen von A sind linear abhängig. Dann hat aber A nicht den Rang n im Widerspruch zur Annahme.
QED.

6.1.3 Die Matrix einer Bilinearform

Seien

$$b: V\times V \to K$$

eine Bilinearform und

$$v = (v_1,...,v_n)$$

eine Basis des Vektorraums V. Dann heißt die n×n-Matrix

$$M_v(b) := (b(v_i,v_j))_{i,j=1,..,n}$$

Matrix von b bezüglich der Basis v.

6.1.4 Verhalten bei Koordinatenwechsel

Seien

$$b: V\times V \to K$$

eine Bilinearform und

$$v = (v_1,...,v_n)$$

eine Basis des Vektorraums V. Dann gilt

(i) Identifiziert man mit Hilfe der gegebenen Basis den Vektorraum V mit K^n so bekommt die Bilinearform b die Gestalt

$$b(v', v'') = v'^T A v''$$

mit $A= M_v(b)$, $v' = \sum_{i=1}^{n} x'_i v_i = \begin{pmatrix} x'_1 \\ ... \\ x'_n \end{pmatrix}$ und $v'' = \sum_{i=1}^{n} x''_i v_i = \begin{pmatrix} x''_1 \\ ... \\ x''_n \end{pmatrix}$.

(ii) Ist $v'_1,...,v'_n \in V$ eine zweite Basis und bezeichnet $B := M^v_{v'}(Id)$ die Basiswechselmaltrix, für den Übergang von v nach v' so gilt

$$M_v(b) = B^T \cdot M_{v'}(b) \cdot B$$

Beweis.
Wir setzen

$$A := M_v(b) = (a_{ij})$$

$$A' := M_{v'}(b) = (a'_{ij})$$
$$B := (b_{ij}).$$

Dann gilt

$$v_i = Id(v_i) = \sum_{\alpha=1}^{n} b_{\alpha i} v'_\alpha$$

also

$$a_{ij} = b(v_i, v_j) = b\Big(\sum_{\alpha=1}^{n} b_{\alpha i} v'_\alpha, \sum_{\beta=1}^{n} b_{\beta j} v'_\beta\Big) = \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} b_{\alpha i}\, b(v'_\alpha, v'_\beta) b_{\beta j}$$

$$= \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} b_{\alpha i} a'_{\alpha\beta} b_{\beta j}$$

d.h. es ist $A = B^T A' B$.
QED.

6.1.3 Beispiele, das Standard-Skalarprodukt

Beispiel 1
Im Fall des obigen Beispiels ist die zur Matrix A = Id gehörige Bilinearform

$$b(x,y) = x^T \cdot Id \cdot y = x^T \cdot y$$

ein Skalarprodukt. Dieses heißt Standard-Skalarprodukt und wird mit

$$\langle x,y\rangle := x^T \cdot y$$

bezeichnet.

Beispiel 2
Der komplexe Vektorraum $\mathbb{C}^n$ mit dem Standard-Skalarprodukt ist isotrop. Zum Beispiel gilt für n = 2,

$$\langle (1,i), (1,i)\rangle = 1^2 - i^2 = 0.$$

Beispiel 3
Durch

$$\langle (x_1,x_2), (y_1,y_2)\rangle := x_1 y_1 - x_1 y_2$$

ist für beliebiges K ein Skalarprodukt definiert: es gehört zur symmetrischen Matrix maximalen Rangs

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

Mit diesem Skalarprodukt ist selbst der reelle Vektorraum $\mathbb{R}^2$ isotrop.

Beispiel 4
Das Skalarprodukt der Relativitätstheorie, das zur Matrix

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

gehört ist selbst für $K=\mathbb{R}$ nicht definit. Die drei positiven Eigenwerte der Matrix entsprechen dabei den drei Raumrichtungen und der negative Eigenwert steht für die Zeit.

6.1.4 Kriterium für nicht-entartete Bilinearformen

Seien V ein endlich-dimensionaler K-Vektorraum,

$$b: V\times V \to K$$

eine Bilinearform und $e = (e_1,...,e_n)$ eine Basis von V. Dann sind folgende Aussagen äquivalent.

(i) b ist nicht entartet.

(ii) Die lineare Abbildung

$$V \to V^*, v \mapsto b_v \text{ mit } b_v(x) := b(v,x)),$$

in das Dual von V ist injektiv (also bijektiv).

(iii) Die lineare Abbildung

$$V \to V^*, v \mapsto b^v \text{ mit } b^v(x) := b(x,v),$$

in das Dual von V ist injektiv (also bijektiv).

(iv) Die Matrix von b bezüglich der Basis e hat maximalen Rang n.

Beweis. (i) ⇔ (ii). Bedingung (ii) besagt gerade, für jedes $v' \in V-\{0\}$ ist $b_{v'}$ nicht die Null-Abbildung, d.h. es gibt ein $v'' \in V$ mit

$$b(v',v'') \neq 0.$$

Das bedeutet aber nach Definition gerade, daß b nicht entartet ist.

(iii) ⇔ (iv). Sei

$$A := (b(e_i,e_j))$$

die Matrix von b bezüglich der gegebenen Basis. Wir führen für die Vektoren $v,x \in V$ Koordinaten bezüglich dieser Basis ein,

$$x = \sum_{i=1}^{n} x_i e_i,\ v = \sum_{i=1}^{n} v_i e_i .$$

Dann gilt

$$b^v(x) = b(x,v) = b(\sum_{i=1}^{n} x_i e_i , \sum_{j=1}^{n} v_j e_j)$$

$$= \sum_{i,j=1}^{n} x_i b(e_i,e_j) v_j$$

$$= (x_1,...,x_n)\cdot A\cdot \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix}$$

Für festes v ist b^v genau dann die Null-Abbildung, wenn der letzte Ausdruck Null ist für alle $(x_1,...,x_n)$, d.h. wenn

$$A\cdot \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix}$$

die Null-Matrix ist. Bedingung (iii) ist also gleichbedeutend mit der Implikation

$$A\cdot \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} v_1 \\ ... \\ v_n \end{pmatrix} = 0.$$

Letzteres besagt aber gerade, die Spalten von A sind linear unabhängig. Damit gilt

$$\text{(iii)} \Leftrightarrow \mathrm{rk}(A) = n,$$

d.h. (iii) ist äquivalent zu (iv).

(ii) ⇔ (iv). Wird in derselben Weise wie die Äquivalentz von (iii) und (iv) bewiesen. (Man ersetze im obigen Beweis überall A durch die transponierte Matrix).

QED.

Bemerkungen

(i) Nach dem gerade bewiesenen Ergebnis kann man eine nicht-entartete Bilinearform
$$b: V\times V \to K$$
benutzen, um den Raum V mit seinem Dual V* zu identifizieren (indem man den Vektor v mit der Linearform b^v: $x \mapsto b(x,v)$ identifiziert).

(ii) Die natürliche Paarung
$$V\times V^* \to K, (v,\ell) \mapsto \ell(v),$$
wird durch diese Identifikation mit der folgenden Abbildung identifiziert
$$V\times V \to K, (v', v'') \mapsto (v', b^{v''}) \mapsto b(v',v'').$$

(iii) Mit anderen Worten, die Identifikation von (i) ist gerade so beschaffen, daß die natürliche Paarung identisch wird mit der gegebenen Bilinearform.

(iv) Für die natürliche Paarung verwendet man oft dieselbe Bezeichnung wie für das Standard-Skalarprodukt,
$$<v, \ell > := \ell(v).$$

(v) Unter Verwendung dieser Bezeichnung nimmt die Definition der in 3.4.7 definierten dualen Abbildung eine besonders elegante Form an. Es gilt nämlich für lineare Abbildungen $f:V\to W$ und Vektoren $v\in V$, $w^*\in W^*$:
$$<v,f^*(w^*)> = <v, w^*\circ f> = (w^*\circ f)(v) = w^*(f(v)) = <f(v),w^*>.$$
Mit anderen Worten, die zu $f:V \to W$ duale Abbildung ist diejenige Abbildung $f^*:W^* \to V^*$, für welche

(1) $$<v, f^*(w^*)> = <f(v), w^*> \text{ für } v\in V \text{ und } w^*\in W^*$$

gilt.

(vi) Sind in (v) die Räume V und W mit einem Skalarprodukt versehen, so kann man aus der Relation (1) die dualen Räume vollständig entfernen (indem man sie mit den Ausgangsräumen identifiziert. Genau dies tun wir in der nachfolgenden Definition.

6.1.5 Die adjungierte lineare Abbildung

Seien V und W endlich-dimensionale K-Vektorräume mit Skalarprodukt, d.h. beide Vektorräume seien jeweils mit einem Skalarprodukt versehen,
$$< , >_V: V\times V \to K$$
$$< , >_W: W\times W \to K$$
Dann gibt es zu jeder lineare Abbildung $f: V\to W$ genau eine lineare Abbildung $f^*:W\to V$ mit
$$<v,f^*w>_V = <fv, w>_W$$
für alle $v\in V$ und alle $w\in W$. Diese Abbildung heißt die zu f (bezüglich der gegebenen Skalarprodukte) adjungierte Abbildung.

Beweis. Existenz von f*. Nach Bemerkung 6.1.4(v) gilt
$$<v, f^*(w^*)> = <f(v), w^*> \text{ für } v\in V \text{ und } w^*\in W^*,$$
wenn < , > die natürliche Paarung (auf $V\times V^*$ bzw. $W\times W^*$) bezeichnet und $f^*:W^*\to V^*$ die duale Abbildung. Wir benutzen die gegebenen Skalarprodukte $< , >_V$ bzw. $< , >_W$ um V mit V* und W mit W* zu identifizieren. Dann wird $f^*:W^*\to V^*$ zu einer Abbildung $f':W\to V$ für welche gilt
$$<v, f'(w)>_V = <f(v), w>_W$$
für alle $v\in V$ und alle $w\in W$, d.h. f' ist eine Abbildung der gesuchten Art.

Eindeutigkeit von f*. Wir nehmen an, es gibt zwei lineare Abbilungen $g,h:W\to V$ mit der angegebenen Eigenschaft, d.h. mit
$$<v,g(w)> = <f(v), w> = <v, h(w)>$$
für alle $v\in V$ und alle $w\in W$ und zeigen,daß dann g=h gilt.
Es gilt
$$<v, (g-h)(w)> = <v,g(w)> - <v, h(w)> = 0$$
für alle $v\in V$ und alle $w\in W$. Es reicht also zu zeigen, aus

(1) $$\langle v, g(w)\rangle_V = 0$$
für alle $v\in V$ folgt
(2) $$g(w) = 0.$$
Nun ist g(w) ein Element von V und das Bild von g(w) bei der durch $\langle\ ,\ \rangle_V$ definierten Injektion
(3) $$V \to V^*,\ x \mapsto (y \mapsto \langle y,x\rangle_V)$$
ist die Abbildung
$$V\to K,\ y \mapsto \langle y,g(w)\rangle_V.$$
Nach (1) ist dies die Nullabbildung. Da aber (3) injektiv ist (denn $\langle\ ,\ \rangle_V$ ist als Skalerprodukt nicht entartet), muß bereits g(w) = 0 sein. Es gilt also tatsächlich (2).
QED.

6.1.6 Beispiel: adjungierte Abbildungen und transponierte Matrizen

Seien $A\in K^{m\times n}$ eine Matrix und
$$f_A: K^n \to K^m,\ x \mapsto Ax,$$
die zugehörige lineare Abbildung. Die Räume K^n und K^m seien mit dem Standard-Skalarprodukt versehen. Weiter sei
$$f_{A^T}: K^m \to K^n,\ y \mapsto A^Ty,$$
die zur transponierten Matrix gehörige lineare Abbildung. Dann gilt für die Standard-Skalarprodukte auf K^m bzw. K^n,
$$\langle v, f_{A^T}(w)\rangle_V = \langle f_A(v), w\rangle_W$$
für beliebige $v\in K^n$,$w\in K^m$. Insbesondere sind die linearen Abbildungen f_A und f_{A^T} adjungiert zueinander.
Beweis. Es gilt
$$\langle v, f_{A^T}(w)\rangle_V = v^T\cdot A^T\cdot w = (Av)^T\cdot w = \langle Av, w\rangle_W = \langle f_A(v), w\rangle_W$$
QED.

6.1.7 Selbstadjungierte Operatoren

Seien V ein endlich-dimensionaler K-Vektorraum, der mit einem Skalarprodukt $\langle\ ,\ \rangle$ versehen ist. Ein selbstadjungierter Operator ist dann eine lineare Abbildung $f:V\to V$ mit
$$\langle v, f(w)\rangle = \langle f(v), w\rangle$$
für alle $v\in V$ und alle $w\in W$.
Beispiel
Sei $A\in K^{n\times n}$ eine symmetrische Matrix. Dann ist die zu A gehörige lineare Abbildung $f_A:K^n \to K^n$ selbstadjungiert bezüglich des Standard-Skalarprodukts von K^n.

6.1.8 Selbstadjungierte Operatoren und invariante Unterräume

Seien V ein K-Vektorraum mit anisotropen Skalarprodukt $\langle\ ,\ \rangle$ (und mit $\dim V < \infty$),
$$f: V\to V$$
ein selbstadjungierter Operator und
$$V' \subseteq V$$
ein f-invarianter Unterraum. Dann gilt:
(i) Die Einschränkung des Skalarprodukts $\langle\ ,\ \rangle$ auf V' ist ein (anisotropes) Skalarprodukt.

(ii) Die Einschränkung

$$f|_{V'}$$

von f auf V' selbstadjungiert.

Beweis. Zu (i). Die Abbildung

$$V'\times V' \to K,\ (a,b) \mapsto \langle a,b\rangle,$$

ist offensichtlich bilinear und symmetrisch. Nach Voraussetzung ist sie auch anisotrop, also nicht-entartet, also ein Skalarprodukt.

Zu (ii). Es gilt

$$\langle f(v), w\rangle = \langle v, f(w)\rangle$$

für $v,w\in V'$, denn dies gilt sogar für beliebige $v,w\in V$.

QED.

Bemerkung

Die Einschränkung eines Skalarprodukts auf einen Unterraum ist im allgemeinen kein Skalarprodukt. Zum Beispiel ist die Einschränkung des Skalarprodukt

$$\left\langle \begin{pmatrix} x' \\ x'' \end{pmatrix}, \begin{pmatrix} y' \\ y'' \end{pmatrix} \right\rangle = x'y'' + x''y'$$

mit der Matrix

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

auf den Unterraum $\begin{pmatrix} 1 \\ 0 \end{pmatrix}\cdot K$ identisch Null.

6.1.9 Diagonalisierbarkeit selbstadjungierter Operatoren im anisotropen Fall

Seien V ein endlich-dimensionaler K-Vektorraum mit anisotropen Skalarprodukt $\langle,\rangle$ und $f:V\to V$ ein selbstadjungierter K-linearer Endomorphismus. Falls K genügend groß ist[28], besitzt V eine Eigenbasis bezüglich f (und die Haupträume von f stimmen mit den Eigenräumen von f überein).

Beweis. Sei

(1) $$V = V_1\oplus\ldots\oplus V_r$$

die zu f gehörige Hauptraumzerlegung, wobei V_i der Hauptraum zum Eigenwert c_i sei,

$$V_i := \ker(f-c_i\cdot Id)^n \text{ für n groß.}$$

Die Zerlegung (1) ist eine Zerlegung in f-invariante Unterräume. Weil $\langle\ ,\ \rangle$ anisotrop sein soll, sind die direkten Summanden wieder Räume mit Skalarprodukt und die Einschränkungen von f auf die direkten Summanden selbstadjungierte Operatoren.

Es reicht zu zeigen, für jedes i besitzt die Einschränkung von f auf V_i eine Eigenbasis.

Wir können also annehmen, V ist bereits selbst schon (der einzige) Hauptraum von f. Sei $c\in K$ der einzige Eigenwert. Es reicht zu zeigen,

$$f = c\cdot Id$$

(d.h. der erste stabile Index ist 1). Nach Definition des Hauptraumbegriffs ist $f - c\cdot Id$ ein nilpotenter (selbstadjungierter) Endomorphismus. Es reicht also, die nachfolgende Aussage zu beweisen.

QED.

6.1.10 Lemma

Seien V ein endlich-dimensionaler K-Vektorraum mit anisotropen Skalarprodukt $\langle,\rangle$ und $f:V\to V$ ein nilpotenter selbstadjungierter K-linearer Endomorphismus. Dann gilt

$$f = 0.$$

Beweis. Nach Voraussetzung gilt

$$f^n = 0 \text{ für ein } n\in\mathbb{N}.$$

[28] d.h. das charakteristische Polynom von f zerfalle über K in Linearfaktoren.

Im Fall n = 1 ist die Behauptung trivial. Im Fall n = 2 gilt für jedes $v\in V$

$$<f(v), f(v)> = <v, f^2(v)> = <v, 0> = 0,$$

Wegen < , > anisotrop folgt f(v) = 0. Sei jetzt n >2. Es reicht zu zeigen,

$$f^{n-1} = 0.$$

Nun ist f^{n-1} selbstadjungiert und wegen n>2 gilt

$$2(n-1) = n + (n-2) > n,$$

also

$$(f^{n-1})^2 = 0.$$

Auf Grund es eben behandelten Falles n = 2 ist mit $f^n = 0$ auch $f^{n-1} = 0$.
QED.

6.1.11 Beispiel (der isotrope Fall)

Sei $V = K^2$ der Raum mit dem Skalarprodukt

$$<\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix}> = xy' + yx'$$

Die Matrix dieser Bilinearform bezüglich der Standardbasis ist

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

d.h. die Bilinearform ist nicht entartet. Sei f die lineare Abbildung

$$f:V\to V, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ 0 \end{pmatrix},$$

mit der Matrix $M(f) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ der Standardbasis. Dann ist M(f) nicht diagonalisierbar, d.h. V besitzt keine Eigenbasis bezüglich f.
Es gilt jedoch

$$<f\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix}> = <\begin{pmatrix} y \\ 0 \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix}> = yy'$$

und

$$<\begin{pmatrix} x \\ y \end{pmatrix}, f\begin{pmatrix} x' \\ y' \end{pmatrix}> = <\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} y' \\ 0 \end{pmatrix}> = yy',$$

d.h. f ist selbstadjungiert.

6.1.12 Orthogonale Zerlegungen

Seien V ein K-Vektorraum mit Skalarprodukt und V',V" zwei K-lineare Unterräume mit

$$V =^{29} V'\oplus V''.$$

Diese Zerlegung in direkte Summanden heißt orthogonal, wenn gilt

$$<v',v''> = 0$$

für beliebige $v'\in V'$ und $v''\in V''$.

6.1.13 Orthogonalität der Hauptraumzerlegung selbstadjungierter Operatoren

Sei f: V → V ein selbstadjungierter Operator (auf einem endlich-dimensionalen Raum mit anisotropen Skalarprodukt). Dann ist die Hauptraumzerlegung

$$V = V_1\oplus...\oplus V_r$$

eine orthogonale Zerlegung.
Beweis. Bezeichne c_i den Eigenwert zum Hauptraum V_i. Dann gilt für $i\neq j$, $v_i\in V_i$, $v_j\in V_j$:

[29] d.h. die lineare Abbildung $V'\oplus V'' \to V$, $(v', v'')\mapsto v'+v''$, ist ein Isomorphismus.

$$c_i\langle v_i,v_j\rangle = \langle fv_i,v_j\rangle = \langle v_i,fv_j\rangle = c_j\langle v_i,v_j\rangle$$

und wegen $c_i \neq c_j$ folgt $\langle v_i,v_j\rangle = 0$.

QED.

6.1.14 Orthonormierte Basen

Eine Basis $v_1,...,v_n$ des K-Vektorraums V mit Skalarprodukt < , > heißt orthogonal, wenn

$$\langle v_i , v_j \rangle = 0 \text{ für beliebige } i,j \text{ mit } i\neq j.$$

Sie heißt orthonormiert, wenn außerdem noch

$$\langle v_i, v_i\rangle = 1 \text{ gilt für alle } i = 1, ... , n.$$

6.1.15 Existenz von orthonormierten Basen

Sei V ein K-Vektorraum mit anisotropen Skalarprodukt (und dim V < ∞). Dann gilt

(i) V besitzt eine orthogonale Basis.

(ii) Falls K hinreichend groß ist, besitzt V eine orthonormierte Basis.

Bemerkung

'Hinreichend groß' soll hier bedeuten, daß man im Körper K aus den Skalarprodukten der Gestalt $\langle v,v\rangle$ mit $v\in V$ die Quadratwurzel ziehen kann.

Beweis. Zu (i). Wir wählen eine Basis $v_1,...,v_n$ V und beschreiben ein Verfahren, durch welches diese Basis in endlich vielen Schritten in eine orthogonale Basis überführt wird. Angenommen, die ersten k Vektoren der Basis sind bereits orthogonal,

$$\langle v_i , v_j \rangle = 0 \text{ für } i,j=1,...,k,\ i\neq j.$$

Es reicht, den Vektor v_{i+1} so durch einen Vektor w zu ersetzen, der orthogonal zu den ersten i Vektoren der Basis ist und zusammen mit den Vektoren $v_1,...,v_i, v_{i+2},...,v_n$ eine Basis von V bildet. Wir setzen

(1) $$w = \lambda_1 v_1 +...+ \lambda_i v_i + v_{i+1} .$$

Dabei seien die Koeffizienten λ_j so gewählt, daß gilt

$$\begin{aligned} 0 &= \langle w, v_j \rangle \\ &= \langle \lambda_1 v_1 +...+ \lambda_i v_i + v_{i+1} , v_j \rangle \\ &= \lambda_j \langle v_j, v_j\rangle + \langle v_{i+1}, v_j\rangle \end{aligned}$$

für j = 1, ... , i, d.h. es sei

$$\lambda_j = - \langle v_{i+1}, v_j\rangle/\langle v_j, v_j\rangle.$$

Man beachte, es ist $\langle v_j, v_j\rangle \neq 0$, weil das Skalarprodukt anisotrop sein soll. Wegen (1) erzeugt w zusammen mit den Vektoren $v_1,...,v_i, v_{i+2},...,v_n$ den Raum V, d.h. indem wir v_{i+1} durch w ersetzen erhalten wir wieder eine Basis. Indem wir die obige Konstruktion endlich oft wiederholen, erhalten wir eine orthogonale Basis.

Zu (ii). Zum Beweis genügt es, aus einer orthogonalen Basis $v_1,...,v_n$ eine orthonormierte $v'_1,...,v'_n$ Basis zu konstruieren. Da K hinreichend groß sein soll, können wir setzen

$$v'_i := \frac{1}{\sqrt{\langle v_i,v_i\rangle}} \cdot v_i \text{ für } i = 1, ... , n.$$

Es gilt dann

$$\langle v'_i, v'_i\rangle = \frac{1}{\langle v_i,v_i\rangle}\langle v_i, v_i\rangle = 1$$

für jedes i.
QED.

6.1.16 Existenz von orthonormierten Eigenbasen
Sei f: $V\to V$ ein selbstadjungierter Operator auf dem endlich-dimensionalem K-Vektorraum V mit anisotropen Skalarprodukt < , >. Der Körper K sei groß genug[30]. Dann besitzt f eineine orthonorierte Eigenbasis.
Beweis. Sei
$$V = V_1 \oplus ... \oplus V_r$$
die Hauptraum-Zerlegung. Da diese Zerlegung orthogonal ist, reicht es zu zeigen, jedes V_i besitzt eine orthonormierte Basis aus Eigenvektoren. Auf V_i hat aber f die Gestalt
$$f = c_i \cdot Id$$
(wobei c_i den i-ten Eigenwert von f bezeichnet), d.h. jede Basis von V_i ist Eigenbasis. Es genügt also eine beliebige orthonormierte Basis von V_i zu wählen.
QED.

6.1.17 Beispiel
$A := \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ und $f:=f_A: \mathbb{R}^3 \to \mathbb{R}^3$.

Bestimmung der Eigenwerte:
Es gilt
$$\chi_f(T) = \det \begin{pmatrix} 2\text{-}T & -1 & 1 \\ -1 & 2\text{-}T & 1 \\ 1 & 1 & 2\text{-}T \end{pmatrix} = (2-T)^3 - 2 + 3(T-2)$$
$$= - T^3 + 6T^2 - 12T + 8 - 2 + 3T - 6$$
$$= - T^3 + 6T^2 - 9T$$
$$= (-T)(T^2 - 6T + 9)$$
$$= (-T)(T-3)^2$$
Als Eigenwerte erhalten wir also $c_1 = 0$ (Vielfachheit 1) und $c_2=1$ (Vielfachheit 2).

Bestimmung einer orthonormierten Eigenbasis:
Eigenraum zum Eigenwert 0 = Lösungsmenge des homogenen Systems zu
$$\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{pmatrix} \quad \text{(2×untere Zeile zur oberen)}$$
$$\sim \begin{pmatrix} 0 & 3 & 3 \\ -1 & 2 & 1 \end{pmatrix}$$
$$\sim \begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & -1 \end{pmatrix} \quad \text{(2× obere Zeile zur unteren)}$$
$$\sim \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$
d.h. y+z = 0 und x + z = 0. Der Eigenraum besteht aus den Vektoren der Gestalt
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -z \\ -z \\ z \end{pmatrix} = -z\cdot\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$
Orthonormierte Basis des Eigenraums:

---
[30] im Sinne von 5.1.5 <u>und</u> 6.1.15.

$$\mathbf{x}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\-1\end{pmatrix}$$

Eigenraum zum Eigenwert 3 = Lösungsmenge des homogenen Systems zu

$(-1\ -1\ 1)$

d.h. x + y - z = 0. Der Eigenraum besteht aus den Vektoren der Gestalt

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x\\y\\x+y\end{pmatrix} = x\cdot\begin{pmatrix}1\\0\\1\end{pmatrix} + y\cdot\begin{pmatrix}0\\1\\1\end{pmatrix}.$$

Eine orthogonalisierte Basis des Eigenraums ist

$$\mathbf{x}'_2 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \mathbf{x}'_3 = \begin{pmatrix}0\\1\\1\end{pmatrix} + \lambda\begin{pmatrix}1\\0\\1\end{pmatrix}$$

mit $0 = <\mathbf{x}'_2 , \mathbf{x}'_3> = 1 + \lambda\cdot 2$, d.h. $\lambda = -\frac{1}{2}$ und $\mathbf{x}'_3 = \begin{pmatrix}-1/2\\1\\1/2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}-1\\2\\1\end{pmatrix}$.

Eine orthonormale Basis des Eigenraums ist

$$\mathbf{x}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}, \mathbf{x}_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\2\\1\end{pmatrix}$$

Zusammen bilden die Vektoren $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ eine orthonormierte Eigenbasis des $\mathbb{R}^3$

Bemerkungen

(i) Sei V = K^n mit den Standard-Skalarprodukt versehen. Die Standard-Basis ist dann eine orthonormierte Basis. Um die Aussage 6.1.16 in der Sprache der Matrizen zu beschreiben brauchen wir eine Charakterisierung der Matrizen, die die Standard-Basis in eine orthonormierte Basis überführen, oder allgemeiner, die orthonormierte Basen ineinander überführen. Dieses sind gerade die sogenannten orthogonalen Matrizen. Zuvor behandeln wir jedoch erst den Fall der komplexen Vektorräume.

(ii) Über $\mathbb{C}$ ist das Standard-Skalarprodukt isotrop. Im komplexen Fall muß man deshalb den Begriff des Skalarprodukts modifizieren. Als Ersatz für das Standard-Skalarprodukt werden wir das hermitische Standard-Skalarprodukt einführen: es ist zwar kein Skalarprodukt mehr, dafür jedoch positiv definit.

6.2 Hermitische Skalarprodukte

6.2.1 Hermitische Formen

Sei V ein $\mathbb{C}$-Vektorraum. Eine hermitische Form auf V ist eine Abbildung

$$h : V\times V \to \mathbb{C}, (x,y) \mapsto < x, y>,$$

mit folgenden Eigenschaften.

1. Die Form ist $\mathbb{C}$-linear bezüglich der ersten Variablen, d.h. für jedes $v\in V$ ist die Abbildung

$$V \to \mathbb{C}, x \mapsto h(x, v),$$

eine $\mathbb{C}$-lineare Abbildung.

2. Die Abbildung ist hermitisch, d.h. für je zwei Vektoren $x,y\in V$ gilt

$$h(y, x) = \overline{h(x, y)},$$

d.h. beim Vertauschen der beiden Argumente wird der Wert konjugiert.

Ist $v_1,...,v_n$ eine Basis des Vektorraums V so nennt man die Matrix

$$M_v(h) := (h(v_i,v_j))_{i,j=1,...,n}$$

auch Matrix von h bezüglich der gegebenen Basis.

Bemerkungen

(i) Diese beiden Eigenschaften implizieren eine Eigenschaft der Abbildung bezüglich des zweiten Arguments, welche der Linearität sehr nahe kommt. für $v', v'' \in V$ und $c',c'' \in \mathbb{C}$ gilt nämlich

$$\begin{aligned} h(x, c'v' + c''v'') &= \overline{h(c'v' + c''v'',x)} \\ &= \overline{c'h(v',x) + c''h(v'',x)} \\ &= \overline{c'}\cdot\overline{h(v', x)} + \overline{c''}\cdot\overline{h(v'', x)} \\ &= \overline{c'}\cdot h(x,v') + \overline{c''}\cdot h(x,v''). \end{aligned}$$

Mit anderen Worten, bezüglich des zweiten Arguments hat man in Linearkombinationen die Koeffizienten zusätzlich zu konjugieren. Man nennt Abbildungen mit dieser Eigenschaft auch anti-linear Hermitische Formen sind also antilinear bezüglich des zweiten Arguments.

(ii) Aus der zweiten Bedingung folgt, daß für jeden Vektor $x \in V$ der Wert $h(x,x)$ mit seinem konjugiert komplexen übereinstimmt, d.h. es ist

$$h(x,x) \in \mathbb{R}$$

eine reelle Zahl.

Beispiel

Sei A eine $n\times n$-Matrix mit komplexen Einträgen. Dann ist durch

$$h(x,y) := x^T A\overline{y}$$

genau dann eine hermitische Form definiert, wenn

$$A^T = \overline{A} \tag{1}$$

gilt. Dabei bezeichne $\overline{A}$ die Matrix, die man aus A erhält, indem man alle Einträge durch deren komplexe Konjugiertes ersetzt. Man beachte,

$$h(y, x) - \overline{h(x, y)} = y^T A\overline{x} - \overline{x}^T\overline{A}y = y^T A\overline{x} - y^T\overline{A}^T\overline{x} = y^T (A - \overline{A}^T)\overline{x}$$

ist genau dann identisch Null, wenn die Matrix $A - \overline{A}^T$ identisch Null ist. Matrizen, die der Bedingung (1) genügen, heißen hermitische Matrizen. Insbesondere ist die Einheitsmatrix eine hermitische Matrix. Insbesondere gilt für A = Id,

$$h(x,x) := x^T\overline{x} = \sum_{i=1}^{n} x_i\overline{x}_i = \sum_{i=1}^{n} |x|_i^2$$

wenn man die Koordinaten des Vektors x mit x_i bezeichnet.

6.2.2 Hermitische Skalarprodukte

Sei V ein komplexer Vektorraum. Eine hermitische Form

$$h: V \times V \to \mathbb{C}$$

heißt positiv definit, wenn

$$h(x,x) > 0$$

gilt für alle von Null verschiedenen Vektoren $x \in V$. Eine positiv definite hermitische Form heißt auch hermitisches Skalarprodukt.

Bemerkungen

(i) Hermitische Skalarprodukte sind eigentlich keine Skalarprodukte im von uns definiertem Sinne. Sie stellen eine spezielle Anpassung des Skalarprodukt-Begriffs an den Fall der komplexen Zahlen dar. Für die komplexe Analysis ist deshalb der Begriff der hermitischen Form bedeutsamer als der des Skalarprodukts im hier verwendeten Sinne.

(ii) Der Begriff des hermitischen Skalarprodukt läßt sich auf den Fall beliebiger Körper verallgemeinern (zum Begriff der sesqui-linearen Form, der $1\frac{1}{2}$ - fach linearen Form).

(iii) Wir werden zeigen, für hermitsche Skalarprodukte sind die sind von uns betrachteten Phänomene im wesentlichen dieselben wie für gewöhnliche Skalarprodukte über den reellen Zahlen.

(iv) Hermitische Skalarprodukte werden wir meistens mit < , > bezeichnen. Zwei Vektoren x,y heißen orthogonal bezüglich eines gegebenen hermitischen Skalarproduktes, wenn gilt

$$\langle x, y \rangle = 0.$$

Die Zahl

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

heißt auch Länge des Vektors x. Man beachte, $\|x\|$ ist eine wohldefinierte reelle Zahl, da $\langle x,x\rangle$ stets nicht-negativ ist.

(vi) Wie im Fall gewöhnlicher Skalarprodukte definiert man für jedes Teilmenge $M\subseteq V$ eines komplexen Vektorraums mit Skalarprodukt deren orthogonales Komplement

$$M^{\perp} := \{v\in V \mid \langle v,m\rangle = 0 \text{ für alle } m\in M\}$$

Beispiel

Die hermitische Form zur Einheitsmatrix ist ein hermitisches Skalarprodukt. Es heißt auch hermitisches Standard-Skalarprodukt des $\mathbb{C}^n$

6.2.3 Orthonormalisierung

Sei V ein endlich-dimensionaler komplexer Vektorraum mit hermitischen Skalarprodukt. Dann gibt es eine Basis $v_1,...,v_n$ von V mit

$$\langle v_i,v_j\rangle = \delta_{ij}$$

für i,j = 1,...,n. Eine solche Basis heißt orthonormiert.

Beweis. Man wende das Orthogonormalisierungsverfahren von 6.1.15 auf irgendeine Basis von V an.

QED.

Bemerkung: Orthonormalität und Standard-Skalarprodukt

Ist $v_1,...,v_n$ eine orthonormierte Basis, so läßt sich das Skalarprodukt von $x = \sum_{i=1}^{n} x_i v_i$ und $y = \sum_{i=1}^{n} y_i v_i$ wie folgt schreiben.

$$\langle x,y\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j \langle v_i,v_j\rangle = \sum_{i=1}^{n} x_i\bar{y}_i = x^T\bar{y},$$

d.h. $\langle x,y\rangle$ läßt sich mit dem Standard-Skalarprodukt der zugehörigen Koordinatenvektoren identifizieren.

6.2.4 Der adjungierte Operator, Selbstadjungiertheit

Seien V und W ein endlich-dimensionale komplexe Vektorräume mit hermitischen Skalarprodukt < , > und

$$f: V \to W$$

eine $\mathbb{C}$-lineare Abbildung. Dann gibt es genau eine $\mathbb{C}$-lineare Abbildung $f^*: W \to V$ mit

$$\langle f(x), y\rangle = \langle x, f^*(y)\rangle$$

für alle $x,y\in V$. Diese Abildung heißt zu f gehöriger adjungierter Operator (bezüglich des gegebenen hermitischen Skalarprodukts). Die Abbildung f heißt selbstadjungiert, falls sie mit ihrem adjungierten Operator übereinstimmt,

$$\langle f(x), y \rangle = \langle x, f(y)\rangle$$

für alle $x,y\in V$ (und insbesondere ist V = W).

Beweis. Zum Beweis können wir V und W durch irgendwelche zu V bzw. W isomorphe Vektorräume ersetzen. Insbesondere können wir annehmen, es ist

$$V = \mathbb{C}^m \text{ und } W = \mathbb{C}^n$$

und die Skalarprodukte sind gerade die Standard-Skalarprodukte. Das hat den Vorteil, daß wir auf V und W Abbildungen, Konjugation genannt,

$$V \to V,\ x \mapsto \bar{x},$$

zur Verfügung haben, die jedem Vektor x den Vektor $\bar{x}$ mit den konjugierten Koordinaten zuordnen. Wir setzen

$$b_V(x,y) = \langle x, \bar{y} \rangle \text{ für } x, y \in V$$

und

$$b_W(x, y) = \langle x, \bar{y} \rangle \text{ für } x, y \in W.$$

Dies sind gerade die gewöhnlichen (nicht-hermitischen) Standard-Skalarprodukte auf V bzw. W. Die Bedingung, daß f* der zu f adjungierte Operator ist, bekommt dann die Gestalt

$$b(f(x), y) = b(x, \varphi(y))$$

wobei

$$\varphi(y) = \overline{f}^*(\bar{y})$$

die Zusammensetzung von f* mit der komplexen Konjugation ist. Mit anderen Wort, φ ist gerade der adjungierte Operator bezüglich des gewöhnlichen Skalarprodukts. Aus der Existenz und Eindeutigkeit von φ ergibt sich aber die von f*,

$$f^*(y) = \bar{\varphi}(\bar{y}).$$

QED.

6.2.5 Selbstadjungierte Operatoren und invariante Unterräume

Seien V ein $\mathbb{C}$-Vektorraum mit hermitischem Skalarprodukt < , > (und mit dim V < ∞),

$$f: V \to V$$

ein selbstadjungierter Operator und

$$V' \subseteq V$$

ein f-invarianter Unterraum. Dann gilt:

(i) Die Einschränkung des Skalarprodukts < , > auf V' ist ein hermitisches Skalarprodukt.

(ii) Die Einschränkung

$$f|_{V'},$$

von f auf V' selbstadjungiert.

Beweis: siehe 6.1.8.

6.2.6 Diagonalisierbarkeit selbstadjungierter Operatoren im definiten Fall

Seien V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit hermitischem Skalarprodukt <,> und f:V→V ein selbstadjungierter $\mathbb{C}$-linearer Endomorphismus. Dann besitzt V eine Eigenbasis bezüglich f.

Beweis. siehe 6.1.9

QED.

6.2.7 Eigenwerte und Eigenvektoren hermitisch selbstadjungierter Operatoren

Seien V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit hermitischen Skalarprodukt < , > und

$$f: V \to V$$

ein selbstadjungierter Operator. Dann gilt

(i) Die Hauptraumzerlegung von V bezüglich f ist orthogonal, d.h. je zwei Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander,

$$x' \in V_{c'} \text{ und } x'' \in V_{c''} \text{ und } c' \neq c'' \Rightarrow x' \perp x''.$$

(ii) Die Nullstellen des charakteristischen Polynoms von f sind reell.

Beweis. Zu (i). siehe 6.1.13

Zu (ii). Sei eine Nullstelle des charakteristischen Polynoms von f. Dann ist c eine komplexe Zahl. Da V ein Vektorraum über $\mathbb{C}$ ist, ist damit c auch ein Eigenwert von f, d.h. es gibt ein

$$v \in V - \{0\}$$

mit

$$f(v) = cv.$$

Damit gilt aber auch

$$c\langle v,v\rangle = \langle cv, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, cv\rangle = \bar{c}\cdot\langle v,v\rangle,$$

also

$$(c-\bar{c})\langle v,v\rangle = 0.$$

Wegen $v\neq 0$ ist auch $\langle v,v\rangle \neq 0$, d.h. es muß $c-\bar{c} = 0$ gelten, d.h.

$$c = \bar{c}.$$

Mit anderen Worten, c ist eine reelle Zahl.

QED.

6.2.8 Eigenbasen hermitisch selbstadjungierter Operatoren

Seien V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit hermitischen Skalarprodukt $\langle\, ,\, \rangle$ und $f:V\rightarrow V$ ein selbstadjungierter Operator. Dann gibt es eine orthonormierte Basis von V, welche aus lauter Eigenvektoren von f besteht.

Beweis. Wie wir bereits wissen, fällt die Hauptraumzerlegung mit der Eigenraumzerlegung zusammen und diese ist eine orthogonale Zerlegung. Es genügt in jedem Hauptraum irgendeine orthonormierte Basis zu wählen. Zusammen bilden die so gewonnenen Vektoren eine Basis von V der gesuchten Art.

QED.

6.3. Orthogonale Transformationen

6.3.1 Ein Problem: Normalformen symmetrischer Bilinearformen

(i) In den vorangehenden Abschnitten haben wir gesehen, die Matrix einer selbstadjungierten Abbildung läßt sich auf eine sehr einfache Gestalt bringen, indem man die Basis geeignet wählt.

(ii) Es ist deshalb naheliegend, zu fragen, ob eine analoge Aussage auch für die Matrizen von symmetrischen Bilinearformen gilt.

(iii) Im Fall definiter Skalarprodukte kennen wir die Antwort bereits: auf Grund des Orthogonalisierungsverfahrens wissen wir, daß man stets dafür sorgen kann, daß die Matrix Diagonalgestalt bekommt.

(iv) Im allgemeinen Fall können wir durch Wahl einer Basis (d.h. durch Identifikation von V mit dem K^n) dafür sorgen, daß die Bilinearform die Gestalt

$$b(x,y) = x^T Ay \text{ für } x,y\in K^n.$$

(v) Wechselt man die Basis, so besteht zwischen alten und neuen Koordinaten die Relation

$$y = C'y'$$
$$x = C'x'$$

mit einer umkehrbaren Matrix C. In den neuen Koordinaten hat die Bilinearform also die Gestalt

$$b(x',y') = (Cx')^T A(Cy') = x'C^T ACy'.$$

Für die Matrix A' in den neuen Koordinaten gilt also

$$A' = CAC^T$$

mit einer umkehrbaren Matrix $C = C'^T$.

(vi) Die Frage nach einer Normalform für symmetrisch Bilinearformen ist also eine ganz ähnliche Frage wie die nach den Normalformen selbstadjungierter Abbildungen. Anstelle der inversen Matrix tritt hier die transponierte Matrix.

Problem: Gibt es zu jeder reellen symmetrischen Matrix eine umkehrbare Matrix C derart, daß

$$A' = CAC^T$$

symmetrisch ist?

Die Antwort lautet ja. Man löst das Problem mit Hilfe des Begriffs der orthogonalen Transformation, der es uns gestattet zu beweisen, daß es sich in Wirklichkeit um das bereits behandelte Eigenbasis-Problem handelt.

6.3.2 Der Begriff der Isometrie (vorläufige Definition)

Seien V und V' zwei endlich-dimensionale K-Vektorräume mit Skalarprodukt und

$$f: V \rightarrow V'$$

eine K-lineare Abbildung. Die Abbildung f heißt isometrisch, wenn sie das "Längenquadrat" der Vektoren von V erhält, d.h. wenn gilt

$$\langle f(x),f(x)\rangle = \langle x,x\rangle$$

für alle $x \in V$.

6.3.3 Die Invarianz des Skalarprodukts

Sei K ein Körper der Charakteristik $\neq 2$. Dann ist das Skalarprodukt eines K-Vektorraums bereits eindeutig festgelegt, wenn man das Längenquadrat jedes Vektor kennt. Genauer gilt

$$\langle x,y\rangle = \frac{1}{2}(\langle x+y,x+y\rangle - \langle x,x\rangle - \langle x,x\rangle)$$

Insbesondere erhält also jede lineare Abbildung f, die die Längenquadrate erhält, auch das Skalarprodukt,

$$\langle f(x),f(y)\rangle = \langle x,y\rangle.$$

Speziell werden also auch orthogonale Vektoren in orthogonale Vektoren abgebildet.

Beweis. Es gilt

$$\langle x+y,x+y\rangle = \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle$$

also

$$2\langle x,y\rangle = \langle x+y,x+y\rangle - \langle x,x\rangle - \langle y,y\rangle.$$

QED.

Vereinbarung: endgültige Definition von 'isometrisch'

Im folgenden wollen wir unter einer Isometrie eine Abbildung verstehen, die das Skalarprodukt erhält.

6.3.4 Der Begriff der orthogonalen Transformation

Sei V ein endlich-dimensionaler K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$. Eine orthogonale Transformation von V ist eine isometrische lineare Abbildung $f:V \rightarrow V$, d.h. eine lineare Abbildung mit

$$\langle f(x), f(y)\rangle = \langle x,y\rangle$$

für alle $x,y \in V$. Die Menge der orthogonalen Transformatione von V wird mit

$$\mathbf{O}(V) = \mathbf{O}(V, \langle, \rangle)$$

bezeichnet.

6.3.5 Eigenschaften orthogonaler Transformationen

Seien V ein endlich-dimensionaler K-Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$,

$$v_1,...,v_n$$

eine orthonormierte Basis von V und

$$f: V \to V$$

eine K-lineare Abbildung. Dann sind die folgenden Aussagen äquivalent.

(i) $f: V \to V$ ist eine orthogonale Transformation.

(ii) $f(v_1),\ldots,f(v_n)$ ist eine orthonormierte Basis von V.

(iii) Die Matrix $A = M_v^v(f)$ genügt der Bedingung

$$A^T A = \mathrm{Id}$$

(iv) Die Matrix $A = M_v^v(f)$ genügt der Bedingung

$$AA^T = \mathrm{Id}.$$

(v) Die Matrix $A = M_v^v(f)$ genügt der Bedingung

$$A^{-1} = A^T.$$

Beweis. (i) $\Rightarrow$ (ii). Nach Voraussetzung gilt

$$\langle f(v_i), f(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}.$$

Falls also die $f(v_i)$ eine Basis von V bilden, so bilden sie auch eine orthonormierte Basis.

Es reicht also zu zeigen, die $f(v_i)$ sind linear unabhängig. Sei

(1) $$c_1 f(v_1)+\ldots+c_n f(v_n) = 0.$$

Wir haben zu zeigen, daß sämtliche Koeffizienten c_i gleich Null sind. Es gilt

$$\begin{aligned} 0 &= \langle c_1 f(v_1)+\ldots+c_n f(v_n), f(v_i)\rangle \\ &= c_1\langle f(v_1), f(v_i)\rangle+\ldots+c_n\langle f(v_n), f(v_i)\rangle \\ &= c_i\langle f(v_i), f(v_i)\rangle \\ &= c_i\cdot 1 \\ &= c_i \end{aligned}$$

(ii) $\Rightarrow$ (i). Für je zwei Vektoren $x = \sum_{i=1}^n x_i v_i$ und $y = \sum_{i=1}^n y_i v_i$ gilt

$$\begin{aligned} \langle f(x), f(y)\rangle &= \sum_{i,j=1}^n x_i\, y_j \langle f(v_i), f(v_j)\rangle \\ &= \sum_{i,j=1}^n x_i\, y_j\, \delta_{ij} \\ &= \sum_{i,j=1}^n x_i\, y_j \langle v_i, v_j\rangle \\ &= \langle x, y\rangle. \end{aligned}$$

Die Abbildung f ist also orthogonal.

(ii) $\Leftrightarrow$ (iii). Mit $A = (a_{ij})$ gilt nach Definition von A,

$$f(v_i) = \sum_{k=1}^n a_{ki} v_k .$$

Also ist

$$\langle f(v_i), f(v_j)\rangle = \sum_{k,\ell=1}^n a_{ki}\, a_{\ell j} \langle v_k, v_\ell\rangle$$

$$= \sum_{k=1}^{n} a_{ki}\, a_{kj}$$

$$= \sum_{k=1}^{n} a^T_{ik}\, a_{kj}$$

Bedingung (ii), d.h. $\delta_{ij} = \langle f(v_i), f(v_j)\rangle$ ist somit äquivalent zur Bedingung $A^TA = Id$, d.h. zur Bedingung (iii).

(iii) ⇔ (iv) ⇔ (v). Alle drei Bedingungen bedeuten gerade, daß die Matrizen A und A^T zueinander invers sind.

QED.

Bemerkungen

(i) Eine Matrix $A \in K^{n\times n}$ heißt orthogonal, wenn sie einer der drei äquivalenten Bedingungen (iii), (iv) bzw. (v) genügt, d.h.

$$AA^T = Id \ (\Leftrightarrow A^TA = Id \Leftrightarrow A^{-1} = A^T)$$

vgl. 3.4. 5.

(ii) Wegen Bedingung (iii) bedeutet die Orthogonalität von A gerade, daß die Spalten von A eine orthonormierte Basis von K^n bilden (bezüglich des Standard-Skalarprodukts).

(ii) Wegen Bedingung (iv) bedeutet die Orthogonalität von A gerade, daß die Zeilen von A eine orthonormierte Basis von K^n bilden (bezüglich des Standard-Skalarprodukts).

(iii) Die Äquivalenz von (iii) und (iv) bedeutet auch, daß die transponierte einer orthogonalen Matrix wieder orthogonal ist,

$$A \text{ orthogonal} \Rightarrow A^T \text{ orthogonal.}$$

Da Transponieren und Invertieren für orthogonale Matrizen dieselbe Operation ist, gilt außerdem auch

$$A \text{ orthogonal} \Rightarrow A^{-1} \text{ orthogonal.}$$

(v) Bedingung (v) ist der Schlüssel für die Lösung der eingangs gestellten Klassifizierungsaufgabe für symmetrische Bilinearformen.

(vi) Aus den obigen Aussagen, z.B. aus 6.2.5(iv), folgt unmittelbar, daß orthogonale Transformationen umkehrbar sind. Beim Beweis haben wir allerdings die Existenz einer orthonormierten Basis vorausgesetzt. Die Umkehrbarkeit einer orthogonalen Transformationen kann man aber auch im allgemeinen Fall beweisen. Es gilt sogar die folgende Aussage.

6.3.6 Die orthogonale Gruppe

Sei V ein Vektorraum mit Skalarprodukt < , >. Dann ist die Menge der orthogonalen Transformationen von V,

$$\mathbf{O}(V) = \mathbf{O}(V, \langle,\rangle) = \{ f \in \mathbf{GL}(V) \mid \langle fx, fy\rangle = \langle x,y\rangle \text{ für } x,y \in V\},$$

eine Untergruppe der allgemeinen linearen Gruppe von V.

Beweis. Zeigen wir zunächst, es gilt

$$\mathbf{O}(V) \subseteq \mathbf{GL}(V),$$

d.h. jede orthogonale Transformation

$$f: V \to V$$

ist umkehrbar. Dazu genügt es zu zeigen, f ist injektiv. Sei $v \in V$ ein Vektor mit $f(v) = 0$. Wir haben zu zeigen, dann gilt auch $v = 0$. Angenommen, $v \neq 0$. Dann gibt es, da die Bilinearform < , > nicht entartet ist, einen Vektor v' mit

$$\langle v, v'\rangle \neq 0.$$

Weil nun f orthogonal ist, folgt

$$\langle fv, fv'\rangle \neq 0,$$

d.h. f(v) kann nicht Null sein, im Widerspruch zu unserer Annahme. Damit ist gezeigt,

$$\mathbf{O}(V) \subseteq \mathbf{GL}(V).$$

Auf Grund des Untergruppenkriteriums reicht es, wenn wir die folgenden Aussagen beweisen.

1. $\mathbf{O}(V) \neq \emptyset$.
2. $f,g\in \mathbf{O}(V) \Rightarrow f\circ g\in \mathbf{O}(V)$.
3. $f\in \mathbf{O}(V) \Rightarrow f^{-1}\in \mathbf{O}(V)$.

Zu 1. Die identische Abbildung ist offensichtlich orthogonal, d.h. O(V) ist nicht leer.
Zu 2. Sind f und g orthogonal, so gilt für je zwei Vektoren v' und v",

$$\begin{aligned} \langle f(g(v')), f(g(v''))\rangle &= \langle g(v'), g(v'')\rangle && \text{(weil f orthogonal ist)} \\ &= \langle v',v''\rangle && \text{(weil g orthogonal ist)} \end{aligned}$$

Damit ist aber auch $f\circ g$ orthogonal.
Zu 3. Für je zwei Vektoren v' und v" gilt nach Voraussetzung

$$\langle f(v'), f(v'')\rangle = \langle v',v''\rangle.$$

Speziell für $f^{-1}(v')$ anstelle von v' und $f^{-1}(v'')$ anstelle von v", erhalten wir

$$\langle f(f^{-1}(v')), f(f^{-1}(v''))\rangle = \langle f^{-1}(v'), f^{-1}(v'')\rangle,$$

also

$$\langle v',v''\rangle = \langle f^{-1}(v'), f^{-1}(v'')\rangle.$$

Mit anderen Worten, auch f^{-1} ist orthogonal.
QED.

6.3.7 Diagonalisierung reeller symmetrischer Matrizen mit Hilfe orthogonaler Matrizen

Seien $A\in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann gibt es eine orthogonale Matrix $B\in \mathbb{R}^{n\times n}$ derart, daß

$$BAB^{-1} = BAB^{T}$$

eine Diagonalmatrix ist.
Bemerkungen

(i) Anstelle des Grundkörpers $\mathbb{R}$ kann man auch einen beliebigen Grundkörper K nehmen, vorausgesetzt es gilt
 1. K ist hinreichend groß.
 2. Das Standard-Skalarprodukt von K^n ist definit.

(ii) Bedingung 2 läuft im wesentlichen aber darauf hinaus, daß die Einträge der Matrix A reell sind (d.h. K in $\mathbb{R}$ liegt).

Beweis von 6.3.7. Wir versehen den $\mathbb{R}^n$ mit dem Standard-Skalarprodukt. Dann definiert die Matrix A einen selbstadjungierten Operator f, d.h.es gibt eine orthonormierte Eigenbasis $v_1,...,v_n$ des $\mathbb{R}^n$ bezüglich A.

Bemerkungen

1. Man beachte, der Körper $\mathbb{R}$ ist im Sinne von 6.1.16 hinreichend groß: zunächst liegen die Nullstellen des charakteristischen Polynoms von A in den komplexen Zahlen. Da die reelle symmetrischen Matrix aber (als komplexe Matrix) auch hermitisch ist (bezüglich des hermitischen Standard-Skalarprodukts), sind alle diese Nullstellen sogar reell. Das charakteristische Polynom zerfällt also über $\mathbb{R}$ in Linearfaktoren, d.h. $\mathbb{R}$ ist hinreichend groß.
2. Weiter ist der Körper $\mathbb{R}$ im Sinne von 6.1.15 hinreichend groß, da man aus jeder nicht-negativen reellen Zahl die Quadratwurzel ziehen kann und die Skalarskalarprodukte der Gestalt $\langle v,v\rangle$ nicht-negativ sind.

Es gibt also tatsächlich eine orthonormierte Eigenbasis bezüglich f. Die Matrix von f bezüglich dieser Eigenbasis v hat Diagonalgestalt,

$$M_v^v(f) = M_v^e(Id)M_e^e(f)M_e^v(Id) = BAB^{-1}$$

mit $B = M_v^e(Id)$. Dabei bezeichne e die Standardbasis des $\mathbb{R}^n$. Zum Beweis der Behauptung ist noch zu zeigen, B ist eine orthogonale Matrix (denn dann gilt nach 6.2.5(v) automatisch auch $B^{-1} = B^T$). Dazu wiederum genügt es zu zeigen, B^{-1} ist orthogonal. Sei

$$B^{-1} = (b_{ij}).$$

Dann gilt wegen $B^{-1} = M_e^v(Id)$

$$v_i = Id(v_i) = \sum_{j=1}^n b_{ji}e_j = \begin{pmatrix} b_{1i} \\ \dots \\ b_{ni} \end{pmatrix}.$$

Mit anderen Worten, v_i ist der i-te Spaltenvektor der Matrix B^{-1}. Insbesondere bilden die Spalten der Matrix B^{-1} eine orthonormierte Basis. Nach 6.3.5 ist B^{-1} eine orthogonale Matrix. Dann ist aber auch B orthogonal.

QED.

Bemerkungen

(iii) Ein analoges Ergebnis ist auch für komplexe hermitische Matrizen richtig. Anstelle des Begriffs der orthogonalen Transformation bzw. Matrix muß man den der unitären Transformation bzw. Matrix benutzen. Die Beweise sind dieselben.

(iv) Sei V ein komplexer Vektorraum mit hermitischen Skalarprodukt < , >. Eine $\mathbb{C}$-lineare Abbildung f: V $\to$ V heißt unitär, wenn für je zwei Vektoren $x,y\in V$ gilt

$$<f(x), \overline{f(y)}> = <x,y>.$$

(v) Eine Matrix $A\in\mathbb{C}^{n\times n}$ heißt unitär, wenn die zugehörige Abbildung $f_A:\mathbb{C}^n\to\mathbb{C}^n$ bezüglich des hermitischen Standard-Skalarprodukts unitär ist, d.h. wenn gilt

$$A\bar{A}^T = Id.$$

6.3.8 Diagonalisierung komplexer hermitischer Matrizen mit Hilfe unitärer Matrizen

Seien $A\in\mathbb{C}^{n\times n}$ eine hermitische Matrix. Dann gibt es eine unitäre Matrix $B\in\mathbb{C}^{n\times n}$ derart, daß

$$BAB^{-1} = BA\bar{B}^T$$

eine Diagonalmatrix ist.

Beweis. Man verwendet dieselben Argumente wie beim Beweis von 6.3.7 mit dem hermitischen Standard-Skalarprodukt anstelle des gewöhnlichen Standard-Skalarprodukts. Der Körper $\mathbb{C}$ ist trivialerweise hinreichend groß.

QED.

6.3.9 Normalformen symmetrischer Bilinearformen

Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und

$$b: V\times V \to \mathbb{R}$$

eine symmetrische Bilinearform. Dann gibt es eine Basis $v_1,...,v_n$ von V, bezüglich welcher die Matrix von b die folgende Gestalt hat

$$M(b) = \begin{pmatrix} Id & 0 & 0 \\ 0 & -Id & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

wobei die angegebenen Blöcke auch die Zeilen- bzw. Spaltenzahl 0 haben dürfen.
Mit anderen Worten, identifiziert man V mit Hilfe der Basis $v_1,...,v_n$ mit dem $\mathbb{R}^n$ so bekommt b die Gestalt

$$b(\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ ... \\ y_n \end{pmatrix}) = x_1y_1 + ... + x_ry_r - x_{r+1}y_{r+1} - ... - x_sy_s$$

Die Zahl s der Glieder auf der rechten Seite heißt Rang der Bilinearform b und wird mit

$$\mathrm{rk}\ (b)$$

bezeichenet. Die Differenz

$$\mathrm{sig}(b) = r - (s-r)$$

aus der Anzahl der positiven und der Anzahl der negativen Vorzeichen heißt Signatur von b.

Beweis. Nach 6.4.6 gibt es zumindest eine Basis $v_1,...,v_n$ bezüglich welcher die Matrix von b Diagonalgestalt hat. Identifiziert man mit Hilfe dieser Basis den Raum V mit dem $\mathbb{R}^n$, so bekommt b die Gestalt

$$b(\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ ... \\ y_n \end{pmatrix}) = a_1x_1y_1 + ... + a_rx_ry_r - a_{r+1}x_{r+1}y_{r+1} - ... - a_sx_sy_s$$

mit gewissen reellen Zahlen $a_i = b(v_i,v_i)$. Durch eine geeignete Wahl der Reihenfolge der Basisvektoren und geeignete Wahl von r und s können wir erreichen, daß sämtliche a_i positiv sind. Wir ersetzen jetzt den Basisvektor v_i durch

$$\frac{1}{\sqrt{a_i}}\cdot v_i.$$

Dann hat b in den so abgeänderten Koordinaten die gesuchte Gestalt.

QED.

6.3.10 Trägheitssatz von Sylvester

Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und

$$b: V\times V \to \mathbb{R}$$

eine symmetrische Bilinearform. Dann gibt es $\mathbb{R}$-lineare UnterräumeV_0, V_+ und V_- mit den folgenden Eigenschaften.

(i) $V = V_0 \oplus V_+ \oplus V_-$
(ii) Die Einschränkung von b auf V_+ ist positiv definit.
(iii) Die Einschränkung von b auf V_- ist negativ definit.
(iv) Für je zwei Vektoren v', v" aus zwei verschiedenen direkten Summanden gilt
$$b(v', v") = 0.$$
(v) V_0 =[31] $\{v\in V \mid b(v,v') = 0 \text{ für alle } v'\in V\}$.

Die Dimension der Räume V_+ und V_- ist unabhängig von der spezielen Wahl der Zerlegung. Insbesondere ist der Rang von b,

$$\mathrm{rk}\ (b) = \dim V_+ + \dim V_-$$

und die Siganatur von b

[31] Dieser Raum heißt auch Entartungsraum von b.

$$\mathrm{sig}(b) = \dim V_+ - \dim V_-$$
unabhängig von der speziellen Wahl der Unterräume.

Beweis. 1. Schritt: Existenz der Zerlegung.
Mit den Bezeichnungen von 6.3.9 setzen wir

$$V_+ := Kv_1+...+Kv_r$$
$$V_- := Kv_{r+1}+...+Kv_s$$
$$V' := Kv_s+...+Kv_n$$

Dann gelten die Aussagen (i)-(iv) des Satzes mit V' anstelle von V_0, und es ist

$$V' \subseteq V_0.$$

Sei jetzt

$$v = x_1v_1 + ... + x_nv_n \in V_0.$$

Dann gilt insbesondere

$$0 = b(v, v_i) = x_i \text{ für } i = 1, ... r$$

und

$$0 = b(v, v_i) = -x_i \text{ für } i = r+1, ... s$$

d.h. es gilt $v \in V'$. Wir haben gezeigt,

$$V' = V_0,$$

d.h. es gelten die Aussagen (i)- (v). Nach Wahl der Räume V_+ und V_- gelten außerdem die angegebenen Formeln für den Rang und die Signatur von b,

$$\mathrm{rk}\,(b) = s = \dim V_+ + \dim V_-$$
$$\mathrm{sig}(b) = r - (s-r) = \dim V_+ - \dim V_-$$

2. Schritt: Unabhängigkeit der Dimensionen.
Sei

$$V = V_0 \oplus V'_+ \oplus V'_-$$

eine zweite Zerlegung der angegebenen Art. Für

$$v \in V'_+ \cap (V_0 + V_-) - \{0\}$$

gilt dann

$$b(v,v) > 0 \text{ (wegen } v \in V'_+)$$

und

$$b(v,v) \le 0 \text{ (wegen } v \in V_0 \oplus V_-)$$

was nicht möglich ist. Also gilt

$$V'_+ \cap (V_0 + V_-) = \{0\}.$$

Damit ist aber die Abbildung

$$V_0 \oplus V'_+ \oplus V_- \to V,\ (a,b,c) \mapsto a+b+c,$$

injektiv, d.h. es gilt

$$\dim V_0 + \dim V'_+ + \dim V_- \le \dim V = \dim V_0 + \dim V_+ + \dim V_-$$

also

$$\dim V'_+ \le \dim V_+$$

Aus Symmetriegründen muß auch die umgekehrte Ungleichung bestehen , d.h. es gilt

$$\dim V'_+ = \dim V_+$$

Dann ist aber auch

$$\dim V'_- = \dim V - \dim V_0 - \dim V'_+$$

$= \dim V - \dim V_0 - \dim V_+$
$= \dim V_-$

QED.

6.3.11 Beispiel: die Hauptachsentransformation

Sei $C \subseteq \mathbb{R}^2$ die ebene Kurve mit der Gleichung

$$f(x,y) := x^2 + 6xy + y^2 + 8x + 8y + 1 = 0,$$

d.h. $C := \{(x,y)\in\mathbb{R}^2 \mid f(x,y) = 0\}$. Handelt es sich um eine Ellipse?

Wir schreiben die Gleichung von f in der Gestalt

$$f(x,y) = (x,y)A\begin{pmatrix} x \\ y \end{pmatrix} + b\cdot\begin{pmatrix} x \\ y \end{pmatrix} + 1$$

mit

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \text{ und } b = (8,8)$$

und führen eine solche orthogonale Transformation durch, daß die Matrix A Diagonalgestalt bekommt.

Die Eigenwerte von A:

$$\det(A-T\cdot Id) = \det\begin{pmatrix} 1-T & 3 \\ 3 & 1-T \end{pmatrix} = (T-1)^2 - 9 = T^2 - 2T - 8.$$

$c_{1,2} = 1 \pm \sqrt{1+8} = 1 \pm 3$

$c_1 = 4,\ c_2 = -2$

Die Eigenvektoren von A:

Der Eigenraum V_{c_1} zum Eigenwert $c_1 = 4$ hat die Gleichung

$$-3x + 3y = 0,$$

d.h. es gilt

$$V_{c_1} = \mathbb{R}\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Der Eigenraum V_{c_2} zum Eigenwert $c_2 = -2$ hat die Gleichung

$$3x + 3y = 0,$$

d.h. es gilt

$$V_{c_2} = \mathbb{R}\cdot\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Als orthonormierte Eigenbasis können wir zum Beispiel die folgende verwenden.

$$v_1 := \frac{1}{\sqrt{2}}\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ und } v_2 := \frac{1}{\sqrt{2}}\cdot\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Bezeichen wir mit x' und y' die Koordinaten bezüglich dieser neuen Basis. Dann gilt

$$\begin{pmatrix} x \\ y \end{pmatrix} = xe_1 + ye_2 = x'v_1 + y'v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} x'+y' \\ x'-y' \end{pmatrix},$$

d.h.

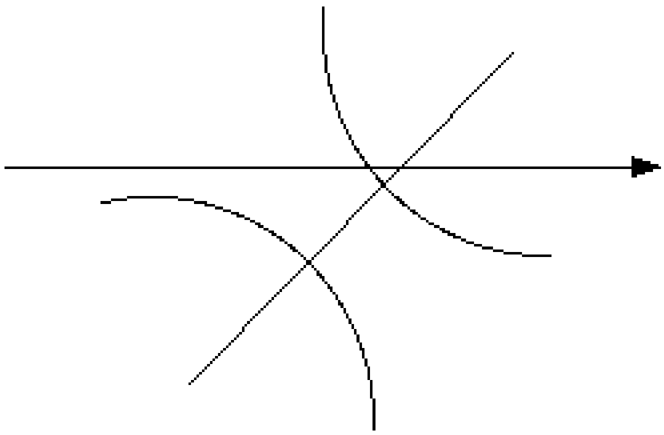
$$x = \frac{1}{\sqrt{2}}(x'+y') \text{ und } y = \frac{1}{\sqrt{2}}(x'-y').$$

In den neuen Koordinaten hat die Kurve C also die Gleichung

$$0 \quad = \frac{1}{2}(x'+y')^2 + 3(x'+y')(x'-y') + (x'-y')^2 + \frac{16}{\sqrt{2}}x' + 1$$

$= \frac{1}{2}(x'^2 + 2x'y' + y'^2)$
$+3(x'^2 - y'^2)$
$+ \frac{1}{2}(x'^2 - 2x'y' + y'^2) + \frac{16}{\sqrt{2}}x' + 1$
$= 4x'^2 - 2y'^2 + \frac{16}{\sqrt{2}}x' + 1$
$= 4\,(x' + \frac{8}{\sqrt{2}})^2 - 2y'^2 + 1 - 32$
$= 4x''^2 - 2y''^2 - 31$

Also ist C eine Hyperbel.



6.3.12 Hyperflächen

Seien K ein Körper der Charakteristik ≠ 2 und V der K-Vektorraum

$$V = K^n$$

und

$$f(x_1,...,x_n) = \sum_{a_1+...+a_n \le m} c_{a_1,...,a_n} \cdot x_1^{a_1} \cdot ... \cdot x_n^{a_n}$$

ein Polynom des Grades m. Multi-Index-Schreibweise:

$$f(x) = \sum_{a \in \mathbb{N}^n,\, |a| \le m} c_a \cdot x^a,$$

d.h. für $x = (x_1,...,x_n)$ und $a = (a_1,...,a_n)$ setzen wir

$$x^a := x_1^{a_1} \cdot ... \cdot x_n^{a_n} \text{ und } |a| = a_1 + ... + a_n.$$

Die Menge

$$V(f) := \{x \in K^n \mid f(x) = 0\}$$

heißt dann Hyperfläche des Grades

$$\deg f := \max\{|a| \mid c_a \neq 0\}.$$

Im Fall des Grades 2 hat f die Gestalt

(1) $$f = \sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{k=1}^{n} b_kx_k + c$$

Wegen

$$a_{ij}x_ix_j + a_{ji}x_jx_i = \frac{a_{ij}+a_{ji}}{2}(x_ix_j + x_jx_i)$$

können wir annehmen, in (1) gilt $a_{ij} = a_{ji}$, d.h. $A := (a_{ij})$ ist eine symmetrische Matrix. Die allgemeine Gleichung zweiten Grades läßt sich also in der Gestalt

$$f = (x_1,...,x_n)A\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix} + b\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix} + c$$

schreiben mit einer symmetrischen Matrix A und einem Spaltenvektor b.

6.3.13 Normalformen quadratischer Gleichungen

Jede Gleichung zweiten Grades f(x) mit Koeffizienten aus $\mathbb{R}$ läßt sich durch eine orthogonale Transformation und eine Verschiebung in die folgende Gestalt bringen.

$$f(x_1,...,x_n) = a_1x_1^2+...+a_rx_r^2 + b_{r+1}x_{r+1} + ... +b_sx_s + c.$$

Beweis.Wir schreiben f in der Gestalt

$$f(x) = x^TAx + bx + c$$

mit einer symmetrischen Matrix A und einem Spaltenvektor b. Wir wählen eine orthogonale Matrix B derart, daß

$$A' = BAB^T$$

Diagonalgestalt hat. Wir führen die Transformation

$$x' = Ax$$

aus. Wir setzen $x = B^{-1}x' = B^Tx'$ in f ein und erhalten

$$\begin{aligned} f'(x') &= f(B^Tx') = (B^Tx')^TAB^Tx' + bB^Tx' + c \\ &= x'^TBAB^Tx' + bB^Tx' + c \\ &= x'A'x' + b'x' + c \end{aligned}$$

In neuen Koordinaten, die wir wieder mit x anstelle von x' bezeichnen wollen, bekommt f also die Gestalt

$$f(x) = a_1x_1^2+...+a_rx_r^2 + b_1x_1 + ... +b_sx_s + c,$$

wobei wir annehmen können, daß alle

$$a_1,...,a_r \in \mathbb{R}$$

von Null verschieden sind. Wir führen die folgende Koordinatentransformation aus.

$$\begin{aligned} x'_i &= x_i - \frac{b_i}{2} \text{ für } i=1,...,r \\ x'_i &= x_i \end{aligned}$$

und erhalten

$$f(x) = a_1x'^2_1+...+a_rx'^2_r + b_{r+1}x'_{r+1} + ... +b_sx'_s + c'$$

QED.

6.3.14 Die Lösungen von Gleichungen zweiten Grades in der reellen Ebene

Die Gleichung einer reellen ebenen Kurve läßt sich stets auf die folgende Gestalt bringen.

$$f(x_1,x_2) = a_1x_1^2+...+a_rx_r^2 + b_{r+1}x'_{r+1} + ... +b_sx'_s + c'$$

mit $1\le r \le s \le 2$, wobei all a_i und b_j von Null verschieden sind.

1. Fall: r = s = 2.

$$f(x, y) = ax^2+ by^2 + c \text{ mit } a \ne 0 \text{ und } b \ne 0$$

Wenn a und b dasselbe Vorzeichen haben ist die Kurve eine Ellipse, ein (doppelt zu zählender) einzelner Punkt oder die leere Menge (je nachdem, welchen Wert die Konstante c hat).

Haben a und b unterschiedliches Vorzeichen, so ist die Kurve eine Hyperbel oder ein Paar sich schneidender Geraden (je nachdem, ob c von Null verschieden ist oder nicht).

2. Fall: r = 1, s = 2.

$$f(x,y) = ax^2 + by + c \text{ mit } a \ne 0 \text{ und } b \ne 0.$$

Durch eine weitere Verschiebung kann man erreichen, daß c = 0 gilt

$$f(x) = ax^2 + by \text{ mit } a \ne 0 \text{ und } b \ne 0.$$

Die Kurve ist dann eine Parabel.

3. Fall: r = s = 1.

$$f(x) = ax^2 + c \text{ mit } a \ne 0$$

Die Kurve besteht je nach Vorzeichen und Wert von c aus zwei parallelen Geraden, einer (doppelt zu zählenden) Geraden bzw. der leeren Menge.

Bemerkung

Im Fall von quadratischen Gleichungen in drei Unbestimmten läßt sich die Lösungsmenge in analoger Weise beschreiben. Man erhält unter anderen Ellipsoide, Paraboloide, einschalige Hyperboloide und zweischalige Hyperboloide.

6.4 Schiefsymmetrische Bilinearformen

6.4.1 Definitionen

Sei V ein K-Vektorraum. Eine Bilinearform auf V,

$$\omega: V \times V \to K$$

heißt antisymmetrisch oder schiefsymmetrisch, wenn gilt

$$\omega(v', v'') = - \omega(v'', v') \text{ für beliebige } v', v''\in V.$$

Sie heißt symplektisch, wenn sie außerdem nicht-entartet ist. Ein symplektischer Vektorraum ist ein K-Vektorraum, der mit einer symplektischen Bilinearform versehen ist.

Seien V ein symplektischer Vektorraum mit der symplektischen Bilinearform ω und

$$W \subseteq V$$

ein K-linearer Unterraum. Dann heißt der K-lineare Unterraum

$$W^\perp := \{ v\in V \mid \omega(v,w) = 0 \text{ für beliebige } w\in W\}$$

orthogonales Komplement von W in V.

Eine Matrix $A\in K^{n\times n}$ heißt schiefsymmetrisch, wenn gilt $A^T = -A$.

6.4.2 Schiefsymmetrische Bilinearformen und Matrizen

Seien V ein K-Vektorraum mit der Basis

$$v_1,...,v_n \in V$$

und

$$\omega: V \times V \to K$$

eine Bilinearform. Dann sind folgende Aussagen ´äquivalent.

(i) ω ist schiefsymmetrisch.

(ii) Die Matrix $M_v(\omega)$ von ω bezüglich der gegebenen Basis ist schiefsymmetrisch.

Beweis. (i) ⇒ (ii). Es gilt

$$\omega(v_i, v_j) = - \omega(v_j, v_i),$$

d.h. die Matrix $M_v(\omega)$ der $\omega(v_i, v_j)$ ist schiefsymmetrisch.

(ii) ⇒ (i). Seien $v' = \sum_{i=1}^{n} x'_i v_i$ und $v'' = \sum_{i=1}^{n} x''_i v_i$ aus V. Dann gilt

$$\begin{aligned}\omega(v', v'') &= \sum_{i=1}^{n}\sum_{j=1}^{n} x'_i\, x''_j\, \omega(v_i, v_j)\\ &= - \sum_{i=1}^{n}\sum_{j=1}^{n} x'_i\, x''_j\, \omega(v_j, v_i)\\ &= - \omega(v'', v'),\end{aligned}$$

d.h. ω ist schiefsymmetrisch.
QED.

6.4.3 Beispiel: der symplektische Standardraum der Dimension 2

$$V = K^2,\ A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in K^{2\times 2} \text{ im } \omega(x,y) = x^T A y.$$

Der Vektorraum V ist mit der Bilinearform ω ein symplektischer Raum. Er heißt 2-dimensionaler symplektischer Standardraum.

6.4.4 Beispiel: direkte Summe von symplektischen Räumen

Seien V' und V" zwei endlich-dimensionale symplektische K-Vektorräume mit den symplektischen Bilinearformen ω' bzw. ω''. Dann ist

$$V := V' \oplus V''$$

mit der Bilinearform

$$\omega: V\times V \to K,\ ((a', a''), (b', b'')) \mapsto \omega'(a', b') + \omega''(a'', b''),$$

ein symplektischer K-Vektorraum, welcher direkte Summe der beiden symplektischen Vektorräume V' und V" heißt (und mit $V' \oplus V''$ bezeichnet wird).
Beweis. Seien v' bzw. v" eine Basis von V' bzw. V". Dann bilden die beiden Basen zusammene eine Basis v von V, und durch direktes Nachrechnen sieht man

$$M_v(\omega) = \begin{pmatrix} M_{v'}(\omega') & 0 \\ 0 & M_{v''}(\omega'') \end{pmatrix}$$

Insbesondere ist ω nicht-entartet und schiefsymmetrisch, also symplektisch.
QED.

6.4.5 Zerlegung in symplektische Standardräume

Sei V ein symplektischer K-Vektorraum der Dimension

$$n = \dim V\ (< \infty)$$

mit der symplektischen Bilinearform ω. Die Charakteristik von K sei ungleich 2,

$$\text{char } K \neq 2.$$

Dann gibt es eine Basis v von V derart, daß $M_v(\omega)$ direkte Summe von Exemplaren der Matrix

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

ist. Insbesondere ist n gerade und V eine direkte Summe von 2-dimensionalen symplektischen Standardräumen.
Beweis. Wir können annehmen, $V \neq \{0\}$. Sei $v' \in V - \{0\}$ beliebig. Da ω nicht entartet ist, gibt es ein $v'' \in V$ mit

$$\omega(v', v'') \neq 0.$$

Durch Multiplikation von v” mit einem von Null verschiedenen Faktor können wir erreichen, daß gilt

(1) $$\omega(v', v'') = 1.$$

Dann gilt

(2) $$\omega(v'', v') = -1$$

und $\omega(v', v') = - \omega(v',v')$. Da die Charakteristik von K ungleich zwei sein soll, folgt

(3) $$\omega(v', v') = 0$$

und analog

(4) $$\omega(v'', v'') = 0$$

Aus (1), (3) und (4) ergibt sich insbesondere, daß v’ und v” nicht proportional sein können, d.h. es gilt

(5) v’ und v” sind linear unabhängig.

Wir setzen

$$W = Kv'+Kv''$$

und

$$U := W^{\perp} = \{v\in V \mid \omega(v,w) = 0 \text{ für } w\in W\} = \{v\in V \mid \omega(v,v') = \omega(v,v'') = 0 \}$$

Dann gilt

$$\dim W = 2$$

und U ist gerade der Kern der linearen Abbildung

$$f: V \to K^2, v \mapsto (\omega(v,v''), \omega(v,v')),$$

Wegen $f(v') = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $f(v'') = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ ist f surjektiv, also

$$\dim U = \dim \ker f = \dim V - \dim \operatorname{im} f = \dim V - 2.$$

Die lineare Abbildung

(6) $$g: W\oplus U \to V, (w, u) \mapsto w+u,$$

ist somit eine Abbildung zwischen Räumen gleicher Dimension. Für $(w,v)\in\ker(g)$ gilt

$$\begin{aligned} v = -w \in U\cap W &= \{u = x'v' + x''v'' \mid x',x''\in K,\ 0 = \omega(u,v') = \omega(u,v'') \} \\ &= \{ u = x'v' + x''v'' \mid x',x''\in K,\ 0 = -x'' = x'\} \\ &= \{0\}. \end{aligned}$$

Der Kern von g ist somit trivial, d.h. g ist injektiv, d.h. g ist ein Isomorphismus.
Es reicht zu zeigen, die

Einschränkung von ω auf U ist symplektisch,

denn dann gibt es nach Induktionsvoraussetzung eine Basis u von U derart, daß gilt

$$M_u(\omega|_U) = \text{direkte Summe von Exemplaren von} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Die Vektoren v’ und v” bilden dann zusammen mit den Vektoren von u eine Basis von V mit

$$M(\omega) = \begin{pmatrix} 0 & 1 & \dots \\ -1 & 0 & \dots \\ \dots & \dots & M_u(\omega|_U) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \oplus M_u(\omega|_U)$$

Beweisen wir also (6). Mit ω ist auch die Einschränkung $\omega|_U$ schiefsymmetrisch. Wir haben noch zu zeigen,

$$\omega|_U: U\times U \to K$$

ist nicht entartet. Sei $x\in U - \{0\}$. Weil ω nicht entartet ist, gibt es ein $y\in V$ mit

(7) $$\omega(x, y) \neq 0.$$

Wir haben zu zeigen, y kann sogar aus U gewählt werden. Weil (6) ein Isomorphismus ist, können wir y zumindest in der folgenden Gestalt schreiben.

$$y = w + u \text{ mit } w\in W \text{ und } u\in U.$$

Es gilt

$$0 \neq \omega(x,y) = \omega(x, w + u) = \omega(x,w) + \omega(x,u)$$

$= \omega(x, u)$ (wegen $x\in U = W^{\perp}$ und $w\in W$).

d.h. wir können $y = u\in U$ wählen.

QED.

6.4.6 Der Rang einer schiefsymmetrischen Bilinearform

Seien V ein endlich-dimensionaler K-Vektorraum und

$$\omega: V \times V \to K$$

eine Bilinearform. Dann ist der Rang von ω definiert als Rang der Matrix von ω bezüglich irgendeiner Basis v von V:

$$\mathrm{rk}(\omega) = \mathrm{rk}\, M_v(\omega).$$

Bemerkungen

(i) Diese Definition ist unabhängig von der Wahl der Basis v. Das folgt zum Beispiel aus Bemerkung (ii) von 6.4.2, denn für jede weitere Basis v' von V hat man

$$M_v(\omega) = B^T \cdot M_{v'}(\omega) \cdot B$$

mit einer umkehrbaren Matrix B.

(ii) Wir zeigen als nächstes, der Rang einer schiefsymmetrischen Bilinearform ist deren einzige Invariante.

6.4.7 Klassifikation der schiefsymmetrischen Bilinearformen

Seien V ein endlich-dimensionaler K-Vektorraum und

$$\omega: V \times V \to K$$

eine schiefsymmetrische Bilinearform. Bezeichne

$$V_0 := \{v\in V \mid \omega(v, v') = 0 \text{ für jedes } v'\in V\}$$

den Entartungsraum von ω. Dann gibt es einen K-linearen Unterraum $V'\subseteq V$ mit

1. $V = V_0 \oplus V'$
2. $\omega|_{V'}: V'\times V' \to K$ ist symplektisch.

Insbesondere gibt es eine Basis von v von V mit

$$M_v(\omega) = \begin{pmatrix} 0 & \mathrm{Id}_r & 0 \\ -\mathrm{Id}_r & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

mit einer $r\times r$-Einheitsmatrix Id_r. Die Zahl r ist unabhängig von der speziellen Wahl der Basis:

$$r= \frac{1}{2}\mathrm{rk}(\omega).$$

Beweis. Wir ergänzen eine Basis $v_1. \dots , v_s$ des Entartungsraums V_0 zu einer Basis

$$v_1. \dots , v_n \in V$$

von V und setzen

$$V' := v_{s+1}K + \dots + v_nK.$$

Dann ist Bedingung 1 erfüllt. Ist $v\in V'$ ein von Null verschiedenes Element, so liegt dieses Element wegen Bedingung 1 nicht im Entartungsraum, d.h. es gibt ein $\tilde{v}\in V$ mit

$$\omega(v, \tilde{v}) \neq 0.$$

Wegen Bedingung 1 gilt

$$\tilde{v} = \tilde{v}' + v' \text{ mit } \tilde{v}'\in V_0 \text{ und } v'\in V'.$$

Damit ist

$$0 \neq \omega(v, \tilde{v}) = \omega(v, \tilde{v}') + \omega(v, v')$$

$$= -\omega(\tilde{v}',v) + \omega(v, v') \quad \text{(weil } \omega \text{ schiefsymmetrisch ist)}$$
$$= \omega(v, v') \qquad\qquad \text{(wegen } \tilde{v}' \in V_0\text{).}$$
Damit ist aber ω auf V' nicht entartet, also symplektisch. Damit ist Bedinung 2 erfüllt. Insbesondere zerfällt V' in eine direkte Summe von symplektischen Standard-Räumen der Dimension 2. Sei
$$v'_1, v'_2, ..., v'_{2r-1}, v'_{2r}$$
eine entsprechende Basis von V'. Dann ist
$$v'_1, v'_3, ..., v'_{2r-1}, v'_2, v'_4, ..., v'_{2r}, v_1, ..., v_s$$
eine Basis, bezüglich welcher die Matrix von ω die angegebene Gestalt hat. Der Rang dieser Matrix ist trivialerweise gleich
$$rk(\omega) = 2r.$$
QED.

6.4.8 Normalformen schiefsymmetrischer Matrizen

Für jede schiefsymmetrische Matrix $A \in K^{n\times n}$ gibt es eine umkehrbare Matrix B mit
$$B^T AB = \begin{pmatrix} 0 & Id_r & 0 \\ -Id_r & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$
Dabei hängt der Rang r der Einheitsmatrix Id_r nicht von der speziellen Wahl von B ab und ist gleich
$$r = \frac{1}{2} rk(A).$$
Beweis. Folgt aus 6.4.7 und Bemerkung (ii) von 6.4 2.
QED.

6.5 Zur Bestimmung einer Jordanbasis

6.5.1 Zum Inhalt dieses Abschnitts

(i) Der Gegenstand dieses Abschnitts ist die Berechnung einer Jordanbasis eines K-linearen Endomorphismus
$$f: V \to V$$
eines endlich-dimensionalen Vektorraums V über einem hinreichend großen Körper. Mit Hilfe der Hauptraumzerlegung reduziert man das Problem auf die Bestimmung der Jordan-Basen für jeden der Haupträume und damit auf den Fall, daß f nilpotent ist.

(ii) Wir beginnen mit der Beschreibung eines Algorithmus, der sich aus dem Existenzbeweis für die Zerlegung in zyklische Unterräume ergibt.

(iii) Bei der effektiven Bestimmung der Basis ist es nützlich auf den Räumen geeignete Skalarprodukte einzuführen. Die Rechnung in Faktorrämen wird dann durch die Rechnung in Unterräumen ersetzt. Wie man dabei vorgeht, beschreiben wir in den nachfolgenden Abschnitten.

6.5.2 Bestimmung einer Jordan-Basis eines nilpotenten Endomorphismus (allgemeiner Fall)

Sei $f: V \to V$ eine nilpotente K-lineare Abbildung mit $\dim V < \infty$ und
$$u = stab(f).$$
Wir setzen
$$W^i := ker(f^i) \text{ für } i = 1, ... , u-1$$

und bezeichnen mit

$$\rho_i: V \to V/W^i,\ v \mapsto v + W^i,$$

die zugehörigen natürlichen Abbildungen.
Zur Bestimmung einer Jordanbasis von V reicht es, die Hauptvektoren einer zyklischen Basis zu bestimmen für jeden direkten Summanden einer Zerlegung von V in eine direkte Summe zyklischer Unterräume.
Die höchste auftretende Ordung eines Hauptvektors ist dabei gleich u.
1. Schritt. Bestimmung der Hauptvektoren der Ordnung u.
Man wähle beliebige Vektoren

(1) $$v_{u,1}, \dots, v_{u,r_u} \in V$$

derart, daß die natürlichen Bilder in V/W^{u-1} eine Basis dieses Faktorraums bilden. Diese sind gerade die Hauptvektoren der Ordnung u.
2. Scbritt: Bestimmung der Hauptvektoren der Ordnung u-1.
Man wähle beliebige Vektoren

(2) $$v_{u-1,1}, \dots, v_{u-1,r_{u-1}} \in W^{u-1}$$

derart, daß die natürlichen Bilder dieser Vektoren in V/W^{u-2} zusammen mit den natürlichen Bildern der Vektoren

$$v_{u,1}, \dots, v_{u,r_u}, f(v_{u,1}), \dots, f(v_{u,r_u})$$

eine Basis dieses Faktorraums bilden. Die Vektoren (2) sind dann gerade die Hauptvektoren der Ordnung u-1 der gesuchten Jordanbasis.
3. Schritt: Bestimmung der Hauptvektoren der Ordnung n-2.
Man wähle beliebige Vektoren

(3) $$v_{u-2,1}, \dots, v_{u-2,r_{u-2}} \in W^{u-2}$$

derart, daß die natürlichen Bilder dieser Vektoren in V/W^{u-3} zusmmen mit den natürlichen Bildern der Vektoren

$$v_{u,1}, \dots, v_{u,r_u}, f(v_{u,1}), \dots, f(v_{u,r_u}), f^2(v_{u,1}), \dots, f^2(v_{u,r_u})$$
$$v_{u-1,1}, \dots, v_{u-1,r_{u-1}}, f(v_{u-1,1}), \dots, f(v_{u-1,r_{u-1}})$$

eine Basis diese Faktorraums bilden. Die Vektoren (3) sind dann gerade die Hauptvektoren der Ordnung u-2 der gesuchten Jordanbasis.

Seien jetzt bereits die Hauptvektor bis zur Ordnung u-i+1 bestimmt:

$$v_{u,1}, \dots, v_{u,r_u}$$
$$v_{u-1,1}, \dots, v_{u-1,r_{u-1}}$$
$$\dots$$
$$v_{u-i+1,1}, \dots, v_{u-i+1,r_{u-i+1}}$$

i+1. Schritt: Besimmung der Hauptvektoren der Ordnung u-i.
Man wähle beliebige Vektoren

(i+1) $$v_{u-i,1}, \dots, v_{u-i,r_{u-i}} \in W^{u-i}$$

derart, daß die natürlichen Bilder dieser Vektoren in V/W^{u-i-1} zusammen mit den natürlichen Bildern der Vektoren

$$v_{u,1}, \dots , v_{u,r_u}, f(v_{u,1}), \dots , f(v_{u,r_u}),\dots, f^i(v_{u,1}), \dots , f^i (v_{u,r_u})$$
$$v_{u-1,1}, \dots , v_{u-1,r_{u-1}},\dots, f^{i-1}(v_{u-1,1}), \dots ,f^{i-1}(v_{u-1,r_{u-1}})$$
$$\dots$$
$$v_{u-i+1,1}, \dots , v_{u-i+1,r_{u-i+1}}, f(v_{u-i+1,1}), \dots , f(v_{u-i+1,r_{u-i+1}})$$

eine Basis dieses Faktorraums bilden. Die Vektoren von (i+1) sind dann gerade die Hauptvektoren der Ordnung u-i der gesuchten Jordanbasis.

Bemerkungen

(i) Nach dem (u-1)-ten Schritt hat man sämtliche Hauptvektoren sämtlicher zyklischer Basis bestimmt.

(ii) Im Fall K = $\mathbb{R}$ bzw. K = $\mathbb{C}$ kann man die beteiligen Vektorräume mit Räumen der Gestalt $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ identifizieren (durch Wahl einer Basis) und somit als Räume mit definiten Skalarprodukt bzw. hermitischen Skalarprodukt betrachten. In einen solchen Raum ist insbesondere das orthogonale Komplement $W^\perp$ eines Unterraums $W\subseteq V$ definiert. Die Einschränkung der natürlichen Abbildung $V \to V/W$ auf dieses orthogonale Komplement

(*) $$W^\perp \subseteq V \to V/W$$

hat den Kern $W^\perp \cap W = \{0\}$, d.h. die Abbildung ist injektiv. Wegen

$$\dim W^\perp = \dim V - \dim W = \dim V/W$$

ist die Abbildung sogar ein Isomorphismus, d.h. man kann V/W mit $W^\perp$ identifizieren.

(iii) In der Situation von (ii) ist die lineare Abbildung

$$W\oplus W^\perp \to V, (a,b) \mapsto a+b,$$

injektiv und bildet zwischen Räumen gleicher Dimension ab, d.h. sie ist ein Isomorphismus:

Aus a+b = 0 folgt a = -b $\in W\cap W^\perp = \{0\}$, also a = b = 0.

(iv) Im nächsten Abschnitt beschreiben wir die Berechnung der Hauptvektoren einer Jordan-Basis in der eben betrachteten Situation.

6.5.3 Bestimmung einer Jordan-Basis eines nilpotenten Endomorphismus auf einem Raum mit definiten oder hermitischen Skalarprodukt

Sei $f: V \to V$ eine nilpotente K-lineare Abbildung auf einem endlich-dimensionalen K-Vektorraum mit definieten oder hermitischen Skalarprodukt und sei

$$u = \mathrm{stab}(f).$$

Wir setzen

$$W^i := \ker(f^i) \text{ für } i = 1, \dots , u-1.$$

Die Unterräume W^i und ihre orthogonalen Komplemente sind dann f-invariant.

Wegen $W^1\subseteq W^2\subseteq\dots\subseteq W^{u-1} \subseteq W^u = V$ gilt

$$0 = (W^u)^\perp \subseteq (W^{u-1})^\perp \subseteq (W^{u-1})^\perp \subseteq \dots \subseteq (W^1)^\perp.$$

Zur Bestimmung einer Jordanbasis von V reicht es, die Hauptvektoren einer zyklischen Basis zu bestimmen für jeden direkten Summanden einer Zerlegung von V in eine direkte Summe zyklischer Unterräume.

Die höchste auftretende Ordung eines Hauptvektors ist dabei gleich u.

1. Schritt. Bestimmung der Hauptvektoren der Ordnung u.
Man wähle eine beliebige Basis

(1) $$v_{u,1}, \dots , v_{u,r_u} \in (W^{u-1})^{\perp}$$

des orthogonalen Komplements rechts (in V). Die Vektoren dieser Basis sind gerade die Hauptvektoren der Ordnung u.
2. Scbritt: Bestimmung der Hauptvektoren der Ordnung u-1.
Man ergänze die orthogonalen Projectionen Vektoren der

(2) $$v_{u,1}, \dots , v_{u,r_u}, f(v_{u,1}), \dots , f(v_{u,r_u})$$

auf

$$(W^{u-2})^{\perp}$$

zu einer Basis dieses Raumes durch Hinzufügen von Vektoren

$$v_{u-1,1}, \dots , v_{u-1,r_{u-1}} \in W^{u-1}$$

Die neuen Vektoren sind dann gerade die Hauptvektoren der Ordnung u-1 der gesuchten Jordanbasis.
3. Schritt: Bestimmung der Hauptvektoren der Ordnung n-2.
Man ergänze die orthogonalen Projectionen der Vektoren

(3) $$\begin{array}{l} v_{u,1}, \dots , v_{u,r_u}, f(v_{u,1}), \dots , f(v_{u,r_u}), f^2(v_{u,1}), \dots , f^2(v_{u,r_u}) \\ v_{u-1,1}, \dots , v_{u-1,r_{u-1}}, f(v_{u-1,1}), \dots ,f(v_{u-1,r_{u-1}}) \end{array}$$

auf

$$(W^{u-3})^{\perp}$$

zu einer Basis dieses Raumes durch Hinzufügen von Vektoren

$$v_{u-2,1}, \dots , v_{u-2,r_{u-2}} \in W^{u-2}$$

Die neuen Vektoren sind dann gerade die Hauptvektoren der Ordnung u-2 der gesuchten Jordanbasis.

Seien jetzt bereits die Hauptvektor bis zur Ordnung u-i+1 bestimmt:

$$\begin{array}{l} v_{u,1}, \dots , v_{u,r_u} \\ v_{u-1,1}, \dots , v_{u-1,r_{u-1}} \\ \dots \\ v_{u-i+1,1}, \dots , v_{u-i+1,r_{u-i+1}} \end{array}$$

i+1. Schritt: Besimmung der Hauptvektoren der Ordnung u-i.
Man ergänze die orthogonalen Projectionen der Vektoren

(i+1) $$\begin{array}{l} v_{u,1}, \dots , v_{u,r_u}, f(v_{u,1}), \dots , f(v_{u,r_u}),\dots, f^i(v_{u,1}), \dots , f^i (v_{u,r_u}) \\ v_{u-1,1}, \dots , v_{u-1,r_{u-1}},\dots, f^{i-1}(v_{u-1,1}), \dots ,f^{i-1}(v_{u-1,r_{u-1}}) \\ \dots \\ v_{u-i+1,1}, \dots , v_{u-i+1,r_{u-i+1}}, f(v_{u-i+1,1}), \dots , f(v_{u-i+1,r_{u-i+1}}) \end{array}$$

auf

$$(W^{u-i-1})^{\perp}$$
zu einer Basis dieses Raumes durch Hinzufügen von Vektoren
$$v_{u-i,1}, \dots , v_{u-i,r_{u-i}} \in W^{u-i}$$

Die neuen Vektoren sind dann gerade die Hauptvektoren der Ordnung u-i der gesuchten Jordanbasis.

6.5.4 Beispiel

$$A := \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Charakteristisches Polynom

$\chi_A(T) = \det (A - T\cdot Id) = 1-6T+15T^2-20T^3+15T^4-6T^5+T^6 = (1-T)^6$

Einziger Eigenwert ist 1

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

rk A - Id = 3

$$(A-Id)^2 = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

rk $(A-Id)^2 = 1$

Der Rang der dritten Potenz muß noch kleiner sein:

rk $(A-Id)^3 = 0$

Damit erhalten wir

$$\begin{aligned} \rho_1+2\rho_2+3\rho_3+4\rho_4+\dots &= 6 \\ \rho_2+2\rho_3+3\rho_4 &= 3 \\ \rho_3+2\rho_4 &= 1 \end{aligned}$$

also

$\rho_4 = 0, \rho_3 = 1, \rho_2 = 1, \rho_1 = 1$

Die Jordansche Normalform von A ist gleich

$$J_3(1)\oplus J_2(1)\oplus J_1(1) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Bestimmung der Hauptvektoren der Ordnung 3:

W^2 = $\ker(f_{(A\text{-}Id)^2})$ hat die Dimension 6 - rk $(A\text{-}Id)^2$ = 5

Es genügt also, 5 lineare unabhängige Vektoren aus W^2 zu finden:

W^2 = Lösung des Gleichungssystems $(A\text{-}Id)^2x = 0$

$= \mathbb{R}e_1+\mathbb{R}(e_2+e_3) + \mathbb{R}e_4+\mathbb{R}e_5+\mathbb{R}e_6$

$(W^2)^\perp = \mathbb{R}(e_2-e_3)$

Einziger Hauptvektor der Ordnung 3:

$$v_{3,1} = e_2-e_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Bestimmung der Hauptvektoren der Ordnung 2:

W^1 = $\ker(f_{A\text{-}Id})$ hat die Dimension 6 - rk (A-Id) = 3

Es genügt also, 3 lineare unabhängige Vektoren aus W^1 zu finden:

W^1 = Lösung des Gleichungssystems (A-Id) x = 0

$= \mathbb{R}e_1 + \mathbb{R}(e_4+e_5)+\mathbb{R}e_6$

$(W^1)^\perp$ $= \mathbb{R}e_2 + \mathbb{R}e_3+\mathbb{R}(e_4-e_5)$

(der Raum ist 3-dimensional, die angegebenen Erzeugenden sind orthogonal zu W^1 und linear unabhängig)

Wir müssen die orthogonalen Projektionen auf $(W^1)^\perp$ von

$$v_{3,1} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = e_2 - e_3$$

und

$$(A\text{-}Id)\, v_{3,1} = \begin{pmatrix} 1 \\ -2 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} = e_1 - 2\, e_2 - 2e_3$$

zu einer Basis von $(W^1)^\perp$ ergänzen, d.h.

$e_2 - e_3$ und $-2(e_2 + e_3)$

sind zu einer Basis von $(W^1)^\perp$ zu ergänzen (durch Hinzufügen eines Vektors aus W^2), d.h. gesucht ist das orthogonale Komplement dieser Vektoren in

$$\begin{aligned}(W^1)^\perp \cap W^2 &= (\mathbb{R} e_2 + \mathbb{R} e_3 + \mathbb{R}(e_4 - e_5)) \cap (\mathbb{R} e_1 + \mathbb{R}(e_2+e_3) + \mathbb{R} e_4 + \mathbb{R} e_5 + \mathbb{R} e_6) \\ &= \mathbb{R}(e_2+e_3) + \mathbb{R}(e_4-e_5)\end{aligned}$$

Einziger Hauptvektor der Ordnung 2:

$$v_{2,1} = e_4 - e_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

Bestimmung der Hauptvektoren der Ordnung 1:

Wir müssen die folgenden Vektoren zu einer Basis von $\mathbb{R}^6$ ergänzen:

$$v_{3,1} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = e_2 - e_3$$

$$(A\text{-}Id)\, v_{3,1} = \begin{pmatrix} 1 \\ -2 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} = e_1 - 2\, e_2 - 2e_3$$

$$(A\text{-}Id)^2\, v_{3,1} = \begin{pmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = -2e_1$$

$$v_{2,1} = \begin{pmatrix} 0\\0\\0\\1\\-1\\0 \end{pmatrix} = e_4 - e_5$$

$$(A\text{-}Id)\, v_{2,1} = \begin{pmatrix} 0\\0\\0\\-2\\-2\\0 \end{pmatrix} = 2(e_4+e_5)$$

(durch hinzufügen eines Vektors aus W^1):

$e_2 - e_3$, $e_1 - 2e_2 - 2e_3$, $-2e_1$, e_4-e_5, $2(e_4+e_5)$.

e_1, e_2, e_3, e_4, e_5

d.h. gesucht ist das orthogonale Komplement dieser Vektoren in

$W^1 = \mathbb{R}e_1 + \mathbb{R}(e_4+e_5) + \mathbb{R}e_6$.

Einziger Hauptvektor der Ordnung 1:

$$v_{2,1} = e_6 = \begin{pmatrix} 0\\0\\0\\0\\0\\1 \end{pmatrix}$$

Die gesuchte Jordanbasis (in umgekehrter Reihenfolge):

$$\begin{pmatrix} 0\\1\\-1\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\-2\\-2\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} -2\\0\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\0\\1\\-1\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\-2\\-2\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\1 \end{pmatrix}$$

6.6 Das Tensorprodukt

Das Tensorprodukt ist der theoretisch anspruchvollste Gegenstand dieser Vorlesung. Bei der Einführung des Tensorprodukts ist man in einer ähnlichen Situation wie bei der Einführung der Determinante: man kann zwar eine geschlossene Formel dafür angeben, aber diese ist so aufwendig in der Handhabe, daß man sie in praktischen Situationen nie benutzt, man verwendet sie nur für theoretische Betrachtungen.

Beim Tensorprodukt ist die explizite Beschreibung nicht einmal für theoretische Zwecke sinnvoll. Sie ist nur von Nutzen beim Beweis für die Existenz des Tensorprodukts.

6.6.0 Vorbemerkungen

(i) Unser nächstes Ziel ist die Betrachtung von bilinearen Abbildungen

$$b: U\times V \to W$$

mit beliebigen K-Vektorräumen U, V und W, d.h. von Abbildungen f mit

$f(c'u'+c''u'', v) = c'\cdot f(u',v) + c''\cdot f(u'',v)$

$f(u, c'v' + c''v'') \quad = c'\cdot f(u,v') + c''\cdot f(u,v'')$
für beliebige $u,u',u''\in U$, $v,v',v''\in V$ und $c',c''\in K$.

(ii) Genauer, wir wollen die allgemeinste Art von Abbildung dieser Gestalt finden, die es für diese Räumen geben kann.

(iii) Dabei ist die konkrete Konstruktion dieser Abbildung nicht besonders schön, relativ kompliziert und genaugenommen nicht so wichtig. Wichtiger ist ihre Eigenschaft, die ‘allgemeinste’ Abbildung zu sein und die Tatsache, daß es eine solche Bilinearform gibt.

(iv) Wir werden deshalb wie folgt vorgehen.

1. Wir beschreiben zunächst, was wir unter der ‘allgemeinsten’ bilinearen Abbildung verstehen wollen, indem wir deren sogenannte Universalitätseigenschaft angeben.
2. Wir zeigen, durch diese Eigenschaft ist die Konstruktion bis auf Isomorphie eindeutig bestimmt.
3. Wir beweisen unter der Annahme, daß das Konstrukt stets existiert, dessen wichtigste Eigenschaften.
4. Erst ganz zum Schluß werden wir zeigen, daß das Konstrukt tatsächlich existiert.

Zunächst wollen wir zur Illustration unserer Vorgehensweise eine ähnlich gelagerte Problemstellung betrachten, deren Lösung wir im wesentlichen bereits kennen.

6.6.1 Beispiel für eine Universalitätseigenschaft

Für jede K-lineare Abbildung $f:U \to V$ wollen wir eine K-lineare Abbildung

$$\rho:V \to \text{Coker}(f)$$

konstruieren, welche natürliche Abbildung auf den Kokern von f heißt. Dabei sollen folgende Bedingungen erfüllt sein.

1. $\rho\circ f = 0$.
2. Für jede K-lineare Abbildung $g:V\to W$ mit $\rho\circ g = 0$ soll es genau eine K-lineare Abbildung $\tilde{g}$: Coker(f) $\to$ W geben mit $g = \tilde{g}\circ f$, mit anderen Worten, eine solche lineare Abbildung, daß das folgende Diagramm kommutativ ist.

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "g"'] & \text{Coker}(f) \arrow[dl, "\tilde{g}"] \\ W & \end{tikzcd}$$

Bemerkungen zum Begriff der Universalitätseigenschaft

(i) Nach Bedingung 1 ist ρ eine Abbildung mit $\rho\circ f = 0$. Offensichtlich hat jede Zusammensetzung

$$V \xrightarrow{\rho} \text{Coker}(f) \xrightarrow{h} W$$

mit einer beliebigen (linearen) Abbildung h ebenfalls diese Eigenschaft,.

(ii) Bedingung 2 besagt gerade, daß man durch Zusammensetzen mit solchen Abbildungen h sämtliche Abbildungen bekommt deren Zusammensetzung mit f Null ist. Und zwar auf genau eine Weise (d.h. verschiedene h liefern verschiedene Zusammensetzungen).

(ii) Bedingung 2 bedeutet also gerade, daß die Abbildung ρ in dem Sinne ‘universell’ ist, daß man aus ihr jede andere Abbildung mit der Eigenschaft 1 gewinnen kann, und zwar auf genau eine Weise. Man sagt in einer solche Situation, ρ ist universell bezüglich Eigenschaft 1. Anders ausgedrückt, 2 ist eine Universalitätseigenschaft.

(iii) Die obige Beschreibung des Raumes

$$\text{Coker}(f)$$

entspricht gerade dem ersten Schritt, wie wir ihn für das Tensorprodukt angekündigt haben. Wir zeigen hier zunächst, daß Coker(f) durch die obigen beiden Eigenschaften bereits bis auf Isomorphie eindeutig bestimmt ist. Anschließend beweisen wir die Existenz von (ρ und) Coker(f).

<u>Die Eindeutigkeit von Coker(f) bis auf Isomorphie</u>. Sei eine weitere K-lineare Abbildung

$$\rho': V \to C'$$

gegeben, für welche die Bedingungen 1 und 2 mit C' anstelle von C := Coker(f) erfüllt sind.

Dann gilt $\rho'\circ f = 0$. Wir wenden 2 an mit $g := \rho'$ und erhalten die Existenz einer K-linearen Abbildung

$$\tilde{\rho}':C \to C' \text{ mit } \rho' = \tilde{\rho}'\circ f.$$

Indem wir die obige Argumentation bezüglich ρ und ρ' mit vertauschten Rollen wiederholen,, erhalten wir die Existenz einer linearen Abbildung

$$\tilde{\rho}:C' \to C \text{ mit } \tilde{\rho}' = \tilde{\rho}\circ f.$$

Damit erhalten wir kommutative Diagramme

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "\rho'"'] & C \arrow[dl, "\tilde{\rho}'"] \\ C' & \end{tikzcd} \quad \text{und} \quad \begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "\rho'"'] & C \\ C' \arrow[ur, "\tilde{\rho}"'] & \end{tikzcd}$$

und durch Zusammensetzen dieser Dreiecke weiterhin kommutative Diagramme

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "\rho"'] & C \arrow[dl, "u"] \\ C & \end{tikzcd} \quad \text{und} \quad \begin{tikzcd} V \arrow[r, "\rho'"] \arrow[d, "\rho'"'] & C' \arrow[dl, "u'"] \\ C' & \end{tikzcd}$$

mit $u := \tilde{\rho}\circ\tilde{\rho}'$ und $u' := \tilde{\rho}'\circ \tilde{\rho}$. Auf Grund von Eigenschaft 2 (mit g:=f) ist die Abbildung u durch die Kommutativität des ersten Diagramms aber eindeutig bestimmt. Da das Diagramm kommutativ bleibt, wenn man u durch die identische Abbildung ersetzt, folgt

$$\tilde{\rho}\circ\tilde{\rho}' = u = \mathrm{Id}.$$

Dieselbe Argumentation mit dem zweiten Diagramm liefert

$$\tilde{\rho}'\circ \tilde{\rho} = u = \mathrm{Id}.$$

Mit anderen Worten, $\tilde{\rho}$ und $\tilde{\rho}'$ sind zueinander inverse Isomorphismen und die Räume C und C' sind isomorphe (sogar in eindeutig bestimmter Weise!).

<u>Existenz von Coker(f)</u>. Wir setzen

$$\mathrm{Coker}(f) := V/\mathrm{im}(f)$$

und verwenden für ρ die natürliche Abbildung

$$\rho:V \to V/\mathrm{im}(f),\ v \mapsto v + \mathrm{im}(f),$$

auf den Faktorraum,

$$\rho(v) := v + \mathrm{im}(f).$$

Dann ist Bedingung 1 offensichtlich erfüllt. Beweisen wir, daß auch 2 gilt. Sei also eine K-lineare Abbildung

$$g: V \to W$$

gegeben mit $g\circ f = 0$. Wir haben zu zeigen, es gibt genau eine K-lineare Abbildung

$$\tilde{g}:V/\mathrm{im}(f) \to W$$

mit $\tilde{g}\circ f = g$. Falls $\tilde{g}$ existiert, so muß gelten,

$$\tilde{g}(v+im(f)) = \tilde{g}(f(v)) = g(v),$$

mit andern Worten, der Wert von $\tilde{g}$ an der Stelle v+im(f) ist eindeutig bestimmt.

Wir haben noch die Existenz von $\tilde{g}$ zu beweisen. Wir setzen

(*) $$\tilde{g}(v+im(f)) := g(v).$$

Falls wir zeigen können, daß diese Definition korrekt ist, so sind wir fertig, denn dannn gilt

$$\tilde{g}(f(v)) = g(v) \text{ für alle } v\in V$$

(und offensichtlich ist $\tilde{g}$ eine lineare Abbildung).
Beweisen wir die Korrektheit der Definition (*). Seien v,v'∈V zwei Vektoren mit

$$v + im(f) = v' + im(f).$$

Wir haben zu zeigen, daß dann g(v) = g(v') gilt. Auf Grund der Voraussetzung gilt

$$v-v' \in im(f).$$

Wegen $g\circ f = 0$ gilt $g|_{Im(f)} = 0$, d.h.

$$0 = g(v-v') = g(v) - g(v'),$$

also g(v) = g(v').
QED.

6.6.2 Definition des Tensorprodukts zweier K-Vektorräume

Seien V und W zwei K-Vektorräume. Das Tensorprodukt von V und W ist ein K-Vektoraum

$$V\otimes W = V\otimes_K W$$

zusammen mit einer K-bilinearen Abbildung

$$\rho=\rho_{V,W}:V\times W \to V\otimes W,\ (v,w) \mapsto \rho(v,w) =: v\otimes w$$

wobei folgende Bedinung erfüllt ist.

(⊗) Für jede K-bilineare Abbildung $b:V\times W\to U$ mit Werten in einem K-Vektorraum U gibt es genau eine K-lineare Abbildung $\tilde{b}:V\otimes W\to U$ mit $b = \tilde{b}\circ\rho$, d.h. deart, daß das folgende Diagramm kommutative ist.

$$\begin{array}{ccc} V\times W & \xrightarrow{b} & U \\ \rho\downarrow & \nearrow\tilde{b} & \\ V\otimes W & & \end{array}$$

Die Elemente von V⊗W heißen <u>Tensoren</u>.

Bemerkungen

(i) Bedingung (⊗) besagt gerade, jede bilineare Abbildung $b:V\times W\to U$ soll sich eindeutig über V⊗W faktorisieren.

(ii) Setzt man die bilineare Abbildung $\rho:V\times W\to V\otimes W$ mit einer linearen Abbildung $V\otimes W\to U$ zusammen, so erhält man trivialerweise eine bilineare Abbildunge $V\times W\to U$. Bedingung (⊗) besagt gerade, daß man auf diese Weise jede auf V×W definierte bilineare Abbildung erhält erhält, und zwar jede auf genau eine Weise.

(iii) Bedingung (⊗) ist äquivaltent zu der Aussage, daß die folgende lineare Abbildung bijektiv ist.

$$Hom_K(V\otimes W,U) \to L(V,W,U),\ \tilde{b} \mapsto \rho\circ \tilde{b}.$$

Dabei bezeichne

$$L(V,W;U)$$

den Vektorraum der K-bilinearen Abbildungen $V\times W\to U$.

(iv) Wir zeigen als nächstes, daß das Tensorprodukt, falls es existiert, bis auf Isomorphie eindeutig bestimmt ist. Danach werden wir die wichtigsten

Eigenschaften des Tensorprodukt unter der Annahme, daß es existiert, ableiten. Seine Existenz beweisen wir ganz zum Schluß.

6.6.3 Eindeutigkeit des Tensorprodukts bis auf Isomorphie

Seien V,W zwei K-Vektorräume und

$$b: V\times W \to U \text{ und } b': V\times W \to U'$$

zwei bilineare Abbildungen, welche die Eigenschaft (⊗) eines Tensorprodukts besitzen. Dann gibt es genau einen K-linearen Isomorphismus $f:U\to U'$, für welche das folgende Diagramm kommutativ ist.

$$\begin{tikzcd} V\times V \arrow[r, "b"] \arrow[d, "b'"'] & U \arrow[dl, "\cong"', "f"] \\ U' & \end{tikzcd}$$

d.h. es gilt $b' = f\circ b$.

Beweis. Auf Grund der Universalitätseigenschaft (⊗) von b gibt es zumindest eine K-lineare Abbildung $f:U\to U'$ mit der geforderten Eigenschaft,

$$\begin{tikzcd} V\times V \arrow[r, "b"] \arrow[d, "b'"'] & U \arrow[dl, "f"] \\ U' & \end{tikzcd}$$

Weiterhin gibt es aber auch (auf Grund der Universaltätseigenschaft (⊗) von b') eine K-lineare Abbildung f' sodaß

$$\begin{tikzcd} V\times V \arrow[r, "b"] \arrow[d, "b'"'] & U \\ U' \arrow[ur, "f'"'] & \end{tikzcd}$$

kommutativ ist. Durch Zusammensetzen dieser kommutativen Dreiecke erhalten wir kommutative Diagramme

$$\begin{tikzcd} V\times V \arrow[r, "b'"] \arrow[d, "b'"'] & U' \arrow[dl, "u'"] \\ U' & \end{tikzcd} \qquad \begin{tikzcd} V\times V \arrow[r, "b"] \arrow[d, "b"'] & U \arrow[dl, "u"] \\ U & \end{tikzcd}$$

mit $u := f\circ f'$ und $u' := f'\circ f$. Auf Grund der Eindeutigkeitsaussage von (⊗) sind die linearen Abbildungen u und u' durch die Kommutativität dieser Diagramme eindeutig festgelegt. Da die Diagramme aber kommutative bleiben, wenn man u und u' durch die identischen Abbildungen ersetzt, so gilt

$$f\circ f' = u = Id \text{ und } f'\circ f = u' = Id.$$

Die Abbildungen f und f' sind folglich zueinander inverse Isomorphismen.
QED.

Bemerkungen

(i) Aus der obigen Argumentation ergibt sich, daß jede K-lineare Abbildung f, für welche das Diagramm

$$\begin{tikzcd} V\times V \arrow[r, "b"] \arrow[d, "b'"'] & U \arrow[dl, "f"] \\ U' & \end{tikzcd}$$

kommutativ ist, automatisch ein Isomorphismus ist.

(ii) Im folgenden nehmen wir an, daß das Tensorprodukt zweier K-Vektorräume V und W stets existiert. Das Bild des Paares $(v,w)\in V\times W$ bei der natürlichen Abbildung

$$V\times W \to V\otimes W$$

bezeichnen wir wie in der Definition mit $v\otimes w$, d.h. die natürliche Abbildung hat gerade die Abbildungsvorschrift

$$V\times W \to V\otimes W,\ (v,w) \mapsto v\otimes w.$$

Die Bilinearität der natürlichen Abbildung bedeutet gerade, daß die folgenden Rechenregeln gelten.

(a) $(v'+v'')\otimes w = v'\otimes w + v''\otimes w$

(b) $v\otimes(w'+w'') = v\otimes w' + v\otimes w''$

(c) $(cv)\otimes w = c(v\otimes w) = v\otimes(cw)$

für beliebige $v,v',v''\in V$, $w,w',w''\in W$, $c\in K$.

Wir beweisen als nächstes die wichtigsten Eigenschaften des Tensorprodukts

6.6.4 Ein Erzeugendensystem für V⊗W

Seien $(v_i)_{i\in I}$ und $(w_j)_{j\in J}$ Erzeugendensysteme der K-Vektorräume V bzw. W. Dann bilden die Vektoren der Gestalt

$$v_i\otimes w_j$$

mit $i\in I$ und $j\in J$ ein Erzeugendensystem des Vektorraums $V\otimes W$.

Beweis. Sei $U \subseteq V\otimes W$ der von den Vektoren $v_i\otimes w_j$ erzeugte K-lineare Unterraum. Jeder Vektor $v\in V$ ist Linearkombination der v_i und jeder Vektor $w\in W$ ist Linearkombination der w_j . Also ist $v\otimes w$ Linearkombination der $v_i\otimes w_j$. Mit anderen Worten, für jedes $v\in V$ und jedes $w\in W$ gilt

$$v\otimes w \in U.$$

Die Abbildungsvorschrift der natürlichen Abbildung

$$\rho: V \times W \to V\otimes W,\ (v,w) \mapsto v\otimes w$$

definiert also auch eine bilineare Abbildung

$$b: V\times W \to U,\ (v,w) \mapsto v\otimes w.$$

Insbesondere hat man ein kommutatives Diagramm

(1)
$$\begin{tikzcd} V\times W \arrow[r, "b"] \arrow[d, "\rho"'] & U \arrow[dl, "i"] \\ V\otimes W & \end{tikzcd}$$

wenn $i: U \to V\otimes W$ die natürliche Einbettung bezeichnet. Zum Beweis der Behauptung reicht es zu zeigen, die natürliche Einbettung i ist surjektiv, denn dann gilt

$$V\otimes W = U = \langle u_i\otimes v_j \mid i\in I, j\in J\rangle.$$

Weil b bilinear ist, ergibt sich aus der Universalitätseigenschaft von ρ die Existenz einer linearen Abbildung $\tilde{b}$, für welches das Diagramm

(2)
$$\begin{tikzcd} V\times W \arrow[r, "b"] \arrow[d, "\rho"'] & U \\ V\otimes W \arrow[ur, "\tilde{b}"'] & \end{tikzcd}$$

kommutativ ist, d.h. mit

$$\tilde{b}(v\otimes w) = v\otimes w.$$

Durch Zusammensetzen der Diagramme (1) und (2) erhalten wir ein kommutatives Diagramm

$$
\begin{array}{ccc}
V\times W & \xrightarrow{\rho} & V\otimes W \\
\rho\downarrow & \nearrow i\circ\tilde{b} & \\
V\otimes W & &
\end{array}
\qquad (2)
$$

Auf Grund der Eindeutigkeitsaussage der Universalitätseigenschaft von ρ muß dann aber

$$i\circ\tilde{b} = Id$$

gelten. Für jedes $t \otimes V\otimes W$ gilt also

$$t = Id(t) = i(\tilde{b}(t)) \in Im(i),$$

d.h. es ist

$$Im(i) = V\otimes W.$$

Die natürliche Einbettung $i : U \to V\otimes W$ ist somit surjektiv.
QED

6.6.5 Eigenschaften des Tensorprodukts von Räumen

Seien U, V, W beliebige K-Vektorräume. Dann gelten die folgenden Aussagen.

(i) Es gibt genau eine K-lineare Abbildung $V\otimes K\to V$ mit

$$v\otimes c \mapsto cv.$$

Diese ist ein Isomorphismus

$$V\otimes K \cong V.$$

(ii) Es gibt genau eine K-lineare Abbildung $V\otimes W \to W\otimes V$ mit

$$v\otimes w \mapsto w\otimes v.$$

Diese ist ein Isomorphismus,

$$V\otimes W \cong W\otimes V$$

(iii) Es gibt genau eine K-lineare Abbildung $U\otimes(V\otimes W) \to (U\otimes V)\otimes W$ mit

$$u\otimes(v\otimes w) \mapsto (u\otimes v)\otimes w.$$

Diese ist ein Isomorphismus

$$U\otimes(V\otimes W) \cong (U\otimes V)\otimes W..$$

(iv) Es gibt genau eine bilineare Abbildung $U\otimes(V\oplus W) \to (U\otimes V)\oplus(U\otimes W)$ mit

$$u\otimes(v,w) \mapsto (u\otimes v, u\otimes w).$$

Diese ist ein Isomorphismus

$$U\otimes(V\oplus W) \cong (U\otimes V)\oplus(U\otimes W)..$$

Beweis. Zu (i). Wir betrachten die bilineare Abbildung

$$m: V\times K \to V, (v,c) \mapsto cv.$$

Können wir zeigen, daß m die Eigenschaft ($\otimes$) des Tensorprodukts hat, so gibt es nach 6.6.3 genau eine K-lineare Abbildung $f:V\otimes K \to V$ mit

(1) $$m(v,c) = f(\rho(v,c)) = f(v\otimes c),$$

und diese Abbildung f ist ein Isomorphismus. Wegen (1) ist

$$f(v\otimes c) = m(v,c) = cv,$$

d.h. f ist gerade die Abbildung, von der in (i) die Rede ist. Es genügt also, wenn wir zeigen, die Abbildung m hat die Eigenschaft ($\otimes$).

Sei also $b: V\times K \to P$ eine bilineare Abbildung. Dann ist die folgende Abbildung K-linear.

$$\tilde{b}: V \to P, v \mapsto b(v,1).$$

Weiter gilt

$$b(v,c) = c\cdot b(v,1) = c\cdot\tilde{b}(v) = \tilde{b}(cv) = \tilde{b}(m(v,c)),$$

mit anderen Worten b faktorisiert sich über m,

(2) $$b = \tilde{b}\circ m.$$

Wir haben noch zu zeigen, die lineare Abbildung $\tilde{b}$ ist durch Bedingung (2) eindeutig festgelegt. Aus (2) folgt für jedes $v\in V$,

$$\tilde{b}(v) = \tilde{b}(1\cdot v) = \tilde{b}(m(v,1)) = (\tilde{b}\circ m)(v,1) = b(1,v),$$

d.h. der Wert von $\tilde{b}$ an der Stelle v ist eindeutig festgelegt.
Zu (ii). Betrachten wir die Abbildung

$$b:V\times W \to W\otimes V,\ (v,w) \mapsto w\otimes v.$$

Nach der Bemerkung von 6.4.2 ist diese Abbildung bilinear. Auf Grund der Eigenschaft ($\otimes$) des Tensorprodukts

$$V\times W \to V\otimes W,\ (v,w) \mapsto v\otimes w,$$

gibt es damit genau eine K-lineare Abbildung

$$f: V\otimes W \to W\otimes V$$

mit $b(v,w) = f(v\otimes w)$, d.h. mit

$$f(v\otimes w) = w\otimes v.$$

Wir haben noch zu zeigen, f ist ein Isomorphismus. Aus Symmetriegründen gibt es aber auch genau eine K-lineare Abbildung

$$f': W\otimes V\to V\otimes W \text{ mit}$$

$f'(w\otimes v) = v\otimes w$. Damit gilt

$$(f\circ f')(w\otimes v) = w\otimes v$$

und

$$(f'\circ f)(v\otimes w) = v\otimes w,$$

d.h. die folgenden Diagramme sind kommutativ.

$$\begin{array}{ccc} V\times W & \xrightarrow{\otimes} & V\otimes W \\ \otimes\downarrow & \swarrow f'\circ f & \\ V\otimes W & & \end{array} \qquad \begin{array}{ccc} W\times V & \xrightarrow{\otimes} & W\otimes V \\ \otimes\downarrow & \swarrow f\circ f' & \\ W\otimes V & & \end{array}$$

Durch die Kommutativität dieser Diagramme sind die Abbildungen zu den schrägen Pfeilen aber eindeutig festgelegt. Also gilt

$$f'\circ f = Id \text{ und } f\circ f' = Id,$$

d.h. f und f' sind zueinander inverse Isomorphismen.
Zu (iii). Betrachten wir die Abbildung

$$U\times V\times W \to U\otimes(V\otimes W),\ (u,v,w) \mapsto u\otimes(v\otimes w).$$

Für jedes feste $w\in W$ ist diese Abbildung bilineare in u und v, faktorisiert sich also eindeutig über das Tensorprodukt

$$U\times V \to U\otimes V,\ (u,v) \mapsto u\otimes v.$$

Mit anderen Worten, es gibt genau eine (von w abhängige) K-lineare Abbildung

$$f_w: U\otimes V \to U\otimes(V\otimes W)$$

mit $f_w(u\otimes v) = u\otimes(v\otimes w)$. Betrachten wir die Abbildung

$$f:(U\otimes V)\times W \to U\otimes(V\otimes W),\ (t, w) \mapsto f_w(t).$$

Diese Abbildung ist linear in t. Zeigen wir, daß sie auch linear in w ist, d.h. daß gilt

$$f_{c'w'+c''w''}(t) = c'f_{w'}(t) + c''f_{w''}(t).$$

Zumindest stehen auf beiden Seiten der zu beweisenden Identität K-lineare Abbildungen

$$\varphi:U\otimes V \to U\otimes(V\otimes W)$$

und die Abbildung auf der linken Seite ist durch die folgende Bedinungung eindeutig festgelegt:

$$\varphi(u\otimes v) = u\otimes(v\otimes(c'w'+c''w'')).$$

Zum Beweis der Gleichheit reicht es folglich, wenn wir zeigen, die Abbildung auf der rechten Seite genügt derselben Bedingung. Sei also φ die Abbildung auf der rechten Seite. Dann gilt

$$\begin{aligned}\varphi(u\otimes v) &= c'f_{w'}(u\otimes v) + c''f_{w''}(u\otimes v)\\ &= c'\cdot u\otimes(v\otimes w') + c''u\otimes(v\otimes w'')\\ &= u\otimes c'\cdot(v\otimes w') + u\otimes c''\cdot(v\otimes w'')\\ &= u \otimes (v\otimes c'w') + u \otimes (v\otimes c''w'')\\ &= u \otimes (v\otimes c'w' + v\otimes c''w'')\\ &= u \otimes (v\otimes(c'w'+c''w''))\end{aligned}$$

Damit ist Bilinearität der Abbildung f gezeigt. Die Abbildung f faktorisiert sich damit über das Tensorprodukt

$$(U\otimes V)\times W \to (U\otimes V)\otimes W,\ (t, w) \mapsto t\otimes w,$$

d.h. es gibt genau eine lineare Abbildung

$$g: (U\otimes V)\otimes W \to U\otimes(V\otimes W)$$

mit $f_w(t) = g(t\otimes w)$. Da die Abbildung f_w bereits durch ihre Werte in den Vektoren der Gestalt $t = u\otimes v$ eindeutig festgelegt ist, gilt dasselbe für g, wobei

$$g((u\otimes v)\otimes w) = f_w(u\otimes v) = u\otimes(v\otimes w)$$

ist. Wir haben damit gezeigt, es gibt genau eine K-lineare Abbildung

$$g: (U\otimes V)\otimes W \to U\otimes(V\otimes W)$$

mit $g((u\otimes v)\otimes w) = u\otimes(v\otimes w)$. Wir haben noch zu zeigen, diese Abbildung ist ein Isomorphismus.

Ein ganz ähnliche Argumentation wie die eben angeführte zeigt, es gibt genau eine Abbildung

$$h: U\otimes(V\otimes W) \to (U\otimes V)\otimes W$$

mit

$$h(u\otimes(v\otimes w)) = (u\otimes v)\otimes w$$

für alle $u\in U$, $v\in V$, $w\in W$. Die beiden Formeln für g und h zeigen, die beiden Zusammensetzungen sind lineare Abbildungen

$$g\circ h: U\otimes(V\otimes W) \to U\otimes(V\otimes W)$$

$$h\circ g: (U\otimes V)\otimes W \to (U\otimes V)\otimes W$$

mit

$$g\circ h(u\otimes(v\otimes w)) = u\otimes(v\otimes w)$$

und

$$h\circ g((u\otimes v)\otimes w) = (u\otimes v)\otimes w$$

für alle $u\in U$, $v\in V$, $w\in W$. Nach 6.6.4 bilden aber die Vektoren der Gestalt

$$u\otimes(v\otimes w) \text{ bzw. } (u\otimes v)\otimes w$$

ein Erzeugendensystem des Vektorraums $U\otimes(V\otimes W)$ bzw. $(U\otimes V)\otimes W$. Die beiden Zusammensetzungen stimmen also auf einem Erzeugendensystem mit der identischen Abbildung überein, sind also gleich der identisdchen Abbildung. Wir haben gezeigt, g und h sind zueinander inverse Isomorphismen.

<u>Zu (iv)</u>. Betrachten wir die Abbildung

$$f: U\times(V\oplus W) \to (U\otimes V)\oplus(U\otimes W),\ (u, (v,w)) \mapsto (u\otimes v, u\otimes w).$$

Auf Grund der Bilinearität von $\otimes$ ist f ebenfalls bilinear:

$$\begin{aligned}f(c'u' + c''u'',(v,w)) &= ((c'u' + c''u'')\otimes v,\ (c'u' + c''u'')\otimes w)\\ &= (c'u'\otimes v + c''u''\otimes v,\ c'u'\otimes w + c''u''\otimes w)\\ &= (c'u'\otimes v,\ c'u'\otimes w) + (c''u''\otimes v,\ c''u''\otimes w)\\ &= c'f(u',(v,w)) + c''f(u'', (v,w))\\ f(u,\ c'(v',w') + c''(v'',w'')) &= f(u,\ (c'v'+c''v'',c'w'+c''w''))\\ &= (u\otimes(c'v'+c''v''),\ u\otimes(c'w'+c''w''))\\ &= (c'u\otimes v' + c''u\otimes v'',\ c'u\otimes w' + c''u\otimes w'')\\ &= c'(u\otimes v',\ u\otimes w') + c''(u\otimes v'',\ u\otimes w'')\\ &= c'f(u, (v',w')) + c''f(u, (v'',w''))\end{aligned}$$

Damit ist die Bilinearität von f bewiesen. Die Abbildung faktorisiert sich also eindeutig über die natürliche Abbildung ins Tensorprodukt, d.h. über

$$U\times(V\oplus W) \to U\otimes(V\oplus W),\ (u,(v,w)) \mapsto u\otimes(v,w).$$
Es gibt also genau eine lineare Abbildung

(3) $\tilde{f}: U\otimes(V\oplus W) \to (U\otimes V)\oplus(U\otimes W)$ mit $u\otimes(v,w) \mapsto (u\otimes v, u\otimes w)$.

Wir haben noch zu zeigen, diese Abbildung ist eine Isomorphismus. Dazu reicht eis zu zeigen, es gibt eine lineare Abbildung

(4) $\tilde{g}: (U\otimes V)\oplus(U\otimes W) \to U\otimes(V\oplus W)$ mit $u\otimes(v,w) \mapsto (u\otimes v, u\otimes w)$,

und die beiden Zusammensetzungen

(5) $\tilde{g}\circ\tilde{f}$; $U\otimes(V\oplus W) \mapsto U\otimes(V\oplus W),\ u\otimes(v,w) \mapsto u\otimes(v,w)$

(6) $\tilde{f}\circ\tilde{g}$: $(U\otimes V)\oplus(U\otimes W) \mapsto (U\otimes V)\oplus(U\otimes W),\ (u\otimes v, u\otimes w) \mapsto (u\otimes v, u\otimes w)$

sind gerade die Identischen Abbildungen.

Abbildung (5) ist dabei ganz offensichtlich die identische Abbildungen, denn bei (5) wird ein Erzeugendensystem von $U\otimes(V\oplus W)$ genauso abgebildet wie bei der identischen Abbildung. Um zu zeigen, auch (6) ist die identische Abbildung, reicht es zu zeigen, die Einschränkung von (6) auf jeden der beiden direkten Summanden ist die identische Abbildung. Diese beiden Einschränkungen bilden aber jeweils ein Erzeugendensystem so ab wie die identische Abbildung:

$$U\otimes V \to U\otimes V,\ u\otimes v \mapsto u\otimes v \text{ bzw. } U\otimes W\to U\otimes W,\ u\otimes w \mapsto u\otimes w.$$

Wir haben somit nur noch zu zeigen, daß die lineare Abbildung (4) existiert. Dazu wiederum reicht es zu zeigen, daß die beiden Einschränkungen auf die beiden direkten Summanden existieren:

$$U\otimes V \to U\otimes(V\oplus W),\ u\otimes v \mapsto (u\otimes v, 0),$$

$$U\otimes W \to U\otimes(V\oplus W),\ u\otimes w \mapsto (0, u\otimes w).$$

Diese beiden letzten Abbildungen existieren aber wegen der Universalitätseingenschaft des Tensorprodukts und der Bilinearität der beiden folgenden Abbildungen.

$$U\times V \to U\otimes(V\oplus W),\ (u, v) \mapsto (u\otimes v, 0),$$

$$U\times W \to U\otimes(V\oplus W),\ (u, w) \mapsto (0, u\otimes w).$$

QED.

6.6.6 Eigenschaften des Tensorprodukts von Elementen

Seien V und W zwei K-Vektorräume und $(v_i)_{i\in I}$ und $(w_j)_{j\in J}$ zwei Familien von Elementen auf V bzw. W. Dann gelten die folgenden Aussagen.

(i) Sind die v_i in V und die w_j in W lineare unabhängig, so sind es auch die $v_i\otimes w_j$ in $V\otimes W$.

(ii) Bilden die v_i ein Erzeugendensystem von V und die w_j eines von W, so bilden die $v_i\otimes w_j$ eines $V\otimes W$.

(iii) Bilden die v_i eine Basis von V und die w_j eine von W, so bilden die $v_i\otimes w_j$ eine $V\otimes W$.

Bemerkungen

Aussage (iii) bietet die Möglichkeit, die Existenz des Tensorproduktes zu beweisen: Man wähle in V und W jeweils eine Basis (v_i) bzw. (w_j) und definere $V\otimes W$ als den Raum mit der Basis $\{v_i\otimes w_j\}$. Man hat dann allerdings die Unabhängigkeit der Konstruktion von der Wahl der Basen zu beweisen. Unser Beweis wird von vornherein von unabhängig von jeder Basis sein.

Beweis. Aussage (ii) wurde bereit bewiesen (vgl. 6.6.4). Trivialerweise folgt (iii) aus (i) und (ii). Es würde also reichen, (i) zu beweisen. Umgekehrt folgt aber auch (i) aus (iii), denn jedes linear unabhängige Familie von Elementen läßt sich zu einer Basis ergänzen. Es reicht also, Aussage (iii) zu beweisen.

Zum Beweis von (iii) benutzen wir eine Konstruktion, die wir vor langer Zeit eingeführt haben: den von einer Menge frei erzeugte Vektorraum.

Wiederholung:

In 3.2.8 haben wir gezeigt, zu jeder beliebigen Menge M gibt es einen K-Vektorraum, der diese Menge als Basis besitzt und welcher mit

$$F(M) := F_K(M)$$

bezeichnet wird.

Als Menge verwenden wir die Menge aller Paare

$$M := \{(v_i, w_j) \mid i\in I, j\in J\}$$

von Vektoren unserer Ausgangsbasen. Der Vektorraum

$$F(M) = \{ \sum_{i\in I,j\in J} c_{ij}(v_i, w_j) \mid c_{ij}\in K, \text{fast alle } c_{ij} = 0\}$$

besteht dann aus allen endliche Linearkombinationen

$$u = \sum_{i\in I,j\in J} c_{ij}(v_i, w_j)$$

von solchen Paaren, wobei die Koeffizienten c_{ij} durch die Linearkombination u eindeutig festgelegt sind. Betrachten wir die Abbildung

$$\varphi: V\times W \to F(M), (\sum_{i\in I} c_i v_i , \sum_{j\in J} d_j w_j) \mapsto \sum_{i\in I,j\in J} c_i d_j(v_i, w_j).$$

Diese Abbildung ist bilinear. Als bilineare Abbildung faktorisiert sich φ über das Tensorprodukt, d.h. es gibt genau eine lineare Abbildung

$$\tilde{\varphi}: V\otimes W \to F(M) \text{ mit } v_i\otimes w_j \mapsto \varphi(v_i,w_j) = (v_i,w_j)$$

Wären nun die Vektoren $v_i\otimes w_j$ linear abhängig in $V\otimes W$, so müßten es auch deren Bilder bei der lineare Abbildung $\tilde{\varphi}$ sein, d.h. die Vektoren

$$\tilde{\varphi}(v_i\otimes w_j) = (v_i,w_j)$$

von F(M). Diese sind aber nach Konstruktion linear unabhängig.

QED.

6.6.7 Die Koordinaten eines Tensors

Seien V und W zwei K-Vektorräume und

$$(v_i)_{i\in I} \text{ und } (w_j)_{j\in J}$$

Basen von V bzw. W. Dann läßt sich jeder Tensor

$$t \in V\otimes W$$

in der Gestalt

$$t = \sum_{i\in I,J\in J} c^{ij} v_i\otimes w_j$$

schreiben mit eindeutig bestimmten $c_{ij}\in K$. die c_{ij} heißen <u>Koordinaten des Tensors</u> t bezüglich der gegebenen Basen,.

Sind endlich viele K-Vektorräume

$$V_1 , ... , V_r$$

gegeben und für jedes i eine Basis

$$\{ v_{j,i} \}_{j\in J_i}$$

von V_i so hat man für jeden Tensor

$$t \in V_1 \otimes ... \otimes V_r$$

eindeutig bestimmte Elemente

$$c_{i_1 i_2 ... i_r} \in K$$

mit

$$t = \sum_{i_1 \in J_1,...,i_r \in J_r} c^{i_1 i_2 ... i_r} v_{i_1,1} \otimes ... \otimes v_{i_r,r} .$$

Die $c_{i_1 i_2 ... i_r}$ heißen dann <u>Koordinaten des Tensors</u> t bezüglich der gegebenen Basen.

6.6.8 Das Verhalten der Koordinaten bei Basiswechsel

Seien

$$V_1 , ... , V_r$$

endlich viele K-Vektorräume und seien für jedes $i \in \{1,...,r\}$ zwei Basen

$$v_i := \{ v_{j,i} \}_{j \in J_i} \qquad v'_i := \{ v'_{j,i} \}_{j \in J_i}$$

von V_i gegeben. Wir betrachten die Koordinaten eines Tensors

$$t \in V_1 \otimes ... \otimes V_r$$

bezüglich der beiden Familien von Basen:

$$t = \sum_{i_1 \in J_1,...,i_r \in J_r} c^{i_1 i_2 ... i_r} v_{i_1,1} \otimes ... \otimes v_{i_r,r} = \sum_{i_1 \in J_1,...,i_r \in J_r} c'^{i_1 i_2 ... i_r} v'_{i_1,1} \otimes ... \otimes v'_{i_r,r} .$$

Bezeichne $A := M(Id) = (a_{ij,\ell})$ die Basiswechelsmatrix für den Übergang der Basis v_ℓ zur Basis v'_ℓ,

$$v_{j,\ell} = \sum_{j' \in J_\ell} a^{j'}_{j\ ,\ell} v'_{j',\ell}$$

Dann besteht zwischen den gestrichenen und den ungestrichenen Koordinaten von t die folgende Relation.

$$c'^{i'_1 i_2 ... i'_r} = \sum_{i_1 \in J_1,...,i_r \in J_r} a^{i'_1}_{i_1\ ,1} \cdot ... \cdot a^{i'_r}_{i_r\ ,r} c^{i_1 i_2 ... i_r}$$

Beweis. Es gilt

$$\begin{aligned} t &= \sum_{i_1 \in J_1,...,i_r \in J_r} c^{i_1 i_2 ... i_r} v_{i_1,1} \otimes ... \otimes v_{i_r,r} \\ &= \sum_{i_1 \in J_1,...,i_r \in J_r} c^{i_1 i_2 ... i_r} \Big(\sum_{i'_1 \in J_1} a^{i'_1}_{i_1\ ,1} v'_{i'_1,1} \Big) \otimes ... \otimes \Big(\sum_{i'_r \in J_r} a^{i'_r}_{i_r\ ,r} v'_{i'_r,r} \Big) \\ &= \sum_{i_1 \in J_1,...,i_r \in J_r} \sum_{i'_1 \in J_1,...,i'_r \in J_r} a^{i'_1}_{i_1\ ,1} \cdot ... \cdot a^{i'_r}_{i_r\ ,r} c^{i_1 i_2 ... i_r} v'_{i'_1,1} \otimes ... \otimes v'_{i'_r,r} \end{aligned}$$

Koeffizientenvergleich liefert die Behauptung.
QED.

6.6.9 Bemerkungen zum physikalische Fall

Seien V ein endlich-dimensionaler K-Vektorraum,

(1) $$v_1,...,v_n \in V$$

eine Basis von V und

(2) $$v^1,...,v^n \in V^*$$

die zugehörige duale Basis. Bezeichne

$$V^{\otimes r} \text{ bzw.} V^{*\otimes s}$$

das Tensorprodukt von r Exemplaren des Raumes V (bzw. s Exemplaren des Raumes V*). Ein r-fach <u>kovarianter</u> und s-fach <u>kontravarianter</u> <u>Tensor</u> im Sinne der Physik ist ein Element von

$$V^{\otimes r} \otimes V^{*\otimes s},$$

wobei man sich den Tensor durch dessen Koordinaten bezüglich der Basen (1) und (2) gegeben denkt.

Bemerkung

Ist

$$v'_1,...,v'_n \in V$$

eine zweite Basis von V,

$$v'^1,...,v'^n \in V^*$$

die zugehörige duale Basis und ist $A = M^v_{v'}(Id) = (a^i_j)$ die Basiswechselmatrix für den Übergang von v nach v', d.h.

$$v_i = \sum_{j=1}^n a^j_i v'_j .$$

Bezeichne $B = (b^i_j)$ die zu B inverse Matrix (d.h. die Basiswechselmatrix $B = M^{v'}_v(Id)$. Dann gilt

(1) $$v^j = \sum_{i=1}^n b^j_i v'^i$$

Für die Koordinaten eines Tensors $t \in V^{\otimes r} \otimes V^{*\otimes s}$,

$$t = \sum_{i_1\in J_1,...,i_r\in J_r} c^{i_1 i_2...i_r}_{j_1...j_s} v_{i_1} \otimes...\otimes v_{i_r} \otimes v^{j_1} \otimes...\otimes v^{j_s}$$

$$= \sum_{i_1\in J_1,...,i_r\in J_r} c'^{i_1 i_2...i_r}_{j_1...j_s} v'_{i_1} \otimes...\otimes v'_{i_r} \otimes v'^{j_1} \otimes...\otimes v'^{j_s}.$$

besteht die folgende Relation.

$$c'^{i_1 i'_2...i'_r}_{j'_1...j'_s} = \sum_{i_1\in J_1,...,i_r\in J_r} a^{i'_1}_{i_1} \cdot...\cdot a^{i'_r}_{i_r} b^{j_1}_{j'_1} \cdot...\cdot b^{j_r}_{j'_r} c^{i_1 i_2...i_r}_{j_1...j_s}$$

Beweis. Es reicht, (1) zu beweisen. Es reicht zu zeigen, die rechte Seite von (1) genügt den definierenden Bedingungen für die duale Basis. Es gilt

$$<v_i, \sum_{\alpha=1}^n b^j_\alpha v'^\alpha> = (\sum_{\alpha=1}^n b^j_\alpha v'^\alpha)(v_i) = (\sum_{\alpha=1}^n b^j_\alpha v'^\alpha)(\sum_{\beta=1}^n a^\beta_i v'_\beta)$$

$$= \sum_{\alpha=1}^n \sum_{\beta=1}^n b^j_\alpha a^\beta_i v'^\alpha(v'_\beta)$$

$$= \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} b^j_\alpha a^\beta_i \delta^\alpha_\beta$$

$$= \sum_{\alpha=1}^{n} b^j_\alpha a^\alpha_i$$

$$= \delta^j_i$$

QED.

Bemerkungen

(i) In der Physik betrachtet man im allgemeinen Koordinatenwechsel zwischen "krummlinigen Koordinaten", sagen wir

$$(x'^1, \ldots, x'^n) = f(x^1,\ldots,x^n).$$

An die Stelle der Matrix $A = (a^i_j)$ tritt dann die Matrix der Linearisierung von f,

$$a^i_j = \frac{\partial x'^i}{\partial x^j}$$

Für die Matrix $B = A^{-1}$ gilt dann (nach der Kettenregel)

$$b^i_j = \frac{\partial x^i}{\partial x'^j}.$$

(ii) Einsteinsche Summenkonvention. In jedem Ausdruck, in welchen ein und derselbe Index sowohl als oberer als auch als unterer Index auftritt, wird über diesen Index summiert (maximaler Summationsbereich).

Unter Verwendung dieser Konvention hätten wir in den obigen Rechnungen sämtliche Summenzeichen weglassen können.

6.6.9 Die Existenz des Tensorprodukts

Für je zwei K-Vektorräume V und W existiert das Tensorprodukt.

Vorbemerkung. Wir haben gesehen, falls das Tensorprodukt $V\otimes W$ existiert, so wird es von den Tensoren $v\otimes w$ mit $v\in V$ und $w\in W$ erzeugt. Mit anderen Worten, $V\otimes W$ ist ein Faktorraum des von den Vektoren $v\otimes w$ frei erzeugten Vektoraums. Wir nutzen jetzt diese Tatsache zur Konstruktion von $V\otimes W$, d.h. wir werden $V\otimes W$ als Faktorraum eines frei erzeugten Vektorraums definieren. Anstelle der Bezeichnung $v\otimes w$ werden wir für die Elemente des frei erzeugten Vektorraum das Symbol (v,w) wählen.

Beweis. Sei M die Menge $V\times W$ der Paare (v,w) mit $v\in V$ und $w\in W$,

$$M = V\times W.$$

Betrachten wir den von M frei erzeugten K-Vektorraum

$$F(M).$$

Dieser Vektorraum besteht aus allen endlichen K-Linearkombinationen von Paaren der Gestalt (v,w) mit $v\in V$ und $w\in W$,

$$c_1\cdot(v_1,w_1) + \ldots + c_r\cdot(v_r,w_r)$$

mit $c_i\in K$, $v_i\in V$, $w_i\in W$ für $i = 1,\ldots,r$.

Bemerkung. Man beachte, F(M) ist unendlich-dimensional, sobald M unendlich viele Elmente enthält, d.h. selbst wenn V und W endliche Dimension haben, kann die Dimension von F(M) unendlich sein, denn nach Konstruktion bilden die Elemente von M gerade eine Basis von F(M),

$$\dim F(M) = \# M.$$

Wir konstruieren jetzt den Unterraum R, nach den wir den Raum F(M) Faktorisieren wollen. Der Raum R werde von allen Vektoren der folgenden Gestalt erzeugt.

(1) $(v, w' + w'') - (v,w') - (v,w'')$ mit $v\in V$, $w',w''\in W$

(2) $(v' + v'', w) - (v',w) - (v'',w)$ mit $v',v''\in V$, $w\in W$
(3) $(cv, w) - c(v,w)$ mit $c\in K$, $v\in V$, $w\in W$
(4) $(v, cw) - c(v,w)$ mit $c\in K$, $v\in V$, $w\in W$.
Dabei haben wir vereinfachend (v,w) anstelle von $1\cdot(v,w)$ geschrieben für $v\in V$, $w\in W$.

Bemerkung. Dieser Unterraum R ist im allgemeinen ebenfalls sehr groß. Wir werden sehen, er ist so groß, daß der Faktoraum F(M)/R im Fall von endlich-dimensionalen Räumen endlich-dimensional wird.

Wir setzen

$$V\otimes W := F(M)/R.$$

Bezeichne

$$\gamma: F(M) \to F(M)/R$$

die natürliche Abbildung. Weiter setzen wir

$$v\otimes w := \gamma((v,w)).$$

Wir haben zu zeigen, die Abbildung

$$\varphi: V\times W \to F(M)/R, (v,w) \mapsto \gamma((v,w)) = v\otimes w$$

ist bilinear und hat die Universalitätseigenschaft des Tensorprodukts.
Linearität von φ bezüglich der ersten Variablen. Wir haben zu zeigen
1. $\varphi(v'+v'', w) - \varphi(v',w) - \varphi(v'',w) = 0$.
2. $\varphi(cv,w) - c\cdot\varphi(v,w) = 0$.
Die linke Seite von 1. ist gleich

$$\varphi(v'+v'', w) - \varphi(v',w) - \varphi(v'',w) \qquad = \gamma((v'+v'',w) - (v',w) - (v',w))$$

Das Argument von γ auf der rechten Seite ist gerade ein Element der Gestalt (2), liegt also im Unterraum R. Da der Kern von γ gerade der Unterraum R ist, steht auf der rechten Seite der Nullvektor.
Die linke Seite von 2. ist gleich

$$\varphi(cv,w) - c\cdot\varphi(v,w) = \gamma((cv,w) - c\cdot(v,w))$$

Das Argument von γ auf der rechten Seite ist gerade ein Element der Gestalt (3), liegt also im Unterraum R. Da der Kern von γ gerade der Unterraum R ist, steht auf der rechten Seite der Nullvektor.
Linearität von φ bezüglich der zweiten Variablen. Man verwendet dieselben Argumente wie beim Beweis der Linearität bezüglich der ersten Variablen, wobei man die Elemente der Gestalt (1) und (4) von R (anstelle der Elemente der Gestalt (2) und (3)) benutzt.
Die Unversalitätseigenschaft der Abbildung φ. Wir haben zu zeigen, jede K-lineare Abbildung

$$b: V\times W \to U$$

faktorisiert sich eindeutig über die Abbildung φ, d.h. zu gegebenen b gibt es genau ein lineare Abbildung $\tilde{b}: F(M)/R \to U$ mit

$$b(v,w) = \tilde{b}(\varphi(v,w)) \tag{5}$$

für alle $v\in V$ und alle $w\in W$.

Beweis der Eindeutigkeit von $\tilde{b}$.Nach Konstruktion bilden die Vektoren der Gestalt (v,w) ein Erzeugendensystem von F(M). Deshalb bilden die Bilder der (v,w) bei der natürlichen Surjektion

$$\gamma: F(M) \to F(M)/R$$

ein Erzeugendensystem von F(M)/R, d.h. die Elemente

$$\gamma((v,w)) = \varphi(v,w) \text{ mit } v\in V \text{ und } w\in W$$

bilden ein Erzeugendensystem von F(M)/R. Bedingung (5) (rückwärts gelesen) legt daher die Werte von $\tilde{b}$ auf einem Erzeugendensystem von F(M)/R fest. Da $\tilde{b}$ linear sein soll, ist damit die gesamte Abbildung $\tilde{b}$ festgelegt.

Bemerkungen. (i) Zum Beweis der Existenz von $\tilde{b}$ könnten wir (5) als Definition verwenden und dann die Korrektheit der Defintion beweisen. Obwohl man auf diese

Weise durchaus zum Ziel kommt, wollen wir hier anders vorgehen, um die bereits bewiesenen Aussagen etwas effektiver nutzen zu können.

(ii) Aus der Existenz der Abbildung $\tilde{b}$: F(M)/R $\to$ U folgt natürlich auch die Existenz der Zusammensetzung von $\tilde{b}$ mit der natürlichen Abbildung γ:F(M) $\to$ F(M)/R. Wir werden zunächst diese Zusammensetzung konstruieren und mit deren Hilfe die Existenz von $\tilde{b}$ beweisen.

Beweis der Existenz von $\tilde{b}$. Betrachten wir die K-lineare Abbildung

$$b': F(M) \to U \text{ mit } (v,w) \mapsto b(v,w).$$

Da die Paare der Gestalt (v,w) eine Basis von F(M) bilden, gibt es genau eine lineare Abbildung, die für jedes $v \in V$ und jedes $w \in W$ in (v,w) den vorgegebenen Wert b(v,w) annimmt. Es reicht zu zeigen, der Unterraum $R \subseteq F(M)$ liegt im Kern von b',

(6) $$R \subseteq \ker(b'),$$

denn auf Grund der von uns bewiesenen Universalitätseigenschaft des Faktorraum faktorisiert sich dann b' in eindeutiger Weise über die natürliche Abbildung γ:F(M)$\to$F(M)/R. Genauer, es gibt genau eine lineare Abbildung

$$\tilde{b}': F(M)/R \to U$$

mit $b'(x) = \tilde{b}'(\gamma(x))$ für alle $x \in F(M)$. Speziell für Elemente der Gestalt x = (v,w) erhalten wir

$$b(v,w) = b'((v,w)) = \tilde{b}'(\gamma((v,w)) = \tilde{b}'(\varphi(v,w)).$$

Die Abbildung $\tilde{b}'$ ist somit gerade die von uns gesuchte Abbildung $\tilde{b}$.

Wir haben noch (6) zu beweisen. Dazu reicht es zu zeigen, die Abbildung b' bildet die Elemente eines Erzeugendensystems von R in die Null ab. Es genügt somit zu zeigen, daß die Elemente der Gestalt (1), (2), (3), (4) bei b' in die Null abgebildet werden. Das ist aber gerade eine Folge der Bilinearität von b. Zum Beispiel ist

$$\begin{aligned} b'((v, w' + w'') - (v,w') - (v,w'')) &= b(v, w' + w'') - b(v,w') - b(v,w'') \\ &= 0. \end{aligned}$$

Das erste Gleichheitszeichen besteht nach Definition von b', das zweite wegen der Liniearität von b bezüglich des zweiten Arguments.

QED.

Bemerkung

Im Fall endlich-dimensionaler Vektorräume V und W ist der Existenzbeweis für das Tensorprodukt einfacher. Die Definition des Tensorprodukt besagt gerade, $V\otimes W$ ist ein Vektorraum mit

$$\mathrm{Hom}_K(V\otimes W, K) = L(V,W;K).$$

Mit anderen Worten, der zu $V\otimes W$ duale Vektorraum ist gerade L(V,W,K). Da man im Fall von endlich-dimensionalen Räumen das doppelte Dual eines Vektorraum mit dem Ausgangsraum identifizieren kann, folg

$$V\otimes W = L(V,W;K)^*$$

für dim V $< \infty$ und dim W $< \infty$. Dies kann man im endlich-dimensionalen Fall als Definition verwenden. Es ist nicht schwer, einzusehen, daß $L(V,W;K)^*$ bezüglich der bilinearen Abbildung

$$V\times W \to L(V,W; K)^*, (v,w) \mapsto (b \mapsto b(v,w))$$

tatsächlich die Universalitätseigenschaft des Tensorprodukts hat.

6.6.10 Die Funktorialität des Tensorprodukts

(i) Seien f: V$\to$V' und g: W$\to$W' zwei K-lineare Abbildungen. Dann gibt es genau eine K-lineare Abbildung $f\otimes g$, welche das folgenden Diagramm kommutativ ist.

$$\begin{array}{ccc} V\times W & \xrightarrow{f\times g} & V'\times W' \\ \rho_{V,W}\downarrow & & \downarrow\rho_{V',W'} \\ V\otimes W & \xrightarrow{f\otimes g} & V'\otimes W' \end{array} \qquad \begin{array}{ccc} (v,w) & \mapsto & (f(v),g(w)) \\ \downarrow & & \downarrow \\ v\otimes w & \mapsto & f(v)\otimes g(w) \end{array}$$

Mit anderen Worten, $f\otimes g$ ist die eindeutig bestimmte lineare Abbildung mit

$$(f\otimes g)(v\otimes w) = f(v)\otimes g(w) \text{ für } v\in V \text{ und } w\in W.$$

(ii) Für beliebige K-lineare Abbildungen $f: V\to V'$, $f':V'\to V''$, $g:W\to W'$, $g':W'\to W''$ gilt

$$(f'\otimes g')\circ(f\otimes g) = (f'\circ f)\otimes(g'\circ g).$$

(iii) Das Tensorprodukt der beiden identischen Abbildungen

$$Id_V : V \to V \text{ und } Id_W: W \to W$$

ist die identische Abbildung von $V\otimes W$,

$$Id_V\otimes Id_W = Id_{V\otimes W} .$$

Beweis. Zu (i). Da f und g linear sind und $\rho_{V',W'}$ bilinear ist, ist die Zusammensetzung

$$\rho_{V',W'}\circ(f\times g)$$

bilinear. Die Existenz und Eindeutigkeit von $f\otimes g$ folgt deshalb aus der Universalitätseigenschaft von $\rho_{V,W}$.

Zu (ii). Für $v\in V$ und $w\in W$ gilt

$$\begin{aligned} ((f'\otimes g')\circ(f\otimes g))(v\otimes w) &= (f'\otimes g')(f(v)\otimes g(w)) \\ &= f'(f(v))\otimes g'(g(w)) \\ &= (f'\circ f(v))\otimes(g'\circ g(w)) \\ &= ((f'\circ f)\otimes(g'\circ g))(v\otimes w). \end{aligned}$$

Die beiden linearen Abbildungen $(f'\otimes g')\circ(f\otimes g)$ und $(f'\circ f)\otimes(g'\circ g)$ haven für alle Vektoren der Gestalt $v\otimes w$ mit $v\in V$ und $w\in W$ denselben Wert. Da die $v\otimes w$ eine Basis von $V\otimes W$ bilden, folgt

$$(f'\otimes g')\circ(f\otimes g) = (f'\circ f)\otimes(g'\circ g).$$

Zu (iii). Für $v\in V$ und $w\in W$ gilt

$$(Id_V\otimes Id_W)(v\otimes w) = (Id_V(v))\otimes(Id_W(w)) = v\otimes w,$$

d.h. die lineare Abbildung $Id_V\otimes Id_W$ hat auf allen Vektoren des Erzeugendensystems

$$\{v\otimes w\}_{v\in V,w\in W}$$

dieselben Werte wie die identische Abbildung $Id_{V\otimes W}$. Deshalb gilt

$$Id_V\otimes Id_W = Id_{V\otimes W} .$$

QED.

6.6.11 Exake Sequenzen

Eine exakte Sequenz von Vektorräumen und linearen Abbildungen ist eine Folge von linearen Abbildungen

$$\ldots \to V_i \xrightarrow{f_i} V_{i+1} \xrightarrow{f_{i+1}} V_{i+2} \to \ldots$$

mit

$$\text{im } f_i = \ker f_{i+1} \text{ für alle i.}$$

Beispiel 1

Die Sequenz von linearen Abbildungen

$$0 \to V' \xrightarrow{f'} V \xrightarrow{f''} V'' \to 0$$

ist genau dann exakt, wenn folgenden Bedingungen erfüllt sind.
1. f' ist injektiv (Exaktheit an der Stelle V')
2. im f' = ker f'' (Exaktheit an der Stelle V)
3. f'' ist surjektiv.
Man spricht in dieser Situation von einer kurzen exakten Sequenz.

Beispiel 2
Die Sequenz

$$0 \to V' \xrightarrow{f'} V'\oplus V'' \xrightarrow{f''} V'' \to 0$$

mit $f'(v') = (v', 0)$ und $f''(v', v'') = v''$ ist exakt.

6.6.12 Exaktheit des Tensorprodukts

Seien W ein K-Vektorraum und

$$\ldots \to V_i \xrightarrow{f_i} V_{i+1} \xrightarrow{f_{i+1}} V_{i+2} \to \ldots$$

eine exakte Sequenz von K-linearen Abbildungen. Dann ist die folgende Sequenz ebenfalls exakt.

$$\ldots \to V_i\otimes W \xrightarrow{f_i\otimes Id} V_{i+1}\otimes W \xrightarrow{f_{i+1}\otimes Id} V_{i+2}\otimes W \to \ldots$$

Beweis. Wir haben zu zeigen

$$\text{im } (f_i\otimes Id) = \ker(f_{i+1}\otimes Id).$$

1 Schritt. Beweis von "$\subseteq$".
Nach Voraussetzung gilt

$$\text{im } f_i = \ker f_{i+1} ,$$

also

$$f_{i+1}\circ f_i = 0,$$

Also

$$(f_{i+1}\otimes Id)\circ(f_i\otimes Id) = (f_{i+1}\circ f_i)\otimes Id = 0\otimes Id$$

Für $v\in V_i$ und $w\in W$ gilt also

$$(f_{i+1}\otimes Id)\circ(f_i\otimes Id)(v\otimes w) = 0\otimes w = 0\cdot(0\otimes Id) = 0.$$

Die Abbildung überführt alle Vektoren des Erzeugendensystems $\{v\otimes w\}$ in den Nullvektor, ist also die Nullabbildung,

$$(f_{i+1}\otimes Id)\circ(f_i\otimes Id) = 0.$$

Also gilt

$$\text{im } (f_i\otimes Id) \subseteq \ker(f_{i+1}\otimes Id).$$

2. Schritt. Beweis von "$\supseteq$".
Wir wählen Basen

$\{w_\gamma\}_{\gamma\in L}$ von W

$\{v_\alpha\}_{\alpha\in I}$ von $\ker f_{i+1}$

Die Basis des Kerns ergänzen wir zu einer Basis von V_{i+1}, d.h. es sei

$$\{v_\alpha\}_{\alpha\in I} \cup \{v_\beta\}_{\beta\in J} \text{ Basis von } V_{i+1}.$$

Dann gilt:

$\{v_\alpha\otimes w_\gamma\}_{\alpha\in I\gamma\in L}$ ist Basis von $\ker(f_{i+1}\otimes Id)$

$\{v_\alpha\otimes w_\gamma\}_{\alpha\in I\gamma\in L} \cup \{v_\beta\otimes w_\gamma\}_{\beta\in J\gamma\in L}$ ist Basis von V_{i+1}.

Sei jetzt

$$t \in \ker((f_{i+1} \otimes \mathrm{Id}).$$

Wir haben zu zeigen, t liegt im Bild von $f_i \otimes \mathrm{Id}$. Weil t in $V_{i+1} \otimes W$ liegt, können wir t in der folgenden Gestalt schreiben.

$$t = \sum_{\alpha \in I \cup J, \gamma \in L} c_{\alpha\gamma} v_\alpha \otimes w_\gamma \text{ mit } c_{\alpha\gamma} \in K.$$

Für $\alpha \in I$ gilt $v_\alpha \in \ker f_{i+1} = \mathrm{im}\, f_i$, d.h.es gibt ein $v'_\alpha \in V_i$ mit

(1) $$v_\alpha = f_i(v'_\alpha) \text{ für jedes } \alpha \in I.$$

Deshalb läßt sich t in der folgenden Gestalt schreiben.

(2) $$t = \sum_{\alpha \in I, \gamma \in L} c_{\alpha\gamma} v_\alpha \otimes w_\gamma + \sum_{\beta \in J, \gamma \in L} c_{\beta\gamma} v_\beta \otimes w_\gamma$$

Die erste Summe läßt sich wegen (1) schreiben als

$$\sum_{\alpha \in I, \gamma \in L} c_{\alpha\gamma} f_i(v'_\alpha) \otimes \mathrm{Id}(w_\gamma) = (f_i \otimes \mathrm{Id})(\sum_{\alpha \in I, \gamma \in L} c_{\alpha\gamma} v'_\alpha \otimes w_\gamma).$$

Mit anderen Worten, die erste Summe liegt im Bild von $f_i \otimes \mathrm{Id}$. Es reicht also zu zeigen, die zweite Summe ist Null,

$$\sum_{\beta \in J, \gamma \in L} c_{\beta\gamma} v_\beta \otimes w_\gamma = 0.$$

Dazu reicht es zu zeigen,

(3) $$c_{\beta\gamma} = 0 \text{ für alle } \beta \in J \text{ und alle } \gamma \in L.$$

Nach Voraussetzung liegt t im Kern von $f_{i+1} \otimes \mathrm{Id}$. Dasselbe gilt von der ersten Summe (das sie sogar im Bild von $f_i \otimes \mathrm{Id}$ liegt). Wir wenden $f_{i+1} \otimes \mathrm{Id}$ auf (2) and und erhalten

$$0 = 0 + \sum_{\beta \in J, \gamma \in L} c_{\beta\gamma} f_{i+1}(v_\beta) \otimes w_\gamma$$

Zum Beweis von (3) reicht es also zu zeigen,

die Vektoren $f_{i+1}(v_\beta) \otimes w_\gamma$ mit $\beta \in J$, $\gamma \in L$ sind linear unabhängig.

Dazu wiederum reicht es zu zeigen,

die Vektoren $f_{i+1}(v_\beta)$ mit $\beta \in J$ sind linear unabhängig

(vgl. 6.6.6(i)). Seien also Koeffizienten $c_\beta \in K$ ($\beta \in J$) gegeben mit

$$\sum_{\beta \in J} c_\beta f_{i+1}(v_\beta) = 0.$$

Wir haben zu zeigen, $c_\beta = 0$ für jedes β. Es gilt

$$0 = \sum_{\beta \in J} c_\beta f_{i+1}(v_\beta) = f_{i+1}(\sum_{\beta \in J} c_\beta v_\beta),$$

also

$$\sum_{\beta \in J} c_\beta v_\beta \in \ker(f_{i+1})$$

Da $\{v_\alpha\}_{\alpha \in I}$ eine Basis von $\ker f_{i+1}$ ist, gibt es Element $d_\alpha \in K$ ($\alpha \in I$) mit

$$\sum_{\beta\in J} c_\beta v_\beta = \sum_{\alpha\in I} d_\alpha v_\alpha ,$$

d.h.

$$0 = \sum_{\beta\in J} c_\beta v_\beta + \sum_{\alpha\in I} (-d_\alpha) v_\alpha$$

Da $\{v_\alpha\}_{\alpha\in I} \cup \{v_\beta\}_{\beta\in J}$ eine Basis von V_{i+1} ist, ist die Menge der v_α und v_β linear unabhängig, d.h. aus der obigen Identität folgt insbesodere

$$c_\beta = 0 \text{ für jedes } \beta\in J.$$

QED.

6.6.13 Begriff der Algebra über einem Ring

Seien R kommutativer Ring mit 1. Eine R-Algebra ist ein Ring, S mit 1 zusammen mit einem Ring-Homomorphismus

$$R \to S$$

von Ringen mit 1. Der Homomorphismus heißt dann Struktur-Homomorphismus von S. Seien S und S' zwei R-Algebren. Ein Homomorphismus von R-Algebren

$$f: S \to S'$$

ist ein Ring-Homomorphismus f, für welchen das folgende Diagramm kommutativ ist.

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

Dabei sollen die schrägen Pfeile gerade die Struktur-Homomorphismen bezeichnen.

Beispiel: die Tensor-Algebra eines K-Vektorraums

Sei K ein Körper und V ein K-Vektorraum. Wir setzen

$$T(V) := T_K(V) := \oplus_{n=0}^{\infty} V^{\otimes n} = V^{\otimes 0} \oplus V^{\otimes 1} \oplus V^{\otimes 2} \oplus \dots$$

Dabei sei $V^{\otimes 0} = K$, $V^{\otimes 1} = V$ und $V^{\otimes n}$ das Tensorprodukt von n Exemplaren des Raumes V. Nach Konstruktion ist $T_K(V)$ ein K-Vektorraum. Der direkte Summand

$$V^{\otimes n} \subseteq T(V)$$

heißt homogener Bestandteil des Grades n von T(V), dessen Elemente t heißen homogene Elemente des Grades n und man schreibt

$$\deg t = n.$$

Seien

$$t' = \sum_{n=1}^{\infty} t'_n \text{ und } t'' = \sum_{n=1}^{\infty} t''_n$$

zwei Elemente von T(V) mit t'_n und t''_n homogen vom Grad n. Wir definieren das Produkt von t' und t" wie folgt

$$t' \cdot t'' = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} t'_m \otimes t''_n$$

Dabei wird $t'_m \otimes t''_n \in V^{\otimes m} \otimes V^n \cong V^{\otimes(m+n)}$ als Element von $V^{\otimes(m+n)}$ aufgefaßt. Speziel für n = 0 haben wir

$$t'_m \otimes t''_0 \in V^{\otimes m} \otimes K \cong V^{\otimes m},\ x\otimes y \mapsto yx,$$

d.h. $t'_m \otimes t''_0$ wird mit $t''_0 \cdot t'_m$ identifiziert. Analog wird für m = 0 das Element

$$t'_0 \otimes t''_n$$

mit $t'_0 \cdot t''_n$ identifiziert. Die Multiplikation von zwei Elementen 0-ten Grades entspricht damit der gewöhnlichen Multiplikation von K.

Die Abbildung

$$K \to T(V),\ c \mapsto c \text{ aufgefaßt als homogenes Element des Grades 0 von } T(V)$$

heißt natürliche Einbettung von K in die Tensoralgebra T(V).

Die Abbildung

$$V \to T(V),\ v \mapsto v \text{ aufgefaßt als homogenes Element des Grades 1 von } T(V)$$

heißt natürliche Einbettung von V in die Tensoralgebra T(V).

Bemerkungen

(i) Mit der oben definierten Multiplikation ist T(V) eine K-Algebra.
(ii) Die natürliche Einbettung $K \to T(V)$ ist ein Homomorphismus von Ringen mit 1.
(iii) Die natürliche Einbettung $V \to T(V)$ ist K-linear.

6.6.14 Lemma: die Universalitätseigenschaft der Tensorpotenz $V^{\otimes n}$

Sei V ein K-Vektorraum. Dann ist die Abbildung

$$\rho_n: V^n = V\times\ldots\times V \to V^{\otimes n},\ (v_1,\ldots,v_n) \mapsto v_1\otimes\ldots\otimes v_n$$

des direkten Produkts von n Exemplaren von V nach $V^{\otimes n}$ n-linear, d.h. linear in jeder der n Variablen. Für jede n-lineare Abbildung

$$m: V\times\ldots\times V \to U$$

gibt es genau eine lineare Abbildung $\tilde m: V^{\otimes n} \to U$ mit $m = \tilde m \circ \rho_n$.

Beweis. Eindeutigkeit von $\tilde m$. Falls $\tilde m$ existiert, so gilt für beliebige $v_1,\ldots,v_n \in V$:

$$\tilde m\,(v_1\otimes\ldots\otimes v_n) = \tilde m\,(\rho_n(v_1,\ldots,v_n)) = m(v_1,\ldots,v_n).$$

Die Werte von $\tilde m$ auf den Elementen der Gestalt $v_1\otimes\ldots\otimes v_n$ sind also eindeutig festgelegt. Da diese Elemente ein Erzeugendensystem von $V^{\otimes n}$ bilden und da $\tilde m$ linear sein soll, ist somit die gesamte Abbildung $\tilde m$ festgelegt.

Existenz von $\tilde m$. Beweis durch Induktion nach n.

Im Fall n = 1 ist eine lineare Abbildunge

$$V = V^{\otimes 1} \to U$$

gegeben und wir können $\tilde m = m$ setzen. Sei jetzt m > 1. Für jedes feste $v_0 \in V$ ist die Abbildung

$$m_{v_0}: V^{n-1} \to U,\ (v_1,\ldots,v_{n-1}) \mapsto m(v_1,\ldots,v_{n-1}, v_0),$$

linear in jeder ihrer n-1 Argumente. Nach Induktionsvoraussetzung gibt es eine lineare Abbildung

$$\tilde m_{v_0}: V^{\otimes(n-1)} \to U$$

mit $m_{v_0} = \tilde m_{v_0} \circ \rho_{n-1}$, d.h.mit

$$\tilde{m}_{v_0}(v_1\otimes\ldots\otimes v_{n-1}) = \tilde{m}_{v_0}(\rho_{n-1}(v_1,\ldots,v_{n-1})) = m_{v_0}(v_1,\ldots,v_{n-1}) = m(v_1,\ldots,v_{n-1},v_0).$$

Die lineare Abbildung $\tilde{m}_{v_0}$ ist durch die Bedingung

(1) $$\tilde{m}_{v_0}(v_1\otimes\ldots\otimes v_{n-1}) = m(v_1,\ldots,v_{n-1},v_0)$$

eindeutig festgelegt (da die $v_1\otimes\ldots\otimes v_{n-1}$ ein Erzeugendensystem von $V^{\otimes(n-1)}$ bilden). Betrachten wir die Abbildung

(2) $$\tilde{m}':V^{\otimes(n-1)}\times V \to U,\ (t, v) \mapsto \tilde{m}_v(t).$$

Nach Konstruktion ist sie linear im ersten Argument t. Zeigen wir, sie ist auch linear im zweiten Argument v, d.h. , zeigen wir, es gilt

$$\tilde{m}_{c'v'+c''v''}(t) = c'\tilde{m}_{v'}(t) + c''\tilde{m}_{v''}(t)$$

für alle $v',v''\in V$, alle $c',c''\in K$ und alle $t\in V^{\otimes(n-1)}$, d.h.

$$\tilde{m}_{c'v'+c''v''} = c'\tilde{m}_{v'} + c''\tilde{m}_{v''}.$$

Auf beiden Seiten stehen lineare Abbildungen. Zum Beweis ihrer Gleichheit reicht es zu zeigen, sie haben dieselben Werten in allen Vektoren eines Erzeugendensystems von $V^{\otimes(n-1)}$. Es reicht also zu zeigen,

$$\tilde{m}_{c'v'+c''v''}(v_1\otimes\ldots\otimes v_{n-1}) = c'\tilde{m}_{v'}(v_1\otimes\ldots\otimes v_{n-1}) + c''\tilde{m}_{v''}(v_1\otimes\ldots\otimes v_{n-1})$$

für beliebige $v_1,\ldots,v_{n-1}\in V$. Wegen (1) gilt

$$\begin{aligned} \text{LHS} &= m(v_1\otimes\ldots\otimes v_{n-1}, c'v'+c''v'') \\ &= c'm(v_1\otimes\ldots\otimes v_{n-1},v') + c''m(v_1\otimes\ldots\otimes v_{n-1},v'') \\ &= \text{RHS} \end{aligned}$$

Wir haben gezeigt, die Abbildung (2) ist bilinear. Auf Grund der Universalitätseigenschaft des Tensorprodukts gibt es genau eine lineare Abbildung

$$\tilde{m}: V^{\otimes n} = V^{\otimes(n-1)}\otimes V \to U$$

mit $\tilde{m}' = \tilde{m}\circ\rho$, d.h. mit

$$\tilde{m}_v(t) = \tilde{m}(t\otimes v)$$

für alle $t\in V^{\otimes(n-1)}$ und alle $v\in V$. Speziell für $t = v_1\otimes\ldots\otimes v_{n-1}$ erhalten wir

$$\tilde{m}(v_1\otimes\ldots\otimes v_{n-1}\otimes v) = \tilde{m}_v(v_1\otimes\ldots\otimes v_{n-1}) = m(v_1,\ldots,v_{n-1},v).$$

Mit anderen Worten, es gilt

$$m = \tilde{m}\circ\rho_n,$$

d.h. $\tilde{m}$ ist gerade die Abbildung, deren Existenz wir beweisen wollen.

QED.

6.6.15 Die Universalitätseigenschaft der Tensor-Algebra

Sei V ein K-Vektorraum. Dann gibt es für jede K-Algebra S und jede K-lineare Abbildung

$$f:V \to S$$

genau einen Homomorphismus

$$\tilde f : T(V) \to S$$
von K-Algebren derart, daß das folgende Diagramm kommutativ ist.

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "f"'] & T(V) \arrow[dl, "\tilde f"] \\ S & \end{tikzcd}$$

dabei sei ρ die natürliche Einbettung.

Beweis. Eindeutigkeit von $\tilde f$. Wir nehmen an, $\tilde f$ existiert und leiten eine Formel für $\tilde f$ her, aus der hervorgeht, daß $\tilde f$ eindeutig festgelegt ist. Sei

$$t \in T(V).$$

Dann ist t eine Summe von endlich vielen homogenen Elementen. Ein homogenes Element des Grades n (aus $V^{\otimes n}$) wiederum ist eine Summe aus endlich vielen Element der Gestalt $v_1 \otimes ... \otimes v_n$. Mit anderen Worten, t hat die Gestalt

$$t = v_1^1 \otimes \dots \otimes v_{n_1}^{r_1} + v_1^2 \otimes \dots \otimes v_{n_2}^{r_2} + \dots + v_1^s \otimes \dots \otimes v_{n_s}^{r_s}$$

Damit gilt da $\tilde f$, linear ist und Produkte in Produkte überführt,

$$\begin{aligned} \tilde f(t) &= \tilde f(v_1^1 \otimes \dots \otimes v_{n_1}^{r_1}) + \dots + \tilde f(v_1^s \otimes \dots \otimes v_{n_s}^{r_s}) \\ &= \tilde f(v_1^1) \dots \tilde f(v_{n_1}^{r_1}) + \,..+\, \tilde f(v_1^s \dots \tilde f(v_{n_s}^{r_s}) \\ &= \tilde f\,(\rho(v_1^1)) \dots \tilde f(\rho(v_{n_1}^{r_1})) + \dots + \tilde f(\rho(v_1^s) \dots f(\rho(v_{n_s}^{r_s})). \\ &= f(\rho(v_1^1)) \dots f(\rho(v_{n_1}^{r_1})) + \dots + f(\rho(v_1^s) \dots f(\rho(v_{n_s}^{r_s})). \end{aligned}$$

Damit ist $\tilde f$ eindeutig festgelegt.

Existenz von $\tilde f$. Betrachten wir die Abbildung

$$V^r \to S,\ (v_1,...,v_n) \mapsto f(v_1)\cdot...\cdot f(v_n).$$

Auf Grund des Distributivgesetzes für den Ring S ist sie linear in jedem der n Argumente. Nach 6.6.14 gibt es genau eine lineare Abbildung

$$\tilde f_n : V^{\otimes n} \to V$$

mit

$$\tilde f_n(v_1,...,v_n) = f(v_1)\cdot...\cdot f(v_n)$$

für alle $v_1,...,v_n \in V$. Wir definieren $\tilde f: T(V) \to R$, indem wir setzen

$$\tilde f(t) = \tilde f_0(t_0) + \tilde f_1(t_1) + ... + \tilde f_s(t_s)$$

falls $t = t_0 + t_1 + ... + t_s$ gilt mit t_i homogen vom Grad i. Dabei sei

$$\tilde{f}_0(t_0) = t_0 \cdot 1_S$$

Die Abbildung ist wohldefiniert und linear. Für homogene Elemente des Grades 1 gilt

$$\tilde{f}(\rho(v)) = \tilde{f}_1(v) = f(v).$$

Auf Grund der Defintion von $\tilde{f}_0$ gilt $\tilde{f}(1_K) = 1_S$. Durch direktes Nachrechnent sieht man, daß $\tilde{f}$ ein Ringhomomorphismus ist.

QED.

6.6.17 Das von einer Menge erzeugte Ideal

Seien K ein Körper, S eine K-Algebra und $M \subseteq S$ Teilmenge. Dann heißt der von der Menge

$$\{ s_1 \cdot m_1 \cdot ... \cdot m_{r-1} \cdot s_r \mid s_1,...,s_r \in S \text{ und } m_1,...,m_{r-1} \in M , r=1,2,3, ...\}$$

erzeugte K-lineare Unterraum von S das von M erzeugte Ideal und wird mit

$$\langle M \rangle \text{ oder } \langle m \mid m \in M \rangle$$

bezeichnet. Im Fall $M = \{ m_1 , ... , m_r \}$ schreibt man auch

$$\langle m_1 , ... , m_r \rangle := \langle M \rangle$$

Bemerkung

$a \in S$ und $b \in \langle M \rangle \Rightarrow ab \in \langle M \rangle$ und $ba \in \langle M \rangle$

6.6.18 Der Faktorraum nach einem Ideal

Seien S eine K-Algbra und $I \subseteq S$ ein Ideal. Dann ist durch

$$(a + I) \cdot (b + I) = ab + I$$

ein Produkt auf S/I definert und der Vektorraum S/I ist mit diesem Produkt ein Ring (mit Eins). Die natürliche Abbildung

$$\rho: S \to S/I$$

ist ein Homomorphismus von Ringen (mit Eins)

Beweis. Seien $a + I = a' + I$ und $b + I = b' + I$. Wir haben zu zeigen, es gilt

$$ab + I = a'b' + I.$$

Es gilt

$a-a' \in I$ und $b-b' \in I$

also

$(a-a')b \in I$ und $a'(b-b') \in I$

also

$ab - a'b' \in I$

also

$ab + I = a'b' + I.$

Nach Konstruktion ist

$$\rho(ab) = ab + I = (a+I)(b+I) = \rho(a)\rho(b).$$

Daraus ergibt sich, daß S/I ein Ring ist und ρ ein Ringhomomorphismus. Bezeichnet 1 das Einselement von S, so spielt $\rho(1)$ die Rolle des Einselements von S/I. Durch die Zusammensetzung der Ringhomomorphismen

$$K \to S \to S/I$$

bekommt S/I die Struktur einer K-Algebra.

QED.

6.6.19 Die symmetrische Algebra

Seien V ein K-Vektorraum und

$$I'(V) \subseteq T(V)$$

das von den Elementen der Gestalt

$$a \otimes b - b \otimes a \text{ mit } a,b \in V$$

erzeugte Ideal. Dann heißt

$$S(V) = S_K(V) = T(V)/I'(V)$$

symmetrische Algebra von V über K. Die Zusammensetzung

$$V \to T(V) \to S(V)$$

der natürlichen Einbettung (im Grad 1) mit der natürliche Abbildung auf den Faktorring heißt wieder natürliche Einbettung (im Grad 1). Analog definiert man die natürliche Einbettung im Grad 0.

Bemerkung

Die natürlichen Einbettungen sind injektiv. Sie gestatten es somit V und K mit Teilmengen von S(V) zu identifizieren. Den Beweis der Injektivität werden wir nur im Fall, daß V endlich-dimensional ist angeben. Der unendlich-dimensionale Fall unterscheidet sich jedoch nicht wesentlich vom endlich-dimensionalen.

6.6.20 Die äußere Algebra

Seien V ein K-Vektorraum und

$$I''(V) \subseteq T(V)$$

das von den Elementen der Gestalt

$$v\otimes v \text{ mit } v\in V$$

erzeugte Ideal. dann heißt

$$\wedge(V) = \wedge_K(V) = T(V)/I''(V)$$

äußere Algebra von V über K. Die Zusammensetzung

$$V \to T(V) \to \wedge(V)$$

der natürlichen Einbettung (im Grad 1) mit der natürliche Abbildung auf den Faktorring heißt wieder natürliche Einbettung (im Grad 1). Analog definiert man die natürliche Einbettung im Grad 0.

Bemerkung

Die natürlichen Einbettungen sind injektiv. Sie gestatten es somit V und K mit Teilmengen von S(V) zu identifizieren.

6.6.21 Die Universalitätseigenschaft der symmetrischen Algebra

Seien V ein K-Vektorraum, S eine kommutative K-Algebra und

$$f: V \to S$$

eine K-lineare Abbildung. Dann gibt es genau eine Fortsetzung

$$\tilde{f}: S(V) \to S$$

von f zu einen Homomorphismus von K-Algebren, d.h. es gibt genau einen Homomorphismus von K-Algebren $\tilde{f}: S(V) \to S$, dessen Zusammensetzung mit der der natürlichen Einbettung

$$V \to T(V) \to T(V)/I'(V) = S(V)$$

gleich f ist.

Beweis. Auf Grund der Universalitätseigenschatt der Tensor-Algebra gibt es genau eine Fortsetzung

$$f': T(V) \to S$$

zu einen Homomorphismus von K-Algebren. Für beliebige Vektoren $v',v''\in V$ und beliebige Tensoren $t',t''\in T(V)$ gilt

$$f'(t'(v'\otimes v'' - v''\otimes v')t'') = f'(t')(f'(v')f'(v'') - f'(v'')f'(v'))f'(t'') = 0,$$

da S kommutativ ist. Mit anderen Worten, ein Erzeugendensystem des definierenden Ideals I'(V) liegt im Kern von f';

$$I'(V) \subseteq \ker(f').$$

Deshalb ist die folgenden Abbildung wohldefiniert.

$$\tilde{f}: S(V) = T(V)/I'(V) \to S,\ t + I'(V) \mapsto f'(t).$$

Aus $t'+I'(V) = t'+I'(V)$ folgt nämlich $t'-t \in I'(V)$, also

$$f'(t') - f'(t) = f'(t'-t)\in f'(I'(V)) = \{0\},$$

also $f'(t') = f(t)$. Da f' ein Homomorphismus von K-Algebren ist, gilt dasselbe für $\tilde{f}$. Für $v \in V$ gilt

$$\tilde{f}(v + I'(V)) = f'(v) = f(v),$$

d.h. die Zusammensetzung von $\tilde{f}$ mit der natürlichen Einbettung

$$V \to S(V),\ v \mapsto v + I'(V),$$

ist gerade f.
QED.

6.6.22 Vergleich mit den Polynom-Algebren

Seien V ein K-Vektorraum mit der Basis $v_1,...,v_n \in V$ und

$$S := K[x_1,...,x_n]$$

die K-Algebra der Polynome in $x_1,...,x_n$ mit Koeffizienten aus K. Dann gibt es genau einen K-Algebra-Homomorphismus

$$K[x_1,...,x_n] \to S(V) \text{ mit } x_i \mapsto v_i .$$

Dieser Homomorphismus ist sogar ein Isomorphismus.
Bemerkungen

(i) Das natürliche Bild

$$S_k(V) = \{ \sum_i s_{1,i} \otimes ... \otimes s_{k,i} + I'(V) \mid v_{ij} \in V\}$$

von $V^{\otimes k}$ in S(V) entspricht dabei gerade den Polynomen k-ten Grades in $x_1,...,x_n$, d.h.

$$\dim S_k(V) = \binom{k+n-1}{n-1}.$$

(ii) Da die $V^{\otimes k}$ den Vektorraum S(V) erzeugen, erzeugen die $S_k(V)$ den Faktorraum S(V),

$$S(V) = \sum_{k=0}^{\infty} S_k(V).$$

Da sich jedes Polynom auf genau eine Weise als Summe homogener Polynome schreiben läßt, ist diese Summe sogar direkt.

$$S(V) = \oplus_{k=0}^{\infty} S_k(V).$$

(iii) $S_k(V)$ heißt k-te <u>symmetrische Potenz</u> von V.

(iv) Da die natürlichen Einbettungen

$$K \to K[x_1,...,x_n] \text{ und } Kx_1+...+Kx_n \to K[x_1,...,x_n]$$

in den Graden 0 und 1 injektiv sind und S(V) durch V bis auf Isomorphie eindeutig bestimmte ist, gilt dasselbe auch für die natürliche Einbettungen

$$K \to S(V) \text{ und } V \to S(V).$$

Beweis. Betrachten wir die K-lineare Abbildung

$$i{:}V \to K[x_1,...,x_n],\ c_1v_1+...+c_1v_n \mapsto c_1x_1+...+c_1x_n .$$

<u>1. Schritt</u>. Die Polynom-Algebra besitzt die Universalitätseingenschaft der Symmetrischen Algebra.

Seien S ein kommutative K-Algebra und

$$f{:}\ V \to S$$

eine K-lineare Abbildung. Wir haben zu zeigen, es gibt genau einen Homomorphismus von K-Algebren

$$\tilde{f}: K[x_1,...,x_n] \to S$$

mit $\tilde{f} \circ i = f$.

Existenz von $\tilde{f}$. Für jedes Polynom $p \in K[x_1,...,x_n]$ setzen wir

$$\tilde{f}(p) = p(f(v_1),...,f(v_n)),$$

d.h. wir ordnen jedem Polynom p den Wert an der Stelle $(f(v_1),...,f(v_n))$ zu. Auf diese Weise ist ein Homomorphismus von K-Algebren definiert,

$$\tilde{f}: K[x_1,...,x_n] \to S$$

Nach Konstruktion gilt

$$\begin{aligned}
\tilde{f} \circ i(c_1v_1+...+c_1v_n) &= \tilde{f}(c_1x_1+...+c_1x_n) && \text{(nach Definition von i)}\\
&= c_1f(v_1)+...+c_1f(v_n) && \text{(nach Definition von } \tilde{f}\text{)}\\
&= f(c_1v_1+...+c_1v_n) && \text{(weil f linear ist)}
\end{aligned}$$

d.h. es gilt $\tilde{f} \circ i = f$.

Eindeutigkeit von $\tilde{f}$. Falls $\tilde{f}$ existiert, so gilt für jedes Polynom p

$$\begin{aligned}
\tilde{f}(p) = \tilde{f}(p(x_1,...,x_n)) &= p(\tilde{f}(x_1),...,\tilde{f}(x_n)) && (\tilde{f} \text{ ist Algebra-Homomorphismus})\\
&= p(\tilde{f}(i(v_1)),...,\tilde{f}(i(v_n))) && \text{(nach Definition von i)}\\
&= p((f(v_1),...,f(v_n)) && (\text{wegen } \tilde{f} \circ i = f).
\end{aligned}$$

2. Schritt. Konstruktion des Algebra-Homomorphismus $K[x_1,...,x_n] \to S(V)$.

Wir wenden die eben bewiesene Universaltiätseigenschaft auf die kommutative K-Algebra

$$S = S(V)$$

und die natürliche Einbettung

$$j: V \to S(V)$$

(im Grad 1) an. Wie eben gezeigt, gibt es genau einen K-Algbra-Homomorphismus

$$\tilde{j}: K[x_1,...,x_n] \to S(V)$$

mit $\tilde{j} \circ i = j$, d.h. mit

$$\tilde{j} \circ i(v_\alpha) = j(v_\alpha) \text{ für } \alpha = 1, ... , n,$$

d.h. mit

$$\tilde{j}(x_\alpha) = v_\alpha \text{ für } \alpha = 1, ... , n.$$

Mit anderen Worten, $\tilde{j}$ ist der K-Algebra-Homomorphismus, desssen Existenz und Eindeutigkeit im Satz behauptet wurde. Wir haben noch zu zeigen, $\tilde{j}$ ist ein Isomorphismus.

3. Schritt. Konstruktion der inversen Abbildung.

Wir betrachten die K-lineare Abbildung

$$i: V \to K[x_1,...,x_n].$$

Auf Grund der Universalitätseigenschaft der symmetrischen Algebra S(V) gibt es genau einen Homomorphismus

$$\tilde{i}: S(V) \to K[x_1,...,x_n]$$

mit $\tilde{i}\circ j = i$., d.h. mit $\tilde{i}\circ j(v_\alpha) = i(v_\alpha)$ für $\alpha = 1, ... , n$, d.h. mit

$$\tilde{i}(x_\alpha) = i(v_\alpha) \text{ für } \alpha = 1, ... , n,$$

Insbesondere gilt

$$\tilde{j}(\tilde{i}(x_\alpha)) = \tilde{j}(v_\alpha) = x_\alpha$$

$$\tilde{i}(\tilde{j}(v_\alpha)) = \tilde{i}(x_\alpha) = v_\alpha$$

für alle α. Aus den Eindeutigkeits-Aussagen der Universalitätseigenschaften von S(V) und $K[x_1,...,x_n]$ folgt damit

$$\tilde{j}\circ\tilde{i} = \mathrm{Id}.$$

$$\tilde{i}\circ\tilde{j} = \mathrm{Id},$$

d.h. $\tilde{j}$ und $\tilde{i}$ sind zueinander inverse Isomorphismen. die beiden letzten Identitäten lassen sich auch durch direktes Nachrechnen überprüfen. Zum Beispiel folgt aus

$$\tilde{j}(\tilde{i}(x_\alpha)) = x_\alpha$$

für alle α und der Tatsache, daß $\tilde{i}$ und $\tilde{j}$ Algebra-Homomorphismen sind, daß

$$\tilde{j}(\tilde{i}(p)) = p$$

gilt für jedes Polynom p von $K[x_1,...,x_n]$, d.h. $\tilde{j}\circ\tilde{i} = \mathrm{Id}$.

QED.

Beweis der Dimensionsformel von Bemerkung (i). Für n = 1 und n = 2 ist die Formel trivial. Für allgemeines n > 2 benutze man die Standardeigenschaften der Binomialkoeffizienten.

QED.

6.6.23 Die Universalitätseigenschaft der äußeren Algebra

Seien V ein K-Vektorraum, S eine K-Algebra und

$$f: V \to S$$

eine K-lineare Abbildung mit

$$f(v)f(v) = 0 \text{ für alle } v \in V.$$

Dann gibt es genau eine Fortsetzung

$$\tilde{f}: \wedge(V) \to S$$

von f zu einen Homomorphismus von K-Algebren, d.h. es gibt genau einen Homomorphismus von K-Algebren $\tilde{f}: \wedge(V) \to S$, dessen Zusammensetzung mit der der natürlichen Einbettung

$$V \to T(V) \to T(V)/I''(V) = \wedge(V)$$

gleich f ist.

Beweis. Auf Grund der Universalitätseigenschatt der Tensor-Algebra gibt es genau eine Fortsetzung

$$f': T(V) \to S$$

zu einen Homomorphismus von K-Algebren. Für beliebige Vektoren $v \in V$ und beliebige Tensoren $t',t'' \in T(V)$ gilt

$$f'(t'(v\otimes v)t'') = f'(t')(f'(v)f'(v) = 0,$$

Mit anderen Worten, ein Erzeugendensystem des definierenden Ideals I”(V) liegt im Kern von f’;

$$I''(V) \subseteq \ker(f').$$

Deshalb ist die folgenden Abbildung wohldefiniert.

$$\tilde{f}: \wedge(V) = T(V)/I''(V) \to S, t + I'(V) \mapsto f'(t).$$

Aus t'+I'(V) = t'+I'(V) folgt nämlich t'-t $\in$ I'(V), also

$$f'(t') - f'(t) = f'(t'-t) \in f'(I'(V)) = \{0\},$$

also f'(t') = f(t). Da f' ein Homomorphismus von K-Algebren ist, gilt dasselbe für $\tilde{f}$. Für $v \in V$ gilt

$$\tilde{f}(v + I'(V)) = f'(v) = f(v),$$

d.h. die Zusammensetzung von $\tilde{f}$ mit der natürlichen Einbettung

$$V \to S(V),\ v \mapsto v + I'(V),$$

ist gerade f.

QED.

6.6.24 Vergleich mit den Graßmann-Algebren

Seien V ein K-Vektorraum mit der Basis $v_1,...,v_n \in V$. Für festes $k \in \mathbb{N}$ und jede echt aufsteigende Folge

$$i_1 < i_2 < ... < i_k$$

von natürlichen Zahlen aus {1,...,n} führen wir ein Symbol

(1) $$e_{i_1} \wedge e_{i_2} \wedge ... \wedge e_{i_k}$$

ein. Bezeichne

$$\wedge^k K^n = \oplus K e_{i_1} \wedge e_{i_2} \wedge ... \wedge e_{i_k}$$

den von der Menge dieser Symbole frei erzeugte K-Vektorraum. Für k = 1 erhalten wir

$$\wedge^1 K^n = Ke_1 + ... + Ke_n = K^n.$$

Für k = n erhalten wir den 1-dimensionalen K-Vektorraum

$$\wedge^n K^n = K\, e_1 \wedge .... \wedge e_n$$

und für k > n ist $\wedge^k K^n = 0$. Für k = 0 wollen wir

$$\wedge^0 K^n = K$$

setzen.. Schließlich sei

(2) $$K<e_1,...,e_n> := \oplus_{k=0}^{\infty} \wedge^k K^n = \wedge^0 K^n \oplus \wedge^1 K^n \oplus ..... \oplus \wedge^n K^n$$

die direkte Summe aller $\wedge^k K^n$ Wir wollen jetzt auf dem Vektorraum (2) eine Multiplikation einführen. Dazu ist es nützlich, die Symbole (1) auch zu definieren, wenn die indizes i_ν keine echt aufsteigende Folge bilden. Wir setzen wir für jede Permutation

$\pi \in S_k$

$$e_{i_{\pi(1)}} \wedge e_{i_{\pi(2)}} \wedge ... \wedge e_{i_{\pi(k)}} := \operatorname{sign}(\pi) \cdot e_{i_1} \wedge e_{i_2} \wedge ... \wedge e_{i_k}$$

Weiter vereinbaren wir, daß ein Symbol der Gestalt (1) mit mehrfach auftretenden Indizes den Nullvektor bezeichnen soll. Mit diesen Vorbereitungen können wir nun eine Multiplikation in

$$K<e_1,...,e_n>$$

einführen, inden wir setzen

$$\Big(\sum_{i_1...i_k} c_{i_1...i_k} e_{i_1} \wedge ... \wedge e_{i_k}\Big) \cdot \Big(\sum_{j_1...j_\ell} d_{j_1...j_k} e_{j_1} \wedge ... \wedge e_{j_\ell}\Big)$$

$$:= \sum_{i_1 \dots i_k j_1 \dots j_\ell} c_{i_1 \dots i_k} d_{j_1 \dots j_k} e_{i_1} \wedge \dots \wedge e_{i_k} \wedge e_{j_1} \wedge \dots \wedge e_{j_\ell}$$

Dann gibt es genau einen K-Algebra-Homomorphismus

$$K<e_1,\dots,e_n> \to S(V) \text{ mit } x_i \mapsto v_i \,.$$

Dieser Homomorphismus ist sogar ein Isomorphismus.

Bemerkungen

(i) Das natürliche Bild

$$\wedge_k(V) = \{ \sum_i v_{1,i} \otimes \dots \otimes v_{k,i} + I''(V) \mid v_{ij} \in V\}$$

von $V^{\otimes k}$ in $\wedge(V)$ entspricht dabei gerade dem K-Vektorraum $\wedge^k K^n$, d.h. es gilt

$$\dim S_k(V) = \binom{n}{k}.$$

Für beliebige Vektoren $v'_1,\dots,v'_n$ aus V bezeichnet man das Bild von $v'_1 \otimes \dots \otimes v'_n$ bei der natürlichen Abbildung $T(V) \to \wedge(V)$ mit

$$v'_1 \wedge \dots \wedge v'_n$$

(ii) Da die $V^{\otimes k}$ den Vektorraum S(V) erzeugen, erzeugen die $\wedge_k(V)$ den Faktorraum $\wedge(V)$,

$$\wedge(V) = \sum_{k=0}^{\infty} \wedge_k(V).$$

Da sich jedes Element von $K<e_1,\dots,e_n>$ auf genau eine Weise als Summe schreiben läßt mit jeweils einem Summanden aus einem $\wedge^k K^n$, so ist diese Summe sogar direkt.

$$\wedge(V) = \oplus_{k=0}^{\infty} \wedge_k(V).$$

(iii) $\wedge_k(V)$ heißt k-te äußere Potenz von V.

(iv) Da die natürlichen Einbettungen

$$K \to K<e_1,\dots,e_n> \text{ und } Ke_1+\dots+Ke_n \to K<e_1,\dots,e_n>$$

in den Graden 0 und 1 injektiv sind und $\wedge(V)$ durch V bis auf Isomorphie eindeutig bestimmte ist, gilt dasselbe auch für die natürliche Einbettungen

$$K \to \wedge(V) \text{ und } V \to \wedge(V).$$

Beweis. Wir betrachten die lineare Abbildung

$$i: V \to K<e_1,\dots,e_n>, \ \sum_{\alpha=1}^{n} c_\alpha v_\alpha \mapsto \sum_{\alpha=1}^{n} c_\alpha e_\alpha$$

1. Schritt. $K<e_1,\dots,e_n>$ genügt der Universalitätseigenschaft der äußeren Algebra.

Seien S eine K-Algebra und

$$f: V \to S$$

eine K-lineare Abbildung mit

$$f(v)\cdot f(v) = 0 \text{ für jedes } v \in V.$$

Wir haben zu zeigen, es gibt genau einen K-Algebra-Homomorphismus

$$\tilde{f}: K<e_1,\dots,e_n> \to S$$

mit $\tilde{f}\circ i = f$, d.h. mit

$$\tilde{f}(e_\alpha) = \tilde{f}(i(v_\alpha)) = f(v_\alpha).$$

<u>Eindeutigkeit von</u> $\tilde{f}$.
Jedes Element von $K< e_1,\dots,e_n>$ hat die Gestalt

$$c = \sum_{k=0}^{n} \sum_{i_1<\dots<i_k} c^k_{i_1,\dots,i_k} e_{i_1}\wedge e_{i_2}\wedge\dots\wedge e_{i_k}$$

Falls $\tilde{f}$ existiert, so gilt

$$\tilde{f}(c) = \sum_{k=0}^{n} \sum_{i_1<\dots<i_k} c^k_{i_1,\dots,i_k} \tilde{f}(e_{i_1}\wedge e_{i_2}\wedge\dots\wedge e_{i_k}) \qquad (\tilde{f} \text{ ist K-Algebra-Homomorphismus})$$

$$= \sum_{k=0}^{n} \sum_{i_1<\dots<i_k} c^k_{i_1,\dots,i_k} \tilde{f}(e_{i_1})\wedge e_{i_2}\wedge\dots\wedge \tilde{f}(e_{i_k})$$

$$= \sum_{k=0}^{n} \sum_{i_1<\dots<i_k} c^k_{i_1,\dots,i_k} f(e_{i_1})\wedge e_{i_2}\wedge\dots\wedge f(e_{i_k}) \qquad (\text{wegen } \tilde{f}\circ i = f).$$

Damit ist die Eindeutigkeit von $\tilde{f}$ bewiesen.

<u>Existenz von</u> $\tilde{f}$. Wir setzen

$$\tilde{f}(c) := \sum_{k=0}^{n} \sum_{i_1<\dots<i_k} c^k_{i_1,\dots,i_k} f(e_{i_1})\wedge e_{i_2}\wedge\dots\wedge f(e_{i_k})$$

für

$$c = \sum_{k=0}^{n} \sum_{i_1<\dots<i_k} c^k_{i_1,\dots,i_k} e_{i_1}\wedge e_{i_2}\wedge\dots\wedge e_{i_k}.$$

Diese Definition ist korrect, weil die Vektoren

$$e_{i_1}\wedge e_{i_2}\wedge\dots\wedge e_{i_k}.$$

mit $i_1 < \dots < i_k$ eine Basis des K-Vektorraums $K<e_1,\dots,e_n>$ bilden, d.h. weil die Koeffizienten $c^k_{i_1,\dots,i_k}$ durch c eindeutig bestimmt sind. Nach Definition gilt dann insbesondere

$$\tilde{f}(i(v_\alpha)) = \tilde{f}(e_\alpha) = f(e_\alpha),$$

d.h. es ist $\tilde{f}\circ i = f$. Nach Konstruktion ist $\tilde{f}$ K-linear. Wir haben noch zu zeigen,

$$\tilde{f}(c'\cdot c'') = \tilde{f}(t')\cdot \tilde{f}(t'') \text{ für } c',c''\in K< e_1,\dots,e_n>.$$

Das sieht man durch direktes Nachrechnen.
QED.

7. Ergängungen

7.1 Moduln

7.1.1 Definition

7.1.2 Beispiele und besondere Phänome

Minimale Erzeugendensysteme brauchen keine Basen zu sein
Basen brauchen nicht zu existieren

7.1.3 Verallgemeinerungen des Dimensionsbegriffs

rk M = Rang eines Moduls (= Maximalzahl linear unabhängiger Vektoren)
$\mu(M)$ = Minimale Anzahl der Elemente eines EZS
$\ell(M)$ = Länge eines Moduls

7.1.4 Kerne, Kokerne, exakte Sequenzen

7.1.5 Noethersche Ringe und Moduln

7.2 Darstellungen endlicher Gruppen (über $\mathbb{C}$)

vgl. das Buch von Serre

7.3 Kategorien und Funktoren

7.3.1 Der Begriff der Kategorie

Eine Kategorie $\mathfrak{C}$ besteht

(i) aus einer Klasse

$$|\mathfrak{C}| = \mathrm{ob}(\mathfrak{C}),$$

welche Objekte der Kategorie heißen,

(ii) aus einer Menge

$$\mathrm{Hom}(X, Y) = \mathrm{Hom}_{\mathfrak{C}}(X, Y)$$

für je zwei Objekte X, Y der Kategorie, deren Elemente Morphismen mit der Quelle X und dem Ziel Y heißen. Für jedes Element f aus Hom(X, Y) schreibt man

$$\mathrm{source}(f) = X \text{ und } \mathrm{target}(f) = Y$$

und nennt X Quelle von f und Y Ziel von f. Man verwendet dann auch die Bezeichnung

$$f\colon X \to Y \text{ oder } X \xrightarrow{f} X\,.$$

und sagt, f ist ein Morphismus von X nach Y.

(iii) aus einer Abbildung

$$\mathrm{Hom}(X, Y) \times \mathrm{Hom}(Y, Z) \to \mathrm{Hom}(X, Z),\ (f, g) \mapsto g\circ f\,,$$

für je drei Objekte X, Y, Z der Kategorie, welche Morphismen-Komposition heißt.

Dabei wird gefordert, daß die folgenden Bedingungen erfüllt sind.

(a) Die Morphismen Komposition ist assozialtiv, d.h. es gilt

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

für je drie Morphismen f, g, h mit

$$\mathrm{target}(h) = \mathrm{source}(g) \text{ und } \mathrm{target}(g) = \mathrm{source}(f).$$

(b) Existenz der identischen Morphismen. Für jedes Objekt $X \in |\mathcal{C}|$ gibt es einen Morphimus

$$id_X \in Hom(X, X)$$

mit

$$f \circ id_X = f$$

für jeden Morphismus mit der Quelle X und

$$id_X \circ f = f$$

für jeden Morphismus mit dem Ziel X.

(c) Die Hom-Mengen sind paarweis disjunkt[32]

$$Hom(X, Y) \cap Hom(X', Y') = \emptyset \text{ für } (X, Y) \neq (X', Y').$$

7.3.2 Beispiele

Ens - die Kategorie der Mengen
$Vect_K$ - die Kategorie der Vektorräume über einem Körper K
Groups - die Kategorie der Gruppen
Rings - die Kategorie der Ringe
Ab - die Kategorie der abelschen Gruppen
$Filt_{K,I}$ - die Kategorie der filtrierten K-Vektorräume über der linear geordneten Menge I
A-Mod - die Kategorie der linken A-Moduln
Die Kategorie der metrischen Räume und kontrahierenden Abbildungen
Die Kategorie der topologischen Räume und stetigen Abbildungen
Interpretation der Gruppen als Kategorie
Interpretation der halbgeordneten Mengen als Kategorie
Interpretation der topologischen Räume als Kategorien
Cat - die Kategorie der Kategorien
Die duale Kategorie

7.3.3 Spezielle Morphismen: Monomorphismen, Epimorphismen, Isomorphismen, Automorphismen

Ein Endomorphismus einer Kategorie $\mathcal{C}$ ist ein Morphismus $f: X \to Y$ mit $X = Y$.

Ein Morphismus $f: X \to Y$ der Kategorie $\mathcal{C}$ heißt Monomorphismus, wenn für je zwei unterschiedliche Morphismen g', g" mit dem Ziel X auch die Kompositionen $f \circ g'$ und $f \circ g''$ verschieden sind, d.h. die die folgende Implikation besteht

$$f \circ g' = f \circ g'' \Rightarrow g' = g''.$$

Ein Morphismus $f: X \to Y$ der Kategorie $\mathcal{C}$ heißt Epimorphismus, wenn für je zwei unterschiedliche Morphismen g', g" mit der Quelle Y auch die Kompositionen $g' \circ f$ und $g'' \circ f$ verschieden sind, d.h. die die folgende Implikation besteht

$$g' \circ f = g'' \circ f \Rightarrow g' = g''.$$

Ein Morphismus $f: X \to Y$ der Kategorie $\mathcal{C}$ heißt Isomorphismus, wenn es in $\mathcal{C}$ einen Morphismus $g: Y \to X$ gibt mit

$$g \circ f = id_X \text{ und } f \circ g = id_Y.$$

Ein Automorphismus ist ein Isomorphismus $f: X \to Y$ mit $X = Y$.

---

[32] Diese Bedingung reflektiert die Ansicht, daß Abbildungen, die zwischen verschiedenen Mengen abbilden, als verschieden angesehen werden sollten, auch wenn deren Abbildungsvorschriefen dieselben sind.

7.3.4 Beipiele: Epimorphie und Surjektivität, Bijektvität und Isomorphie

Sei X eine Kategorie, deren Morphismen Abbildungen sind und deren Morphismen-Komposition die gewöhnliche Zusammensetzung von Abbildungen ist. Dann sind surjektive Abbildungen Epimorphismen und injektive Abbildungen Monomorphismen.
Beispiel:
In der Kategorie der topologischen Räume (oder der Kategorie der metrischen Räume) ist die natürliche Einbettung

$$\mathbb{Q} \to \mathbb{R},\ x \mapsto x,$$

der rationalen in die reellen Zahlen ein Epimorphismus (welcher nicht surjektiv ist).
Beispiel:
Wir betrachten die Kategorie $\mathrm{Filt}_{K,I}$ mit I = {0,1} (versehen mit der natürlichen Ordnung). Bezeichne X den Vektorraum K mit der Filtration

$$0 \subseteq K$$

und Y den Vektorraum K mit der Filtration

$$K \subseteq K.$$

Die identischen Abbildung definiert dann einen Morphismus

$$f: X \to Y.$$

Dieser ist kein Isomorphismus.

7.3.5 Funktoren

7.3.6 Beispiele für Funktoren

7.3.7 Funktorielle Morphismen (natürliche Transformationen)

7.3.8 Additive Kategorien (und Beispiele)

7.3.9 Abelschen Kategorien (und Beispiele)

Index

—C—

—D—

—E—

—F—

—G—

—H—

—I—

—J—

—K—

—L—

—M—

—N—

—O—

Inhalt