Professor Dr. Jürgen Ritter
Lehrstuhl für Algebra und Zahlentheorie
Institut für Mathematik · Universität Augsburg

Vorlesung Lineare Algebra 1 , Sommersemester 2005

In das Gebiet *Algebra* gehören alle Untersuchungen, die auf *algebraische Gleichungen* führen, also auf Gleichungen, in denen Koeffizienten aus bekannten Bereichen und *Unbekannte* über die Vorschriften $+, -, \cdot, :$ verbunden sind. Die berühmte *Fermatsche Gleichung* $x^n + y^n = z^n$, $n = 3, 4, 5, \dots$, x, y, z positive ganze (unbekannte) Zahlen, ist ein Beispiel einer solchen Gleichung. Eine Minimalanforderung an die Koeffizientenbereiche ist dabei, daß sie gegenüber $\pm$ und $\cdot$ abgeschlossen sind.

In der *Linearen Algebra* beschränkt man sich auf *lineare Gleichungen*; linear besagt dabei, daß die Unbekannten isoliert vorkommen, also weder miteinander multipliziert noch durcheinander dividiert werden. Darüber hinaus fordert man die Abgeschlossenheit des Koeffizientenbereichs gegenüber der Division (außer durch Null).

Lineare Abhängigkeit in einem endlichen System kann in allgemeinster Weise durch das gleichzeitige Erfülltsein von den m Gleichungen

$$
(1) \qquad \begin{array}{ccccccc}
\alpha_{11}x_1 & + & \alpha_{12}x_2 & + \dots + & \alpha_{1n}x_n & = & \beta_1 \\
\alpha_{21}x_1 & + & \alpha_{22}x_2 & + \dots + & \alpha_{2n}x_n & = & \beta_2 \\
\vdots & & \vdots & & \vdots & & \vdots \\
\alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + \dots + & \alpha_{mn}x_n & = & \beta_m
\end{array}
$$

in den n Unbekannten $x_1, \dots, x_n$ ausgedrückt werden; hier sind die α_{ij} und β_ν konstante Parameter aus dem Koeffizientenbereich – etwa reelle Zahlen.

Die Vorlesung hat drei Ziele, nämlich Antworten zu geben auf

- Wie hängen die Lösungen von (1) von den Parametern α_{ij}, β_ν ab; wann gibt es überhaupt Lösungen; können die Lösungen explizit angegeben werden?
- Welche mathematischen Probleme sind linear angreifbar?
- Gibt es gute geometrische Anschauungen, die uns helfen, Ideen für die Angreifbarkeit tatsächlicher Probleme zu entwickeln?

Ist zum Beispiel $n = 3$, so beschreibt i. allg. (1) für $m = 1$ eine Ebene im dreidimensionalen Raum, für $m = 2$ den Schnitt zweier Ebenen und für $m = 3$ den dreier Ebenen.

Zur zweiten Frage oben vergleiche das folgende Beispiel, auf das wir bei späterer Gelegenheit des öfteren mit (B) zurückverweisen werden.

V bezeichne die Gesamtheit aller 3×3 *magischen Quadrate* x: das sind Schemata der Form

$$x = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{pmatrix},$$

in denen alle Zeilen, Spalten- und die beiden Diagonalsummen übereinstimmen; die Einträge x_i seien dabei reelle Zahlen.
Die Aufgabe, die sich sofort stellt, heißt: Schreibe alle solchen x explizit auf!

Um dies zu tun, beobachten wir:

1. Magische Quadrate können addiert und mit reellen Zahlen skaliert werden, nämlich jeweils komponentenweise. Die Menge V trägt damit eine Struktur; solche besonderen Mengen werden später *Vektorräume* genannt.

2. Die Zuordnung $\sigma : x \mapsto s_x := x_1 + x_2 + x_3$ ist eine reellwertige Funktion auf V, die die Struktur erhält
$$\begin{aligned} \sigma(x+y) &= \sigma(x) + \sigma(y) \\ \sigma(rx) &= r\sigma(x) \end{aligned}$$
Hierbei sind x, y magische Quadrate und r ist eine beliebige reelle Zahl.
Man nennt σ eine *lineare Abbildung* und die Teilmenge W von V, die aus allen x mit $\sigma(x) = 0$ besteht, den *Kern* von σ.

3. Ist x aus V, so gilt $x = \tilde{x} + y$ mit $\tilde{x}$ als dem Quadrat mit dem konstanten Eintrag $s_x/3$ und mit einem geeigneten $y \in W$. Um also V zu kennen, reicht es, W zu kennen.

4. Die x aus W erfüllen $x_5 = 0$, da $2s_x + 3x_5 = 3s_x$.

5. Die Abbildung $f : x \mapsto (x_1, x_4)$ identifiziert W mit allen Paaren (r, s) reeller Zahlen r, s. Sie ist im obigen Sinne linear.

6. Es folgt, daß jedes magische Quadrat x eindeutig als Linearkombination
$$x = r_0 w_0 + r_1 w_1 + r_2 w_2 \text{ mit reellen } r_0, r_1, r_2$$
geschrieben werden kann, wobei w_0 das Quadrat mit dem konstanten Eintrag 1 ist und
$$w_1 = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{pmatrix}, \; w_2 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$
Umgekehrt ist auch jede solche Linearkombination ein magisches Quadrat. Man beachte, daß w_1 ein Urbild von $(1, 0)$ und w_2 eines von $(0, 1)$ unter f ist. Derartige *freie Erzeugende* von V wie w_0, w_1, w_2 werden später eine *Basis* von V genannt werden.

7. Der Zusammenhang mit dem System (1) ist der: (1) übersetzt sich in

$$\begin{array}{ccccccccccccccccccc}
x_1 & + & x_2 & + & x_3 & & & & & & & & & & & & & = & s \\
& & & & & & x_4 & + & x_5 & + & x_6 & & & & & & & = & s \\
& & & & & & & & & & & & x_7 & + & x_8 & + & x_9 & = & s \\
x_1 & + & & & & & & & x_5 & + & & & & & & & x_9 & = & s \\
x_1 & + & & & & & x_4 & + & & & & & x_7 & & & & & = & s \\
& & x_2 & + & & & & & x_5 & + & & & & & x_8 & & & = & s \\
& & & & x_3 & + & & & & & x_6 & + & & & & & x_9 & = & s \\
& & & & x_3 & + & & & x_5 & + & & & x_7 & & & & & = & s
\end{array}$$

mit s beliebig reell.

Literatur

Artin, E.	*Analytische Geometrie I, II* (Univ. Hamburg)
Brieskorn, E.	*Lineare Algebra und Analytische Geometrie 1,2* (Vieweg, 1983, 1985)
Bröcker, T.	*Lineare Algebra und Analytische Geometrie* (Birkhäuser Verlag, 2003)
Fischer, G.	*Lineare Algebra* (Vieweg 1980)
Greub, W.	*Lineare Algebra* (Heidelberger Taschenbücher - Band 179, Springer-Verlag)
Jänich, K.	*Lineare Algebra* (Springer-Verlag, 1991)
Jacobson, N.	*Lectures in Abstract Algebra* (Springer-Verlag, 1975)
Kaplansky, I.	*Linear Algebra and Geometry* (Chelsea, 1969)
Klingenberg, W.	*Lineare Algebra und Geometrie* (Springer-Verlag, 1984)
Koecher, M.	*Lineare Algebra und Analytische Geometrie* (Grundwissen der Mathematik 2, Springer-Verlag, 1983)
Kostrikin, A./Manin, Y.	*Linear Algebra and Geometry* (Gordon & Breach, 1989)
Kowalsky, H.-J.	*Lineare Algebra* (Walter de Gruyter & Co., 1970)
Lamprecht, E.	*Lineare Algebra* (UTB 1224)
Lang, S.J.	*Linear Algebra* (Addison-Wesley Publishing Company, 1972)
Lorenz, F.	*Lineare Algebra I u. II* (Bibliographisches Institut Mannheim, Hochschultaschenbücher 601, 605)
Pickert, G.	*Analytische Geometrie* (Leipzig, 1961)
Smith, L.	*Linear Algebra* (Springer-Verlag, 1978)
Sperner, E.	*Einführung in die Analytische Geometrie und Algebra 1* (Vandenhoeck u. Ruprecht, Göttingen, 1969)
Stammbach, U.	*Lineare Algebra* (Teubner, 1980)

1. Kapitel: Sprache

Eine *Menge* ist eine wohlbestimmte Gesamtheit von gewissen Objekten [1]; diese Objekte nennt man die *Elemente* der Menge. Im gegenwärtigen Kapitel werden Mengen mit großen lateinischen

[1] Das ist ein grammatikalisch richtiger deutscher Satz, der vielleicht sogar dem Leser ein Gefühl für den zu erklärenden Begriff *Menge* geben mag. Mathematisch gesehen taugt er allerdings nichts. Dennoch wollen wir uns hier an dieser Stelle nicht weiter ins Zeug legen: die Mengenlehre ist nicht unser zu studierendes Objekt. Alle unsere Mengen lassen sich aus der Menge der natürlichen Zahlen (s.u.) durch die im folgenden genannten und als zulässig bezeichneten Operationen gewinnen. Davon nicht Erfaßtes lassen wir außer acht.

Buchstaben bezeichnet. Man schreibt $M = \{a, b, c, \ldots\}$, wenn $a, b, c, \ldots$ die (nicht notwendig verschiedenen) Elemente der Menge M sind.
$a \in M$ heißt: a ist Element von M; $a \notin M$ heißt: a ist kein Element von M.
Eine auch gebräuchliche Schreibweise für eine Menge M ist: $M = \{z : z$ erfüllt eine gewisse Eigenschaft$\}$.

Eine Menge T heißt *Teilmenge* von M, wenn jedes $t \in T$ auch Element von M ist; Notationen: $T \subset M$, oder auch: $t \in T \Longrightarrow t \in M$, gelesen: $t \in T$ impliziert $t \in M$.

Spezielle Mengen mit fixierter Bezeichnung:

$\emptyset$ = leere Menge: sie enthält gar kein Element; sie ist Teilmenge in jeder Menge

$\mathbb{N} = \{1, 2, 3, \ldots\}$ = Menge der *natürlichen* Zahlen

$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ = Menge der *ganzen* Zahlen

$\mathbb{Q} = \{m/n : \; m, n \in \mathbb{Z}, n \neq 0\}$ = Menge der *rationalen* Zahlen

$\mathbb{R} = \{$ unendliche Dezimalbrüche $\}$ = Menge der *reellen* Zahlen.

M heißt endlich, wenn M nur endlich viele Elemente besitzt; sonst *unendlich.* $|M|$ bezeichnet die Anzahl aller Elemente von M; falls die nicht endlich ist, wird $|M| = \infty$ gesetzt.

Operationen mit Teilmengen

T_1 und T_2 seien Teilmengen von M.

Vereinigung:	$T_1 \cup T_2$	$=$	$\{m \in M : m \in T_1$ oder $m \in T_2\}$
Durchschnitt:	$T_1 \cap T_2$	$=$	$\{m \in M : m \in T_1$ und $m \in T_2\}$
Differenz:	$T_1 \backslash T_2$	$=$	$\{m \in T_1 : m \notin T_2\}$
Komplement:	$\complement T_1$	$=$	$\{m \in M : m \notin T_1\}$

Beliebige Vereinigungen und Durchschnitte

Gegeben seien M und eine "Indexmenge" I, die aus Teilmengen T von M besteht. Dann ist
$\cup_{T \in I} T \; = \; \{m \in M : m$ gehört zu wenigstens einem $T \in I\}$ die Vereinigung und
$\cap_{T \in I} T \; = \; \{m \in M : m$ gehört zu allen $T \in I\}$ der Durchschnitt der $T \in I$.

Zulässige Mengenbildungen aus gegebenen Mengen

1. *Teilmengenbildung*

2. *Produktmengenbildung aus M und N:* $M \times N = \{(m, n) : m \in M, n \in N\}$; in $M \times N$ gilt $(m, n) = (m_1, n_1)$ genau dann, wenn sowohl $m = m_1$ als auch $n = n_1$. Folglich

$$|M \times N| = |M| \cdot |N| \, .$$

Allgemeiner : Sind T_i Mengen, wobei i eine Indexmenge I durchläuft, so ist

$$\prod_{i \in I} T_i = \{(\ldots, t_i, \ldots) : t_i \in T_i\} \, .$$

3. *Potenzmengenbildung*: $\mathfrak{P}(M) = \{T : T \subset M\}$ ist die Menge aller Teilmengen von M. Ist M endlich, so auch $\mathfrak{P}(M)$, und es gilt $|\mathfrak{P}(M)| = 2^{|M|}$.

Beweismethodik und Mengen

Für gewisse Elemente der Menge M sei eine Aussage $\mathfrak{A}$ zu beweisen. Man unterscheidet zwischen der *direkten* und *indirekten* Methode. Zu $\mathfrak{A}$ gehört eine Teilmenge A von M, die "Richtigkeitsmenge", $A = \{x \in M : x \text{ erfüllt } \mathfrak{A}\}$. Die direkte Methode nachzuprüfen, ob $\mathfrak{A}$ für ein $m \in M$ zutrifft, besteht nun im Beweis, daß $m \in A$, die indirekte Methode darin, aus der Annahme $m \notin A$ einen Widerspruch herzuleiten. Entsprechend bedeutet eine *Implikation* $\mathfrak{A} \Longrightarrow \mathfrak{B}$, daß $A \subset B$, wozu $\complement B \subset \complement A$ gleichwertig ist, also (nicht $\mathfrak{B}$) $\Longrightarrow$ (nicht $\mathfrak{A}$).
Am Beispiel (B) sei das kurz erläutert. M sei die Menge

$$M = \left\{ x = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{pmatrix} : x_i \in \mathbb{R} \ (1 \le i \le 9) \right\}.$$

Dann ist $V = A$ diejenige Teilmenge, die durch die Aussage

$$\mathfrak{A} : \ x \text{ ist ein magisches Quadrat}$$

definiert ist.

Ein ganz wichtiges Beispiel in diesem Zusammenhang ist die Methode der *vollständigen Induktion* bei Beweisen. Sie beruht auf dem Induktionsargument, nach dem eine Teilmenge $T \subset \mathbb{N}$ genau dann $= \mathbb{N}$ ist, wenn folgendes gilt:

$$1 \in T \quad \& \quad [\ n \in T \Longrightarrow n+1 \in T\ ]\ .$$

Danach ist eine Aussage $\mathfrak{A}$ über natürliche Zahlen richtig, falls $1 \in A = \{x \in \mathbb{N} : x \text{ erfüllt } \mathfrak{A}\}$ und falls $[\ x \in A \Longrightarrow x+1 \in A\ ]$ gilt. Hieraus folgt etwa ohne weiteres die binomische Formel

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \quad \text{mit} \quad \binom{n}{i} = \frac{n!}{i!(n-i)!} \in \mathbb{N}\ ,$$

dem Binomialkoeffizienten.

Gleichheit

Für den Begriff $=$ verweisen wir einmal mehr nur zurück auf unser Sprachempfinden, halten jedoch folgende Regeln für ihn fest: Sind x, y, z Elemente einer Menge, so gilt

$x = x$ (Reflexivität)

$x = y \Longrightarrow y = x$ (Symmetrie)

$x = y \ \&\ y = z \Longrightarrow x = z$ (Transitivität)

Funktionen

Gegeben sind zwei Mengen X und Y. Eine Zuordnung f, die jedem $x \in X$ ein *bestimmtes* $y \in Y$ zuweist, heißt *Funktion oder Abbildung* von X nach Y. Schreibweise: $f : X \longrightarrow Y$, oder $X \xrightarrow{f} Y$, oder $f : x \mapsto y$, oder $f(x) = y$. X heißt der *Definitionsbereich* von f; die Teilmenge $\mathrm{im}(f) = f(X) = \{f(x) : x \in X\}$ von Y heißt *Bild-* oder *Wertebereich* von f. Die Funktion f heißt *eineindeutig* oder *injektiv*, wenn $x_1 = x_2$ aus $f(x_1) = f(x_2)$ folgt, *auf* oder *surjektiv*, wenn $f(X) = Y$, und *bijektiv*, wenn f injektiv und surjektiv ist.
Im Beispiel (B) tauchten zwei Funktionen auf: das nicht injektive, aber surjektive $\sigma : V \to \mathbb{R}$ und das bijektive $f : W \to \{(r, s) : r, s \in \mathbb{R}\}$.

Abkürzungen

$\exists$ heißt "es gibt" oder "es existiert"
$\exists'$ heißt "es gibt genau ein"
$\forall$ heißt "für alle" oder "für jedes".

Gruppe, Ring, Körper

1. Eine *Gruppe* ist eine Menge G mit einer (durch eine Verknüpfung gegebenen) Struktur:

   Jedem Paar $a, b \in G$ von Elementen aus G ist ein neues Element, das mit ab bezeichnet wird, in G gemäß folgender Gesetze zugeordnet

   $\exists 1 \in G$ mit $1a = a1 = a$ für alle $a \in G$
   $\forall a \in G\ \exists a^{-1} \in G$ mit $aa^{-1} = a^{-1}a = 1$
   $\forall a, b, c \in G$ gilt: $a(bc) = (ab)c$ (Assoziativgesetz)

   Man schreibt manchmal auch $a + b$ statt ab; in dem Falle wird das Neutralelement nicht 1 sondern 0 und das Inverse nicht a^{-1} sondern $-a$ genannt.
   Gilt $ab = ba$ für alle $a, b, \in G$, so heißt die Gruppe G *kommutativ* oder *abelsch*.

   Einfache Beobachtungen zur Definition: Die 1 ist eindeutig bestimmt; a^{-1} ist durch a eindeutig bestimmt; $(a^{-1})^{-1} = a$; $[ac = ab \Longrightarrow c = b]$.

2. Ein *Ring* ist eine abelsche Gruppe R bezüglich $+$, die zudem noch eine zweite Verknüpfung trägt, die wie oben mit ab bezeichnet wird. Folgendes wird für die zweite Verknüpfung und ihre Relation zur ersten verlangt:
   $ab \in R$, $[\exists 1 \in R$ mit $1a = a1 = a]$, $\forall a, b, c \in R$ gilt: $a(bc) = (ab)c$, $a(b + c) = ab + ac$, $(a + b)c = ac + bc$.

   Die letzten beiden Gesetze heißen *Distributivgesetze.*

   Beobachtungen zur Definition: $a0 = 0a = 0$; $(-a)b = -ab$; $[a + c = a + b \Longrightarrow c = b]$.

   Gilt $ab = ba$ für alle $a, b \in R$, so heißt der Ring *kommutativ.*

3. Ein *Körper* ist ein Ring K, dessen von Null verschiedene Elemente bezüglich der Verknüpfung "mal" eine Gruppe bilden. Also:

$$\forall a \neq 0\ \exists a^{-1} \in K : aa^{-1} = a^{-1}a = 1\ .$$

Gilt stets $ab = ba$, heißt der Körper kommutativ, sonst *schief.*

Beobachtung zur Definition: $ab = 0 \Longrightarrow a = 0$ oder $b = 0$. *Aber:* In einem Ring R folgt i. allg. nicht aus $ab = 0$, daß wenigstens einer der Faktoren, a oder b, verschwindet:

$R = \mathbb{Z}/4 = \{0, 1, a_2, a_3\}$ mit den folgenden Regeln ist ein Ring, in dem $a_2 \cdot a_2 = 0$ ist:
0 ist das Neutralelement bezüglich $+$, 1 das bezüglich $\cdot$; R ist kommutativ.
$1 + a_2 = a_3,\ 1 + a_3 = 0,\ a_2 + a_2 = 0,\ a_2 + a_3 = 1,\ a_3 + a_3 = a_2;$
$a_2 \cdot a_2 = 0,\ a_2 \cdot a_3 = a_2,\ a_3 \cdot a_3 = 1.$
Diese Regeln spiegeln übrigens die gewöhnliche Addition und Multiplikation in $\mathbb{Z}$ wider, wenn man jede Zahl hier mit dem kleinsten nichtnegativen Rest nach Division durch 4 identifiziert (a_2, a_3 entsprechen also den Resten 2 beziehungsweise 3).

Die Parameterbereiche in dieser Vorlesung sind stets Körper. Es schadet übrigens nicht, wenn man sich zunächst auf $\mathbb{Q}$ oder $\mathbb{R}$ (oder gar $\mathbb{C}$, den Körper der komplexen Zahlen) beschränkt.

Beispiele

- $\mathbb{Z}$ ist eine Gruppe bezüglich $+$ und ein Ring bezüglich $+$ und $\cdot$.
- $\mathbb{N}$ ist keine Gruppe. $\mathbb{Q}$, $\mathbb{R}$ sind Körper; $\mathbb{Z}$ nicht.
- $\mathbb{Z}/2 = \{0, 1\}$ mit den Vereinbarungen $1 + 1 = 0, 1 \cdot 1 = 1, 1 \cdot 0 = 0 \cdot 1 = 0$ ist ein Körper mit zwei Elementen (hier rechnet man also in Wahrheit wieder nur mit Resten, nämlich nach Division durch 2 diesmal – genauso wie es ein Computer tut).
- Die *symmetrische Gruppe* S_n ist die Menge der bijektiven Selbstabbildungen der endlichen Menge $M = \{m_1, \ldots, m_n\}$ mit n Elementen; n ist hier eine natürliche Zahl. Der Menge S_n wird eine Multiplikation durch Hintereinanderausführung der Abbildungen aufgeprägt:

$$\sigma, \tau \in S_n : (\sigma\tau)(m_i) = \sigma(\tau(m_i)) \ .$$

  Das 1-Element ist die identische Abbildung;
  σ^{-1} ist durch $\sigma^{-1}(m_i) = m_j \Longleftrightarrow \sigma(m_j) = m_i$ erklärt.

  Beobachtungen:

  $S_1 = \{1\}\,, \quad S_2 = \{1, \tau\}$ mit $\tau(m_1) = m_2, \tau(m_2) = m_1$, also $\tau^2 = 1$,
  $S_3 = \{1, \sigma, \sigma^2, \tau, \sigma\tau, \sigma^2\tau\}$ mit $\sigma(m_1) = m_2, \sigma(m_2) = m_3, \sigma(m_3) = m_1$, also $\sigma^3 = 1$; $\tau(m_1) = m_1, \tau(m_2) = m_3, \tau(m_3) = m_2$, also $\tau^2 = 1$.

  S_3 (und allgemein S_n für $n \geq 3$) ist nicht kommutativ: $\sigma\tau \neq \tau\sigma = \sigma^2\tau$.

  Die Ordnung $|S_n|$ von S_n, d.i. die Anzahl der Elemente (der *Permutationen*) in S_n, ist $= n! = \prod_{i=1}^{n} i$.
  Es gibt eine Abbildung $\mathrm{sgn} : S_n \to \{\pm 1\}$, die $\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\tau)$ erfüllt.
  Sie ist so definiert: Bilde den Ausdruck $\Delta = \prod_{1 \leq i < j \leq n}(i - j)$. Wende hierauf $\sigma \in S_n$ wie folgt an

$$\sigma\Delta := \prod_{1 \leq i < j \leq n} (\sigma(i) - \sigma(j)) \ ;$$

  dabei sei $\sigma(i) = k$, falls $\sigma(m_i) = m_k$. Offenbar ist $\sigma\Delta = \pm\Delta$; wir setzen $\sigma\Delta = (\mathrm{sgn}\sigma) \cdot \Delta$.

Weitere Begriffe

G sei eine Gruppe (multiplikativ geschrieben). Eine *Untergruppe* U von G ist eine Teilmenge in G, die

$$1 \in U \ ; \ a \in U \Longrightarrow a^{-1} \in U \ ; \ a, b \in U \Longrightarrow ab \in U$$

erfüllt. Notation: $U \leq G$. Die Untergruppe $U \leq G$ heißt *Normalteiler*, falls

$$\forall x \in G \ \forall u \in U \ : \ x^{-1}ux \in U$$

gilt. Notation: $U \lhd G$.

Sind G und H Gruppen, so heißt eine Abbildung $f : G \to H$ ein *Homomorphismus*, wenn

$$f(ab) = f(a)f(b) \qquad (\forall a, b \in G)$$

gilt. Insbesondere ist $f(1) = 1$. - Die Signumsabbildung sgn: $S_n \to \{\pm 1\}$ ist ein Beispiel.

$\ker f = \{x \in G : f(x) = 1\}$ heißt der *Kern* des Homomorphismus f; er ist ein Normalteiler in G. Der Kern der Signumsabbildung sgn: $S_n \to \{\pm 1\}$ heißt die *alternierende Gruppe vom Grad* n und wird mit A_n bezeichnet; z.B. $A_3 = \{1, \sigma, \sigma^2\}$.

Der Homomorphismus heißt *Monomorphismus*, falls f injektiv, *Epimorphismus*, falls f surjektiv, und *Isomorphismus*, falls f bijektiv ist. Im letzten Fall ist die Umkehrabbildung f^{-1} automatisch ein Homo-, also ein Isomorphismus.

Auch hierzu ein paar Bemerkungen:

1. Das Bild $\mathrm{im}(f)$ des Homomorphismus f ist eine Untergruppe von H, i. allg. aber kein Normalteiler.

2. Ist G abelsch, so stimmen die Begriffe *Normalteiler* und *Untergruppe* überein.

3. In der Definition des Normalteilers reicht es, für x nur die Elemente in G außerhalb U zu nehmen.

4. Ist A eine abelsche Gruppe (bez. $+$), so ist

$$R = \mathrm{Hom}(A, A) = \{f : f \text{ ist Homomorphismus von } A \text{ nach } A\}$$

ein Ring. Dabei sind Addition und Multiplikation in R so erklärt

$$(f_1 + f_2)(a) = f_1(a) + f_2(a)\,, \quad (f_1 f_2)(a) = f_1(f_2(a))\,.$$

Homomorphismen einer Gruppe in sich nennt man übrigens auch *Endomorphismen* und bijektive solche *Automorphismen*.

Sind R und S Ringe, so heißt ein $f : R \to S$ *Homomorphismus*, wenn für alle $a, b, \in R$

$$f(a + b) = f(a) + f(b)\,, \qquad f(ab) = f(a)f(b)$$

erfüllt ist. Mono-, Epi- und Isomorphismus haben m.m. die gleiche Bedeutung wie oben. Der Kern von f ist hier

$$\ker f = \{x \in R : f(x) = 0\} \ ;$$

dies ist ein *zweiseitiges Ideal* von R, also eine Teilmenge I von R, für die gilt

$$0 \in I\,; \quad a \in I \implies -a \in I\,; \quad a, b \in I \implies a + b \in I \ ;$$

$$a \in I, x \in R \implies ax \in I \ \& \ xa \in I \ .$$

Ist R ein Körper und $f : R \to S$ ein Homomorphismus, so ist entweder $f(x) = 0$ für jedes $x \in R$ oder f ist injektiv.

Beispiel eines nichtkommutativen Körpers – der Quaternionenschiefkörper $\mathbb{H}$

Er entsteht so: Als Menge nehmen wir alle Quadrupel $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ mit reellen Einträgen. Gleichheit zweier solcher Quadrupel bedeute komponentenweise Gleichheit. Wir definieren:

$$\lambda(\alpha_0, \alpha_1, \alpha_2, \alpha_3) = (\lambda\alpha_0, \lambda\alpha_1, \lambda\alpha_2, \lambda\alpha_3) = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)\lambda, \ \lambda \in \mathbb{R}$$

$$(\alpha_0, \alpha_1, \alpha_2, \alpha_3) + (\beta_0, \beta_1, \beta_2, \beta_3) = (\alpha_0 + \beta_0, \alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3)$$

$$1 := (1, 0, 0, 0), \ i = (0, 1, 0, 0), \ j = (0, 0, 1, 0), \ k = (0, 0, 0, 1) \ .$$

Dann ist $\mathbb{H}$ eine Gruppe bezüglich $+$, und jedes Element in $\mathbb{H}$ läßt sich eindeutig als

$$\lambda_0 1 + \lambda_1 i + \lambda_2 j + \lambda_3 k$$

mit reellen λ_i schreiben. Wir definieren jetzt die Multiplikation durch

1 ist das Einselement

$ij = -ji = k, i^2 = j^2 = -1$

und durch die Gültigkeit des Assoziativgesetzes und der Distributivgesetze.

Damit ist $\mathbb{H}$ ein Schiefkörper; das multiplikative Inverse zu $\lambda_0 1 + \lambda_1 i + \lambda_2 j + \lambda_3 k \neq 0$ ist

$$\frac{1}{\lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_3^2} \ (\lambda_0 1 - \lambda_1 i - \lambda_2 j - \lambda_3 k) \ .$$

Bemerkung: Für $\lambda_0 1$ schreiben wir fortan nur λ_0.

Der Quaternionenschiefkörper ist in der Geometrie, Zahlentheorie und in der Physik von Bedeutung. Er wurde von William R. Hamilton (1. Hälfte des 19. Jahrhundert) zur koordinatenfreien Beschreibung von längen- und winkeltreuen Selbstabbildungen des 3-dimensionalen Raumes eingeführt und ist der einzige ‘echte’ Schiefkörper von endlicher Dimension über $\mathbb{R}$.

Bemerkung: Die Teilmenge $\mathbb{C} = \{\lambda_0 1 + \lambda_1 i \ : \ \lambda_0, \lambda_1 \in \mathbb{R}\}$ von $\mathbb{H}$ wird mit der Addition und Multiplikation in $\mathbb{H}$ ein kommutativer Körper, der $\mathbb{R}$ als Teilkörper $\{\lambda_0 1 : \lambda_0 \in \mathbb{R}\}$ enthält. $\mathbb{C}$ heißt der Körper der komplexen Zahlen; $i \in \mathbb{C}$ ist eine Quadratwurzel aus -1 (die es in $\mathbb{R}$ selbst nicht gibt).

2. Kapitel: Vektorraum, Lineare Abhängigkeit, Basis und Dimension

Für die gesamte weitere Vorlesung steht K für einen fest gewählten (Schief-)Körper, z.B. $K = \mathbb{Q},\ \mathbb{R},\ \mathbb{C},\ \mathbb{H},\ \mathbb{Z}/2,\ \ldots$

Definition 2.1. *Ein Vektorraum V über K ist eine abelsche Gruppe (bez. $+$), in der zu jedem $v \in V$ und $\lambda \in K$ ein neues Element $v\lambda \in V$ erklärt ist, so daß gilt:*

$$\begin{aligned} v1 &= v \\ v(\lambda_1 + \lambda_2) &= v\lambda_1 + v\lambda_2 \\ (v_1 + v_2)\lambda &= v_1\lambda + v_2\lambda \\ v(\lambda_1\lambda_2) &= (v\lambda_1)\lambda_2\ ; \end{aligned}$$

*hier sind v, v_1, v_2 beliebige Elemente (*Vektoren*) aus V und $\lambda, \lambda_1, \lambda_2$ beliebige Elemente (*Skalare*) aus K.*

Man beachte, daß $v_1 \cdot v_2$ für $v_1, v_2 \in V$ *nicht* erklärt ist – und genausowenig λv für $\lambda \in K, v \in V$. Ist allerdings K kommutativ, so können und wollen wir λv durch $v\lambda$ definieren; es gelten dann die zu obiger Definition entsprechenden Regeln für diese vertauschte Multiplikation.

Beobachtungen und Beispiele

- $v \cdot 0 = 0$ (hier ist die linke Null das Nullelement von K, die rechte das von V); $0 \cdot \lambda = 0$ (hier sind beide Nullen die von V); $v\lambda = 0 \Longrightarrow v = 0 \in V$ oder $\lambda = 0 \in K$; $v(-\lambda) = (-v)\lambda = -(v\lambda)$
- Sowohl das V als auch das W aus Beispiel (B) sind Vektorräume über $\mathbb{R}$
- $K = V$ wird über die Multiplikation ein Vektorraum über K
- $\mathbb{H}$ ist ein Vektorraum über $K = \mathbb{R}$
- V_n, der n-dimensionale Standardvektorraum über K, ist wie folgt erklärt:

  $V_n = \left\{ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} : \alpha_i \in K \right\}$ mit komponentenweiser Addition, i.e.,

  $$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}$$

  und skalarer Multiplikation

  $$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \cdot \lambda = \begin{pmatrix} \alpha_1\lambda \\ \vdots \\ \alpha_n\lambda \end{pmatrix}\ .$$

Die Vektoren $\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ aus V_n nennen wir *Spaltenvektoren.* Natürlich bedeute Gleichheit zwischen zwei solchen Gleichheit an allen Komponenten.

Definition 2.2. *$v_1, \ldots, v_m$ seien m Vektoren aus V. Sie heißen linear abhängig, falls es Skalare $\lambda_1, \ldots, \lambda_m \in K$ gibt, die nicht alle $= 0$ sind, so daß*

$$v_1\lambda_1 + \ldots + v_m\lambda_m = 0 \text{ ist .}$$

Andernfalls heißen sie linear unabhängig.

Beobachtungen

Ist $m = 1$, so ist v_1 genau dann linear abhängig, wenn $v_1 = 0$ ist. Kommt unter den $v_1, \ldots, v_m$ der Nullvektor vor, so sind $v_1, \ldots, v_m$ linear abhängig.

Ist $V = V_1$ (oder $V = K$) und $m \geq 2$, so sind je m Vektoren $v_1, \ldots, v_m$ linear abhängig.

In $V = \mathbb{H}$ sind $1, i, j, k$ linear unabhängig; hingegen sind

$$v_1 = 1 + j + k,\; v_2 = 1/2 + i + 3j + 4k,\; v_3 = 3 + 4i + 13j + 17k$$

linear abhängig: $v_1 + 4v_2 - v_3 = 0$. Mehr als vier Quaternionen sind stets linear abhängig, wie wir später einsehen werden.

In $V = V_n$ sind die n Vektoren (im folgenden *Standardbasisvektoren* genannt)

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

die 1 an der i-ten Stelle (von oben gelesen), linear unabhängig und, wie wir bald sehen werden, mehr als n Vektoren stets linear abhängig.

Im Beispiel (B) sind w_0, w_1, w_2 linear unabhängig.

Definition 2.3. *Die nichtleere Teilmenge U von V heißt Teilraum, in Zeichen $U \leq V$, wenn*

$$\begin{aligned} u_1, u_2 \in U &\implies u_1 \pm u_2 \in U \\ u \in U, \lambda \in K &\implies u\lambda \in U \;. \end{aligned}$$

U ist also selbst ein Vektorraum.

Beispiele

$U=\{0\}$; $U=V$; $U=W$ im Beispiel (B); $U=\left\{\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} : \alpha_i\in K,\ \alpha_2=0\right\}\le V_n$. Allerdings ist $U=\left\{\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} : \alpha_i\in K,\ \alpha_2=1\right\}$ kein Teilraum.

Aus zwei Teilräumen U_1 und U_2 von V können die folgenden neuen Teilräume konstruiert werden: der Durchschnitt $U_1\cap U_2\le V$ und die Summe $U_1+U_2=\{v=u_1+u_2 : u_1\in U_1,\ u_2\in U_2\}\le V$; jedoch ist $U_1\cup U_2$ i. allg. kein Teilraum mehr. Die Summe heißt *direkt*, in Zeichen $U_1\oplus U_2$, wenn $U_1\cap U_2=0$, wenn also die Darstellung $v=u_1+u_2$ eindeutig ist.

Ein weiteres wichtiges **Beispiel** entsteht so. Es seien $v_1,\dots,v_m$ Vektoren aus V. Der *Aufspann* $\langle v_1,\dots,v_m\rangle$ ist die Teilmenge

$$\{\sum_{i=1}^m v_i\alpha_i : \alpha_i\in K\} \text{ in } V\,;$$

dies ist offenkundig ein Teilraum von V.

Konstruktion von neuen Vektorräumen aus gegebenen; Homomorphismen

1. Sind V_1,V_2 zwei K-Vektorräume, so wird $V_1\times V_2=\{(v_1,v_2) : v_1\in V_1, v_2\in V_2\}$ über

$$\begin{aligned}(v_1,v_2)+(\tilde v_1,\tilde v_2) &= (v_1+\tilde v_1, v_2+\tilde v_2)\\ (v_1,v_2)\lambda &= (v_1\lambda, v_2\lambda)\end{aligned}$$

ebenfalls ein K-Vektorraum. Allgemeiner: I sei eine Indexmenge und V_i, für jedes $i\in I$, ein K-Vektorraum. Dann ist

$$\prod_{i\in I} V_i=\Big\{(\dots,v_i,\dots) : v_i\in V_i\Big\}$$

analog zu obigem $V_1\times V_2$ ein K-Vektorraum. Darin liegt der Unterraum $\bigoplus_{i\in I} V_i$, der aus allen Vektoren $(\dots,v_i,\dots)$ besteht, in denen nur endlich viele $v_i\neq 0$ sind. Ist I endlich, so gibt es keinen Unterschied zwischen dem *direkten Produkt* $\prod_{i\in I} V_i$ und und der *direkten Summe* $\bigoplus_{i\in I} V_i$ der V_i.

2. Es sei W ein Teilraum von V. Zu W betrachten wir die *Nebenklassen*

$$\overline{v}\overset{\text{def}}{=} v+W=\{v+w : w\in W\}\subset V\ .$$

Es gilt

$$v_1+W=v_2+W \Longleftrightarrow v_1-v_2\in W$$
$$(v_1+W)\cap(v_2+W)\neq\emptyset \Longrightarrow v_1+W=v_2+W\ .$$

Die Menge aller Nebenklassen wird mit $\overline{V}$ oder mit V/W bezeichnet und heißt der Faktorraum von V nach W. V/W wird ein Vektorraum über

$$\begin{aligned}(v_1+W)+(v_2+W) &= (v_1+v_2)+W\\ (v+W)\lambda &= v\lambda+W\ .\end{aligned}$$

Im Sinne von Aufgabe 1 gilt also $V/W = V/\sim$ mit der Äquivalenzrelation $[\, v_1 \sim v_2 \iff v_1 - v_2 \in W\,]$ auf V. Wenn kein Zweifel über den Teilraum W besteht, schreiben wir auch $\overline{V}$ für V/W und $\overline{v}$ für $v + W$: also

$$\overline{V} = \{\overline{v} : v \in V\}\,,\ \overline{v}_1 = \overline{v}_2 \iff v_1 - v_2 \in W\,,\ \overline{v}_1 + \overline{v}_2 = \overline{v_1 + v_2}\,,\ \overline{v}\lambda = \overline{v\lambda}\,.$$

Die Begriffe *Homomorphismus, Mono-, Epi-, Endo-, Auto- und Isomorphismus* haben die offenkundige Bedeutung für eine Abbildung $f : V \to U$ zwischen zwei K-Vektorräumen V und U. Also: f ist Homomorphismus, falls

$$f(v_1 + v_2) = f(v_1) + f(v_2)\,, f(v\lambda) = f(v) \cdot \lambda\,.$$

Man spricht in diesem Fall auch von einer *linearen (oder K-linearen) Abbildung f*. Insbesondere gilt $f(0) = 0$ und $[\, f$ ist Isomorphismus $\iff$ f^{-1} ist Isomorphismus$\,]$. Wie früher bezeichnet $\ker f = \{v \in V : f(v) = 0\}$ den Kern von f und $\mathrm{im}\,(f) = \{f(v) : v \in V\}$ das Bild von f; dies sind beides Unterräume (= Teilvektorräume) in V bzw. U. Im Beispiel (B) ist $\sigma : V \to \mathbb{R}$ ein Epimorphismus mit $\ker \sigma = W$ und $f : W \to \{(r, s) : r, s \in \mathbb{R}\}$ ein Isomorphismus. Des weiteren ist $(r, s) \mapsto \binom{r}{s}$ ein Isomorphismus $\{(r, s) : r, s \in \mathbb{R}\} \to V_2$. Wir nennen $f : V \xrightarrow{v \mapsto \overline{v}} \overline{V}$ den *kanonischen Homomorphismus* von V auf $\overline{V}$.

Satz 2.A. *Ist $f : V \to W$ ein Homomorphismus, so gilt*

1. $f(v_1) = f(v_2) \iff v_1 - v_2 \in \ker f$*; insbesondere ist f genau dann ein Monomorphismus (i.e., injektiv), wenn $\ker f = 0$ ist*
2. $\varphi : V/\ker f \to \mathrm{im}\,(f)$, $\varphi(v + \ker f) = f(v)$, *ist ein Isomorphismus. Insbesondere ist f die Hintereinanderausführung* $V \xrightarrow{v \mapsto \overline{v}} \overline{V} \xrightarrow{\varphi} \mathrm{im}\,(f) \hookrightarrow W$.

So gilt im Beispiel (B), daß $V/W = \mathbb{R}$ unter der Identifizierung $x + W = s_x$.

Gleichungssysteme und Homomorphismen

Zurück zum Gleichungssystem (1) auf Seite 1.

1. Es seien im Standardvektorraum V_m Vektoren $v_i = \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{mi} \end{pmatrix}$, $1 \le i < n$, gegeben. Eine lineare Abhängigkeit zwischen diesen besteht genau dann, wenn das Gleichungssystem (1) mit $\beta_1 = \beta_2 = \ldots = \beta_m = 0$ eine nichttriviale Lösung besitzt.

2. Gegeben sei das Gleichungssystem (1). Wir definieren einen Homomorphismus $f : V_n \to V_m$, indem wir $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in V_n$ den Vektor $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in V_m$ zuordnen, der sich gemäß (1) mit λ_i für x_i eingesetzt berechnet. Dann gilt:
$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} \in V_m$ gehört zu $\mathrm{im}\,(f)$, genau wenn das System (1) mit den γ_i anstelle der β_i eine

Lösung besitzt.

$\ker f$ faßt alle Lösungen des zugehörigen *homogenen* Systems zusammen (d. h. alle $\beta_i = 0$).

Basis und Dimension

DEFINITION 2.3. *Die Vektoren* $\{v_1, \dots, v_m\}$ *bilden eine Basis von* V*, falls* $v_1, \dots, v_m$ *linear unabhängig sind und* $V = \langle v_1, \dots, v_m \rangle$ *ist.*

Ist $\{v_1, \dots, v_m\}$ *eine Basis von* V*, so heißt* m *die Dimension von* V*; in Zeichen:* $\dim V = m$.

$\{v_1, \dots, v_m\}$ bilden also genau dann eine Basis von V, wenn jedes $v \in V$ eindeutig als eine Linearkombination $v = \sum_{i=1}^{m} v_i \lambda_i$ mit Skalaren $\lambda_i \in K$ geschrieben werden kann. In unseren Beispielen zu Beginn des Kapitels ist $\{1\}$ eine Basis für $V = K$; $\{1, i, j, k\}$ eine für $V = \mathbb{H}$, aufgefaßt als Vektorraum über $\mathbb{R}$; $\{e_1, \dots, e_n\}$ eine für V_n, und, für V in (B), w_0, w_1, w_2.

SATZ 2.B. *1. Sind* $v_1, \dots, v_m$ *linear unabhängig und ist* $v = v_1\lambda_1 + \ldots + v_m\lambda_m$ *ein Vektor in* V *mit* $\lambda_i \in K$ *und* $\lambda_1 \neq 0$*, so sind auch* $v, v_2, \dots, v_m$ *linear unabhängig und es gilt:* $\langle v_1, \dots, v_m \rangle = \langle v, v_2, \dots, v_m \rangle$*. ("Austauschsatz")*

2. Falls V *eine Basis* $\{v_1, \dots, v_m\}$ *besitzt, sind je* n *Vektoren mit* $n > m$ *linear abhängig.*

FOLGERUNGEN .

1. *Hat* V *eine Basis* $\{v_1, \dots, v_m\}$*, so hat jede andere Basis auch* m *Elemente, und jeder Satz von* m *linear unabhängigen Vektoren ist eine Basis. Insbesondere ist* $\dim V$ *durch* V *allein bestimmt.*

   Vereinbarung: *Gibt es zu jeder natürlichen Zahl* m *in* V m *linear unabhängige Vektoren, so setzen wir* $\dim V = \infty$.

2. *Ist* $U \leq V$*, so ist* $\dim U \leq \dim V$*; falls überdies* $\dim V < \infty$*, erzwingt* $\dim U = \dim V$*, daß* $U = V$ *ist.*

3. *Ist* $\dim V$ *endlich, so kann jeder Satz von linear unabhängigen Vektoren zu einer Basis von* V *ergänzt werden (Ergänzungssatz).*

4. *Zu* $W \leq V$ *existieren Teilräume* $U \leq V$ *mit* $V = W \oplus U$*. Der kanonische Epimorphismus* $V \twoheadrightarrow V/W$ *bildet jedes solche* U *isomorph auf* V/W *ab.*

5. $U_1, U_2 \leq V \Longrightarrow \dim(U_1 \cap U_2) + \dim(U_1 + U_2) = \dim U_1 + \dim U_2$.

6. *Für eine lineare Abbildung, d.i. ein Homomorphismus*

   $$f : V \to W$$

   zwischen den Vektorräumen V *und* W *über* K*, gilt*

   $$\dim(V/\ker f) = \dim(\operatorname{im} f) \quad \& \quad \dim(\ker f) + \dim(\operatorname{im} f) = \dim V\ .$$

   Insbesondere: $\dim V < \infty$ & $V = W \Longrightarrow [\, f$ *injektiv* $\iff f$ *surjektiv* $\iff f$ *bijektiv*$]$.

In (B) haben wir $\dim V = \dim W + \dim(\operatorname{im}\sigma) = \dim W + 1 = \dim V_2 + 1 = 2 + 1 = 3$.

Satz 2.C. *Je zwei K-Vektorräume derselben (endlichen) Dimension sind isomorph.*

Anders ausgedrückt: $\dim V = m \Longrightarrow V \simeq V_m$ [2]. Der Isomorphismus hängt von der Wahl einer Basis $e_1, \ldots, e_m$ ab und ordnet $v = \sum_{i=1}^{m} e_i\lambda_i$ den Spaltenvektor mit den Koordinaten $\lambda_1, \ldots, \lambda_m$ zu.

Der Begriff *Vektorraum* birgt also an sich überhaupt keine Schwierigkeit.

3. Kapitel: Explizite Angaben von Basen in Unterräumen und explizites Lösen von linearen Gleichungssystemen

Das sei die Situation

> V ist ein K-Vektorraum mit vorgegebener Basis $e_1, \ldots, e_m$
>
> $v_1, \ldots, v_k$ sind k feste Vektoren aus V; $U = \langle v_1, \ldots, v_k\rangle$ sei ihr Aufspann in V
>
> W ist ein K-Vektorraum mit vorgegebener Basis $w_1, \ldots, w_n$
>
> $f : V \to W$ ist eine lineare Abbildung von V nach W.

Die Daten v_l $(1 \leq l \leq k)$ und f sollen uns explizit gegeben sein, nämlich in Abhängigkeit von den genannten Basen:

1. $v_l = \sum_{i=1}^{m} e_i\alpha_{il}, \ \alpha_{il} \in K$,
2. $f(e_i) = \sum_{j=1}^{n} w_j\beta_{ji}, \ \beta_{ji} \in K$.

In beiden Fällen erhalten wir *Matrizen*, nämlich

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2k} \\ \vdots & & & \\ \alpha_{m1} & \ldots & \ldots & \alpha_{mk} \end{pmatrix},$$

deren Spalten sozusagen die Vektoren $v_1, \ldots, v_k$ sind [3], bzw.

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2m} \\ \vdots & & & \\ \beta_{n1} & \ldots & \ldots & \beta_{nm} \end{pmatrix},$$

[2] Das Zeichen $\simeq$ bedeutet die Existenz eines Isomorphismus zwischen den zu seinen beiden Seiten stehenden Objekten. Wir sprechen darüber hinaus von einer *kanonischen* Isomorphie, wenn ein natürlich gegebener Isomorphismus vorliegt, der von keinen willkürlichen Wahlen abhängt.

[3] An diesem Sprachgebrauch halten wir fest. Beachte hier, daß die Matrix A und die Vektoren $v_1, \ldots, v_k$ sich eineindeutig entsprechen: aus A erhält man die v_i durch 1. zurück

deren Spalten die Vektoren $f(e_1), \ldots, f(e_m)$ sind. Im ersten Fall ist die Anzahl der Zeilen die Dimension von V und die Anzahl der Spalten die Zahl k der gegebenen Vektoren v_i; im zweiten Fall ist die Anzahl der Zeilen die Dimension von W, also die des Bildraumes, und die Anzahl der Spalten die Dimension von V, also die des Urbildraumes.

Für $v = \sum_{i=1}^m e_i\gamma_i \in V$ ist

3. $f(v) = \sum_{j=1}^n w_j(\sum_{i=1}^m \beta_{ji}\gamma_i)$; in Kurzschreibweise

$$\begin{pmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1m} \\ \beta_{21} & & & \\ \beta_{n1} & \ldots & \ldots & \beta_{nm} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m \beta_{1i}\gamma_i \\ \vdots \\ \sum_{i=1}^m \beta_{ni}\gamma_i \end{pmatrix} ;$$

dies heißt die "Zeile $\times$ Spalte"-Regel.

Satz 3.A. *Es gilt*

(1) $U = \langle v_{\sigma(1)}, \ldots, v_{\sigma(k)}\rangle$ *für jede Permutation* $\sigma \in S_k$ *der Zahlen* $1, 2, \ldots, k$.

(2) $U = \langle v_1\delta_1, \ldots, v_k\delta_k\rangle$ *für jede Skalarenwahl* $\delta_l \in K$ *mit* $\delta_l \neq 0$ $(1 \leq l \leq k)$.

(3) In $U = \langle v_1, \ldots, v_k\rangle$ *darf* v_{i_0} *(für* $1 \leq i_0 \leq k$*) durch* $v_{i_1}\alpha + v_{i_0}$ *für beliebiges* $i_1 \neq i_0$ *zwischen* 1 *und* k *und beliebiges* $\alpha \in K$ *ausgetauscht werden.*

Bei der Matrix A nennt man die zu (1), (2), (3) analogen Bildungen: Vertauschung von Spalten, Skalierung von Spalten bzw. Addition eines Vielfachen einer Spalte zu einer anderen. Nach einer kurzen technischen Unterbrechung werden wir mithilfe dieser Spaltenoperationen einen Algorithmus für die Bestimmung von Basen in U und im f, sowie für die Entscheidung, ob ein vorgebenes $v \in V$ oder $w \in W$ zu U bzw. zu im f gehört, vorstellen.

Multiplikation von Matrizen, Hintereinanderausführung linearer Abbildungen

Definition 3.1. *Zu einer Matrix* $A = (\alpha_{il})$ *vom Format* $m \times k$ *(d.h.* A *hat* m *Zeilen und* k *Spalten) und zu* $B = (\beta_{ji})$*, einer* $n \times m$ *Matrix (dasselbe* m*), bilde die* $n \times k$ *Matrix* $C = (\gamma_{jl})$ *gemäß*

$$\gamma_{jl} = \sum_{i=1}^m \beta_{ji}\alpha_{il} \ .$$

$C = BA$ *heißt das Produkt von* B *und* A*; es ist nach der Regel "Zeile* $\times$ *Spalte" gebildet.*

Man rechnet nun leicht nach: Ist, wie weiter oben, B die Matrix zu $f : V \to W$ (bezüglich der Basen v_i, w_j) und gehört, analog, A zu einer linearen Abbildung $g : Z \to V$, wobei Z ein dritter K-Vektorraum mit Basis $z_1, \ldots, z_k$ und also $g(z_l) = \sum_{i=1}^m v_i\alpha_{il}$ $(1 \leq l \leq k)$ gesetzt ist, so gehört zu der zusammengesetzten Abbildung

$$f \circ g : Z \to W\,,\ Z \xrightarrow{g} V \xrightarrow{f} W$$

die Matrix C bezüglich der Basen $z_1, \ldots, z_k$ in Z und $w_1, \ldots, w_n$ in W.

Wir schreiben im folgenden $K_{n\times m}$ als Abkürzung für die Gesamtheit aller Matrizen mit n Zeilen, m Spalten und Einträgen aus K; $K_{n\times m}$ wird dann bei komponentenweiser Addition eine abelsche Gruppe und ist als solche isomorph zu $\mathrm{Hom}_K(V,W)$, der Gruppe aller K-linearen Abbildungen von V nach W, wobei V ein m-dimensionaler und W ein n-dimensionaler K-Vektorraum ist [4]. Der Isomorphismus wird durch die Basen $e_1,\ldots,e_m$ von V und $w_1,\ldots,w_n$ von W gemäß 2. und 3. zu Beginn des Kapitels vermittelt:

Satz 3.B. *$K_{n\times m} \simeq \mathrm{Hom}_K(V,W)$ vermöge $B=(\beta_{ji}) \leftrightarrow f : V \to W$ mit*

$$f(e_i) = \sum_{j=1}^{n} w_j\beta_{ji} \quad \textit{und} \quad f(\sum_{i=1}^{m} e_i\gamma_i) = \sum_{j=1}^{n} w_j(\sum_{i=1}^{m} \beta_{ji}\gamma_i) \ .$$

Die weiter oben erklärte Multiplikation von Matrizen liefert eine Abbildung $K_{n\times m}\times K_{m\times k} \to K_{n\times k}$, $(B,A)\mapsto C = BA$, die folgende Regeln erfüllt

$$(B_1+B_2)A = B_1A + B_2A\,,\ B(A_1+A_2) = BA_1 + BA_2\,,\ (BA)D = B(AD)$$

wobei zuletzt $D\in K_{k\times s}$ mit beliebigem s ist. Wir notieren mit 0 die Nullmatrix, d.i. die Matrix mit sämtlichen Einträgen $=0$ (unbeschadet ihrer Rahmengröße); wir schreiben 1 für jede *quadratische* Matrix (also mit gleichviel Zeilen und Spalten), deren Diagonaleinträge [5] alle $=1$ und deren sämtliche anderen Einträge $=0$ sind, und nennen 1 die Einheitsmatrix. Es gilt: $0\cdot A = 0$, $A\cdot 0 = 0$ und $1\cdot A = A$, $A\cdot 1 = A$ (bei zueinander passenden Rahmengrößen von A und 0 bzw. 1).
Man beachte, daß $K_{m\times m}$ nicht nur eine abelsche Gruppe, sondern über die Matrixmultiplikation sogar ein Ring wird; 1 ist das Einselement. $K_{m\times m}$ ist genau dann kommutativ, wenn $m=1$ und K selbst kommutativ ist.

Die Matrizen $\tilde{A}$ und $B^{\approx}$

Definition 3.2. *Die $m\times k$ Matrix A sei gegeben. Falls $A=0$, setzen wir $\mathrm{rg}=\mathrm{rg}(A)=0$. Ist $A\neq 0$, so existiert eine Zahl $\mathrm{rg}=\mathrm{rg}(A)$ zwischen 1 und k, so daß A mittels der im Anschluß an Satz 3.A genannten Bildungen (1), (2), (3) in die neue $m\times k$ Matrix $\tilde{A}$ transformiert werden kann, deren Spalten mit den Nummern $\mathrm{rg}+1,\ldots,k$ alle nur die Null als Eintrag haben, deren rg-Spalte aber einen Eintrag $\alpha'_{i_0,\mathrm{rg}}=1$ hat, der dazu der einzige Eintrag $\neq 0$ in der i_0-ten Zeile ist. Genauso weist die Spalte $\mathrm{rg}-1$ einen Eintrag $\alpha'_{i_1,\mathrm{rg}-1}=1$ auf, der der einzige Eintrag $\neq 0$ in der Zeile i_1 ist, usw.*

Satz 3.C. *Die "Spaltenvektoren" $\tilde{v}_1,\ldots,\tilde{v}_{\mathrm{rg}}$ von $\tilde{A}$ sind linear unabhängig. Gehört, wie zu Beginn des Kapitels, A zu den Vektoren $v_1,\ldots,v_k\in V$ mit Aufspann U (und alles bezüglich der Basis $e_1,\ldots,e_m$ von V), so gilt*

$$\mathrm{rg}(A)=\dim(U) \quad \text{und} \quad \tilde{v}_1,\ldots,\tilde{v}_{\mathrm{rg}} \quad \text{ist Basis von } U\,, \quad \text{sowie} \quad \mathrm{rg}A\le k,m\ .$$

Man erkennt, daß $\mathrm{rg}(A)$ durch A bestimmt ist, die Matrix $\tilde{A}$ allerdings nicht. Man nennt $\mathrm{rg}(A)$ den *Rechtsspaltenrang* von A.

[4] Die Addition in $\mathrm{Hom}_K(V,W)$ ist die übliche, also $(f+g)(v)=f(v)+g(v)$ für $f,g\in\mathrm{Hom}_K(V,W)$ und $v\in V$.
[5] das sind die α_{ii} in einer Matrix (α_{ij})

Satz 3.D. *Ist* $v = \sum_{i=1}^{m} e_i\delta_i \in V$, *so ist* $v \in U$ *genau dann, wenn* $\mathrm{rg}(A) = \mathrm{rg}(A|v)$ *ist, wobei* $(A|v)$ *die Matrix* A *erweitert um die Spalte* $\begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix}$ *bedeutet.*

Beispiel

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \Longrightarrow \mathrm{rg}(A) = 2$$

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in U \iff \mathrm{rg}\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 3 \end{pmatrix} = 2\,, \quad \text{aber}$$

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

$$\text{i.e., rg}\left(A \left| \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \right.\right) = 3\,,\ \text{also} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \notin U\,.$$

Es ist wohl leicht zu sehen, welche der drei Operationen (1), (2) oder (3) bei einem Pfeil $\longrightarrow$ durchgeführt wurden.

Folgerung. $w = \sum_{j=1}^{n} w_j\delta_j \in W$ *gehört zum Bild von* f, *genau wenn* $(B|w)$ *und* B *denselben Rang haben. Es gilt:* $\dim(\mathrm{im}\, f) = \mathrm{rg}(B)$.

Soweit können wir nun explizit Basen (und also Dimensionen) von Unterräumen eines Vektorraumes bestimmen, die durch aufspannende Vektoren gegeben sind. Wir können des weiteren entscheiden, ob ein $v \in V$ zu U gehört; wir können schließlich entscheiden, ob ein $w \in W$ im Bild von f vorkommt.

Im folgenden gehen wir daran, eine Basis von $U_1 \cap U_2$ anzugeben, falls die Unterräume $U_1, U_2 \leq V$ durch sie aufspannende Vektoren beschrieben sind, sowie eine Basis von $\ker f$. Mit letzterem beginnen wir.
Analog zu (1),(2),(3) für die Spalten von A definieren wir Operationen für Zeilen von $B \in K_{n\times m}$:

(Z1) bedeute die Vertauschung zweier Zeilen von B

(Z2) bedeute die Skalierung der j-ten Zeile von B mit dem Faktor $\gamma_i \neq 0$ aus K *von links* $(1 \leq j \leq n)$

(Z3) bedeute die Addition des γ-Linksvielfachen der j-ten Zeile zur i-ten Zeile von B für $i \neq j$.

Die Definition 3.2 kann jetzt mit folgenden Änderungen wiederholt werden: Ersetze A durch B und rg durch zrg, den sogenannten *Linkszeilenrang von* B; ersetze das Wort *Spalte* überall durch *Zeile*, $\tilde{A}$ durch $B^{\approx}$, $\alpha'_{i_t,\mathrm{rg}-t}$ durch $\beta'_{\mathrm{zrg}-t,i_t}$ $(t \geq 0)$.

Der Übersichtlichkeit halber wiederholen wir, wie ein solches $B^{\approx}$ aus B rechnerisch gewonnen werden kann:

1. bringe alle Nullzeilen von B nach unten und erhalte B_1
2. in Zeile 1 von B_1 suche den ersten Eintrag (von links), der $\neq 0$ ist; der stehe in Spalte i'_1
3. ziehe geeignete Linksvielfache der 1. Zeile von den anderen Zeilen ($\neq 0$) in B_1 ab und erhalte so die Matrix B'_1 mit, außer bei $(1, i'_1)$, nur Nullen in Spalte i'_1
4. bringe alle Nullzeilen von B'_1 nach unten und erhalte B''_1
5. in Zeile 2 von B''_1 suche den ersten Eintrag (von links), der $\neq 0$ ist; der stehe in Spalte i'_2
6. ziehe geeignete Vielfache der 2. Zeile von den anderen Zeilen ($\neq 0$) ab, etc.pp.
7. die so schließlich erreichte Endmatrix heiße $B^{\approx}$ und ihre unterste Zeile $\neq 0$ erhalte die Nummer "zrg".

Ein Vergleich mit der vor diesem rechnerischen Verfahren genannten Numerierung zeigt $i'_{\mathrm{zrg}-t} = i_t$ für $t \geq 0$.

Definition 3.3. $R = \{1, \dots, m\} \setminus \{i'_1, \dots, i'_{\mathrm{zrg}}\}$

P_1 bezeichne eine Matrix, die aus $1 \in K_{n\times n}$ durch Vertauschung zweier Zeilen entstehe.

P_2 sei eine $n \times n$-Matrix mit Einträgen $\neq 0$ in der Diagonalen und sonst überall Nullen.

P_3 sei eine $n \times n$-Matrix vom Typ $\begin{pmatrix} 1 & & \gamma \\ & \ddots & \\ & & 1 \end{pmatrix}$; *$\gamma \in K$ steht an einer Stelle $j \neq i$, alle Leerstellen sind Nullen.*

Die Matrizen P_ν, $1 \leq \nu \leq 3$, sind in $K_{n\times n}$ invertierbar, nämlich:

$$P_1^{-1} = P_1\,,\ P_2^{-1} = \begin{pmatrix} \gamma_1^{-1} & & \\ & \ddots & \\ & & \gamma_n^{-1} \end{pmatrix},\ P_3^{-1} = \begin{pmatrix} 1 & & -\gamma \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Multiplikationen von links mit den P_ν an B spiegeln die weiter obengenannten Zeilenoperationen in B wider.

Satz 3.E. *Für $r \in R$ sei $y_r = \sum_{i=1}^m e_i\lambda_{ir} \in V$ der durch folgende Bedingung eindeutig bestimmte Vektor mit*

$$\lambda_{rr} = 1;\ \lambda_{sr} = 0 \text{ für } s \in R,\ s \neq r\,;\ B^{\approx} \begin{pmatrix} \lambda_{1r} \\ \vdots \\ \lambda_{mr} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Die y_r bilden dann eine Basis von $\ker f$ *und insbesondere gilt*

$$\dim(\ker f) = \dim V - \mathrm{zrg}(B)$$

und damit $\mathrm{zrg}(B) = \dim(\mathrm{im}\, f) = \mathrm{rg}(B)$; *i.e., "Linkszeilenrang = Rechtsspaltenrang"* [6].

Im Falle $R = \emptyset$ ist $\ker f = \{0\}$.

Satz 3.E folgt so: $v = \sum_{i=1}^k e_i\lambda_i \in V$ ist in $\ker f$ genau dann, wenn $B \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ (diese Gleichung entspricht übrigens der Matrixmultiplikation $K_{n\times m} \times K_{m\times 1} \to K_{n\times 1}$). Zufolge des Zusammenhangs der erlaubten Zeilenoperationen für B mit Linksmultiplikationen mit Matritzen vom Typ P_ν ist dazu ebenfalls gleichbedeutend

$$B^{\approx} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Beispiel

$$B = \begin{pmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 2 \end{pmatrix} \xrightarrow{(a)} \begin{pmatrix} 1 & 2 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} \xrightarrow{(b)} \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \xrightarrow{(c)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = B^{\approx}$$

mit

$$(a), (b), (c) = \text{Linksmultiplikationen mit } \begin{pmatrix} 1 & & \\ & 1 & \\ -1 & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 1/2 & \\ & & 1 \end{pmatrix} \text{ bzw. } \begin{pmatrix} 1 & -2 & \\ & 1 & \\ & & 1 \end{pmatrix} .$$

Die Menge R ist hier leer.

Die vorgestellte Basisbestimmung von $\ker f$ kann nun auch zu einer *Basisbestimmung von $U_1 \cap U_2$* ausgenützt werden:

Seien $U_1 = \langle v_1, \ldots, v_s\rangle$, $U_2 = \langle v_{s+1}, \ldots, v_q\rangle$ Unterräume in V; $e_1, \ldots, e_m$ sei wieder Basis von V. Sei ferner jetzt $f : V \to V/U_2$ der kanonische Homomorphismus von f auf den Faktorraum V/U_2 : $\quad f(v) = v + U_2$.

Schritt 1: Finde eine Basis $u_1, \ldots, u_d$ in U_2 und ergänze sie mit Hilfe des Austauschsatzes zu einer Basis $u_1, \ldots, u_d, u_{d+1}, \ldots, u_m$ von V. Dann ist $u_{d+1} + U_2, \ldots, u_m + U_2$ Basis von V/U_2.

Schritt 2: Finde Basis $y_1, \ldots, y_p$ von U_1.

Schritt 3: Berechne die Koeffizienten $\beta_{\nu i}$ in $f(y_i) = \sum_{\nu=d+1}^m (u_\nu + U_2)\beta_{\nu i}$, $1 \le i \le p$, und bilde daraus die Matrix $B_1 \in K_{(m-d)\times p}$.

Schritt 4: Mit Hilfe der Matrix B_1 berechne eine Basis in $\ker f_1$, wenn f_1 die Einschränkung von f auf U_1 bezeichnet; *beachte nun, daß* $\ker f_1 = U_1 \cap U_2$.

[6] Man bemerke an dieser Stelle jedoch, daß $\left(\begin{smallmatrix} i & 1 \\ j & k \end{smallmatrix}\right) \in \mathbb{H}_{2\times 2}$ den Rechtsspaltenrang 1 und den Rechtszeilenrang 2 hat.

Beispiel: Basisbestimmung von $U_1 \cap U_2$: Sei $V = V_4$ und $K = \mathbb{R}$; des weiteren

$$U_1 = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle , \; U_2 = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} , \begin{pmatrix} 1 \\ 3 \\ 5 \\ 6 \end{pmatrix} \right\rangle .$$

(1) Berechnung einer Basis von U_2:

$$A_2 := \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 0 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 4 \\ 0 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 0 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 2 \\ -3 & 3 \end{pmatrix} = \tilde{A}_2 \; .$$

Also: Basis von U_2 ist $\begin{pmatrix} 1 \\ 0 \\ -1 \\ -3 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$.

Diese kann offenbar [7] durch die Standardbasisvektoren e_3, e_4 zu einer Basis von V_4 ergänzt werden. Somit: $\{\overline{e_3}\, , \overline{e_4}\}$ ist Basis von V/U_2.

(2) Bestimmung einer Basis von U_1:

$$A_1 := \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \\ 4 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 2 \\ -2 & 3 \end{pmatrix} = \tilde{A}_1 \; .$$

Also: Basis von U_1 ist $\begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$.

f sei die kanonische Abbildung $V \to V/U_2$ und $f_1 : U_1 \to V/U_2$ deren Restriktion auf U_1. Die Matrix B zu f_1 bezüglich der Basen $\begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$ von U_1 und $\overline{e_3}\, , \overline{e_4}$ von V/U_2 ist $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, weil

i) $f(e_1 - e_3 - 2e_4) = f(e_1) - \overline{e_3} - 2\overline{e_4} = \overline{e_3} + 3\overline{e_4} - \overline{e_3} - 2\overline{e_4} = \overline{e_4}$ denn

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ -3 \end{pmatrix} + e_3 + 3e_4$$ — wir benützen die Basis aus (1) von V

[7] rechnerisch etwa über den Austauschsatz

ii) $f(e_2+2e_3+3e_4) = f(e_2)+2\overline{e_3}+3\overline{e_4} = -2\overline{e_3}-3\overline{e_4}+2\overline{e_3}+3\overline{e_4} = 0$ denn

$$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} - 2e_3 - 3e_4 \,.$$

Wir bestimmen schließlich $\ker f_1$.

$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = B^{\approx}$, i.e., $\operatorname{zrg}(B) = 1$, $R = \{2\}$. Daher hat $\ker f_1$ die Basis $y_2 = \begin{pmatrix} \alpha \\ 1 \end{pmatrix}$ mit $B^{\approx}\begin{pmatrix} \alpha \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, also $\alpha = 0$.

Somit: $U_1 \cap U_2 = \ker f_1 = \left\langle 0 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \\ -2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle.$

Unser Verfahren zur Bestimmung von Basen in Unterräumen, insbesondere von $\ker f$, sowie zur Entscheidung, ob ein gegebener Vektor in einem gegebenen Unterraum U liegt, heißt *Gaußscher Algorithmus*. Es liefert zugleich ein explizites Verfahren zur **Lösung von Gleichungssystemen (1)** wie auf Seite 1. Die zugehörigen Vektorräume sind dabei die Standardvektorräume $V = V_m$, $W = V_n$ und die Basen die entsprechenden Standardbasen. Nämlich:

Ist von jetzt an A die Matrix $(\alpha_{ij}) \in K_{m\times n}$ von Seite 1, so ist (1) genau dann lösbar, wenn

$$\operatorname{rg}\left(A \,\middle|\, \begin{matrix} \beta_1 \\ \vdots \\ \beta_m \end{matrix}\right) = \operatorname{rg}(A)$$

ist. Im Falle der Lösbarkeit sind dies die sämtlichen Lösungen von (1):

$$y_0 + y_1\lambda_1 + \ldots + y_t\lambda_t \,,\ \lambda_i \in K \ (\, 1 \le i \le t = n - \operatorname{rg}(A)\,)\,,$$

wobei y_0 eine sogenannte *spezielle Lösung* von (1) und $y_1, \ldots, y_t$ Basis des Raumes

$$\{v \in V_n : Av = 0\}$$

ist [8]. Unser Verfahren liefert nicht nur die y_i, $1 \le i \le t$, sondern auch das $y_0 = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{pmatrix}$. In der Tat,

$$A\begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} \iff A^{\approx}\begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{pmatrix} = P\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix},$$

[8] Av ist die Matrixmultiplikation $K_{m\times n} \times K_{n\times 1} \to K_{m\times 1}$; beachte, daß $\{v \in V_m : Av = 0\}$ der Kern der durch A über die Standardbasen vermittelten linearen Abbildung $V_n \to V_m$ ist.

wenn P das Produkt aus den Matritzen P_ν mit $PA = A^{\approx}$ ist. Somit bleibt zum Beispiel eine Gleichung des Typs

$$\begin{pmatrix} 0 & * & 0 & 1 & * \\ 0 & * & 1 & 0 & * \\ 1 & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}_{4\times 5} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{pmatrix}_{5\times 1} = \begin{pmatrix} \beta_1' \\ \beta_2' \\ \beta_3' \\ \beta_4' \end{pmatrix}_{4\times 1}$$

übrig, in der $*$ einen auf dem Weg von A nach $A^{\approx}$ erhaltenen Eintrag bezeichne, der uns nicht weiter interessieren soll. Die Lösbarkeit dieses Gleichungssystems ist äquivalent zu $\beta_4' = 0$. Um nun eine spezielle Lösung $y_0 = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_5 \end{pmatrix}$ zu finden, setzen wir $\delta_i = 0$ für $i = 2$ und $i = 5$, also für die i, die Indizes derjenigen Spalten von $A^{\approx}$ sind, in denen ein $*$ steht. Es bleibt: $\delta_4 = \beta_1'$, $\delta_3 = \beta_2'$, $\delta_1 = \beta_3'$.

Es ist nun offensichtlich, wie i.allg. ein y_0 berechnet werden kann: Zunächst erhalten wir die Spaltenindizes $i_1', \ldots, i_{\mathrm{zrg}}'$ bei der Ableitung von $A^{\approx}$ aus A [9]. Setze nun in y_0 die Koordinaten δ_i mit $i \neq i_\nu'$ $(1 \leq \nu \leq \mathrm{zrg})$ gleich Null; dann ist $A^{\approx} \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix} = \begin{pmatrix} \beta_1' \\ \vdots \\ \beta_m' \end{pmatrix}$ direkt und eindeutig lösbar.

Wir fassen zusammen

Satz 3.F. *Das Gleichungssystem (1) ist genau dann lösbar, wenn für* $A = (\alpha_{ij}) \in K_{m\times n}$

$$\mathrm{rg}\Bigl(A \Bigm| \begin{matrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{matrix}\Bigr) = \mathrm{rg}(A) \quad \textit{gilt.}$$

Im Falle der Lösbarkeit ist $\{y_0 + y_1\lambda_1 + \ldots y_t\lambda_t : \lambda_i \in K \ (1 \leq i \leq t)\}$ *mit* $t = n - \mathrm{rg}(A)$ *die Lösungsmenge. Dabei ist* y_0 *eine spezielle Lösung von (1) und* $y_1, \ldots, y_t$ *eine Basis des zugehörigen homogenen Systems, d.i. (1) mit allen* $\beta_l = 0$, $1 \leq l \leq m$. *Sowohl* y_0 *als auch die* y_i , $1 \leq i \leq t$, *sind mittels des Gaußschen Algorithmus berechenbar.*

4. Kapitel: Der Dualraum

In diesem Kapitel 'gaukeln' wir uns einen geometrischen Hintergrund vor, ohne jedoch wirklich einen solchen zu kennen. Immerhin werden wir den Begriff *senkrecht* definieren und geben als Anwendung davon eine zweite Berechnungsmöglichkeit von einer Basis in $U_1 \cap U_2$. Für geeignete Skalarkörper, wie etwa $K = \mathbb{R}$, wird all das aus wirklichen geometrischen Überlegungen resultieren – dazu mehr in einem späteren Kapitel.

[9] Achtung! Im Vergleich mit dem auf S. 19 Gesagten arbeiten wir hier mit einer Matrix $A \in K_{m\times n}$ statt mit dem alten $B \in K_{n\times m}$.

V sei im folgenden ein m-dimensionaler Vektorraum über K mit Basis $e_1, \dots, e_m$.

Wir machen $\operatorname{Hom}_K(V, K)$ zu einem K-*Linksvektorraum* über die Definition

$$(\lambda\varphi)(v) = \lambda \cdot \varphi(v)\,,\ \lambda \in K\,,\ \varphi \in \operatorname{Hom}_K(V, K)\,.$$

Bei der Identifizierung

$$\operatorname{Hom}_K(V, K) \simeq K_{1\times m}\,, \quad \varphi \leftrightarrow (\gamma_1, \dots, \gamma_m)\,,\ \gamma_i = \varphi(e_i)\,.$$

(über die Basis e_i von V und dem Basiselement 1 von K) entspricht diese Vektorraumstruktur von $\operatorname{Hom}_K(V, K)$ derjenigen auf $K_{1\times m}$, die entsteht, wenn man die Zeile $(\gamma_1, \dots, \gamma_m)$ von links komponentenweise mit λ multipliziert. Wir nennen $\operatorname{Hom}_K(V, K) = \hat{V}$ den *Dualraum* zu V; die Elemente von $\hat{V}$ heißen *Linearformen* auf V.

Satz 4.A. *Die Abbildungen φ_j mit*

$$\varphi_j(\sum_{i=1}^m e_i\gamma_i) = \gamma_j\ ,$$

also $\varphi_j(e_i) = \delta_{ji} = \begin{cases} 0 & \text{wenn} \quad i \neq j \\ 1 & \text{wenn} \quad i = j \end{cases}$ *, bilden eine Basis von $\hat{V}$, die die Dualbasis zu $e_1, \dots, e_m$ genannt wird. Insbesondere ist* $\dim \hat{V} = m = \dim V$ [10].

Satz 4.B. *$\hat{\hat{V}}$ ist kanonisch isomorph zu V – indem nämlich $v \in V$ mit derjenigen Abbildung $\hat{V} \to K$ identifiziert wird, die $\varphi \in \hat{V}$ auf $\varphi(v)$ schickt. (Beachte: $V, \hat{\hat{V}}$ sind beides Rechtsvektorräume).*

Satz 4.C. *Ist $U \leq V$ ein Unterraum der Dimension s, so ist*

$$U^\perp = \{\varphi \in \hat{V} : \varphi(u) = 0\ (\forall u \in U)\}$$

ein Unterraum von $\hat{V}$ der Dimension $m - s$. Hat umgekehrt $T \leq \hat{V}$ die Dimension t, so hat der Unterraum

$$T^\perp = \{v \in V : \varphi(v) = 0\ (\forall \varphi \in T)\} \leq V$$

die Dimension $m - t$. Es gilt $U^{\perp\perp} = U$, $T^{\perp\perp} = T$.

Sind U_1, U_2 Unterräume von V, so gilt

$$(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp\ ,\ (U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp\ .$$

Definition 4.1. *Eine Hyperebene H von V ist ein Unterraum der Kodimension 1, also* $\dim H = m - 1$.

Wegen $V/H \simeq K$ [11], sind die Hyperebenen also genau die Räume $\ker \varphi$ für $0 \neq \varphi \in \hat{V}$.

---

[10] Wäre K kommutativ, so bräuchten wir nicht zwischen Links- und Rechtsvektorräumen zu unterscheiden und hätten folglich $V \simeq \hat{V}$. Allerdings gibt es i. allg. keinen kanonischen solchen Isomorphismus.

[11] Dieser Isomorphismus entsteht z.B. so: Ergänze eine Basis $h_1, \dots, h_{m-1}$ von H durch h_m zu einer Basis von V und bilde den Vektor $v = h_1\lambda_1 + \dots + h_m\lambda_m \in V$ auf $\lambda_m \in K$ ab; diese Abbildung ist zugleich auch ein φ mit $\ker \varphi = H$.

SATZ 4.D. *Jeder s-dimensionale Unterraum $U \leq V$ ist Durchschnitt von $m-s$ Hyperebenen (und nicht von weniger).*

Damit kann eine Basisbestimmung von $U_1 \cap U_2$ für $U_1, U_2 \leq V$ auch so durchgeführt werden:

Schritt 1: Bestimme $m-p$ Hyperebenen, deren Schnitt U_1, des weiteren $m-q$, deren Schnitt U_2 ist. Nenne die zugehörigen Linearformen $\varphi_1, \ldots, \varphi_{2m-p-q}$ und den von ihnen in $\hat{V}$ aufgespannten Unterraum T.

Schritt 2: Bestimme $T^\perp$.

Die Durchführung der beiden Schritte läuft auf das Auffinden von Basen in den Lösungsräumen geeigneter homogener linearer Gleichungssysteme hinaus. Hier ist das genaue Rezept.

V habe die Basis $e_1, \ldots, e_m$; U_1 sei erzeugt von $v_1, \ldots, v_s$ mit $v_l = \sum_{i=1}^m e_i \alpha_{il}$, U_2 sei erzeugt von $w_1, \ldots, w_r$ mit $w_j = \sum_{i=1}^m e_i \beta_{ij}$. Schreibe die Koordinaten von $v_1, \ldots, v_s$ (von $w_1, \ldots, w_r$) als *Spalten* einer Matrix A (einer Matrix B). Löse die linearen homogenen Gleichungssysteme

$$(x_1, \ldots, x_m)A = (0, \ldots, 0) \quad \text{und} \quad (y_1, \ldots, y_m)B = (0, \ldots, 0)\,;$$

d.h. berechne Basen der zugehörigen Lösungsräume, also von $U_1^\perp$ und $U_2^\perp$. Diese werden, in Koordinaten bezüglich e_i, im ersten Fall die Zeilen einer Matrix C und im zweiten Fall die Zeilen einer Matrix D. Weiter sei E die Matrix, deren obere Zeilen die von C und deren untere die von D sind. Löse nun das homogene Gleichungssystem $E \begin{pmatrix} z_1 \\ \vdots \\ z_t \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ (t passend). Eine Basis dieses Lösungsraumes ist dann zugleich eine Basis von $U_1 \cap U_2$, da $(U_1^\perp + U_2^\perp)^\perp = U_1 \cap U_2$ (nach Satz 4.C).

Wir schließen das Kapitel mit der folgenden Beobachtung. Die Identifizierung

$$\mathrm{Hom}_K(V, W) = K_{n \times m}$$

geschehe über die Basen $e_1, \ldots, e_m$ von V und $w_1, \ldots, w_n$ von W, also

$$\mathrm{Hom}_K(V, W) \ni f \leftrightarrow B = (\beta_{ji})\,,\ f(e_i) = \sum_{j=1}^n w_j \beta_{ji}\ .$$

Jeder linearen Abbildung f zwischen den K-Rechtsvektorräumen V, W wird nun eine lineare Abbildung $\hat{f} \in \mathrm{Hom}_K(\hat{W}, \hat{V})$ zwischen den K-Linksvektorräumen $\hat{W}, \hat{V}$ gemäß

$$(\psi)\hat{f} = \psi \circ f \quad \text{für} \quad \psi \in \hat{W}\,,$$

zugeordnet, i.e., $(\psi)\hat{f}$ ist die zusammengesetzte Abbildung $V \xrightarrow{f} W \xrightarrow{\psi} K$. Beachte, daß die Abbildung $\hat{f}$ von rechts geschrieben wird. Um eine $\hat{f}$ beschreibende Matrix zu bestimmen, müssen wir zunächst Basen in $\hat{W}$ und in $\hat{V}$ wählen: diese seien die jeweiligen Dualbasen ψ_j , φ_i zu w_j bzw. e_i. Wir schreiben dann $(\psi_j)\hat{f} = \sum_{i=1}^n \gamma_{ji}\varphi_i$.

SATZ 4.E. *Mit obiger Wahl gehört zu $\hat{f}$ die Matrix $(\gamma_{ji}) \in K_{n \times m}$ mit $\gamma_{ji} = \beta_{ji}$. — Im Falle* *eines kommutativen Grundkörpers K unterscheiden wir nicht länger zwischen Rechts- und Linksvektorräumen, sehen also V und den Dualraum $\hat{V}$ als Linksvektorräume an. In diesem Fall müssen wir in Übereinstimmung mit unserer Notation für f und der Gleichheit $v\lambda = \lambda v$ für Vektoren v und Skalare λ die Abbildung $\hat{f}$ so ansetzen*

$$\hat{f}(\psi_j) = \sum_{i=1}^{n} \gamma_{ij}\varphi_i$$

und erhalten dann $\gamma_{ij} = \beta_{ji}$. Beachte, daß wir jetzt natürlich auch die Abbildung $\hat{f}$ von links schreiben! In Matrizenform entspricht dem $\hat{f}(\psi)$ die Multiplikation von (β_{ji}) mit dem Spaltenvektor $\psi = \sum_j \alpha_j \psi_j$.

Die Matrix $B^T \in K_{m\times n}$, deren Zeilen die Spalten von $B \in K_{n\times m}$ (in gleicher Numerierung) sind, heißt die zu B *transponierte Matrix* (also $B^T = (\beta_{ij})$ falls $B = (\beta_{ji})$). Es gilt:

$$B_1, B_2 \in K_{n\times m}\, ,\ C \in K_{m\times r} \Longrightarrow (B_1 + B_2)^T = B_1^T + B_2^T\ ,\ (BC)^T \underset{K\text{ kommutativ}}{=} C^T B^T\ .$$

5. Kapitel : Basisabhängigkeiten

Inzwischen sind wir schon verschiedentlich auf lineare Abbildungen $V \xrightarrow{f} W$ zwischen K-Vektorräumen gestoßen und zugleich auf deren rechnerische Beschreibung durch Matrizen B. Wir wollen den Zusammenhang $f \leftrightarrow B$ jetzt näher studieren. In der Tat ist es nämlich fast immer eine Abbildung f, in der schon alle Information über ein lineares Problem versteckt ist; geschickte Basiswahlen werden deshalb die zum schließlichen Rechnen benötigten Matrizen mit speziellen Einträgen füllen (etwa reichlich mit Nullen) und können uns so die Arbeit erleichtern.

Erinnerung: Ist $e_1, \ldots, e_m$ eine Basis von V und $w_1, \ldots, w_n$ eine von W, so sind $\mathrm{Hom}_K(V, W)$ und $K_{n\times m}$ isomorphe abelsche Gruppen unter

$$f \mapsto B = (\beta_{ji})\, ,\ f(e_i) = \sum_{j=1}^{n} w_j\beta_{ji} \quad \text{und} \quad B \mapsto f\, ,\ f(\sum_{i=1}^{m} e_i\gamma_i) = \sum_{j=1}^{n} w_j(\sum_{i=1}^{m}\beta_{ji}\gamma_i)\, .$$

Identifizieren wir $\sum_{i=1}^{m} e_i\gamma_i$ mit dem Spaltenvektor $\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$, so gilt also $f(v) = Bv$.

Basiswechsel Neben der Basis $e_1, \ldots, e_m$ von V sei eine zweite Basis $g_1, \ldots, g_m$ gegeben. Daraus erhalten wir zwei Matrizen $E, G \in K_{m\times m}$, nämlich

$$E = (\varepsilon_{li})\, ,\ e_i = \sum_{l=1}^{m} g_l\varepsilon_{li} \quad \text{sowie} \quad G = (\gamma_{kl}),\ g_l = \sum_{k=1}^{m} e_k\gamma_{kl}$$

Man sieht sofort, daß $EG = GE = 1 = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \in K_{m\times m}$ gilt. Also: $E = G^{-1}$.

Gehört bezüglich der Basen $\{e_i\}$ von V und $\{w_j\}$ von W zu $f : V \to W$ die Matrix B, so erhalten wir zu f bezüglich der neuen Basen $\{g_i\}$ von V und $\{y_j\}$ von W die Matrix HBG mit G wie oben und, analog, H als derjenigen Matrix $(\eta_{jk}) \in K_{n\times n}$ mit $w_k = \sum_{j=1}^n y_j\eta_{jk}$.

Ein wichtiger Spezialfall ist $V = W$ [12] und $e_i = w_i$, $g_i = y_i$ $(1 \le i \le n = m)$. Gehört dann $B = (\beta_{ji})$ zu f bezüglich der Basis e_i, also $f(e_i) = \sum_{j=1}^n e_j\beta_{ji}$, so gehört $G^{-1}BG$ zu f bezüglich der Basis g_i. Hat man nun über f gewisse zusätzliche (geometrische) Informationen, so wird man eine dieser Information angepaßte Basis von V wählen und dann eine i. allg. recht einfach gebaute Matrix $A = G^{-1}BG$ zu f erwarten dürfen. Das wird im nächsten Kapitel ausgeführt werden.

Wir wollen an dieser Stelle noch vereinbaren, daß im Fall $V = W$ stets dieselbe Basis $e_i = w_i$, bzw. $g_i = y_i$, für das Erstellen der Matrix zu f zu wählen ist.

Definition 5.1. *1.* $GL_n(K) = \Big\{ X \in K_{n\times n} : X$ *ist in* $K_{n\times n}$ *invertierbar, d. h.* $\exists Y \in K_{n\times n}$ *mit* $XY = YX = 1 = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \Big\}$. $GL_n(K)$ *heißt die* allgemeine lineare Gruppe *vom Grad n über K; in der Tat ist $GL_n(K)$ bezüglich der Matrixmultiplikation eine Gruppe (nicht kommutativ für $n \ge 2$)* [13].

2. Zwei Matrizen $A, B \in K_{n\times n}$ heißen konjugiert*, wenn es eine Matrix $G \in GL_n(K)$ mit $A = G^{-1}BG$ gibt, in Zeichen $A \sim B$* [14].

$A \sim B$ ist eine Äqivalenzrelation, i.e., $A \sim A$; $A \sim B \Longrightarrow B \sim A$; $A \sim B$ & $B \sim C \Longrightarrow A \sim C$.

Satz 5.A. $\mathrm{Hom}_K(V, V) \simeq K_{n\times n}$ *ist ein Isomorphismus von Ringen. Den Automorphismen von V entspricht dabei $GL_n(K)$.*

Sei $X \in K_{n\times n}$. Dann gilt: $X \in GL_n(K) \Longleftrightarrow rgX = n$.

Ist B eine zum Endomorphismus f von V gehörige Matrix (etwa bezüglich der Basis $\{e_i\}$), so kann durch Übergang zu neuen Basen von V jede zu B konjugierte Matrix als Matrix für f erreicht werden, aber keine anderen.

Sei $X \in GL_n(K)$. Wie berechnet man X^{-1}? Dazu schreibt man X und die 1-Matrix 1 nebeneinander und führt in dieser großen Matrix $(X|1) \in K_{n\times 2n}$ Zeilenmanipulationen (Z1),(Z2),(Z3) so durch, daß aus X schließlich 1 wird. Damit ist dann zugleich aus der Matrix 1 das gesuchte X^{-1} geworden. Die Begründung beruht darauf, daß solche Manipulationen bedeuten, X, bzw. 1, mit Matrizen des Typs P_1, P_2, P_3 von links zu multiplizieren.

[12] vgl. Satz 2.A.*2.*: $f : V \to W$ ist die Zusammensetzung von $V \overset{v\mapsto\bar v}{\longrightarrow} V/\ker f$, $V/\ker f \overset{\varphi}{\simeq} \mathrm{im}\, f$, $\mathrm{im}\, f \hookrightarrow W$

[13] $GL_1(K) = K^\times \overset{\text{def}}{=} K \setminus \{0\}$ ist nur kommutativ, wenn K selbst kommutativ ist.

[14] $A \sim B \iff \exists\, X = (x_{ij}) \in GL_n(K) : XA = BX$. Die Gleichung stellt für kommutatives K ein (sehr großes) lineares Gleichungssystem in den x_{ij} dar, von dem eine Lösung so zu finden ist, daß sie sich zu einer invertierbaren Matrix X zusammensetzen läßt. Ist allerdings K nicht kommutativ, so handelt es sich hier nicht länger um ein lineares Gleichungssystem für die x_{ij} im Sinne von (1), Seite 1, weil dann Skalare als Faktoren links und rechts von den Unbekannten auftreten.

Beispiel Berechne das Inverse der Matix $X = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right)$$

$\Longrightarrow X^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$.

6. Kapitel: Eigenwerte

Ab jetzt ist K stets ein *kommutativer* Körper. Wir schreiben deshalb die Skalare nun auch von *links* an die Vektoren.

Definition 6.1. *f sei ein Endomorphismus des n-dimensionalen Vektorraums V in sich. Bezüglich der Basis $\{e_i\}$ von V gehöre zu f die Matrix $A \in K_{n\times n}$. Ein* Eigenvektor *von f (oder von A) ist ein $0 \neq v \in V$ mit $f(v) = \lambda v$ für ein $\lambda \in K$ (oder $Av = \lambda v$). λ heißt der zugehörige* Eigenwert *von f (oder von A).*

Beispiel $\begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = 2\begin{pmatrix} 2 \\ -1 \end{pmatrix}$

Satz 6.A. *1. Sind $v_1, \ldots, v_r$ Eigenvektoren von f, deren zugehörige Eigenwerte paarweise verschieden sind, so sind sie linear unabhängig.*

2. λ ist Eigenwert von $A \Longleftrightarrow rg(A - \lambda 1) < n$ [15]

Folgerung. *1) Schreibt man die Matrix von f zu einer Basis $v_1, \ldots, v_r,\ v_{r+1}, \ldots, v_n$, deren erste r Vektoren Eigenvektoren zu paarweise verschiedenen Eigenerten sind, so hat diese die Gestalt*

$$G^{-1}AG = \left(\begin{array}{ccc|c} \lambda_1 & & 0 & \\ 0 & \ddots & & * \\ & & \lambda_r & \\ 0 & & 0 & \end{array}\right).$$

2) 0 ist Eigenwert von A genau wenn $\mathrm{rg}(A) < n$ ist.

In Fall, wenn 0 Eigenwert von A ist, kann also A in eine Matrix konjugiert werden, deren 1. Spalte nur Nullen als Einträge hat.

Polynome und Eigenwerte

Bezeichnet $h(x) = \alpha_d x^d + \alpha_{d-1}x^{d-1} + \ldots + \alpha_2 x^2 + \alpha_1 x + \alpha_0$ ein Polynom in der Unbestimmten x mit Koeffizienten $\alpha_0, \alpha_1, \ldots, \alpha_d$ aus K, so kann man darin den Endomorphismus $f : V \to V$ und ebenso die zugehörige Matrix $A \in K_{n\times n}$ für x einsetzen:

[15] $\lambda \cdot 1 = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix}$

f^i bedeutet die Abbildung $f \circ f \circ \ldots \circ f$ (i-mal hintereinander ausgeführt)

A^i ist die i-te Potenz von A

$(\alpha_i f^i)(v) = \alpha_i \cdot f^i(v)$

$$\alpha_i A^i = \begin{pmatrix} \alpha_i & & \\ & \ddots & \\ & & \alpha_i \end{pmatrix} \cdot A^i .$$

Insbesondere ist der Begriff "f ist Nullstelle von h" oder, was dasselbe bedeutet, "A ist Nullstelle von h", wohldefiniert.

Bemerkung: Polynome $h(x)$ können also als eine Vorschrift für Abbildungen von K nach K, oder von $\mathrm{Hom}_K(V,V)$ nach $\mathrm{Hom}_K(V,V)$, oder von $K_{n\times n}$ nach $K_{n\times n}$, angesehen werden: $\lambda \mapsto h(\lambda)$, $f \mapsto h(f)$, bzw. $A \mapsto h(A)$. Der Addition (Multiplikation) von Polynomen [16] entspricht dann die Addition (Multiplikation oder Hintereinanderausführung) bei den Abbildungen.

Das Polynom $h(x)$ heißt das Nullpolynom, wenn alle Koeffizienten $\alpha_0, \ldots, \alpha_d$ gleich Null sind; es heißt vom Grad d, falls $\alpha_d \neq 0$, und normiert, fall $\alpha_d = 1$ ist.

SATZ 6.B. *Es gibt ein vom Nullpolynom verschiedenes Polynom $h(x)$ vom Grad $\leq n^2$ mit Nullstelle f (mit Nullstelle A).*

Denn die Ringe $\mathrm{Hom}_K(V,V) \simeq K_{n\times n}$ sind zugleich K-Vektorräume der Dimension n^2 und also sind $1, A, A^2, \ldots, A^{n^2}$ linear abhängig.

Für alles weitere ist der folgende Hilfssatz von grundlegender Bedeutung:

Zu Polynomen $h(x)$ und $g(x) \neq 0$ mit Koeffizienten aus dem kommutativen Körper K existieren Polynome $s(x)$, $r(x)$ mit Koeffizienten aus K so, daß $h(x) = s(x) \cdot g(x) + r(x)$ gilt und entweder $r(x) = 0$ ist oder $r(x)$ kleineren Grad als $g(x)$ hat.

Eine unmittelbare Folgerung ist: *Hat $h(x)$ die Nullstelle $\alpha \in K$, so ist das lineare Polynome $x - \alpha$ ein Teiler von $h(x)$, i.e., $\exists\, s(x):\ h(x) = s(x) \cdot (x - \alpha)$.*

SATZ 6.C. *Ist $f \neq 0$ ($A \neq 0$) so gibt es genau ein normiertes Polynom $m(x)$ mit den Eigenschaften:*

$m(f) = 0$ ($m(A) = 0$)

unter allen Polynomen $\neq 0$ mit Nullstelle f (oder A) ist $m(x)$ dasjenige kleinsten Grades .

$m(x)$ heißt das *Minimalpolynom* von f (von A). Beachte: Matrizen A und $G^{-1}AG$ für $G \in GL_n(K)$ haben stets das gleiche Minimalpolynom. Es gilt $\mathrm{grad}(m(x)) = 1 \iff \exists\, \lambda : A = \lambda \cdot 1$.

SATZ 6.D. *1. Gilt $h(f) = 0$ ($h(A) = 0$), so ist $h(x)$ ein Vielfaches von $m(x)$.*

[16] die geschehe natürlich distributiv, wobei die Unbestimmte x mit allen $\alpha \in K$ vertauscht

2. *Ist $h(f) = 0$ ($h(A) = 0$), so ist auch $h(\lambda) = 0$ für jeden Eigenwert λ von f (von A).*
3. *Die Nullstellen des Minimalpolynoms $m(x)$ von f (von A) sind genau die Eigenwerte von f (von A).*

Beispiel $n = 2$

$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. Man rechnet nach, daß $m(x) = x^2 - (\alpha + \delta)x + (\alpha\delta - \beta\gamma)$ ein Polynom mit Nullstelle A ist [17]. Außer für $\alpha = \delta,\ \beta = \gamma = 0$ ist $m(x)$ das Minimalpolynom; im Ausnahmefall ist $x - \alpha$ das Minimalpolynom.

Im Falle, daß $m(x)$ zwei verschiedene Nullstellen λ_1, λ_2 in K hat, und damit die Diskriminante

$$d = (\alpha + \delta)^2 - 4(\alpha\delta - \beta\gamma)$$

ein Quadrat in K ist, existiert ein $G \in GL_2(K)$ mit

$$G^{-1}AG = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \ \lambda_1 + \lambda_2 = \alpha + \delta\,,\ \lambda_1\lambda_2 = \alpha\delta - \beta\gamma\,.$$

Im Falle, daß $m(x)$ genau eine Nullstelle λ in K hat und $A \neq \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$ ist [18], existiert ein G wie oben mit

$$G^{-1}AG = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \ 2\lambda = \alpha + \delta\,.$$

Im Falle, daß $m(x)$ ohne Nullstellen in K, also d kein Quadrat in K ist, existiert ein $0 \neq v \in V$ mit "v, Av ist Basis von V" und folglich ein $G \in GL_2(K)$ mit

$$G^{-1}AG = \begin{pmatrix} 0 & \beta\gamma - \alpha\delta \\ 1 & \alpha + \delta \end{pmatrix}.$$

Ein zweites, recht triviales **Beispiel** in beliebiger (endlicher) Dimension ist dies:
Äquivalent sind

$A \in K_{n\times n}$ hat jeden Vektor $v \neq 0$ als Eigenvektor

$$A = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix} = \lambda \cdot 1 \text{ für ein } \lambda \in K$$

$AB = BA$ für alle $B \in K_{n\times n}$; man sagt: A ist *zentral*.

Im übernächsten Kapitel kommen wir auf die Berechnung von Eigenwerten (über Minimalpolynome) zurück. Dazu werden wir *Determinanten* ausnützen, die wir im folgenden Kapitel behandeln. Wir schließen zuvor dieses Kapitel mit der Einführung einer bestimmten Linearform auf dem K-Vektorraum $K_{n\times n}$ (sic!), die uns oft den zweithöchsten Koeffizienten von Minimalpolynomen liefert (vgl. obiges Beispiel $n = 2$).

[17] Wir lernen bald ein Verfahren kennen, mit dem man das Minimalpolynom eines Endomorphismus oder einer Matrix berechnen kann.

[18] das ist nur möglich, wenn $d = 0$ oder wenn 2=0 in K gilt (also etwa für $K = \mathbb{Z}/2$)

Die Spur

Definition 6.2. *Für* $A = (\alpha_{ij}) \in K_{n\times n}$ *sei* $\mathrm{Sp}A \overset{\text{def}}{=} \sum_{i=1}^n \alpha_{ii}$ *die Spur von A.*

Satz 6.E. *Für* $\lambda \in K,\ A, A_1, A_2, B \in K_{n\times n}$ *gilt*

$$\mathrm{Sp}(\lambda A) = \lambda \mathrm{Sp}A\ ; \quad \mathrm{Sp}(A_1 + A_2) = \mathrm{Sp}A_1 + \mathrm{Sp}A_2$$

$$\mathrm{Sp}(AB) = \mathrm{Sp}(BA).\ \textit{Insbesondere:}\ \mathrm{Sp}(G^{-1}AG) = \mathrm{Sp}A\ \textit{für}\ G \in GL_n(K).$$

Im Falle $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \neq \begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix}$ gilt $m(0) = \alpha\delta - \beta\gamma$. Wir zeigen, daß dieser letzte Wert bis aufs Vorzeichen der Flächeninhalt F des von den Vektoren $\begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$ und $\begin{pmatrix} \beta \\ \delta \end{pmatrix}$ im V_2 aufgespannten Parallelogramms ist.

Es gilt: $F =$ Grundseitenlänge $\times$ Höhenlänge.

Im folgenden arbeiten wir besser mit den Quadraten der einzelnen Größen. Als Grundseite werde der Vektor $\begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$ genommen, also

$$(\text{Grundseitenlänge})^2 = \alpha^2 + \gamma^2$$

(nach Pythagoras). Die Höhe tragen wir im Punkt $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ an. Sie steht senkrecht auf $\begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$. Zunächst drehen wir dazu den Vektor $\begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$ nach links um $90°$ (also dabei $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ nach $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ nach $-e_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$):

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} -\gamma \\ \alpha \end{pmatrix} .$$

Die Höhe h sei nun der Vektor $x_1 \begin{pmatrix} -\gamma \\ \alpha \end{pmatrix}$ mit einem noch zu bestimmenden reellen Skalar x_1. Sie erfüllt offenbar

$$\exists\, x_2 \in \mathbb{R} \quad \text{mit} \quad x_1 \begin{pmatrix} -\gamma \\ \alpha \end{pmatrix} + x_2 \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} \beta \\ \delta \end{pmatrix} .$$

Also lösen wir das lineare Gleichungssystem

$$\begin{pmatrix} -\gamma & \alpha \\ \alpha & \gamma \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \beta \\ \delta \end{pmatrix} .$$

Dazu multiplizieren wir von links mit $\begin{pmatrix} \gamma & -\alpha \\ -\alpha & -\gamma \end{pmatrix}$ und erhalten

$$-(\alpha^2 + \gamma^2) x_1 = \gamma\beta - \delta\alpha .$$

Folglich: $h = \frac{\delta\alpha - \gamma\beta}{\alpha^2+\gamma^2} \cdot \begin{pmatrix} -\gamma \\ \alpha \end{pmatrix}$.

Das Quadrat der Länge von h ist somit

$$\frac{\delta\alpha - \gamma\beta)^2}{(\alpha^2+\gamma^2)^2} \cdot (\gamma^2 + \alpha^2) = \frac{(\alpha\delta - \beta\gamma)^2}{(\alpha^2+\gamma^2)} \,.$$

Dies multipliziert mit dem Quadrat der Grundseitenlänge ergibt $(\alpha\delta - \beta\gamma)^2 = F^2$.

7. Kapitel: Determinanten

In diesem Kapitel ist K kommutativ und V ein n-dimensionaler Vektorraum über K. Mit je n Vektoren $v_1, \ldots, v_n$ in V verbinden wir die geometrische Vorstellung des durch sie aufgespannten Parallelotops, dessen Ecken die Punkte

$$\alpha_1 v_1 + \ldots + \alpha_n v_n \quad \text{mit} \quad \alpha_i = 0 \text{ oder } 1\,,\ 1 \le i \le n\,,$$

sind.

Der Rauminhalt dieses Parallelotops wird offenbar genau dann Null sein, wenn die Vektoren $v_1, \ldots, v_n$ linear abhängig sind: das Parallelotop ist dann sozusagen entartet. Das Ziel dieses Kapitels ist es, eine Funktion D von $v_1, \ldots, v_n$ mit Werten in K zu finden, die den Rauminhalt beschreibt. Die Anschauung im 2-dimensionalen lehrt, an D folgende Bedingungen zu stellen:

(1) D ist K-multilinear, d.h. es gilt für jedes i, $1 \le i \le n$,

$$\begin{aligned} &D(v_1, \ldots, v_{i-1},\ v_i + \tilde{v}_i,\ v_{i+1}, \ldots, v_n) \\ &\quad = D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + D(v_1, \ldots, v_{i-1}, \tilde{v}_i, v_{i+1}, \ldots, v_n)\,, \\ &D(v_1, \ldots, v_{i-1}, \lambda v_i, v_{i+1}, \ldots, v_n) = \lambda D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)\,. \end{aligned}$$

Hier sind $v_1, \ldots, v_{i-1}, v_i, \tilde{v}_i, v_{i+1}, \ldots, v_n$ beliebige Vektoren aus V, und λ ist beliebig aus K.

(2) D ist alternierend, d.h. für alle Paare i, $i+1$ mit $1 \le i \le n-1$ gilt

$$D(v_1, \ldots, v_i,\ v_{i+1}, \ldots, v_n) = 0 \text{ , wenn nur } v_i = v_{i+1} \text{ ist.}$$

Bemerkung: Im Zweidimensionalen haben wir am Ende des letzten Kapitels gesehen, daß i.allg. $m(x) = x^2 - (\alpha + \delta)x - (\alpha\delta - \beta\gamma)$ das Minimalpolynom der Matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ und zugleich (bis aufs Vorzeichen) der Flächeninhalt des von den beiden Vektoren $\begin{pmatrix} \alpha - x \\ \gamma \end{pmatrix}$, $\begin{pmatrix} \beta \\ \delta - x \end{pmatrix}$ aufgespannten Parallelogramms ist. Für bestimmte Werte $\lambda \in K$ von x ist letzterer $= 0$, und zwar offenbar genau für die Eignwerte der Matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. Mit anderen Worten: es ist zu erwarten, daß der Rauminhalt des von den n Spalten einer Matrix $A - x \cdot 1 \in K_{n \times n}$ aufgespannten Parallelotops eng mit dem Minimalpolynom von A zusammenhängt.

Im folgenden untersuchen wir Funktionen

$$D: \underbrace{V \times \ldots \times V}_{n-\text{mal}} \to K\,,\ \text{die (1) und (2) genügen}\,.$$

Eigenschaften von D

i) Vertauscht man in D den Eintrag an der Stelle i mit dem Eintrag an der Stelle $j \neq i$, so ändert sich bei D das Vorzeichen:

$$\begin{aligned} &D(v_1, \ldots, v_{i-1},\, v_i,\, v_{i+1}, \ldots, v_{j-1}, v_j,\; v_{j+1}, \ldots, v_n) \\ &\quad = -D(v_1, \ldots, v_{i-1},\, v_j,\, v_{i+1}, \ldots, v_{j-1},\; v_i,\; v_{j+1}, \ldots, v_n)\,. \end{aligned}$$

ii) $D(v_1, \ldots, v_n) = 0$ falls $v_i = v_j$ für ein Paar $i \neq j$.

iii) Ist $e_1, \ldots, e_n$ eine Basis von V, und gehört zu den v_i bezüglich dieser Basis die Koordinatenmatrix $A = (\alpha_{ij}) \in K_{n\times n}$, also

$$v_i = \sum_{j=1}^{n} \alpha_{ji} e_j\,,$$

so gilt

$$D(v_1, \ldots, v_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \alpha_{1\pi(1)} \cdot \alpha_{2\pi(2)} \cdot \ldots \cdot \alpha_{n\pi(n)} \cdot D(e_1, \ldots, e_n)\,.$$

Folgerungen (im Falle, daß D nicht identisch Null ist)

(a) $D(v_1, \ldots, v_n) \neq 0 \iff v_1, \ldots, v_n$ linear unabhängig.

(b) Erfüllt auch $\tilde{D} : \underbrace{V \times \ldots \times V}_{n-\text{mal}} \to K$ die Bedingungen (1) und (2), so gibt es einen Skalar $\lambda \in K$ mit

$$\tilde{D}(v_1, \ldots, v_n) = \lambda D(v_1, \ldots, v_n) \quad (\forall\, v_1, \ldots, v_n)\,.$$

Im wesentlichen ist also D durch (1) und (2) festgelegt – was uns später tatsächlich erlauben wird, D als Rauminhalt zu interpretieren.

(c) Ist $f \in \operatorname{Hom}_K(V, V)$, so gilt

$$D(f(v_1), \ldots, f(v_n)) \;=\; (\det f)\; D(v_1, \ldots, v_n)$$

mit einem Skalar $\det f \in K$. Für zwei Endomorphismen $f, g \in \operatorname{Hom}_K(V, V)$ gilt $\det(fg) = (\det f)(\det g)$.

Definition 7.1. *$e_1, \ldots, e_n$ sei eine fest gewählte Basis von V. Für eine Matrix $A \in K_{n\times n}$ setze*

$$\det A = D(v_1, \ldots, v_n) D(e_1, \ldots, e_n)^{-1}\,,$$

wenn die Vektoren v_i bezüglich $e_1, \ldots, e_n$ die Matrix A als Koordinatenmatrix haben

$$A = (\alpha_{ji}) \iff v_i = \sum_{j=1}^{n} \alpha_{ji} e_j\ .$$

(Wir nehmen hier wieder an, daß D nicht identisch Null ist.)

Satz 7.A. *Ist D nicht identisch Null, so gilt:*

1. $\det A = 0 \Longleftrightarrow rg\ A < n \Longleftrightarrow A$ *ist nicht invertierbar*
2. $\det f = \det A$, *wenn zu f die Matrix A bezüglich der Basis $e_1, \ldots, e_n$ gehört.*
3. $\det \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = 1\,, \quad \det(AB) = \det A \cdot \det B$
4. $\det A^T = \det A$
5. $\det A = -\det \tilde{A}$, *wenn $\tilde{A}$ aus A durch Vertauschung zweier Spalten oder zweier Zeilen entsteht.*
6. $\det(\alpha A) = \alpha^n \det A\,, \ \alpha \in K$.

Beispiel $\det \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$

$$\begin{aligned} \det \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} &= \alpha_{11}\alpha_{22}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} \\ &\quad -\alpha_{13}\alpha_{22}\alpha_{31} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33} \end{aligned}$$

Satz 7.B. *"Laplacescher Entwicklungssatz"*

Für jeden Zeilenindex i zwischen 1 und n und für jeden Spaltenindex j zwischen 1 und n gilt

$$\det A = \begin{cases} \sum_{k=1}^{n} (-1)^{i+k} \alpha_{ik} \det(A_{ik}) \\ \sum_{k=1}^{n} (-1)^{j+k} \alpha_{kj} \det(A_{kj})\ ; \end{cases}$$

dabei ist $A_{\nu\mu}$ diejenige $(n-1) \times (n-1)$-Matrix, die aus A durch Streichung der ν-ten Zeile und der μ-ten Spalte entsteht.

Anwendung Es existiert ein D, das nicht identisch Null ist. Denn Satz 7.B liefert die Möglichkeit, D induktiv (nach der Dimension n) zu definieren; $\det(\alpha) = \alpha$ ist der Induktionsanfang.

Weitere Rechenregeln für Determinanten (vgl. auch Satz 7.A)

- $\det(\alpha_{ij}) = \sum_{\pi \in S_n} sgn(\pi) \alpha_{1\pi(1)} \ldots \alpha_{n\pi(n)}$.
- $\det \tilde{A} = \alpha \cdot \det A$, wenn $\tilde{A}$ aus A dadurch entsteht, daß in A entweder eine Zeile oder eine Spalte mit dem Skalar α durchmultipliziert wird.
- $\det A^+ = \det A + \det A'$, wenn A , A' und A^+ sich nur in einer Zeile, etwa der i-ten, unterscheiden und diese in A^+ die Summe derjenigen von A und A' ist.

  Entsprechendes gilt mit Spalten statt Zeilen; i ist beliebig zwischen 1 und n.

- $\det \tilde{A} = \det A$, wenn $\tilde{A}$ aus A dadurch entsteht, daß in A die i-te Zeile (Spalte) durch die Summe aus i-ter Zeile (Spalte) und dem α-fachen der j-ten Zeile (Spalte) ersetzt wird. Hier sind $i \neq j$ beliebig zwischen 1 und n.

- Hat $A = \begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_r \end{pmatrix}$ Kästchenform mit quadratischen Kästchen $A_\nu \in K_{n_\nu \times n_\nu}$ mit $n_\nu \leq n$ und $\sum_\nu n_\nu = n$, so gilt

$$\det A = (\det A_1) \cdot (\det A_2) \cdot \ldots \cdot (\det A_r)\,.$$

  Entsprechendes gilt, wenn die Nullen oberhalb der A_ν stehen. Insbesondere

$$\det \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} = \alpha_1 \alpha_2 \cdot \ldots \cdot \alpha_n\ .$$

- "Vandermonde-Determinante" : [15]

$$\det \begin{pmatrix} 1 & \alpha_0 & \alpha_0^2 & & \alpha_0^{n-1} \\ 1 & \alpha_1 & \alpha_1^2 & & \alpha_1^{n-1} \\ \cdot & & & & \\ \cdot & & & & \\ 1 & \alpha_{n-1} & \alpha_{n-1}^2 & \cdots & \alpha_{n-1}^{n-1} \end{pmatrix} = \prod_{0 \leq i < j \leq n-1} (\alpha_j - \alpha_i)\,.$$

Anwendung auf lineare Gleichungssysteme

Ist $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ ein Gleichungssystem mit n Gleichungen und n Unbekannten, und ist $\det A \neq 0$, so gilt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det A} \cdot A^* \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \quad \text{mit } A^* \text{ wie im Satz 7.C weiter unten.}$$

Zur tatsächlichen Berechnung der Unbekannten ist vielleicht die *Cramersche Regel* am vorteilhaftesten:

$$x_i = \frac{\det A^{(i)}}{\det A}\,,$$

wobei $A^{(i)}$ aus A dadurch entsteht, daß für die i-te Spalte in A die neue Spalte $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ eingesetzt wird.

[15] Viele Determinantenberechnungen spezieller Matrizen findet man etwa in den Büchern von Muir und Kowalewski.

Satz 7.C. *Ist* $A^* = (\beta_{ij})$ *diejenige Matrix, die durch* $\beta_{ij} = (-1)^{i+j}\det(A_{ji})$ *definiert ist, so gilt*

$$AA^* = A^*A = (\det A)\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix},$$

insbesondere also $A^{-1} = \frac{1}{\det A} \cdot A^*$, *wenn* $\det A \neq 0$.

Soweit betrifft die Theorie der Determinanten nur quadratische Matrizen. Tatsächlich kann man leicht verallgemeinern.

Definition 7.2. *Sei* $A \in K_{m\times n}$. *Unter einem* r–Minor *von* A, *mit* $1 \leq r \leq \min(m,n)$, *versteht man die Determinante einer Matrix* $\in K_{r\times r}$, *die aus* A *durch Streichung von* $m-r$ *Zeilen und* $n-r$ *Spalten entsteht.*

Satz 7.D. *Für* $A \in K_{m\times n}$ *gilt*

$$\mathrm{rg} A = r \Longleftrightarrow [\,\exists\ r\text{−Minor } \neq 0\ ,\ \nexists (r+1)\text{−Minor } \neq 0\,]\,.$$

Wir schließen dieses Kapitel mit der Einführung einer speziellen, über die Determinante definierten Matrixgruppe.

Definition 7.3. $GL_n(K)$ *sei, wie früher, die Menge aller Matrizen aus* $K_{n\times n}$ *mit von Null verschiedener Determinante. Die Determinante induziert einen surjektiven Gruppenhomomorphismus* $GL_n(K) \to K^\times$ *mit Kern*

$$SL_n(K) = \{A \in K_{n\times n} : \det A = 1\}\,;$$

$SL_n(K)$ *heißt die spezielle lineare Gruppe vom Rang* n.

$E_{ij} \in K_{n\times n}$ *bezeichne die Matrix mit den Einträgen 1 an der Stelle* i,j *und Nullen überall sonst. Matrizen der Gestalt* $1+\alpha E_{ij}$, $\alpha \in K$ *heißen elementare Matrizen.*

Satz 7.E. $GL_n(K)/SL_n(K) \simeq K^\times$.

Die Gruppe $SL_n(K)$ *ist von den elementaren Matrizen* $1+\alpha E_{ij}$, $i \neq j$, $\alpha \in K$ *erzeugt, d.h. jedes* $A \in SL_n(K)$ *ist (endliches) Produkt solcher elementarer Matrizen.*

8. Kapitel: Determinanten und Eigenwerte

Es sei $A = (\alpha_{ij}) \in K_{n\times n}$, K kommutativ. Wir rekapitulieren aus dem 6. Kapitel

$$\lambda \text{ ist Eigenwert von } A \Longleftrightarrow \mathrm{rg}(\lambda\cdot 1 - A) < n \Longleftrightarrow \det(\lambda \cdot 1 - A) = 0.$$

Definition 8.1. $\chi_A(x) = \det(x \cdot 1 - A) =$

$$\det \begin{pmatrix} x-\alpha_{11} & -\alpha_{22} & \cdots & -\alpha_{1n} \\ -\alpha_{21} & x-\alpha_{22} & \cdots & -\alpha_{2n} \\ \vdots & & & \\ -\alpha_{n1} & -\alpha_{n2} & \cdots & x-\alpha_{nn} \end{pmatrix}$$

heißt das charakteristische Polynom von A. Dies ist ein Polynom in der Unbestimmten x mit Koeffizienten aus K.

Satz 8.A. *$\chi_A(x) = x^n - \mathrm{Sp}(A)x^{n-1} + \ldots + (-1)^n \det A$ ist ein normiertes Polynom vom Grad n mit zweithöchstem Koeffizienten $-\mathrm{Sp}(A)$ und mit konstantem Koeffizienten $(-1)^n \det A$.*

Die Nullstellen von $\chi_A(x)$ in K sind genau die Eigenwerte von A.

Konjugierte Matrizen A und $G^{-1}AG$ (für $G \in GL_n(K)$) haben dasselbe charakteristische Polynom.

V sei wieder der n-dimensionale Vektorraum über dem kommutativen Körper K und f ein Endomorphismus von V, zu dem die Matrix $A = (\alpha_{ji}) \in K_{n\times n}$ bezüglich einer Basis $e_1, \ldots, e_n$ von V gehöre:

$$f(e_i) = \sum_{j=1}^{n} \alpha_{ji} e_j \ .$$

Des weiteren sei U ein Unterraum in V, der unter f stabil ist, d.h. $f(u) \in U$ für alle $u \in U$. Die Restriktion f_U von f auf U besitze bezüglich einer Basis $u_1, \ldots, u_r$ von U die Matrix $B = (\beta_{lk}) \in K_{r\times r}$

$$f_U(u_k) = \sum_{l=1}^{r} \beta_{lk} u_l \, .$$

Vermittels $v+U \mapsto f(v)+U$ induziert f einen Endomorphismus $\overline{f} : V/U \mapsto V/U$. Ist $\overline{v}_1, \ldots, \overline{v}_s$ eine Basis von $\overline{V} = V/U$, so schreibe

$$\overline{f}(\overline{v}_i) = \sum_{j=1}^{s} \gamma_{ji} \overline{v}_j$$

und erhalte die Matrix $C = (\gamma_{ji}) \in K_{s\times s}$. Wähle sodann $v_i \in V$ mit $\overline{v}_i = v_i + U \in V/U$ $(1 \le i \le s)$. Dann gilt:

$u_1, \ldots, u_r, v_1, \ldots, v_s$ ist eine Basis von V (insbesondere ist also $r+s=n$), bezüglich der die Matrix von f die folgende Gestalt hat:

$$A_1 = \left(\begin{array}{c|c} B & * \\ \hline 0 & C \end{array}\right) ,$$

somit

$\exists\, G \in GL_n(K)$ mit $G^{-1}AG = A_1$. Tatsächlich ist G die Matrix, deren Spalten $u_1, \ldots, u_r, v_1, \ldots, v_s$ sind.

Folgerung. *Genau wenn $\chi_A(x)$ über K vollständig in Linearfaktoren zerfällt, also*

$$\chi_A(x) = \prod_{i=1}^{n}(x-\lambda_i) \text{ mit } \lambda_i \in K \text{ gilt},$$

$$\text{kann } A \text{ in "obere Dreiecksgestalt" } A_1 = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ konjugiert werden, i.e.,}$$

$$\exists G \in GL_n(K) : G^{-1}AG = A_1 \, .$$

Bemerkung : Nach dem sogenannten *Fundamentalsatz der Algebra* hat jedes Polynom mit komplexen Koeffizienten auch eine komplexe Nullstelle. Daraus folgt sofort, daß $\chi_A(x)$, für $A \in \mathbb{C}_{n\times n}$, vollständig in Linearfaktoren zerfällt, $\chi_A(x) = \prod_{i=1}^{n}(x-\lambda_i)$. Die λ_i sind dann also genau die Eigenwerte von A und es gilt

Summe aller Eigenwerte (Vielfachheiten mitgerechnet) = Spur von A

Produkt aller Eigenwerte (Vielfachheiten mitgerechnet) = Determinante von A.

Satz 8.B. *"Satz von Cayley-Hamilton"*

Das charakteristische Polynom $\chi_A(x)$ ist ein Vielfaches vom Minimalpolynom $m_A(x)$ von A. Insbesondere gilt $\chi_A(A) = 0$ und $\operatorname{grad} m_A(x) \le n$. Des weiteren hat $m_A(x)$ dieselben Nullstellen wie $\chi_A(x)$, möglicherweise mit kleineren Vielfachheiten allerdings.

Der Beweis von Satz 8.B für beliebiges kommutatives K wird erst aus den Überlegungen des späteren Kapitels 14 resultieren; hier werde er nur für $K = \mathbb{C}$ durchgeführt. Hauptbausteine des Beweises sind

1. Gilt $m_A(x) = m_1(x)m_2(x)$ mit teilerfremden Polynomen $m_1(x), m_2(x)$, so zerfällt V_n in die direkte Summe der unter A stabilen Teilräume $U = \ker(m_1(A))$, $U_2 = \ker(m_2(A))$ — nämlich weil es wegen der Teilerfremdheit von m_1 und m_2 Polynome $s(x), t(x) \in K[x]$ mit $1 = s(x)m_1(x) + t(x)m_2(x)$ gibt, woraus erst $1 = s(A)m_1(A) + t(A)m_2(A)$ und dann $v = s(A)m_1(A)v + t(A)m_2(A)v$ für alle $v \in V_n$ folgt.
2. $A \sim \begin{pmatrix} \lambda & & * \\ & \ddots & \\ & & \lambda \end{pmatrix} \in K_{n\times n} \Longrightarrow (\lambda A - 1)^n = 0 \;\&\; m_A(x) = (x-\lambda)^e$ mit $e \le n$.

Beispiel $n = 2$, $K = \mathbb{C}$

1. $\chi_a(x) = (x-\lambda_1)(x-\lambda_2)$, $\lambda_1 \ne \lambda_2$. Somit $m_A(x) = \chi_A(x)$ und $\exists\, G \in GL_2(\mathbb{C})$ mit

(1) $$G^{-1}AG = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} .$$

2. $\chi_A(x) = (x-\lambda)^2$. Dann $\exists G \in GL_2(\mathbb{C})$ mit $G^{-1}AG = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}$.

   Es gilt nun :

   (2a) $$m_A(x) = x - \lambda \iff \exists\, G_1 \in GL_2(\mathbb{C}) \text{ mit } G_1^{-1}AG_1 = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

   (2b) $$m_A(x) = (x-\lambda)^2 \iff \exists\, G_1 \in GL_2(\mathbb{C}) \text{ mit } G_1^{-1}AG_1 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

(1) und (2) werden fortan die *die Normalformen* der Matrizen in $\mathbb{C}_{2\times 2}$ genannt.

Beispiel $n = 3,\ K = \mathbb{C}$

3. $\chi_A(x) = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$, $\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1$. Somit $m_A(x) = \chi_A(x)$ und $\exists\, G \in GL_3(\mathbb{C})$ mit

   (3) $$G^{-1}AG = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}.$$

2. $\chi_A(x) = (x-\lambda)^2(x-\mu)$, $\lambda \neq \mu$.

   Sei v_1 ein Eigenvektor zum Eigenwert μ; setze $U = \ker(A - \lambda \cdot 1)^2 \leq V_3$. Dann
   $$U \cap \langle v_1 \rangle = 0\,,\ \dim U = 2\,.$$
   Ist v_2, v_3 Basis von U wie im zweiten Fall des Beispiels $n = 2$, so ist v_1, v_2, v_3 eine Basis von V_3, bezüglich der die zu A gehörige lineare Abbildung die Matrix $\begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & \alpha \\ 0 & 0 & \lambda \end{pmatrix}$ mit $\alpha = 0$ oder 1 besitzt. Also

   (4) $$\exists\, G \in GL_3(\mathbb{C}) \text{ mit } G^{-1}AG = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & \alpha \\ 0 & 0 & \lambda \end{pmatrix},\ \alpha = 0 \text{ oder } 1\,,$$

   und

   (4a) $$\alpha = 0 \iff m_A(x) = (x-\mu)(x-\lambda)\,.$$

   (4b) $$\alpha = 1 \iff m_A(x) = \chi_A(x)$$

3. $\chi_A(x) = (x-\lambda)^3$. Es gibt jetzt ein $G \in GL_3(\mathbb{C})$ mit $G^{-1}AG = \begin{pmatrix} \lambda & \alpha & \beta \\ 0 & \lambda & \gamma \\ 0 & 0 & \lambda \end{pmatrix}$. Obiges G kann nun so abgeändert werden, daß folgendes gilt :

(5a) $$\mathrm{im}\,(A-\lambda 1)=0 \iff m_A(x)=x-\lambda \iff \alpha=\beta=\gamma=0$$

(5b) $$\dim(\mathrm{im}\,(A-\lambda 1))=1 \iff \dim(\ker(A-\lambda 1))=2 \iff m_A(x)=(x-\lambda)^2 \iff (\alpha,\beta,\gamma)=(0,0,1)$$

(5c) $$\dim(\mathrm{im}\,(A-\lambda 1))=2 \iff \dim(\ker(A-\lambda 1))=1 \iff m_A(x)=(x-\lambda)^3 \iff (\alpha,\beta,\gamma)=(1,0,1)\,.$$

(3) bis (5) heißen die Normalformen der $A \in \mathbb{C}_{3\times 3}$.

Normalformen von Matrizen $\in K_{n\times n}$ für beliebigen kommutativen Körper K und beliebige (endliche) Dimension n werden im Kapitel *Jordansche Normalformen* vorgestellt.

Bevor ab jetzt Geometrisches in den Vordergrund gerückt wird, noch ein Wort zur wohl gschicktesten Art, eine einer vorgegebenen linearen Abbildung $V \xrightarrow{f} V$ gut angepaßte Basis zu konstruieren (also der Wahl eines auf eine gegebene lineare Aufgabe *zugeschnittenen* Koordinatensystems). Wähle dazu irgendein $v_1 \in V$ und wende darauf f an: $f(v_1)=v_2$. Sind v_1 und v_2 linear abhängig, so gilt $v_2=\lambda v_1$ und v_1 ist ein Eigenvektor zum Eigenwert λ. Sind aber v_1, v_2 linear unabhängig, so bilde $v_3=f(v_2)$. Sind auch v_1,v_2,v_3 linear unabhängig, so bilde $v_4=f(v_3)$, u.s.w. Wir erreichen schließlich linear unabhängige Vektoren $v_1, v_2=f(v_1),\ldots,v_r=f(v_{r-1})$ mit $1\le r\le \dim(V)$ und $f(v_r)\in U_1 \overset{\text{def}}{=} \langle v_1,\ldots,v_r\rangle$. Offenbar gilt $f(U_1)\le U_1$, und $f_{|U_1}$ hat bezüglich der Basis $v_1,\ldots,v_r$ von U_1 die Matrix

$$A_1=\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & \alpha_1\\ 1 & 0 & 0 & \ldots & 0 & \alpha_2\\ 0 & 1 & 0 & \ldots & 0 & \alpha_3\\ \vdots & \vdots & \vdots & & \vdots & \alpha_{r-1}\\ 0 & 0 & 0 & \ldots & 1 & \alpha_r \end{pmatrix} \quad \text{mit} \quad f(v_r)=\alpha_1 v_1+\ldots+\alpha_r v_r\ .$$

Deren Minimalpolynom (und zugleich charakteristisches Polynom) ist $x^r-\alpha_r x^{r-1}-\ldots-\alpha_1$ (insbesondere also $\det(A_1)=(-1)^{r-1}\alpha_1$).

Nun setze

$$V_1 \overset{\text{def}}{=} V/U_1\,,\quad f_1(v \bmod U_1) \overset{\text{def}}{=} [\,f(v) \bmod U_1\,]$$

und wähle, sofern möglich (i.e., falls $U_1\ne V$) ein $0\ne w_1\in V_1$ und wiederhole obiges Verfahren mit V_1, f_1, w_1 anstelle von V, f, v_1. Derart erhalte $U_2\le V_1$ und die Matrix A_2 zu $(f_1)_{|U_2}$. Etc.

Alles zusammengenommen liefert die Matrix

$$\begin{pmatrix} A_1 & & \star & \\ & A_2 & & \star\\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

für f bezüglich der Basis $v_1, \ldots, v_r, v_{r+1}, \ldots$ von V mit $[v_{r+1} \bmod U_1] = w_1$ u.s.w. (und insbesondere $\det(f) = \prod_{i=1}^{s}(-1)^{r_i-1}\alpha_i$). Die A_i haben alle die gleiche Form wie A_1, allerdings mit von i abhängigen r und Einträgen in der letzten Spalte.

Wir werden später noch sehen, daß $\star$ rechts oben auch noch vereinfacht werden können.

Beispiel: Die Matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$ beschreibe $V_3 \xrightarrow{f} V_3$ bezüglich der Basis e_1, e_2, e_3.

Wähle $v_1 = e_1$; damit wird $v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $v_3 = \begin{pmatrix} 14 \\ 25 \\ 31 \end{pmatrix}$, $U_1 = \langle v_1, v_2, v_3 \rangle = V_3$ und

$$A \sim A_1 = \begin{pmatrix} 0 & 0 & x_1 \\ 1 & 0 & x_2 \\ 0 & 1 & x_3 \end{pmatrix}$$

mit

$$Av_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} \begin{pmatrix} 14 \\ 25 \\ 31 \end{pmatrix} = \begin{pmatrix} 157 \\ 283 \\ 353 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + 14x_3 \\ 2x_2 + 25x_3 \\ 3x_2 + 31x_3 \end{pmatrix} = x_1 v_1 + x_2 v_2 + x_3 v_3 \;;$$

es bleibt damit noch ein lineares Gleichungssystem mit eindeutiger Lösung x_1, x_2, x_3 zu lösen. Übrigens: wegen $\mathrm{Sp}(A) = \mathrm{Sp}(A_1)$, ist $x_3 = 11$, womit nur

$$x_1 + x_2 + 154 = 157 \quad \text{und} \quad 2x_2 + 275 = 283$$

zu lösen ist, i.e., $x_1 = -1$, $x_2 = 4$.

9. Kapitel: Skalarprodukte auf reellen und komplexen Vektorräumen

W sei der n-dimensionale komplexe Vektorraum mit Basis e_k, $V = \{\sum_{k=1}^{n} \lambda_k e_k \,,\, \lambda_k \in \mathbb{R}\} \subset W$; V ist also ein n-dimensionaler reeller Vektorraum, in dem "wie in W gerechnet wird".

Definition 9.1. *(a) Seien $v = \sum_{k=1}^{n} \gamma_k e_k$, $\tilde{v} = \sum_{k=1}^{n} \tilde{\gamma}_k e_k$ zwei Vektoren aus V ($\gamma_k, \tilde{\gamma}_k \in \mathbb{R}$).*

$$\text{Setze} \quad (v, \tilde{v}) = \sum_{k=1}^{n} \gamma_k \tilde{\gamma}_k \in \mathbb{R}\,.$$

(b) Für die komplexe Zahl $\gamma = r_1 + r_2 i$, $r_1, r_2 \in \mathbb{R}$, heißt $\overline{\gamma} = r_1 - r_2 i$ die zu γ konjugiert komplexe Zahl.

Sind $w = \sum_{k=1}^{n} \gamma_k e_k$, $\tilde{w} = \sum_{k=1}^{n} \tilde{\gamma}_k e_k$ Vektoren aus W ($\gamma_k, \tilde{\gamma}_k \in \mathbb{C}$), so setze

$$(w, \tilde{w}) = \sum_{k=1}^{n} \gamma_k \overline{\tilde{\gamma}_k} \in \mathbb{C}\,.$$

$(\;,\;)$ heißt Skalarprodukt auf V beziehungsweise auf W.

(c) *Ein Endomorphismus f von V (von W) heißt orthogonal, falls* $(\, f(v_1), f(v_2)\,) = (v_1, v_2)$ $(\forall\, v_1, v_2 \in V)$ (bzw. $(\, f(w_1), f(w_2)\,) = (w_1, w_2)$ $(\forall\, w_1, w_2 \in W)$) *gilt.*

Eigenschaften :

$$
\begin{array}{rclcrcl}
 & V & & \qquad & & W & \\
(v,v) & = & 0 \Longleftrightarrow v = 0 & & (w,w) & = & 0 \Longleftrightarrow w = 0 \\
(v,\tilde v) & = & (\tilde v, v) & & (w,\tilde w) & = & \overline{(\tilde w, w)} \\
(v_1+v_2,\tilde v) & = & (v_1,\tilde v)+(v_2,\tilde v) & & (w_1+w_2,\tilde w) & = & (w_1,\tilde w)+(w_2,\tilde w) \\
(\lambda v,\tilde v) & = & (v,\lambda\tilde v) = \lambda(v,\tilde v) & & (\lambda w,\tilde w) & = & \lambda(w,\tilde w) = (w,\overline{\lambda}\tilde w)
\end{array}
$$

Definition 9.2. $|v| = l(v) = +\sqrt{(v,v)}$ *heißt die* Länge *von V.*

v heißt senkrecht zu $\tilde v$, in Zeichen $v \perp \tilde v$, falls $(v,\tilde v) = 0$.

Entsprechendes wird für Vektoren in W definiert; beachte dazu, daß stets $(w,w) \in \mathbb{R}$ gilt und daß $(w,\tilde w) = 0$ automatisch $(\tilde w, w) = 0$ nach sich zieht.

Eine Basis $v_1, \ldots, v_n$ in V heißt orthogonal (orthonormal), falls $v_k \perp v_j$ für $k \neq j$ (und dazu $|v_k| = 1$ für alle k) gilt. Entsprechend wieder in W.

Die Matrix $S_{\{v_k\}} = ((v_l, v_j))_{l,j} \in \mathbb{R}_{n\times n}$ *heißt die die Basis $\{v_k\}$* begleitende *Matrix zu $(\ ,\ )$.*

Beobachtungen :

1. $S_{\{v_k\}}$ ist eine symmetrische Matrix
2. $u_1 = \sum_{k=1}^n \beta_k v_k$, $u_2 = \sum_{k=1}^n \gamma_k v_k \Longrightarrow (u_1,u_2) = u_1^T S_{\{v_k\}} u_2 \stackrel{\text{def}}{=} (\beta_1,\ldots,\beta_n) S_{\{v_k\}} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$
3. $S_{\{v_k\}} = 1 \Longleftrightarrow \{v_k\}$ ist eine Orthonormalbasis.
4. $S_{\{e_k\}} = 1$.
5. Ist $\{b_1,\ldots,b_n\}$ eine Basis von V und $A = (\alpha_{jk})$ durch $v_k = \sum_{j=1}^n \alpha_{jk} b_j \in V$ definiert, so gilt $S_{\{v_k\}} = A^T \cdot S_{\{b_j\}} \cdot A$.

   Entsprechend für W : Hier gilt $S_{\{w_k\}} = A^T \cdot S_{\{c_j\}} \cdot \overline{A}$, wobei $\overline{A}$ aus A durch Ersetzung jedes Eintrages durch den konjugiert komplexen Eintrag entsteht ($\{w_k\}, \{c_j\}$ sind hier Basen von W über $\mathbb{C}$ und ersetzen obige $\{v_k\}, \{b_j\}$).
6. $\det S_{\{v_k\}} > 0$; $\det S_{\{w_k\}} \in \mathbb{R}_{>0}$ für jede Basis w_k von W.

Konstruktion von Orthonormalbasen :

Sei $u_1,\ldots,u_n$ eine gegebene Basis von V. Wir konstruieren daraus eine *Orthonormalbasis* von V.

1. Ersetze u_1 durch $u_1' = u_1/|u_1|$, so daß also jetzt $|u_1'| = 1$ gilt.

2. Sei schon erreicht, daß $u_1, \ldots, u_k$ durch neue $u'_1, \ldots, u'_k$ ersetzt wurden, für die wir wissen, daß

$$|u'_l| = 1 \text{ für } l = 1, \ldots, k\,;\ u'_l \perp u'_j \text{ für } l \neq j\,;\ 1 \leq l, j \leq k\,.$$

Dann ersetze u_{k+1} durch $u''_{k+1} = u_{k+1} + \sum_{l=1}^{k} \alpha_l u'_l$ mit $\alpha_l = -(u_{k+1}, u'_l)$. Jetzt ist u''_{k+1} senkrecht zu $u'_1, \ldots, u'_k$. Ersetze schließlich noch u''_{k+1} durch $u'_{k+1} = u''_{k+1}/|u'_{k+1}|$.

Analog für eine Basis $w_1, \ldots, w_n$ von W.
Die Konstruktion zeigt übrigens, daß jeder Unterraum V_0 von V eine Orthonormalbasis $v_1, \ldots, v_s$ besitzt, und daß

$$V_0 \subset W_0 \stackrel{\text{def}}{=} \mathbb{C}v_1 \oplus \mathbb{C}v_2 \oplus \ldots \oplus \mathbb{C}v_s \subset W$$

ist, wir also mit dem Paar W_0, V_0 in der gleichen Situation wie mit dem Paar W, V sind.

Satz 9.A. *Es gilt $|(v,w)| \leq |v| \cdot |w|$ für alle $v, w \in V$ (Cauchy-Schwartzsche Ungleichung).*
Insbesondere gilt $|v + w| \leq |v| + |w|$ (Dreiecksungleichung) und

$$\exists^1 \varphi\,,\ 0 \leq \varphi < \pi\ :\ \cos\varphi = \frac{(v,w)}{|v| \cdot |w|}\ ;$$

φ heißt der Winkel zwischen den Vektoren v und w (die hier beide als $\neq 0$ vorausgesetzt sind).

Satz 9.B. *$f \in \mathrm{Hom}_{\mathbb{R}}(V,V)$ ($f \in \mathrm{Hom}_{\mathbb{C}}(W,W)$) ist orthogonal, genau wenn f ein Automorphismus ist und die Matrix A von f in Bezug auf eine Orthonormalbasis $\overline{A^{-1}} = A^T$ erfüllt.*

Wir kommen nun zurück zum Dualraum von V (vgl. das 4. Kapitel).

Satz 9.C. *Die Abbildung $v \mapsto \varphi_v \in \hat{V}, \varphi_v(\tilde{v}) = (v, \tilde{v})$ ist ein Isomorphismus $V \simeq \hat{V}$.*

Insbesondere bekommen jetzt die zu Unterräumen $U \leq V$ gehörigen Loträume $U^\perp \leq \hat{V}$ eine geometrische Interpretation in V selbst. Nämlich:

Sei $U \leq V$ und $U^\perp = \{v \in V : v \perp u\ (\forall\, u \in U)\}$. Dann entspricht dieses neue $U^\perp \leq V$ dem alten $U^\perp \in \hat{V}$ unter obiger Identifizierung und folglich gilt $\dim U + \dim U^\perp = \dim V$.

Für den komplexen Raum W bekommen wir nicht ganz dasselbe. Die Abbildung

$$w \mapsto \varphi_w : W \to \hat{W},\ \varphi_w(\tilde{w}) = (\tilde{w}, w)$$

ist noch verträglich mit $+$, aber λw geht nicht mehr auf $\lambda\varphi_w$, sondern stattdessen auf $\overline{\lambda}\varphi_w$. Nichtsdestotrotz ist die Abbildung bijektiv.

Im folgenden ist $f : V \to V$ eine $\mathbb{R}$-lineare Abbildung, die bezüglich der Orthonormalbasis $e_1, \ldots, e_n$ die Matrix $A \in \mathbb{R}_{n \times n}$ besitzt. Entsprechend gehöre (auf dem komplexen Niveau) zu $f : W \to W$ die Matrix $A \in \mathbb{C}_{n \times n}$.

Definition 9.3. *Durch*

$$(f(v), \tilde{v}) = (v, f^a(\tilde{v})) \ (\forall\, v, \tilde{v} \in V)\,,$$

bzw.

$$(f(w), \tilde{w}) = (w, f^a(\tilde{w})) \ (\forall\, w, \tilde{w} \in W)\,,$$

ist eine neue $\mathbb{R}$-lineare ($\mathbb{C}$-lineare) Abbildung $f^a : V \to V$ ($f^a : W \to W$) eindeutig definiert. f^a heißt die zu f adjungierte Abbildung.

Satz 9.D. *Die von $f : V \to V$ kanonisch induzierte Abbildung $\hat{f} : \hat{V} \to \hat{V}$ wird bei der Identifizierung $V \simeq \hat{V}$ aus Satz (9.A) die Abbildung f^a.*

Bezüglich $\{e_k\}$ gehört zu f^a die Matrix A^T (bzw. $\overline{A^T}$).

Zu den symmetrischen reellen Matrizen (d.h. $A = A^T$) gehören also genau die selbstadjungierten (d.h. $f = f^a$) Endomorphismen von V.

Satz 9.E. *A sei eine symmetrische Matrix aus $\mathbb{R}_{n\times n}$. Es existiert eine Matrix $O \in GL_n(\mathbb{R})$ mit*

(i) $O^{-1} = O^T$

(ii) $O^{-1}AO = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

Insbesondere zerfällt das charakteristische Polynom $\chi_A(x)$ von A in n Linearfaktoren; kurz: "jeder Eigenwert einer reellen symmetrischen Matrix ist reell".

Bemerkungen:

1. Die Eigenschaft (i) besagt, daß die neue Basis von V, die über O aus der alten Orthonormalbasis $\{e_i\}$ entsteht (also die Spalten von O), wieder eine Orthonormalbasis ist; der von O vermittelte Endomorphismus ist also orthogonal. Matrizen O, die (i) erfüllen, heißen orthogonal; ihre Gesamtheit ist eine Gruppe, die mit $\mathfrak{O}_n(\mathbb{R})$ bezeichnet wird. Beachte: $\det O = \pm 1$ und $\mathfrak{O}_n(\mathbb{R}) \leq GL_n(\mathbb{R})$.

2. Ist $O \in \mathfrak{O}_n(\mathbb{R})$ und λ ein Eigenwert von O mit Egenvktor v, so gilt

$$(v, v) = (Ov, Ov) = (\lambda v, \lambda v) = \lambda^2 (v, v)\,, \quad \text{also } \lambda = \pm 1\ .$$

   Des weiteren: $V = \langle v \rangle \oplus \langle v \rangle^\perp$ und O induziert eine orthogonale Abbildung auf dem $n-1$-dimensionalen Teilraum $\langle v \rangle^\perp$.

Wir werden später auf Satz 9.E im Zusammenhang mit den sogenannten *Hauptachsentransformationen* zurückkommen und explizit reelle *quadratische* Gleichungen in n Unbekannten lösen.

Beschreibung der orthogonalen Matrizen für $n = 2, 3$:

- $O \in \mathfrak{O}_2(\mathbb{R}), \det O = 1 \ : \ O = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$ mit $0 \leq \varphi < 2\pi$.

- $O \in \mathfrak{O}_2(\mathbb{R}), \det O = -1$: O hat die Eigenwerte ± 1 und kann also orthogonal in $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ konjugiert werden.
- $O \in \mathfrak{O}_3(\mathbb{R})$: O hat den Eigenwert ± 1 und damit bis auf orthogonale Konjugation die Form

$$\begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & \cos\varphi & \sin\varphi \\ & -\sin\varphi & \cos\varphi \end{pmatrix}, \begin{pmatrix} -1 & & \\ & \cos\varphi & \sin\varphi \\ & -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Das Analogon zu Satz 9.E auf dem komplexen Niveau ist

Satz 9.F. *$A \in \mathbb{C}_{n\times n}$ erfülle $A = \overline{A^T}$. Dann hat A nur reelle Eigenwerte.*

Solche A heißen hermitesche Matrizen.

Auf all das werden wir im Kapitel über Bilinearformen wieder zurückkommen.

10. Kapitel: Geometrie, I

Definition 10.1. *Der n-dimensionale affine Euklidische Raum A^n ist ein Paar (P, V_n) bestehend aus einer Menge $P \neq \emptyset$ der sogenannten Punkte von A^n und aus dem n-dimensionalen reellen Standardvektorraum V_n, dem das übliche Skalarprodukt aufgeprägt ist. P und V_n sind dabei über eine Abtragungsabbildung $\theta : V_n \times P \to P$ verknüpft, die folgendes erfüllt:*

1. *zu $p, q \in P$ existiert genau ein $v \in V_n$ mit $\theta(v, p) = q$; wir schreiben $v = pq$*
2. *$pq + qr = pr$ für alle $p, q, r \in P$.*

Der Abstand von p nach q ist $|pq|$. Er hängt mit dem Skalarprodukt über $|pq| = |op - oq|$ zusammen, wobei o ein beliebiger, aber fest gewählter Grundpunkt in P ist.

Beobachtung

$pp = 0$; $pq = 0 \Rightarrow p = q$; $pq = -qp$; $pq_1 = pq_2 \Rightarrow q_1 = q_2$; $pq = rs \Rightarrow pr = qs$ (Parallelogramm).

Satz 10.A. *Sei o ein fest gewählter Grundpunkt in P. Wir identifizieren über ihn V_n und $\{ox : x \in P\}$. Ist nun f eine distanztreue Abbildung von P in sich, also $|f(p)f(q)| = |pq|$ für alle $p, q \in P$, so gibt es eine orthogonale lineare Abbildung $f_0 : V_n \to V_n$ mit*

$$f(x) = \theta(f_0(ox), f(o)) \quad \text{oder, gleichbedeutend:} \quad f(o)f(x) = f_0(ox)\,.$$ [16]

[16] Aufgrund dieses Satzes nennen wir fortan orthogonale Endomorphismen auch *längen- und winkeltreue* Abbildungen.

Interpretation Die distanztreuen Abbildungen sind die orthogonalen Selbstabbildungen von V, gefolgt von einer Translation.

Unter einem *affinen Unterraum* von A^n wird jedes Paar der Gestalt $(y + U \overset{\text{def}}{=} \{\theta(u,y) : u \in U\}, U)$ mit einem $y \in P$ und einem Unterrraum $U \leq V_n$ verstanden. Der Abstand "dist" zweier affiner Unterräume $(y_1 + U_1, U_1)$, $(y_2 + U_2, U_2)$ von A^n wird als das Infimum aller Zahlen $|\theta(u_1, y_1)\,\theta(u_2, y_2)|$ $(u_1 \in U_1,\ u_2 \in U_2)$ erklärt. Der ist offenbar derselbe wie der Abstand $d(v,U) \overset{\text{def}}{=} \inf\{|v-u| : u \in U\}$ des Vektors $v = y_1\,y_2$ vom Unterraum $U = U_1 + U_2$.

Satz 10.B. *Sei $U \leq V_n$ und $a_1, \ldots, a_r$ eine Basis von U.*

Es gilt

$$d(v,U)^2 = |v|^2 - a^T A^{-1} a \ ,$$

wenn A die Matrix $((a_i, a_j)) \in \mathbb{R}_{r\times r}$ bezeichnet und wenn a für den Spaltenvektor mit den Einträgen (v, a_i), $1 \leq i \leq r$, steht.

Ist $u \in U$ so, daß $|v-u| = d(v,U)$, dann steht $v-u$ senkrecht auf U. Umgekehrt, ist $u \in U$ so, daß $v-u$ senkrecht auf U steht, so ist u eindeutig und erfüllt $|v-u| = d(v,U)$.

Sonderfall

U sei Hyperebene in V_n. Wir können dann U als $\langle c\rangle^\perp$ schreiben, indem wir ein $c \in U^\perp$ der Länge 1 wählen. Der Abstand eines Punktes $a \in A^n$ von einem affinen Unterraum $(y + U, U)$ wird dann durch die einfache Hessesche Formel gegeben

$$\operatorname{dist}(a, y+U) = |(c, ay)| \ .$$

Definition 10.2. *Es seien $a_1, \ldots, a_r$ linear unabhängige Vektoren in V_n. Das Volumen des* Würfels

$$W \overset{\text{def}}{=} W(a_1, \ldots, a_r) = \{\lambda_1 a_1 + \ldots + \lambda_r a_r : 0 \leq \lambda_i \leq 1\}$$

mit den Ecken $a_1, \ldots, a_r$ wird als $|W| = \sqrt{\det A_r}$ erklärt; hier ist $A_r = ((a_i, a_j)) \in \mathbb{R}_{r\times r}$.

Beachte, daß $\det A_r > 0$.

Satz 10.C. *Ist $a_{r+1} \in V \setminus \langle a_1, \ldots, a_r\rangle$, so ist*

$$|W(a_1, \ldots, a_{r+1})| = d(a_{r+1}, \langle a_1, \ldots, a_r\rangle) \cdot |W(a_1, \ldots, a_r)| \ .$$

Ist $\langle a_1, \ldots, a_r\rangle = \langle b_1, \ldots, b_r\rangle$ und $a_i = \sum_{j=1}^r \alpha_{ji} b_j$, so ist

$$|W(a_1, \ldots, a_r)| = |\det(\alpha_{ji})| \cdot |W(b_1, \ldots, b_r)| \ .$$

Mit diesem Satz haben wir endlich den Zusammenhang von Determinanten und Volumenberechnung erhalten.

Das Vektorprodukt

Im n-dimensionalen reellen Vektorraum V mit Orthonormalbasis $e_1, \ldots, e_n$ kann man auf die folgende Weise zu $n-1$ vorgegebenen Vektoren $v_i = \sum\limits_{j=1}^{n} \alpha_{ji} e_j \in V$, $1 \leq i \leq n-1$ einen eindeutig bestimmten Vektor $p = \sum\limits_{j=1}^{n} \pi_j e_j \in V$ konstruieren, der folgendes erfüllt

$$p \perp v_i\,, \quad |p| = \text{Volumen des Parallelotops mit den Eckpunkten } v_i\,, \quad \det(v_1, \ldots, v_{n-1}, p) \geq 0\,.$$

Dazu setze einfach

$$\pi_j = \det(v_1', \ldots, v_{n-1}') \text{ mit } v_i' \text{ als Spalte } v_i = \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{ni} \end{pmatrix} \text{ vermindert um den Eintrag bei } (j,i)\,.$$

Für $n = 3$ schreiben wir

$$p = v_1 \times v_2 = \det \begin{pmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} e_1 - \det \begin{pmatrix} \alpha_{11} & \alpha_{13} \\ \alpha_{31} & \alpha_{33} \end{pmatrix} e_2 + \det \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} e_3$$

und nennen p das Vektorprodukt aus v_1 und v_2. Es gilt $v_1 \times v_2 = -v_2 \times v_1$ sowie $(v_1 \times v_2, v_3) = \det(v_1, v_2, v_3)$; insbesondere ist das Vektorprodukt $\mathbb{R}$-linear in beiden Argumenten. Darüber hinaus gelten die Jakobi- und die Graßmann-Identitäten

$$v_1 \times (v_2 \times v_3) + v_2 \times (v_3 \times v_1) + v_3 \times (v_1 \times v_2) = 0$$

$$v_1 \times (v_2 \times v_3) = (v_1, v_3) v_2 - (v_1, v_2) v_3$$

sowie
$$(v_1 \times v_2, v_3 \times v_4) = (v_1, v_3)(v_2, v_4) - (v_1, v_2)(v_3, v_4)\,.$$

Die Jakobi-Identität ist ein Ersatz für das fehlende Assoziativgesetz mit Bezug auf $\times$; sie führt übrigens zu dem wichtigen mathematischen Begriff einer *Liealgebra*.

11. Kapitel: Geometie, II; Hauptachsentransformationen

Eine Quadrik im A^n ist die Menge aller Punkte $p \in P$ mit $op = \sum_{i=1}^{n} x_i e_i$ und

(1)
$$\sum_{1 \leq i \leq j \leq n} \tilde{\alpha}_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i x_i + \rho = 0\,,$$

wobei $\tilde{\alpha}_{ij}$, β_i und ρ reelle Skalare sind.
Wir können (1) zugleich als die allgemeine quadratische Gleichung in den kommutierenden Variablen $x_1, \ldots, x_n$ und den reellen Koeffizienten $\tilde{\alpha}_{ij}$, β_i, ρ ansehen.

Drei Fragen stellen sich. Wann ist (1) lösbar? Wie sehen die Lösungen aus? Welches geometrische Bild gehört zu (1) im A^n?

Für $1 \leq i, j \leq n$ setze $\alpha_{ij} = \begin{cases} \tilde{\alpha}_{ii} & \text{falls } i = j \\ \frac{1}{2}\tilde{\alpha}_{ij} & \text{falls } i < j \\ \frac{1}{2}\tilde{\alpha}_{ji} & \text{falls } i > j \end{cases}$ und nenne A die Matrix (α_{ij}); A ist also symmetrisch. Des weiteren seien x und b die Spaltenvektoren mit den Koordinaten x_i bzw. β_i. Mit diesen Bezeichnungen wird aus (1)

$$x^T Ax + b^T x + \rho = 0\,. \tag{2}$$

G sei eine orthogonale Matrix mit

$$G^{-1}AG = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad G^{-1} = G^T\,.$$

Sie überführt die Orthonormalbasis e_i in die neue Orthonormalbasis e'_i.

Definiere y_i als die Koordinate des Vektors x in bezug auf e'_i, also $Gy = x$ mit $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Es folgt

$$x^T Ax = (Gy)^T AGy = y^T G^{-1}AGy = y^T Dy = \sum_{i=1}^{n} \lambda_i y_i^2$$

und, mit γ_i als der Koordinate von $c = G^T b$ bei e'_i ,

$$b^T x = b^T Gy = (G^T b)^T y = c^T y = \sum_{i=1}^{n} \gamma_i y_i\,.$$

Numeriere gegebenenfalls so um, daß $\lambda_1 \cdot \ldots \cdot \lambda_r \neq 0$ und $\lambda_{r+1} = \ldots \lambda_n = 0$. Für $i \leq r$ schreibe

$$\lambda_i y_i^2 + \gamma_i y_i = \lambda_i (y_i + \frac{\gamma_i}{2\lambda_i})^2 - \frac{\gamma_i^2}{4\lambda_i}\,.$$

Der Übergang von y_i zu den neuen Variablen $z_i = y_i + \frac{\gamma_i}{2\lambda_i}$ $(1 \leq i \leq r)$, $z_i = y_i$ $(r+1 \leq i \leq n)$ ist eine Translation. Indem wir alle vorkommenden $-\frac{\gamma_i^2}{4\lambda_i}$ und das ρ zu ρ_1 zusammenfassen, erhalten wir

$$\sum_{i=1}^{r} \lambda_i z_i^2 + \sum_{i=r+1}^{n} \gamma_i z_i + \rho_1 = 0\,. \tag{3}$$

Deren Lösungsmannigfaltigkeit ist bis auf eine distanztreue Abbildung des Punktraumes P dieselbe wie die von (1).

Sofern nicht alle $\gamma_{r+1}, \ldots, \gamma_n$ Null sind, führen wir eine weitere orthogonale Transformation $O \in \mathfrak{O}_{n-r}(\mathbb{R})$ aus, diesmal im Unterraum $U = \langle e'_{r+1}, \ldots, e'_n \rangle \leq V_n$. Dann nämlich ergänzen wir

$$e''_n \stackrel{\text{def}}{=} \sqrt{(\sum_{i=r+1}^{n} \gamma_i^2)^{-1}} \sum_{i=r+1}^{n} \gamma_i e'_i$$

durch $e''_{r+1}, \ldots, e''_{n-1}$ zu einer Orthonormalbasis von U und schreiben

$$g \stackrel{\text{def}}{=} \textstyle\sum_{i=r+1}^{n} \gamma_i e'_i = \delta_n e''_n \text{ mit } \delta_n = \sqrt{(\sum_{i=r+1}^{n} \gamma_i^2)}$$

$$z \stackrel{\text{def}}{=} \textstyle\sum_{i=r+1}^{n} z_i e'_i = \sum_{i=r+1}^{n} z''_i e''_i\ ,$$

so daß also $\sum_{i=r+1}^{n} \gamma_i z_i = (g, z) = \delta_n z_n''$ wird. Um die Notation nicht ausufern zu lassen, schreiben wir für $\delta_n z_n'' + \rho_1$ einfach $\delta_n z_n$, haben dabei also erneut eine Translation ausgeführt.

Wir fassen zusammen: (1) kann distanztreu in

$$\sum_{i=1}^{n} \lambda_i z_i^2 + \rho_1 = 0 \quad \text{oder} \quad \sum_{i=1}^{r} \lambda_i z_i^2 + \rho_1 = 0 \quad \text{oder} \quad \sum_{i=1}^{r} \lambda_i z_i^2 + \delta_n z_n = 0 \tag{4}$$

übergeführt werden. Dabei sind die λ_i die von Null verschiedenen Eigenwerte von A und δ_n und ρ_1 explizit ableitbar aus b und ρ. Natürlich erzwingen der zweite und dritte Fall $r < n$.

Ist $\rho_1 \neq 0$, so dividieren wir: z.B. $\sum_{i=1}^{r} \frac{\lambda_i}{-\rho_1} z_i^2 - 1 = 0$. Indem wir die Variablen z_i, z_n skalieren, was zwar eine leichte Verzerrung des geometrischen Bildes der Lösungsmannigfaltigkeit bewirkt, erhalten wir folgende Normalformen

$$\begin{gathered} \sum_{i=1}^{s} z_i^2 - \sum_{i=s+1}^{n} z_i^2 = \begin{cases} 0 \\ 1 \end{cases} ; \quad \sum_{i=1}^{s} z_i^2 - \sum_{i=s+1}^{r} z_i^2 = \begin{cases} 0 \\ 1 \end{cases} ; \\ \sum_{i=1}^{s} z_i^2 - \sum_{i=s+1}^{r} z_i^2 + z_n = 0\,,\ r < n \end{gathered} \tag{5}$$

Dabei ist s die Anzahl der positiven Eigenwerte von A. Die Eigenvektoren zu den λ_i heißen die Hauptachsen der Quadrik.

Beispiel: $n = 2$ — dann wird (5) zu

$s = r$	$z_1^2 + z_2^2 = 0$ oder $= 1$ (Punkt bzw. Ellipse)
$s = 1$	$z_1^2 - z_2^2 = 0$ oder $= 1$ (zwei sich schneidende Geraden bzw. Hyperbel)
$r = 1$	$z_1^2 = 0$ oder $= 1$ (Doppelgerade bzw. Geradenpaar)
$s = r = 1$	$z_1^2 \pm z_2 = 0$ (Parabel)

und, mutatis mutandis, im 3-dimensionalen.

Und wie sehen nun tatsächlich die einzelnen durchzuführenden Rechnungen für $n = 2$ aus?

Ausgehend von

$$\alpha_1 x_1^2 + \alpha_2 x_1 x_2 + \alpha_3 x_2^2 + \beta_1 x_1 + \beta_2 x_2 + \rho = 0$$

bilden wir $A = \begin{pmatrix} \alpha_1 & \alpha_2/2 \\ \alpha_2/2 & \alpha_3 \end{pmatrix}$ und $b = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$. Dann folgt

$$\chi_A(x) = x^2 - (\alpha_1 + \alpha_3)x + (\alpha_1\alpha_3 - \frac{\alpha_2^2}{4}) = (x - \lambda_1)(x - \lambda_2)\ ,$$

mit den (reellen) Eigenwerten λ_i von A $(i = 1, 2)$. Ist A nicht schon selbst diagonal, also $\alpha_2 \neq 0$, so muß die Übergangsmatrix $G = (\gamma_{ji}) \in \mathfrak{O}_2(\mathbb{R})$ bestimmt werden: ihre Spalten sind die auf die Länge 1 normierten Eigenvektoren zu λ_i. Löse also

$$\begin{pmatrix} \alpha_1 - \lambda_i & \alpha_2/2 \\ \alpha_2/2 & \alpha_3 - \lambda_i \end{pmatrix} \begin{pmatrix} x_{1i} \\ x_{2i} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

zunächst durch $x_{2i} = 1$ und $x_{1i} = \frac{\lambda_i - \alpha_3}{\alpha_2/2}$ und normiere

$$\gamma_{ji} = \sqrt{(1 + (\frac{\lambda_i - \alpha_3}{\alpha_2/2})^2)^{-1}} \cdot x_{ji}\ , \quad 1 \leq j, i \leq 2\ .$$

Definiere schließlich

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} .$$

Alle weiteren Rechnungen sind dann bereits weiter oben beschrieben. Man beachte, daß die zweite vorkommende orthogonale Transformation O nicht explizit bestimmt werden muß.