A.-S. Elsenhans

Lineare Algebra I und II

WS 2023, SS 2024

Institut für Mathematik der Universität Würzburg

Inhaltsverzeichnis

Vorbemerkungen

Formales

1. Anmeldung zu den Übungen über `WueStudy`
   (Übungen zur linearen Algebra 1)

2. E-learning: WueCampus2

3. Assistent: Benedikt Wolf

4. 40% der Hausaufgabenpunkte sind als Klausurvorleistung erforderlich

5. Fristgerechte Online-Anmeldung zur Klausur

6. Klausurtermin: 23. Februar 2024 von 8 - 10 Uhr, 90 Minuten Bearbeitungszeit

Inhaltliches

Was sind die Ziele der Vorlesung?

Erste Antwort: Das erlernen der Linearen Algebra da diese Grundlage für fast alle weiteren mathematischen Gebiete ist.

Zweite Antwort: Das erlernen grundlegender mathematischer Konzepte und Fähigkeiten. An erster Stelle steht hier das Zusammenspiel aus Definitionen, Sätzen und Beweisen. Insbesondere sollen Sie in dieser Vorlesung das Beweisen mathematischer Aussagen lernen.

Eine Definition legt fest, wie ein Begriff genau zu verstehen ist. Kennt man eine Definition nicht, so ist der Begriff bedeutungslos.

Beispiel aus dem Leben: Unterhalten Sie sich einmal mit einem Amerikaner über die Frage was ein Bier ist.

Beispiel aus der Musik: Was genau ist eine Quinte (im heutigen und im pythagoreischen Sinn)?

Literatur

Die erste Hälfte der Vorlesung basiert auf dem Lehrbuch: Gerd Fischer, Lineare Algebra.

Kapitel 1

Lineare Algebra von Ebene und Raum

1.1 Ebene und Raum als Vektorraum

1.1.1 Definition

Es bezeichnet $\mathbb{R}$ die Menge der reellen Zahlen, $\mathbb{R}^2$ bezeichnet die Paare und $\mathbb{R}^3$ die Tripel reeller Zahlen.

1.1.2 Beispiele

$(1,2), (3/7, \pi) \in \mathbb{R}^2$ und $(1,2,-19), (3/7, \pi, \sqrt{13}) \in \mathbb{R}^3$. Man nennt diese auch Vektoren.

1.1.3 Bemerkung

Die Mengen $\mathbb{R}^2$ und $\mathbb{R}^3$ können als (euklidische) Ebene und als (euklidischer) Raum veranschaulicht werden.

1.1.4 Definition

Es gibt die Verknüpfungen

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) \coloneqq (x_1 + y_1, \ldots, x_n + y_n)$$

sowie

$$\lambda(x_1, \ldots, x_n) \coloneqq (\lambda x_1, \ldots, \lambda x_n)$$

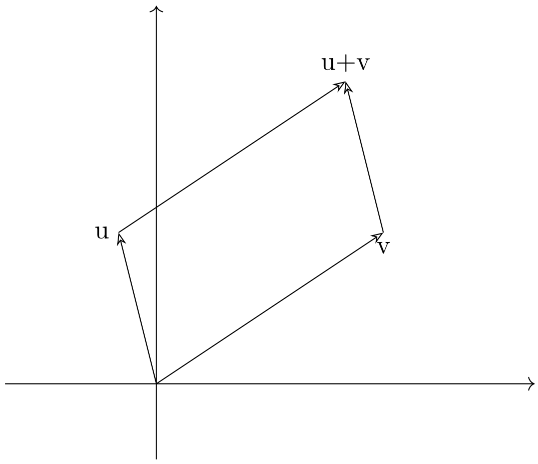
und

$$-(x_1, \ldots, x_n) \coloneqq (-x_1, \ldots, -x_n).$$

Hierbei sind $\lambda, x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$.

Der Punkt $\mathcal{O} = (0, \ldots, 0)$ heißt *Nullvektor* und wird als Ursprung des Koordinatensystems veranschaulicht. Addition und Multiplikation kann graphisch veranschaulicht werden.

1.1.5 Graphische Addition von Vektoren im $\mathbb{R}^2$



1.2 Geraden und Ebenen

1.2.1 Definition

Sind $a_1, a_2, b \in \mathbb{R}$ mit $(a_1, a_2) \neq (0, 0)$, so bezeichnet man die Menge

$$g := \{(x_1, x_2) \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = b\}$$

als *Gerade*.

1.2.2 Beispiel

Es ist $\{(x, y) \in \mathbb{R}^2 \mid x + y = 1\}$ eine Gerade.

1.2.3 Satz

Zu jeder Geraden gibt es $c_1, c_2, d_1, d_2 \in \mathbb{R}$, sodass die Gerade in der Form

$$\{(c_1, c_2) + t(d_1, d_2) : t \in \mathbb{R}\}$$

geschrieben werden kann. Weiterhin ist obige Menge im Fall $(d_1, d_2) \neq (0, 0)$ immer eine Gerade.

Beweis:

Sei zunächst

$$g = \{(x_1, x_2) \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = b\}$$

eine Gerade. Es sind $(b/a_1, 0)$ und $(0, b/a_2)$ Punkte auf der Geraden, sofern die Nenner nicht Null sind. In jedem Fall hat jede Gerade mindestens einen Punkt p.

Ist p ein Punkt von g, so sind alle Punkte der Form $p + t(-a_2, a_1)$ in g enthalten, da

$$a_1(p_1 + t(-a_2)) + a_2(p_2 + ta_1) = a_1p_1 - ta_1a_2 + a_2p_2 + ta_2a_1 = a_1p_1 + a_2p_2 = b$$

gilt. Dies zeigt, dass unsere Menge eine Teilmenge der Geraden ist.

Ist nun (x_1, x_2) ein beliebiger Punkt der Geraden. Im Fall $a_2 \neq 0$ setzen wir $t_1 \coloneqq (x_1 - p_1)/(-a_2)$ und erhalten

$$\begin{aligned} p + t_1(-a_2, a_1) &= (p_1 - a_2(x_1 - p_1)/(-a_2), p_2 + a_1(x_1 - p_1)/(-a_2)) \\ &= (p_1 - (x_1 + p_1), p_2 - a_1x_1/a_2 + a_1p_1/a_2) \\ &= (x_1, p_2 - (b - a_2x_2)/a_2 + a_1p_1/a_2) \\ &= (x_1, p_2 + x_2 - b/a_2 + a_1p_1/a_2) \\ &= (x_1, x_2 + (p_2a_2 - b + a_1p_1)/a_2) = (x_1, x_2)\,. \end{aligned}$$

Somit ist auch (x_1, x_2) ein Punkt der Geraden. Im Fall $a_2 = 0$ kann man die Rechnung mit $t_2 \coloneqq (x_2 - p_2)/(a_1)$ statt t_1 wiederholen und erhält das gleiche Ergebnis.

Seien nun die Zahlen $c_1, c_2, d_1, d_2 \in \mathbb{R}$ beliebig mit $(d_1, d_2) \neq (0, 0)$. Dann ist

$$M \coloneqq \{(c_1, c_2) + t(d_1, d_2) : t \in \mathbb{R}\}$$

die Lösungsmenge der Gleichung $d_2x_1 + (-d_1)x_2 = d_2c_1 - d_1c_2$. (Dies zu bestätigen ist dem Leser als Übungsaufgabe überlassen.) Somit ist M eine Gerade und der letzte Teil der Behauptung ist gezeigt. $\square$

1.2.4 Beispiel

Es gilt $\{(x, y) \in \mathbb{R}^2 \mid x + y = 1\} = \{(1, 0) + t(1, -1) \colon t \in \mathbb{R}\}$.

1.2.5 Beispiel

Bestimme alle Schnittpunkte der Geraden $\{(x, y) \in \mathbb{R}^2 \mid x+y = 1\}$ und $\{(x, y) \in \mathbb{R}^2 \mid x - 2y = -1\}$.

Wir müssen das Gleichungssystem $x+y = 1, x-2y = -1$ lösen. Wir erhalten $-3y = -2$. D.h. $(1/3, 2/3)$ ist der einzige Schnittpunkt.

1.2.6 Definition

Sind $a_1, a_2, a_3, b \in \mathbb{R}$ mit $(a_1, a_2, a_3) \neq (0, 0, 0)$, so wird die Menge

$$E = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid a_1x_1 + a_2x_2 + a_3x_3 = b\}$$

als *Ebene* (im dreidimensionalen Raum) bezeichnet.

1.2.7 Bemerkung

Genau wie oben kann man zeigen, dass es zu jeder Ebene E Zahlen c_1, c_2, c_3, d_1, d_2, d_3, $e_1, e_2, e_3 \in \mathbb{R}$ gibt, sodass

$$E = \{(c_1, c_2, c_3) + t(d_1, d_2, d_3) + u(e_1, e_2, e_3) \colon t, u \in \mathbb{R}\}$$

gilt.

1.2.8 Beispiel

Was sind die gemeinsamen Punkte der Ebenen

$$E_1 \coloneqq \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = -6\}$$
$$E_2 \coloneqq \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid 2x_1 + 3x_2 + 4x_3 = -16\}\,?$$

Eine Äquivalenzumformung liefert das Gleichungssystem $x_1+x_3+x_3 = -6, x_2+2x_3 = -4$. Dies hat die allgemeine Lösung

$$x_3 \text{ beliebig}, \; x_2 = -4 - 2x_3, \; x_1 = -6 - (-4 - 2x_3) - x_3 = -2 + x_3\,.$$

D.h. man kann die Schnittmenge in der Form

$$E_1 \cap E_2 = \{(-2 + t, -4 - 2t, t) \colon t \in \mathbb{R}\}$$

angeben.

1.2.9 Definition

Sei $n > 3$ eine ganze Zahl. Betrachtet man $a_1, \ldots, a_n, b \in \mathbb{R}$ mit $(a_1, \ldots, a_n) \neq (0, \ldots, 0)$ so wird

$$H \coloneqq \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \cdots a_nx_n = b\}$$

als *Hyperebene* im n-dimensionalen Raum $\mathbb{R}^n$ bezeichnet.

1.3 Abbildungen

Vorbemerkung

Wir kommen später zur formalen Definition von Abbildungen bzw. Funktionen, hier geben wir zunächst ein paar Beispiele.

1.3.1 Definition

Ist $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, so heißt die Abbildung $\mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto x + a$ *Translation* (oder Verschiebung) um a.

1.3.2 Beispiel

Die Abbildung $\mathbb{R}^2 \to \mathbb{R}^3$, $(x_1, x_2) \mapsto (x_1, x_2, 0)$ wird auch als *Einbettung* bezeichnet.

1.3.3 Beispiel

Die Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$, $(x_1, x_2) \mapsto (x_2, x_1)$ ist die Spiegelung an der Geraden $x_1 = x_2$.

Kapitel 2

Grundlagen

2.1 Naive Mengenlehre

2.1.1 Definition

Unter einer Menge verstehen wir jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Die wohlunterschiedenen Objekte heißen Elemente der Menge. (1895 Georg Cantor: Beiträge zur Begründung der Mengenlehre)

2.1.2 Beispiel und Definition

Endliche Mengen kann man durch Aufzählen angeben:

$$A := \{1, 2, 3\}, \ B := \{a, b, c\}$$

Hierbei steht $:=$ für *ist definiert als.* Es ist $1 \in A$ und $c \in B$. Hierbei bedeutet $\in$ ist Element von.

Es gilt

$$\{1, 2\} = \{1, 1, 2\}\,.$$

D.h., ein Element in einer Menge hat keine *Vielfachheit.* Zwei Mengen A,B heißen *gleich*, wenn jedes Element der einen Menge auch in der anderen enthalten ist und umgekehrt.

2.1.3 Bemerkung

Die Elemente $x_1, \ldots, x_n \in X$ heißen paarweise verschieden, wenn $x_i \neq x_j$ für jedes Indexpaar i, j mit $i \neq j$ gilt.

Die leere Menge bezeichnen wir mit $\emptyset$.

Eine Menge A heißt *Teilmenge* der Menge B, wenn jedes Element aus A in B enthalten ist. Man schreibt dann $A \subset B$. Es gilt

$$A = B \Longleftrightarrow A \subset B \text{ und } B \subset A\,.$$

2.1.4 Natürliche, ganze, rationale und reelle Zahlen

Die einfachste unendliche Menge ist die Menge

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

Sie heißt die Menge der *natürlichen Zahlen*.

Weiterhin ist

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\} \text{ und } \mathbb{Q} = \left\{\frac{p}{q} \colon p, q \in \mathbb{Z}, q \neq 0\right\}$$

die Menge der ganzen Zahlen und der rationalen Zahlen. Die Menge der reellen Zahlen wird auch mit $\mathbb{R}$ bezeichnet. Es gilt $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Die wichtigste Eigenschaft der natürlichen Zahlen ist die folgende: Ist $M \subset \mathbb{N}$ eine Teilmenge mit

$$1 \in M \text{ und } n \in M \Longrightarrow n + 1 \in M$$

so folgt $M = \mathbb{N}$. Dieser Sachverhalt wird auch als Prinzip der vollständigen Induktion bezeichnet.

2.1.5 Konstruktion von Mengen aus Mengen

Ist X eine Menge, so ist

$$Y \coloneqq \{x \in X \mid x \text{ hat die Eigenschaft } E\}$$

eine Teilmenge von X. Beispiel: $\{u \in \mathbb{Z} \mid u \text{ ist ungerade }\}$.

Sind $X_1, \ldots, X_n$ Mengen, so ist

$$X_1 \cup \ldots \cup X_n \coloneqq \{x \colon \text{Es gibt ein } i \in \{1, \ldots, n\} \text{ mit } x \in X_i\}$$

die *Vereinigungsmenge* und

$$X_1 \cap \ldots \cap X_n \coloneqq \{x \colon \text{Für jedes } i \in \{1, \ldots, n\} \text{ gilt } x \in X_i\}$$

die *Schnittmenge*.

Manchmal schreibt man auch $\bigcap_{i \in I} X_i$ und $\bigcup_{i \in I} X_i$ wenn Schnitt und Vereinigung von einer Familie von Mengen $X_i, i \in I$ betrachtet werden soll.

Weiterhin ist

$$A \setminus B \coloneqq \{a \in A \mid a \notin B\}$$

die *Differenzmenge* (oder das Komplement).

Philosophie: Um mathematische Objekte formal genau zu beschreiben verwendet man sehr oft Mengen.

2.2 Abbildungen und Relationen

Sind X, Y zwei Mengen, so betrachtet man *Funktionen* (oder Abbildungen) von X nach Y. Darunter versteht man eine Zuordnung $f \colon X \to Y$, die jedem Element $x \in X$ genau ein Element $y \in Y$ zuordnet. Man schreibt $f(x)$ für y sowie $x \mapsto f(x)$. Man nennt X die *Argumentmenge (Definitionsmenge)* und Y die *Zielmenge* von f.

Zwei Abbildungen $f, g \colon X \to Y$ heißen *gleich*, wenn $f(x) = g(x)$ für jedes $x \in X$ erfüllt ist. Mit $\mathrm{Abb}(X, Y)$ oder Y^X bezeichnet man die Menge aller Abbildungen $X \to Y$.

2.2.1 Definition

Es seien X, Y zwei Mengen.

1. Es bezeichnet
$$X \times Y := \{(x, y) \colon x \in X, y \in Y\}$$
das *kartesische (oder direkte) Produkt* von X und Y. Es besteht aus den Mengen aller Paare (x, y). Zwei Paare (u, v) und (x, y) sind genau dann gleich, wenn $u = x$ und $v = y$ gilt.

2. Eine *Relation R zwischen den Mengen X und Y* ist eine Teilmenge von $X \times Y$. Es steht $x \in X$ mit $y \in Y$ genau dann in Relation, wenn $(x, y) \in R$ gilt. Eine Relation zwischen X und X wird auch als *Relation auf X* bezeichnet.

3. Eine Relation R auf X heißt *symmetrisch*, wenn $(x_1, x_2) \in R \Longrightarrow (x_2, x_1) \in R$ gilt.

4. Eine Relation R auf X heißt *reflexiv*, wenn $(x, x) \in R$ für alle $x \in X$ gilt.

5. Eine Relation R auf X heißt *transitiv*, wenn $(x, y), (y, z) \in R \Longrightarrow (x, z) \in R$ gilt.

6. Eine symmetrische, reflexive und transitive Relation auf X heißt auch *Äquivalenzrelation auf X*.

7. Eine Funktion $f \colon X \to Y$ ist eine Relation zwischen X und Y, sodass jedes $x \in X$ mit genau einem $y \in Y$ in Relation steht. Dieses y wird dann auch als $f(x)$ bezeichnet.

8. Eine Funktion $f \colon X \to Y$ heißt *injektiv*, wenn $f(x_1) = f(x_2)$ nur im Fall $x_1 = x_2$ erfüllt ist.

9. Eine Funktion $f \colon X \to Y$ heißt *surjektiv*, wenn es zu jedem $y \in Y$ ein $x \in X$ mit $f(x) = y$ gibt.

10. Eine Funktion heißt *bijektiv* wenn sie injektiv und surjektiv ist.

2.2.2 Bemerkung

Ist eine Funktion $f \colon X \to Y$ bijektiv, so schreibt man im Fall $f(x) = y$ oft $x = f^{-1}(y)$. Es wird $f^{-1} \colon Y \to X$ als die Umkehrabbildung von f bezeichnet. f^{-1} ist in dieser Situation eine Funktion, da es zu jedem $y \in Y$ stets genau ein $x \in X$ mit $f(x) = y$ gibt.

2.2.3 Bemerkung

Man schreibt oft xRy oder $x \sim y$ wenn x und y bezüglich R oder $\sim$ in Relation stehen.

2.2.4 Beispiele

- Stellt man sich eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ als Relation vor, so ist dies genau der Graph der Funktion als Teilmenge des $\mathbb{R}^2$.

- Eine typische Anwendung von Relationen finden Sie in einem Geschäft. Haben Sie eine Menge von Aufträgen und eine Menge von Kunden, so steht jeder Auftrag mit dem zugehörigen Kunden in Relation.

- Auf der Menge aller Menschen gibt es die Relationen ≪haben das gleiche Alter≫, ≪ist Tochter von≫ und viele weitere. Das erste ist ein Beispiel für eine Äquivalenzrelation.

- Auf den ganzen Zahlen gibt es die Relation x teilt y. (In Zeichen $x \mid y$.)

- Auf $\mathbb{R}^3$ gibt es die Äquivalenzrelation $(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$ g.d.w. $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ gilt.

Wir zeigen, dass obige Relation $\sim$ eine Äquivalenzrelation auf $\mathbb{R}^3$ ist:

Sei zunächst $(x_1, x_2, x_3) \in \mathbb{R}^3$ beliebig. Dann gilt $x_1^2+x_2^2+x_3^2 = x_1^2+x_2^2+x_3^2$. Es steht also (x_1, x_2, x_3) bezüglich $\sim$ mit sich selbst in Relation. Dies zeigt die Reflexivität.

Sei nun $(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$. Dann gilt $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ und folglich auch $y_1^2+y_2^2+y_3^2 = x_1^2+x_2^2+x_3^2$. Hieraus folgt $(y_1, y_2, y_3) \sim (x_1, x_2, x_3)$. Die Symmetrie ist gezeigt.

Sei nun $(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$ und $(y_1, y_2, y_3) \sim (z_1, z_2, z_3)$, dann gilt $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ und $y_1^2 + y_2^2 + y_3^2 = z_1^2 + z_2^2 + z_3^2$. Hieraus folgt $x_1^2 + x_2^2 + x_3^2 = z_1^2 + z_2^2 + z_3^2$. Dies impliziert $(x_1, x_2, x_3) \sim (z_1, z_2, z_3)$ und zeigt somit die Transitivität. $\square$

2.2.5 Satz

Ist X eine endliche Menge und $f\colon X \to X$ eine Funktion. Dann sind äquivalent: f ist injektiv, f ist surjektiv und f ist bijektiv.

Beweis:

Es sei $X = \{x_1, \ldots, x_n\}$ eine n-elementige Menge.

Angenommen f ist injektiv, dann folgt aus $i \neq j$ sofort $f(x_i) \neq f(x_j)$. Somit sind $f(x_1), \ldots, f(x_n)$ paarweise verschiedene Elemente aus X. Da X genau n Elemente hat gilt $X = \{f(x_1), \ldots, f(x_n)\}$. Somit ist f surjektiv.

Angenommen f ist surjektiv, dann gibt es für jedes $y \in X$ ein $x \in X$ mit $f(x) = y$. Es gilt

$$n = \#X = \# \bigcup_{y \in X} \{x \in X \mid f(x) = y\} = \sum_{y \in X} \#\{x \in X \mid f(x) = y\}\,.$$

Da jedes y Funktionswert von f ist, hat jede der Mengen $\{x \in X \mid f(x) = y\}$ mindestens ein Element. Hätte eine dieser Mengen mehr als ein Element, so führt dies zu dem Widerspruch

$$n = \sum_{y \in X} \#\{x \in X \mid f(x) = y\} > \sum_{y \in X} 1 = \#X = n\,.$$

Somit hat jeder der Mengen $\{x \in X \mid f(x) = y\}$ genau ein Element. Damit ist f injektiv. $\square$

2.2.6 Definition

Sei $f\colon X \to Y$ ein Funktion.

1. Ist $A \subset X$, so heißt $f(A) := \{f(x)\colon x \in A\}$ das *Bild von A*.
2. Ist $B \subset Y$, so heißt $f^{-1}(B) := \{x \in X \mid f(x) \in B\}$ das *Urbild von B*.
3. Es heißt $f(X)$ auch das *Bild (Bildmenge, Wertemenge) von f*.

2.2.7 Definition

Sind $f\colon X \to Y$ und $g\colon Y \to Z$ zwei Funktionen, so ist $g \circ f$ eine Funktion $X \to Z$. Sie ist durch $(g \circ f)(x) := g(f(x))$ gegeben. Man nennt dies die *Komposition* (Verkettung) der Abbildungen.

2.2.8 Bemerkung

Im Allgemeinen gilt $f \circ g \neq g \circ f$.

2.2.9 Satz

Es seien $h\colon X \to Y$, $g\colon Y \to Z$ und $f\colon Z \to W$ Abbildungen. Dann gilt

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

D.h. die Komposition von Abbildungen ist assoziativ.

Beweis:

Für jedes $x \in X$ gilt:

$((f \circ g) \circ h)(x) = ((f \circ g)(h(x)) = f(g(h(x))) = f((g \circ h)(x)) = (f \circ (g \circ h))(x)\,.$ $\square$

2.2.10 Definition

Sei X eine Menge, dann bezeichnen wir mit $\mathrm{id}_X\colon X \to X, x \mapsto x$ die *identische Abbildung*. Die identische Abbildung bildet jedes Element auf sich selber ab.

2.2.11 Satz

Es sei $f\colon X \to Y$ eine Abbildung zwischen den nicht-leeren Mengen X und Y. Dann gilt

1. f ist genau dann injektiv, wenn es eine Funktion $g\colon Y \to X$ mit $g \circ f = \mathrm{id}_X$ gibt.
2. f ist genau dann surjektiv, wenn es eine Funktion $g\colon Y \to X$ mit $f \circ g = \mathrm{id}_Y$ gibt.
3. f ist genau dann bijektiv, wenn es eine Funktion $g\colon Y \to X$ mit $f \circ g = \mathrm{id}_Y$ und $g \circ f = \mathrm{id}_X$ gibt. In diesem Fall heißt g die Umkehrfunktion von f. Oft schreibt man $g = f^{-1}$.

Beweis:

1. Sei f injektiv. Dann gibt es zu jedem $y \in f(X)$ genau ein $x \in X$ mit $f(x) = y$. Wir definieren $g(y) := x$. Mit einem beliebigen $x_0 \in X$ definieren wir $g(y) := x_0$ für alle $y \in Y \setminus f(X)$. Nun gilt $g \circ f = \mathrm{id}_X$.

   Ist umgekehrt $g\colon Y \to X$ mit $g \circ f = \mathrm{id}_X$ gegeben, und ist $f(x) = f(x')$ für $x, x' \in X$, so folgt

   $$x = \mathrm{id}_X(x) = (g \circ f)(x) = g(f(x)) = g(f(x')) = \mathrm{id}_X(x') = x' .$$

   Folglich ist f injektiv.

2. Sei f surjektiv. Dann gibt es zu jedem $y \in Y$ mindestens ein $x \in X$ mit $f(x) = y$. Wir können daher zu jedem $y \in Y$ ein solches $x \in X$ auswählen. Dies definiert eine Funktion $g\colon Y \to X$ mit $g(y) = x$. Sie hat die Eigenschaft $f \circ g = \mathrm{id}_Y$.

   Ist umgekehrt $g\colon Y \to X$ mit $f \circ g = \mathrm{id}_Y$ gegeben, so gilt für jedes $y \in Y$

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

   Somit liegt y im Bild von f und f ist surjektiv.

3. Ist f bijektiv, so erfüllt f^{-1} die beiden Gleichungen. Ist umgekehrt ein solches g gegeben, so ist f nach den zuvor bewiesenen Aussagen injektiv und surjektiv. Folglich ist f bijektiv und $g = f^{-1}$.

$\square$

2.2.12 Bemerkung

Ist $f\colon X \to Y$ eine Funktion, so ist $\Gamma_f := \{(x, f(x))\colon x \in X\} \subset X \times Y$ der Graph von f. Formal ist der Graph genau die mengentheoretische Definition der Funktion.

2.2.13 Bemerkung

Möchte man die Mengen $\mathbb{R}^2, \mathbb{R}^3$ formal exakt definieren, so muss man diese als die Menge aller Funktionen $\{1,2\} \to \mathbb{R}$ und $\{1,2,3\} \to \mathbb{R}$ verstehen.

So ist beispielsweise $(3, 17, -9) \in \mathbb{R}^3$ eine Funktion, die 1 auf 3, 2 auf 17 und 3 auf -9 abbildet.

Sind A, B, C, D beliebige Mengen, so definiert man in gleicher Art $A \times B \times C \times D$ als die Menge aller Funktionen $f\colon \{1,2,3,4\} \to A \cup B \cup C \cup D$ mit $f(1) \in A, f(2) \in B, f(3) \in C$ und $f(4) \in D$.

Darüber hinaus definiert man für eine beliebige Indexmenge I und eine Familie von Mengen $(X_i)_{i \in I}$

$$\bigtimes_{i \in I} X_i$$

als die Menge aller auf I definierten Funktionen f mit $f(i) \in X_i$ für alle $i \in I$. Beispielsweise ist die Menge aller reellen Zahlenfolgen die Menge $\bigtimes_{i \in \mathbb{N}} \mathbb{R}$.

2.2.14 Definition

Sei X eine Menge und $\sim$ eine Äquivalenzrelation auf X. Für $x \in X$ bezeichnen wir mit

$$[x] := \{y \in X \mid x \sim y\}$$

die Menge aller Elemente mit denen x in Relation steht. Diese Menge heißt *Äquivalenzklasse von x*.

Darüber hinaus ist $\{[x] \colon x \in X\}$ die Menge aller Äquivalenzklassen.

2.2.15 Bemerkung

Die Reflexivität von $\sim$ liefert, dass $x \in [x]$ für alle $x \in X$ erfüllt ist.

2.2.16 Satz

Sind A, A' zwei Äquivalenzklassen der Äquivalenzrelation $\sim$ auf X so gilt entweder $A = A'$ oder $A \cap A' = \emptyset$. D.h. zwei Klassen sind gleich oder disjunkt.

Beweis:

Angenommen A, A' sind nicht disjunkt. Dann gibt es $a, a' \in X$ mit $[a] = A$ und $[a'] = A'$ und ein b mit $b \in A, b \in A'$.

Sei t ein beliebiges Element von A. Dann gilt $a \sim t$, $a \sim b$ und $a' \sim b$. Verwenden wir die Transitivität und die Symmetrie von $\sim$ so folgt $a' \sim t$. Folglich ist $t \in A'$. Dies beweist $A \subset A'$.

In gleicher Art zeigt man $A' \subset A$, woraus $A = A'$ folgt. □

2.2.17 Bemerkung

Die Äquivalenzklassen bilden eine disjunkte Zerlegung von X. Ein Element $a \in [x]$ wird auch als Repräsentant der Äquivalenzklasse bezeichnet. Wählt man aus jeder Klasse einen Repräsentanten so erhält man ein *Repräsentantensystem*.

2.2.18 Beispiel

Auf der Menge aller Schülerinnen und Schüler einer Schule kann man die Relation ≪gehen in die gleiche Klasse≫ betrachten. Man sieht leicht, dass dies eine Äquivalenzrelation ist. Versteht man eine Schulklasse als Menge von Schülern, so bilden die Schulklassen genau die Äquivalenzklassen der Relation. Ein typisches Repräsentantensystem ist die Menge aller Klassensprecher.

Bei organisatorischen Fragen (z.B. Stundenplänen) redet man oft von den Schulklassen und nicht von einzelnen Schülern. Dies ist der praktische Einsatz von Äquivalenzrelationen.

2.3 Gruppen

2.3.1 Definition

Es sei M eine Menge. Eine Abbildung $\star \colon M \times M \to M$ wird als *Verknüpfung* (Komposition) bezeichnet.

2.3.2 Beispiele

1. Auf den Mengen $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ gibt z.B. es die Verknüpfungen $+$ und $\cdot$.
2. $-$ ist keine Verknüpfung auf $\mathbb{N}$ aber eine Verknüpfung auf $\mathbb{Z}$.
3. $a \star b := \frac{a+b}{2}$ ist eine Verknüpfung auf $\mathbb{R}$. Sie heißt arithmetisches Mittel.
4. Auf $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ ist $\circ$ eine Verknüpfung.

2.3.3 Definition

1. Ist M eine Menge mit Verknüpfung $\star$, so bezeichnet man $(M, \star)$ als *Magma.*
2. Ist $(M, \star)$ ein Magma und erfüllt $\star$ das Assoziativgesetz, $(a \star b) \star c = a \star (b \star c)$ für alle $a, b, c \in M$, so nennt man $(M, \star)$ eine *Halbgruppe.*
3. Hat die Halbgruppe $(M, \star)$ ein neutrales Element $e \in M$, d.h. es gilt $x \star e = e \star x = x$ für alle $x \in M$, so wird $(M, \star, e)$ sie als *Monoid* bezeichnet.
4. Ist $(M, \star, e)$ ein Monoid und gibt es zusätzlich zu jedem $x \in M$ ein $x' \in M$ mit $x' \star x = e$, so nennt man $(M, \star, e)$ eine *Gruppe.* Man nennt x' das zu *x inverse Element.*
5. Eine Gruppe $(M, \star, e)$ heißt *abelsch* (kommutativ), wenn $a \star b = b \star a$ für alle $a, b \in M$ erfüllt ist.

Definierende Eigenschaften werden auch als *Axiome* bezeichnet. Die definierenden Eingenschaften einer Gruppe nennt man auch *Gruppenaxiome.*

2.3.4 Bemerkung

Oft schreibt man nur M statt $(M, \star, e)$ wenn $\star$ und e aus dem Kontext klar sind. Weiterhin schreibt man oft $a \cdot b$ oder ab statt $a \star b$ sowie a^{-1} für das inverse Element.

Ist die Gruppe kommutativ, so schreibt man auch $a + b$. In diesem Fall schreibt man auch 0 statt e und $-x$ für das inverse Element.

2.3.5 Beispiel

Wir betrachten die Menge $\{-1, 1\}$ mit der gewöhnlichen Multiplikation. Wir erhalten die Verknüpfungstafe:

$\cdot$	1	-1
1	1	-1
-1	-1	1

Man sieht leicht, dass dies eine abelsche Gruppe mit 2 Elementen ist. Prinzipiell kann man jede endliche Gruppe durch eine Verknüpfungstafel beschreiben.

2.3.6 Beispiel

Sei X eine Menge. Wir untersuchen, wie sich $(\mathrm{Abb}(X,X),\circ)$ in obiges Schema einordnet. Da man durch Verkettung zwei beliebigen Abbildungen $f,g\colon X\to X$ wieder eine Abbildung $f\circ g\colon X\to X$ erhält, ist dies mindestens ein Magma. Die Assoziativität von $\circ$ wurde bereits untersucht. Somit erhalten wir eine Halbgruppe.

Weiterhin ist id_X aufgrund von $f\circ\mathrm{id}_X=f=\mathrm{id}_X\circ f$ ein neutrales Element. Es handelt sich also um ein Monoid.

Hat die Menge X mehr als ein Element, so können wir zu festem $x_0\in X$ die Funktion $f_0(x):=x_0$ betrachten. Diese ist nicht invertierbar. Somit ist die Menge keine Gruppe.

Betrachtet man $S(X):=\{f\in\mathrm{Abb}(X,X)\mid f \text{ ist bijektiv }\}$, so erhält man eine Gruppe. Man beachte, dass die Verkettung von bijektiven Abbildungen wieder bijektiv ist und somit $(S(X),\circ)$ erst einmal ein Magma ist. $S(X)$ heißt die *symmetrische Gruppe* auf X.

2.3.7 Satz

Ist G eine Gruppe und $a,x,y,\tilde{x},\tilde{y}\in G$ so gilt $(a^{-1})^{-1}=a$ sowie $ax=x\implies a=e$ und $ax=e\implies a=x^{-1}$. Insbesondere ist das neutrale Element und jedes inverse Element eindeutig bestimmt. Zudem gelten die Kürzungsregeln $ax=a\tilde{x}\implies x=\tilde{x}$ und $ya=\tilde{y}a\implies y=\tilde{y}$.

Anmerkung: Dieser Satz rechtfertigt die Bezeichnungen das neutrale und das inverse Element.

Beweis:

Wir erhalten

$$a=ea=((a^{-1})^{-1}a^{-1})a=(a^{-1})^{-1}(a^{-1}a)=(a^{-1})^{-1}e=(a^{-1})^{-1}\,.$$

Dies zeigt auch $aa^{-1}=(a^{-1})^{-1}a^{-1}=e$. Nun folgt

$$ax=x\implies(ax)x^{-1}=xx^{-1}\implies a(xx^{-1})=e\implies ae=e\implies a=e\,.$$

Die Kürzungsregeln zeigt man wie folgt:

$$\begin{gathered}ax=a\tilde{x}\implies a^{-1}ax=a^{-1}a\tilde{x}\implies ex=e\tilde{x}\implies x=\tilde{x}\\ ya=\tilde{y}a\implies yaa^{-1}=\tilde{y}aa^{-1}\implies ye=\tilde{y}e\implies y=\tilde{y}\end{gathered}$$

□

2.3.8 Definition

Sei $(G,\cdot,e)$ eine Gruppe. Eine nichtleere Teilmenge $U\subset G$ heißt *Untergruppe*, wenn $a,b\in U\implies a\cdot b,a^{-1}\in U$ gilt.

Sind $(G,\cdot)$ und $(H,\star)$ zwei Gruppen, so heißt eine Abbildung $\varphi\colon G\to H$ mit

$$\varphi(a\cdot b)=\varphi(a)\star\varphi(b)\text{ für alle }a,b\in G$$

ein *Gruppenhomomorphismus (Homomorphismus von Gruppen)*.

Ein bijektiver Homomorphismus heißt auch *Isomorphismus*.

2.3.9 Bemerkung

Ist G eine Gruppe und $U \subset G$ eine Untergruppe, so ist U mit der Verknüpfung von G wieder eine Gruppe.

Man nennt die Verknüpfung auch die von G induzierte Verknüpfung.

Beweis:

Die induzierte Verknüpfung ist assoziativ, da sie von G kommt. Da U nicht leer ist, gibt es ein $a \in U$. Es sind a^{-1} und $aa^{-1} = e$ ebenfalls Elemente von U. $\square$

2.3.10 Beispiele

1. Es ist $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ mit der gewöhnlichen Multiplikation eine Gruppe. Zudem ist $P := \{x \in \mathbb{R} \mid x > 0\} = (0, \infty)$ ebenfalls eine Gruppe mit der Multiplikation als Verknüpfung. Es sind $\{\pm 1\}$ und P Untergruppen von $(\mathbb{R}^\times, \cdot)$. Weiterhin sind

$$\begin{aligned} &\varphi \colon \mathbb{R}^\times \to P, x \mapsto |x| \\ &\psi \colon \mathbb{R}^\times \to \{\pm 1\}, x \mapsto \operatorname{sgn}(x) \end{aligned}$$

   Gruppenhomomorphismen.

2. Es sind $(\mathbb{R}, +)$ und $(P, \cdot)$ Gruppen. Die Abbildung $\exp \colon \mathbb{R} \to (0, \infty), x \mapsto e^x$ ist ein Gruppenhomomorphismus, da $e^{x+y} = e^x e^y$ gilt. Da exp bijektiv ist, handelt es sich sogar um einen Isomorphismus.

3. Wir betrachten die Gruppe $(\mathbb{Z}, +)$. Zu $m \in \mathbb{Z}$ konstruieren wir

$$m\mathbb{Z} := \{m \cdot a \colon a \in \mathbb{Z}\}\,.$$

   Es ist $m\mathbb{Z}$ eine Untergruppe da $ma + mb = m(a+b)$ und $-(ma) = m(-a)$ gilt. Zudem ist $\varphi_m \colon \mathbb{Z} \to m\mathbb{Z}, a \mapsto ma$ ein Gruppenhomomorphismus, da

$$\varphi_m(a+b) = m(a+b) = ma + mb = \varphi_m(a) + \varphi_m(b)$$

   gilt.

2.3.11 Satz

Sei $m > 0$ eine ganze Zahl, so bilden die Mengen

$$r + m\mathbb{Z} := \{r + ma \colon a \in \mathbb{Z}\},\ r = 0, 1, \ldots, m-1$$

eine disjunkte Zerlegung der ganzen Zahlen. D.h., es gilt

$$\mathbb{Z} = (0 + m\mathbb{Z}) \cup (1 + m\mathbb{Z}) \cup \ldots \cup ((m-1) + m\mathbb{Z})\,.$$

Die Mengen auf der rechten Seite heißen die Restklassen modulo m.

Beweis:

Sei $b \in \mathbb{Z}$ beliebig. Die Division mit Rest durch m liefert einen Quotienten $q \in \mathbb{Z}$ und einen Teilerrest $r \in \mathbb{Z}$. Es gelten

$$b = q \cdot m + r \text{ und } 0 \leq r < m\,.$$

Somit ist $b \in r + m\mathbb{Z}$. Dies zeigt, dass jede ganze Zahl in einer der Restklassen modulo m enthalten ist, und wie obige Zahl r passend konstruiert werden kann.

Wir haben zu zeigen, dass keine Zahl in zwei verschiedenen Restklassen enthalten ist. Angenommen dies wäre der Fall, so würde

$$b = r_1 + ma_1 = r_2 + ma_2 \text{ mit } a_1, a_2 \in \mathbb{Z}, 0 \leq r_1, r_2 < m, r_1 \neq r_2$$

gelten. Hieraus erhalten wir $r_1 - r_2 = m(a_2 - a_1)$. Folglich ist $r_1 - r_2$ durch m teilbar. Zudem gilt $-m < r_1 - r_2 < m$. Da die einzige durch m teilbare Zahl in diesem Bereich die 0 ist, folgt $r_1 - r_2 = 0$ und somit $r_1 = r_2$. □

2.3.12 Definition

Man nennt zwei Zahlen $b_1, b_2 \in \mathbb{Z}$ *kongruent modulo m*, wenn sie in der selben Restklasse modulo m liegen. Man schreibt dann auch $b_1 \equiv b_2 \pmod{m}$. Man notiert die Restklasse einer Zahl b mit $\overline{b}$ und schreibt auch $\overline{b_1} = \overline{b_2}$ für zwei kongruente Zahlen. Weiterhin schreibt man $\mathbb{Z}/m\mathbb{Z}$ für die Menge aller Restklassen modulo m.

2.3.13 Satz

Sei $m > 0$ eine ganze Zahl. Durch $\overline{a} + \overline{b} := \overline{a+b}$ wird auf $\mathbb{Z}/m\mathbb{Z}$ eine Verknüpfung definiert. Sie ist assoziativ und kommutativ. Das neutrale Element ist $\overline{0}$. Das zu $\overline{a}$ inverse Element ist $\overline{-a}$. D.h., wir haben eine abelsche Gruppe mit m Elementen konstruiert. Man nennt $(\mathbb{Z}/m\mathbb{Z}, +)$ die zyklische Gruppe der Ordnung m.

Beweis:

Wir haben zu zeigen, dass $\overline{a} + \overline{b} := \overline{a+b}$ wohldefiniert ist. Seien hierzu $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ mit $\overline{a_1} = \overline{a_2}$ und $\overline{b_1} = \overline{b_2}$. Dann gilt

$$a_1 = a_2 + km \text{ und } b_1 = b_2 + lm \text{ mit } k, l \in \mathbb{Z},$$

da a_1, a_2 sowie b_1, b_2 den gleichen Rest bei Division durch m lassen. Somit gilt

$$a_1 + b_1 = (a_2 + km) + (b_2 + lm) = (a_2 + b_2) + (k + l)m.$$

Dies zeigt $\overline{a_1 + b_1} = \overline{a_2 + b_2}$. Somit ist $\overline{a+b}$ unabhängig von der Wahl der Repräsentanten von $\overline{a}$ und $\overline{b}$. Weiterhin gilt $-a_1 = -a_2 - km$, was $\overline{-a_1} = \overline{-a_2}$ zeigt.

Alle weiteren Eigenschaften werden von den ganzen Zahlen vererbt. Wir zeigen dies exemplarisch für die Assoziativität:

$$(\overline{a} + \overline{b}) + \overline{c} = \overline{a+b} + \overline{c} = \overline{(a+b)+c} = \overline{a+(b+c)} = \overline{a} + \overline{b+c} = \overline{a} + (\overline{b} + \overline{c}) \quad \square$$

2.3.14 Bemerkung

Man findet die Restklassen in der Kalenderrechnung. Die Wochentage bilden die Tage modulo 7 ab. Die Stundenangabe ist eine Angabe modulo 12 bzw. 24.

D.h., 15 Tagen nach Montag ist Dienstag, 5 Stunden nach 10 Uhr ist 3 Uhr.

2.3.15 Beispiel

Teilt man die ganzen Zahlen in geraden und ungerade Zahlen ein, so ist dies die Zerlegung in Restklassen modulo 2.

2.3.16 Bemerkung

In der Sprache der Relationen haben wir zeigt, dass «Kongruent modulo m» eine Äquivalenzrelation auf den ganzen Zahlen ist, die zudem *verknüpfungsverträglich* ist. D.h., es gilt

$$a_1 \equiv a_2, b_1 \equiv b_2 \Longrightarrow a_1 + b_1 \equiv a_2 + b_2 \pmod{m}.$$

2.4 Ringe, Körper und Polynome

2.4.1 Definition

Eine Menge R zusammen mit zwei Verknüpfungen

$$+\colon R \times R \to R, (a, b) \mapsto a + b$$
$$\cdot\colon R \times R \to R, (a, b) \mapsto a \cdot b$$

heißt *Ring*, wenn gilt:

1. R ist zusammen mit der Addition $+$ eine abelsche Gruppe.
2. Die Multiplikation $\cdot$ auf R ist assoziativ.
3. Es gelten die Distributivgesetze

   $$a \cdot (b + c) = a \cdot b + a \cdot c \text{ und } (a + b) \cdot c = a \cdot c + b \cdot c$$

   für alle $a, b, c \in R$.

Weiterhin heißt ein Ring kommutativ, wenn $a \cdot b = b \cdot a$ für alle $a, b \in R$ gilt. Ein Element $1 \in R$ heißt *Einselement*, wenn $1 \cdot a = a \cdot 1 = a$ für alle $a \in R$ gilt. Hat ein Ring ein Einselement, so wird er auch als *Ring mit Eins* bezeichnet.

Die definierenden Eigenschaften eines Rings heißen auch *Ringaxiome*.

2.4.2 Beispiele

1. Die ganzen Zahlen, die rationalen Zahlen und die reellen Zahlen bilden zusammen mit der gewöhnlichen Addition und Multiplikation kommutative Ringe mit Eins.

2. Ist M eine nicht-leere Menge, so bildet die Menge aller reellen Funktionen auf M
$$R := \{f\colon M \to \mathbb{R}\}$$
einen kommutativen Ring, wenn wir die Verknüpfungen
$$(f+g)(x) := f(x) + g(x), (f \cdot g)(x) := f(x) \cdot g(x)$$
betrachten.
Nachweis: Übung.

2.4.3 Bemerkung

Ist R ein beliebiger Ring und $0 \in R$ das Nullelement, so gilt für alle $a \in R$
$$0 \cdot a = a \cdot 0 = 0\,.$$

Beweis:
$$0 + 0 \cdot a = 0 \cdot a = (0+0) \cdot a = 0 \cdot a + 0 \cdot a \Longrightarrow 0 = 0 \cdot a \quad \square$$

2.4.4 Beispiel

Ist m eine natürliche Zahl, so haben wir auf der Menge der Restklassen $\mathbb{Z}/m\mathbb{Z}$ bereits eine Addition eingeführt. Wir definieren zusätzlich die Multiplikation
$$\overline{a} \cdot \overline{b} := \overline{a \cdot b}$$
auf $\mathbb{Z}/m\mathbb{Z}$. Wir haben zunächst zu zeigen, dass dies eine Definition ist. Seien $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ mit $\overline{a_1} = \overline{a_2}$ und $\overline{b_1} = \overline{b_2}$. Dann gibt es $k, l \in \mathbb{Z}$ mit $a_1 = a_2 + km$ und $b_1 = b_2 + lm$. Es folgt
$$a_1 \cdot b_1 = (a_2 + km)(b_2 + lm) = a_2 \cdot b_2 + m(kb_2 + la_2 + klm)\,,$$
somit ist die Multiplikation unabhängig von der Wahl der Repräsentanten. Die zweite und die dritte Forderung in der Definition des Begriffs Ring übertragen sich von $\mathbb{Z}$ auf $\mathbb{Z}/m\mathbb{Z}$.

Wir stellen die Multiplikationstafeln von $\mathbb{Z}/2\mathbb{Z}$ $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$ auf:

$\cdot$	0	1
0	0	0
1	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Beobachtung: In $\mathbb{Z}/4\mathbb{Z}$ gilt $2 \cdot 1 = 2 \cdot 3$ und $2 \cdot 2 = 0$.
Die Kürzungsregel $a \neq 0$ und $a \cdot b = a \cdot c \Longrightarrow b = c$ gilt hier nicht.
Die Regel $a \cdot b = 0 \Longrightarrow a = 0$ oder $b = 0$ gilt ebenfalls nicht.

2.4.5 Definition

Ein Ring heißt *nullteilerfrei* wenn in ihm die Rechenregel
$$a \cdot b = 0 \Longrightarrow a = 0 \vee b = 0$$
gilt.

Hierbei ist $\vee$ das Symbol für oder.

2.4.6 Beispiel

$\mathbb{R}$ ist nullteilerfrei, $\mathbb{Z}/4\mathbb{Z}$ ist nicht nullteilerfrei.

2.4.7 Bemerkung

Ist $m \geq 2$ eine natürliche Zahl, so ist der Ring $\mathbb{Z}/m\mathbb{Z}$ genau dann nullteilerfrei, wenn m eine Primzahl ist.

Beweis:

Ist m keine Primzahl, so ist m ein Produkt $m = k \cdot l$ mit $1 < k, l < m$. Somit ist $\overline{k}, \overline{l} \neq 0$ aber $\overline{0} = \overline{m} = \overline{kl} = \overline{k} \cdot \overline{l}$.

Ist umgekehrt m eine Primzahl und $\overline{k} \cdot \overline{l} = \overline{0}$, so folgt

$$k \cdot l = r \cdot m$$

für ein passendes $r \in \mathbb{Z}$. Folglich hat mindestens eines von k und l den Primfaktor m. Somit gilt $\overline{k} = 0$ oder $\overline{l} = 0$. $\square$

2.4.8 Definition

Ist R ein Ring und $S \subset R$ eine Teilmenge, so heißt S ein *Unterring*, wenn S bezüglich der Addition eine Untergruppe und bezüglich der Multiplikation $a, b \in S \Longrightarrow a \cdot b \in S$ gilt.

Sind R und T Ringe mit den Verknüpfungen $+, \cdot$ und $\oplus, \odot$ so heißt eine Abbildung $\varphi \colon R \to T$ ein *Homomorphismus von Ringen* (oder Ringhomomorphismus), wenn für alle $a, b \in R$

$$\varphi(a + b) = \varphi(a) \oplus \varphi(b) \text{ und } \varphi(a \cdot b) = \varphi(a) \odot \varphi(b)$$

gilt.

2.4.9 Beispiele

1. Es ist $\mathbb{Z}$ ein Unterring von $\mathbb{R}$.
2. Die Menge der geraden Zahlen $2\mathbb{Z}$ bildet einen Unterring von $\mathbb{Z}$.
3. Die Abbildung $\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, a \mapsto \overline{a}$ ist ein Ringhomomorphismus.

2.4.10 Bemerkung

Sei R ein nullteilerfreier Ring. Dann ist $R \setminus \{0\}$ abgeschlossen unter $\cdot$. Welche algebraische Struktur erhalten wir?

- Das Beispiel $\mathbb{Z}$ zeigt, dass nicht alle Elemente invertierbar sind.
- Das Beispiel $2\mathbb{Z}$ zeigt, dass es kein neutrales Element geben muss.

Wir erhalten somit nur eine Halbgruppe. In Spezialfällen wie $\mathbb{Q}$ und $\mathbb{R}$ gilt jedoch mehr.

2.4.11 Definition

Eine Menge K zusammen mit zwei Verknüpfungen

$$+\colon K \times K \to K, (a,b) \mapsto a+b$$
$$\cdot\colon K \times K \to K, (a,b) \mapsto a \cdot b$$

heißt *Körper*, wenn sie ein nullteilerfreier Ring ist und zudem $K^\times := K \setminus \{0\}$ mit der Multiplikation eine kommutative Gruppe bildet. Das Inverse zu einem Element $a \in K^\times$ bezüglich $\cdot$ wird mit a^{-1} oder $\frac{1}{a}$ bezeichnet.

Die definierenden Eigenschaften eines Körpers nennt man auch *Körperaxiome.*

2.4.12 Beispiele

$\mathbb{Q}$ und $\mathbb{R}$ sind Körper.

2.4.13 Bemerkung

In jedem Körper K gelten die folgenden Rechenregeln:

1. $1 \neq 0$, insbesondere hat jeder Körper mindestens 2 Elemente.
2. $a \cdot 0 = 0 \cdot a = 0$.
3. $a \cdot b = 0 \implies a = 0$ oder $b = 0$.
4. $a(-b) = -(ab)$ und $(-a)(-b) = ab$.
5. Aus $x \cdot a = y \cdot a$ und $a \neq 0$ folgt $x = y$.

Beweis:

1. Es ist $0 \in K$, $0 \notin K^\times$ aber $1 \in K^\times$.
2. Dies wurde bereits für beliebige Ringe gezeigt.
3. Dies ist die Definition von Nullteilerfreiheit.
4. Es gilt $ab + (-a)b = (a + (-a))b = 0 \cdot b = 0$. Hiermit folgt

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

   da das Inverse des Inversen das ursprüngliche Element ist.
5. Da a in $K^\times$ ein Inverses hat, folgt $(xa)a^{-1} = (ya)a^{-1}$ und somit $x = y$.

$\square$

2.4.14 Konstruktion der komplexen Zahlen $\mathbb{C}$

Auf $\mathbb{R} \times \mathbb{R} = \{(a,b) : a,b \in \mathbb{R}\}$ definieren wir die Addition

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$$

und die Multiplikation

$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)\,.$$

Das additiv neutrale Element ist $(0,0)$, das additiv Inverse von (a,b) ist $(-a,-b)$. Die Eins hat die Form $(1,0)$, das multiplikativ inverse Element zu (a,b) ist

$$(a,b)^{-1} := \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right)\,.$$

Wir bezeichnen $\mathbb{R} \times \mathbb{R}$ mit dieser Verknüpfung als *Körper der komplexen Zahlen* $\mathbb{C}$.

Die Abbildung $\mathbb{R} \to \mathbb{C}, a \mapsto (a,0)$ ist injektiv. Da

$$(a_1, 0) + (a_2, 0) = (a_1 + a_2, 0) \text{ und } (a_1, 0) \cdot (a_2, 0) = (a_1 a_2, 0)$$

gilt, identifizieren wir $\mathbb{R}$ mit $\{(a,0) : a \in \mathbb{R}\} \subset \mathbb{C}$. Dann müssen wir auch bezüglich der Verknüpfungen nicht unterscheiden.

Man bezeichnet $i = (0,1)$ als imaginäre Einheit. Es gilt dann $i^2 = -1$ und $(a,b) = (a,0) + (0,b) = a + ib$.

Man nennt $\mathrm{Re}((a,b)) := a$ den Realteil und $\mathrm{Im}((a,b)) := b$ den Imaginärteil. Weiterhin heißt $(a,-b) = \overline{(a,b)}$ die zu (a,b) konjugierte komplexe Zahl. Für $\alpha, \beta \in \mathbb{C}$ gelten die Rechenregeln

$$\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta},\ \overline{\alpha \cdot \beta} = \overline{\alpha} \cdot \overline{\beta} \text{ und } \alpha = \overline{\alpha} \iff \alpha \in \mathbb{R}.$$

Es gilt für $\lambda = (a,b)$

$$\lambda\overline{\lambda} = (a+ib)(a-ib) = a^2 + b^2 \in [0, \infty).$$

Man setzt den *Absolutbetrag*

$$|\lambda| = \sqrt{\lambda\overline{\lambda}} = \sqrt{a^2+b^2}\,.$$

Man kann leicht zeigen, dass

$$|\lambda + \mu| \leq |\lambda| + |\mu| \text{ und } |\lambda \cdot \mu| = |\lambda| \cdot |\mu|$$

für $\lambda, \mu \in \mathbb{C}$ gilt.

2.4.15 Bemerkung

Die Addition in $\mathbb{C}$ entspricht der gewöhnlichen Addition in $\mathbb{R}^2$. Wir veranschaulichen die Multiplikation.

Stellt man eine komplexe Zahl λ in Polarkoordinaten mit Betrag $|\lambda|$ und Winkel α dar, so gilt

$$\mathrm{Re}(\lambda) = \cos(\alpha)|\lambda| \text{ und } \mathrm{Im}(\lambda) = \sin(\alpha)|\lambda|\,.$$

Mann nennt α das Argument $\arg(\lambda) := \alpha \in [0, 2\pi[$. Gemäß der eulerschen Formel gilt dann

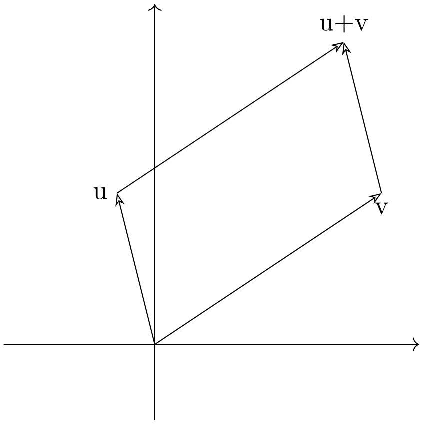
$$\lambda = |\lambda| \cdot e^{i\arg(\lambda)}$$

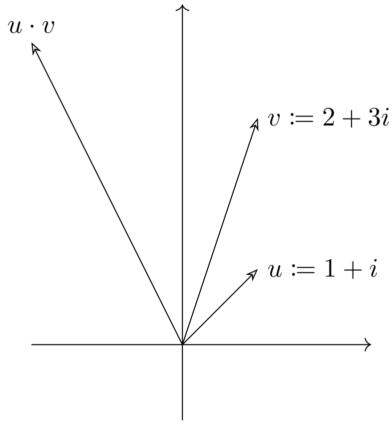
und somit

$$\lambda \cdot \mu = |\lambda| \cdot e^{i\arg(\lambda)} \cdot |\mu| \cdot e^{i\arg(\mu)} = |\lambda| \cdot |\mu| \cdot e^{i(\arg(\lambda)+\arg(\mu))}.$$

D.h. man multipliziert komplexe Zahlen indem man die Beträge multipliziert und die Winkel addiert.

Zeichnung: Summe und Produkt komplexer Zahlen





Beispiele:

$$|1+i| = \sqrt{2},\ \arg(1+i) = \frac{\pi}{4},\ (1+i) + (3-7i) = 4 - 6i$$

$$(5-i)\cdot(2+i) = 10 + 1 + (5-2)i = 11 + 3i$$

$$\frac{8}{1+i} = \frac{8(1-i)}{(1+i)(1-i)} = \frac{8(1-i)}{2} = 4(1-i) = 4 - 4i$$

2.4.16 Beispiel: Der kleinste Körper

Wir wissen, jeder Körper hat mindestens 0 und 1. Dies reicht, wenn man die Verknüpfung

$$\begin{array}{c|c|c} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array} \quad \text{und} \quad \begin{array}{c|c|c} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 0 & 1 \end{array}$$

benutzt. Dieses ist das Rechnen mit den Restklassen in $\mathbb{Z}/2\mathbb{Z}$.

2.4.17 Beispiel: Weitere endliche Körper

Ist m eine Primzahl, so ist $\mathbb{Z}/m\mathbb{Z}$ ein endlicher, nullteilerfreier, kommutativer Ring mit 1. Alle diese Ringe sind auch Körper.

Beweis:

Es ist zu zeigen, dass jedes $a \in \mathbb{Z}/m\mathbb{Z} \setminus \{0\}$ ein Inverses hat. Angenommen es gilt für kein $b \in \mathbb{Z}/m\mathbb{Z}$ die Gleichung $ab = 1$. Dann nimmt die Abbildung

$$\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, x \mapsto ax$$

nicht alle Werte an. Da es m Element in der Argument- und der Zielmenge gibt, muss es ein Element geben, dass mehrfach angenommen wird. Somit gibt es verschiedene Elemente $b_1, b_2 \in \mathbb{Z}/m\mathbb{Z}$ mit $ab_1 = ab_2$. Hieraus folgt $a(b_1 - b_2) = 0$. Wir haben Nullteiler gefunden, dies ist ein Widerspruch.

Ergebnis: Ist m eine Primzahl, so ist $\mathbb{Z}/m\mathbb{Z}$ ein Körper. □.

2.4.18 Definition

Ist R ein Ring mit Eins und $n \in \mathbb{N}$, so definieren wir

$$n \cdot 1 = \underbrace{1 + \cdots + 1}_{n-\text{mal}}$$

und

$$\mathrm{char}(R) := \begin{cases} 0 & \text{falls } n \cdot 1 \neq 0 \text{ für alle } n \geq 1 \\ \min\{n \in \mathbb{N} \setminus \{0\} \colon n \cdot 1 = 0\} & \text{sonst} \end{cases}$$

Es wird $\mathrm{char}(R)$ als *Charakteristik* von R bezeichnet.

2.4.19 Beispiele

Die Charakteristik von $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ sind 0. Die Charakteristik von $\mathbb{Z}/12\mathbb{Z}$ ist 12.

2.4.20 Lemma

Ist K ein Körper, so ist $\mathrm{char}(K)$ entweder Null oder eine Primzahl.

Beweis:

Angenommen, die Charakteristik ist nicht Null und keine Primzahl. Dann hat sie die Form $\mathrm{char}(K) = m = k \cdot l \neq 0$ mit $1 < k, l < m$. Es folgt

$$0 = m \cdot 1 = (k \cdot l) \cdot 1 = (k \cdot 1)(l \cdot 1)\,.$$

Da Körper nullteilerfrei sind, folgt

$$k \cdot 1 = 0 \text{ oder } l \cdot 1 = 0\,.$$

Dies ist ein Widerspruch. □

2.4.21 Definition

Sei K ein Körper und t eine formale Unbestimmte. Ein *Polynom mit Koeffizienten in* K ist ein formaler Ausdruck der Form

$$f(t) = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n .$$

Hierbei sind $a_0, \ldots, a_n \in K$. Mit $K[t]$ bezeichnet man die Menge dieser Polynome. Sind alle Koeffizienten Null, so spricht man vom *Nullpolynom* $f = 0$. Der *Polynomgrad* ist definiert als

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

Ein Polynom heißt normiert, wenn der Leitkoeffizient a_n gleich 1 ist.

2.4.22 Bemerkung

Aus einem Polynom $f = a_0 + a_1t + \cdots + a_nt^n \in K[t]$ kann man die Abbildung

$$\tilde{f}\colon K \to K, \lambda \mapsto \tilde{f}(t) := a_0 + a_1\lambda + \cdots + a_n\lambda^n$$

konstruieren. Diese Konstruktion ist selber eine Abbildung $K[t] \to \mathrm{Abb}(K, K), f \mapsto \tilde{f}$.

2.4.23 Beispiel

Sei $K = \{0, 1\}$ und $f = t^2 + t \in K[t]$. Dann ist

$$f = t^2 + t, \tilde{f}(0) = 0, \tilde{f}(1) = 0.$$

D.h. $\tilde{f}$ ist die Nullabbildung, aber f ist nicht das Nullpolynom.

2.4.24 Rechnen mit Polynomen

Sei $f = a_0 + a_1t + \cdots + a_nt^n, g = b_0 + b_1t + \cdots + b_mt^m \in K[t]$. Wir setzen formal $a_{n+1} = a_{n+2} = \cdots = 0 = b_{m+1}, b_{m+2} = \cdots$. Dann ist

$$f + g := (a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \cdots + \ldots t^{\max(n,m)}$$

die *Summe* und

$$f \cdot g := c_0 + c_1t + \cdots + c_{n+m}t^{n+m}$$

mit $c_k = \sum_{i+j=k} a_ib_j$ das *Produkt* der Polynome. Insbesondere ist

$$c_0 = a_0b_0, c_1 = a_0b_1 + a_1b_0, c_2 = a_0b_2 + a_1b_1 + a_2b_0, \ldots .$$

Gilt $f \cdot g = h$, so heißen f und g *Teiler* von h.

2.4.25 Bemerkung

Obige Verknüpfungen geben $K[t]$ die Struktur eines kommutativen Rings mit Eins. Er heißt *Polynomring* über K. Es gilt zudem

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

für $f, g \in K[t]$. Dabei rechnen wir formal $n + -\infty = -\infty + m = -\infty - \infty = -\infty$.

2.4.26 Satz (Polynomdivision)

Sind $f, g \in K[t]$ und $g \neq 0$, so gibt es eindeutig bestimmte Polynome $q, r \in K[t]$ mit

$$f = qg + r \text{ und } \deg(r) < \deg(g)$$

Beweis:

Zur Eindeutigkeit: Sei $f = q_1 \cdot g + r_1 = q_2 g + r_2$. Dann ist $r_1 - r_2 = g(q_2 - q_1)$. Links steht ein Polynom vom Grad kleiner g, rechts steht entweder Null oder ein Polynom vom Grad mindestens g. Somit steht links und rechts Null. Es folgt $r_1 = r_2$ und $q_1 = q_2$.

Zur Existenz: Diese beweisen wir konstruktiv. Der Fall $f = 0$ ist trivial. Andernfalls sei $f = a_0 + a_1 t + \cdots + a_n t^n$ und $g = b_0 + b_1 t + \cdots + b_m t^m$ mit $a_n \neq 0 \neq b_m$. Im Fall $n < m$ setzen wir $q := 0, r := f$. Im Fall $n \geq m$ setzen wir

$$q_1 := \frac{a_n}{b_m} t^{n-m}$$

und erhalten mit $f_1 := f - q_1 g$ ein Polynom kleineren Grades als f. Mit f_1, g berechnen wir (falls $\deg(f_1) \geq \deg(g)$) in gleicher Art q_2 und erhalten $f_2 := f_1 - q_2 g = f - (q_1 + q_2)g$. Wir wiederholen diesen Schritt $k \leq n - m + 1$ mal, bis wir

$$f_k := f_{k-1} - q_k g$$

mit $\deg(f_k) < \deg(g)$ erhalten. Dies liefert

$$f = (q_1 + q_2 + \cdots + q_k)g + f_k$$

Dies ist die gesuchte Darstellung von f, wenn wir $q := q_1 + q_2 + \cdots + q_k$ und $r := f_k$ setzen. $\square$

2.4.27 Beispiel

$f = 3t^3 + 2t + 1, g = t^2 - 4t \in \mathbb{R}[t]$:

$$(3t^3 + 2t + 1) = (t^2 - 4t)(3t + 12) + (50t + 1)$$

Es ist $q = 3t + 12$ und $r = 50t + 1$.

2.4.28 Beispiel: Nullstellen

Das Polynom $t^2 + 1 \in \mathbb{R}[t]$ hat keine Nullstelle in den reellen Zahlen, da $f(\lambda) \geq 1$ für alle $\lambda \in \mathbb{R}$ gilt.

Ist $K = \{a_0, a_1, \ldots, a_n\}$ ein endlicher Körper, so hat das Polynom

$$f := (t - a_0) \cdots (t - a_n) + 1$$

keine Nullstelle in K. Für jedes $\lambda \in K$ gilt $f(\lambda) = 1$.

2.4.29 Lemma

Ist $\lambda \in K$ eine Nullstelle von $f \in K[t]$, so gibt es ein eindeutig bestimmtes $g \in K[t]$ mit $f = (t - \lambda)g$. Es gilt $\deg(g) = \deg(f) - 1$.

Beweis:

Division mit Rest von f durch $t-\lambda$ liefert eindeutige $g, r \in K[t]$ mit

$$f = (t-\lambda)g + r$$

Hierbei ist r eine Konstante, da $\deg(r) < \deg(t-\lambda) = 1$ gilt. Einsetzen von λ für t liefert $r = 0$. $\square$

2.4.30 Folgerung

Ist $0 \neq f \in K[t]$ ein Polynom über einem Körper, so gilt für die Anzahl der Nullstellen k von f:

$$k \leq \deg(f)\,.$$

Beweis:

Induktion nach $\deg(f)$. Ist $\deg(f) = 0$, so ist f konstant ungleich Null. Somit ist nichts zu zeigen.

Sei $\deg(f) = n+1 > 0$ und die Aussage sei für alle Polynome vom Grad n bereits bewiesen. Ist dann λ eine Nullstelle von f, so gibt es ein Grad n Polynom $g \in K[t]$ mit $f = (t-\lambda)g$. Alle Nullstellen von f ungleich λ sind Nullstellen von g. Somit hat f höchstens $n+1$ Nullstellen. $\square$

2.4.31 Folgerung

Ist K ein unendlicher Körper, so ist die Zuordnung $K[t] \to \mathrm{Abb}(K, K), f \mapsto (x \mapsto f(x))$ injektiv.

Beweis:

Seien f_1, f_2 zwei Polynome mit gleichem Bild. Dann ist $g \coloneqq f_1 - f_2$ ein Polynom welches die Nullfunktion $K \to K$ darstellt. Somit ist g das Nullpolynom, da es andernfalls höchstens $\deg(g)$ Nullstellen hätte. Dies zeigt $f_1 = f_2$. $\square$

2.4.32 Definition

Ist $0 \neq f \in K[t]$ ein Polynom und $\lambda \in K$, so heißt

$$\mu(f, \lambda) \coloneqq \max\{r \in \mathbb{N} \cup \{0\} \colon f = (t-\lambda)^r g \text{ mit } g \in K[t]\}$$

die *Vielfachheit* der Nullstelle λ von f.

2.4.33 Bemerkung

1. Es ist $\mu(f, \lambda) = 0$ g.d.w. $f(\lambda) \neq 0$ gilt.

2. Sind $\lambda_1, \ldots, \lambda_k \in K$ paarweise verschiedene Nullstellen von $f \in K[t]$, so gilt
$$f = (t-\lambda_1)^{\mu(f,\lambda_1)} \cdots (t-\lambda_k)^{\mu(f,\lambda_k)} g$$
wobei $g \in K[t]$ vom Grad $\deg(f) - \mu(f, \lambda_1) - \cdots - \mu(f, \lambda_k)$ ist. Im Fall $\deg(g) = 0$ sagt man, dass f vollständig in Linearfaktoren zerfällt.

2.4.34 Fundamentalsatz der Algebra

In $\mathbb{C}[t]$ zerfällt jedes Polynom vollständig in Linearfaktoren. In $\mathbb{R}[t]$ zerfällt jedes Polynom in lineare und quadratische Faktoren.

2.4.35 Vietascher Wurzelsatz

Sind $\lambda_1, \dots, \lambda_n$ die Nullstellen des Grad n Polynoms $f = t^n + a_{n-1}t^{n-1} + \dots + a_1 t + a_0$, so gilt:

$$\begin{aligned} a_{n-1} &= -\lambda_1 - \dots - \lambda_n \\ a_{n-2} &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \dots + \lambda_1\lambda_n + \dots + \lambda_{n-1}\lambda_n \\ &\vdots \\ a_0 &= (-1)^n \lambda_1 \cdots \lambda_n \,. \end{aligned}$$

(Francois Vieta, Paris, 1540 - 1603)

Beweis:

Dies zeigt man durch ausmultiplizieren von $(t - \lambda_1)\cdots(t - \lambda_n)$. □

2.4.36 Beispiel

$$t^2 - 5t + 6 = (t-2)(t-3)$$

2.4.37 Satz

Es sei $0 \neq a_0 + a_1 t + \dots + a_n t^n = f \in \mathbb{Q}[t]$ ein Polynom vom Grad n mit ganzen Koeffizienten sowie $\frac{p}{q}$ mit p, q teilerfremd in $\mathbb{Z}$ eine Nullstelle von f. Dann gilt

$$p | a_0 \text{ und } q | a_n \,.$$

Beweis:

Es ist $0 = a_0 + a_1 \frac{p}{q} + \dots + a_n \frac{p^n}{q^n}$. Hieraus folgt

$$0 = a_0 q^n + a_1 p q^{n-1} + a_2 p^2 q^{n-2} + \dots + a_{n-1} p^{n-1} q + a_n p^n \,.$$

Somit ist $a_0 q^n$ durch p teilbar und $a_n p^n$ durch q teilbar. Da p und q teilerfremd sind, folgt $p | a_0$ und $q | a_n$. □

2.5 Vektorräume

2.5.1 Definition

Ist K ein Körper, so nennt man eine Menge V mit den Verknüpfungen

$$+ \colon V \times V \to V, (v, w) \mapsto v + w$$

(genannt Addition) und

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

(Multiplikation mit einem Skalar) einen K-Vektorraum, wenn folgendes gilt:

1. V ist zusammen mit der Addition eine abelsche Gruppe. Der Nullvektor 0 ist das neutrale Element, das inverse Element zu v wird mit $-v$ bezeichnet.

2. Die Multiplikation mit einem Skalar erfüllt:

$$\begin{aligned} (\lambda+\mu)\cdot v &= \lambda\cdot v+\mu\cdot v \\ \lambda\cdot(v+w) &= \lambda\cdot v+\lambda\cdot w \\ (\lambda\mu)\cdot v &= \lambda\cdot(\mu\cdot v), \\ 1\cdot v &= v \end{aligned}$$

   für alle $\lambda,\mu\in K$ und $v,w\in V$.

Die obigen definierenden Eigenschaften eines Vektorraums nennt man auch *Vektorraumaxiome*.

2.5.2 Beispiele

1. Der Standardraum $K^n=\{(x_1,\dots,x_n)\colon x_1,\dots,x_n\in K\}$ ist ein Vektorraum über K. Die Verknüpfungen sind durch

$$\begin{aligned} (x_1,\dots,x_n)+(y_1,\dots,y_n) &\coloneqq (x_1+y_1,\dots,x_n+y_n) \\ \lambda\cdot(x_1,\dots,x_n) &\coloneqq (\lambda x_1,\dots,\lambda x_n) \end{aligned}$$

   gegeben.

2. Der Körper $\mathbb{C}$ der komplexen Zahlen wird zu einem $\mathbb{R}$-Vektorraum, indem man die gewöhnliche Addition und die Multiplikation

$$\mathbb{R}\times\mathbb{C}\to\mathbb{C},(\lambda,a+ib)\mapsto\lambda a+i\lambda b$$

   verwendet.

3. Betrachtet man auf dem Polynomring $K[t]$ die gewöhnliche Addition und die Multiplikation mit Konstanten

$$K\times K[t]\mapsto K[t],(\lambda,f)\mapsto\lambda f$$

   die man durch

$$\lambda\cdot(a_0+a_1t+\cdots+a_nt^n)\coloneqq\lambda a_0+\lambda a_1t+\cdots+\lambda a_nt^n$$

   erklären kann, so wird $K[t]$ zu einem Vektorraum über K.

4. Ist M eine beliebige Menge und K ein Körper, so ist

$$V\coloneqq\mathrm{Abb}(M,K)=\{f\colon M\to K\}$$

   ein K-Vektorraum, wenn wir für $f,g\in V$ und $\lambda\in K$

$$(f+g)(x)\coloneqq f(x)+g(x)\text{ und }(\lambda\cdot f)(x)\coloneqq\lambda f(x)$$

   eine Addition und eine Skalarmultiplikation definieren. Im Spezialfall $M=\{1,\dots,n\}$ erhalten wir so den Standardraum K^n.

In allen Fällen kann man leicht nachrechnen, dass die in der Definition von Vektorraum genannten Eigenschaften erfüllt sind.

2.5.3 Bemerkung

In jedem K-Vektorraum V gelten die folgenden Rechenregeln:

1. $0 \cdot v = 0$.
2. $\lambda \cdot 0 = 0$.
3. Aus $\lambda \cdot v = 0$ folgt $\lambda = 0$ oder $v = 0$.
4. $(-1) \cdot v = -v$.

Beweis:

Für den Beweis haben wir neben den definierenden Eigenschaften von Körpern und Vektorräumen bisher nur die Kürzungsregel zur Verfügung.

1. $0+0\cdot v = 0\cdot v = (0+0)\cdot v = 0\cdot v+0\cdot v$. Die Kürzungsregel liefert $0 = 0\cdot v$.
2. $0+\lambda\cdot 0 = \lambda\cdot 0 = \lambda\cdot(0+0) = \lambda\cdot 0+\lambda\cdot 0$. Die Kürzungsregel liefert $0 = \lambda\cdot 0$.
3. Angenommen es gelte $\lambda \cdot v = 0$ und $\lambda \neq 0$. Dann folgt
$$v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1}(\lambda \cdot v) = \lambda^{-1} \cdot 0 = 0\,.$$
4. $$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = 0\,.$$

□

2.5.4 Definition

Sei V ein K-Vektorraum. Eine Teilmenge $W \subset V$ heißt *Untervektorraum* von V, falls folgendes gilt:

1. $W \neq \emptyset$.
2. $v, w \in W \Longrightarrow v + w \in W$. ($W$ ist abgeschlossen unter Addition.)
3. $v \in W, \lambda \in K \Longrightarrow \lambda \cdot v \in W$. ($W$ ist abgeschlossen unter Multiplikation mit Skalaren.)

2.5.5 Beispiele

1. In $V = \mathbb{R}^2$ betrachten wir die Teilmengen
$$\begin{aligned}&W_1 = \{0\},\ W_2 = \{(x,y) \in \mathbb{R}^2 \mid ax + by = c\},\\ &W_3 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\},\ W_4 = \{(x,y) \in \mathbb{R}^2 \mid x, y \geq 0\},\\ &W_5 = \{(x,y) \in \mathbb{R}^2 \mid xy \geq 0\}.\end{aligned}$$

   Man prüft leicht, dass W_3 und W_4 nicht abgeschlossen unter Skalarmultiplikation sind. W_5 ist nicht abgeschlossen unter der Addition. W_1 ist ein Unterraum, W_2 ist nur im Fall $c = 0$ ein Unterraum.

2. Im Vektorraum $V = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ hat man die Unterraumkette

$$\begin{aligned}
&\{\text{Polynomfunktionen } \mathbb{R} \to \mathbb{R} \text{ vom Grad höchstens } d\} \\
&\subset \{\text{alle Polynomfunktionen } \mathbb{R} \to \mathbb{R}\} \\
&\subset \{\text{differenzierbare Funktionen } \mathbb{R} \to \mathbb{R}\} \\
&\subset \{\text{stetige Funktionen } \mathbb{R} \to \mathbb{R}\} \\
&\subset \mathrm{Abb}(\mathbb{R}, \mathbb{R})\,.
\end{aligned}$$

2.5.6 Satz

Jeder Untervektorraum $W \subset V$ ist ein Vektorraum, wenn man Addition und Skalarmultiplikation von V verwendet.

Beweis:

Alle Rechenregel (Kommutativ- und Assoziativgesetz) gelten für W, da sie für alle Elemente von V gelten.

Da W nicht leer ist, existiert ein $w \in W$. Dann folgt $0 = 0 \cdot w \in W$. Somit liegt der Nullvektor in W. Zudem ist mit $w \in W$ auch $-w = (-1) \cdot w \in W$. $\square$

2.5.7 Lemma

Ist V ein Vektorraum, I eine Indexmenge und $W_i, i \in I$ eine Familie von Untervektorräumen, so ist der Durchschnitt

$$W := \bigcap_{i \in I} W_i$$

wieder ein Untervektorraum.

Beweis:

Da 0 in allen W_i enthalten ist, ist auch $0 \in W$, also $\emptyset \neq W$. Sind $v, w \in W$, so sind v, w in allen W_i enthalten. Dann liegt auch $v + w$ in allen W_i und es folgt $v + w \in W$. Ist $v \in W$ und $\lambda \in K$, so ist v in allen W_i und somit auch λv in allen W_i enthalten. Es folgt $\lambda v \in W$. $\square$

2.5.8 Bemerkung

Seien $W_1, W_2 \subset V$ zwei Untervektorräume, sodass $W_1 \cup W_2$ ebenfalls ein Untervektorraum ist. Dann gilt $W_1 \subset W_2$ oder $W_2 \subset W_1$.

Beweis:

Wir zeigen, dass die Annahme $W_1 \not\subseteq W_2$ und $W_2 \not\subseteq W_1$ auf einen Widerspruch führt.

Gilt die Annahme, so gibt es $w_1 \in W_1 \setminus W_2$ und $w_2 \in W_2 \setminus W_1$. Mit diesen gilt $w_1, w_2, w_1 + w_2 \in W_1 \cup W_2$.

Aus $w_1 + w_2 \in W_1$ folgt $w_2 = (w_1 + w_2) - w_1 \in W_1$. In gleicher Art folgt aus $w_1 + w_2 \in W_2$ die Aussage $w_1 = (w_1 + w_2) - w_2 \in W_2$. Zusammen ist dies ein Widerspruch. Somit gilt wenigstens eine der Inklusionen $W_1 \subset W_2$ oder $W_2 \subset W_1$. $\square$

2.5.9 Definition

Sei $M \subset V$ eine Teilmenge eines Vektorraums über K. Sind $v_1, \ldots, v_r \in M$ endlich viele Vektoren und $\lambda_1, \ldots, \lambda_r \in K$ endlich viele Skalare, so nennt man

$$v = \lambda_1 v_1 + \cdots + \lambda_r v_r$$

eine *Linearkombination* von Elementen aus M. Mit $\operatorname{span}_K(M)$ oder $\operatorname{span}(M)$ bezeichnet man die Menge aller Linearkombinationen von Elementen aus M. D.h.:

$$\operatorname{span}(M) = \{\lambda_1 v_1 + \cdots + \lambda_r v_r : r \in \mathbb{N}, \lambda_1, \ldots, \lambda_r \in K, v_1, \ldots, v_r \in M\}$$

Im Fall $M = \emptyset$ setzt man formal $\operatorname{span}(\emptyset) = \{0\} \subset V$.

Ist $M = \{v_1, \ldots, v_r\}$ endlich, so schreibt man auch

$$Kv_1 + \cdots + Kv_r \text{ oder } \operatorname{span}(v_1, \ldots, v_r)$$

für $\operatorname{span}(M)$.

2.5.10 Satz

Ist $M \subset V$ eine Teilmenge eines Vektorraums, so gilt: $\operatorname{span}(M) \subset V$ ist ein Untervektorraum. Ist weiterhin $W \subset V$ ein Unterraum mit $M \subset W$, so gilt auch $\operatorname{span}(M) \subset W$.

Kurz: $\operatorname{span}(M)$ ist der kleinste Unterraum von V der M enthält. Man nennt daher $\operatorname{span}(M)$ auch den von M aufgespannten (oder erzeugten) Unterraum.

Beweis:

$\operatorname{span}(M)$ ist ein Unterraum, da Summen und skalare Vielfache von Linearkombinationen von Elementen aus M wieder Linearkombinationen von Elementen aus M sind. Genauer: Sind $\lambda_1 x_1 + \cdots + \lambda_m x_m$ und $\mu_1 y_1 + \cdots + \mu_n y_n$ zwei Linearkombinationen von Elementen aus M und $\lambda \in K$, so sind auch

$$\begin{aligned}
&\lambda_1 x_1 + \cdots + \lambda_m x_m + \mu_1 y_1 + \cdots + \mu_n y_n\,, \\
&\lambda\,(\lambda_1 x_1 + \cdots + \lambda_m x_m) = (\lambda\lambda_1)x_1 + \cdots + (\lambda\lambda_m)x_m
\end{aligned}$$

Linearkombinationen von Elementen aus M.

Sei nun $v := \lambda_1 v_1 + \cdots + \lambda_r v_r$ eine Linearkombination von Elementen aus M. Sind alle $v_i \in W$, so ist auch $v \in W$, da W als Unterraum abgeschlossen unter Addition und skalarer Multiplikation ist. $\square$

2.5.11 Beispiele

1. Ist $V = \mathbb{R}^3$ und $v \in V$ nicht der Nullvektor. Dann ist $\operatorname{span}(v)$ die Gerade durch 0 und v.

2. Sind $v_1, v_2 \in V = \mathbb{R}^3$, $v_1 \neq 0$ und $v_2 \notin \operatorname{span}(v_1)$, so ist $\operatorname{span}(v_1, v_2)$ die Ebene durch 0, v_1 und v_2.

3. In K^n setzen wir mit $I = \{1, \dots, n\}$ für jedes $i \in I$
$$e_i \coloneqq (0, \dots, 0, 1, 0, \dots, 0)\,,$$
wobei die 1 an der i-ten Stelle steht. Dann gilt
$$K^n = \operatorname{span}(e_1, \dots, e_n) = \operatorname{span}\left((e_i)_{i\in I}\right)\,.$$

4. Ist $V = K[t]$ der Polynomring, und $v_n = t^n$, so ist
$$K[t] = \operatorname{span}\left((v_n)_{n\in\mathbb{N}\cup\{0\}}\right)\,.$$

2.5.12 Beispiel

Der $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ kann auf viele Arten erzeugt werden.
$$\mathbb{R}^2 = \operatorname{span}(e_1, e_2) = \operatorname{span}\{(x, y)\colon x, y \in (0, 1)\} = \operatorname{span}((1, 0), (0, 1), (1, 1))$$
Es gilt $(x, y) = t \cdot (1, 1) + (x - t) \cdot (1, 0) + (y - t) \cdot (0, 1)$ für jedes $t \in \mathbb{R}$. D.h. in diesem Fall kann jeder Vektor auf unendlich viele Arten als Linearkombination dargestellt werden.

2.5.13 Definition

Es sei $(v_1, \dots, v_r)$ eine endliche Familie von Vektoren des K-Vektorraums V. $(v_1, \dots, v_r)$ heißen *linear unabhängig*, wenn gilt: Sind $\lambda_1, \dots, \lambda_r \in K$ mit
$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0$$
gegeben, so folgt $\lambda_1 = \lambda_2 = \dots = \lambda_r = 0$. In anderen Worten: Nur die triviale Linearkombination von $v_1, \dots, v_r$ ergibt den Nullvektor. Eine unendliche Familie von Vektoren $(v_i)_{i\in I}$ heißt linear unabhängig, wenn jede endliche Teilfamilie linear unabhängig ist. Ist eine Familie nicht linear unabhängig so heißt sie *linear abhängig*. Man sagt oft „$v_1, \dots, v_n \in V$ ist linear (un-)abhängig“.

2.5.14 Lemma

Für eine Familie von Vektoren $(v_i)_{i\in I}$ eines K-Vektorraums sind äquivalent:

1. (v_i) ist linear unabhängig.
2. Jeder Vektor $v \in \operatorname{span}(v_i)$ lässt sich in eindeutiger Weise aus Vektoren der Familie (v_i) linear kombinieren.

Beweis:

2) $\Longrightarrow$ 1) ist klar, da sich der Nullvektor nur als triviale Linearkombination darstellen lässt.

Zu 1) $\Longrightarrow$ 2): Sei ein $v \in \operatorname{span}(v_i)$ gegeben. Angenommen dieses hätte zwei verschiedene Darstellungen als Linearkombination, also
$$v = \sum_{i\in I} \lambda_i v_i = \sum \mu_i v_i,$$

wobei in jeder der beiden Summen nur endlich viele Skalare λ_i, μ_i ungleich Null sind. Für mindestens ein i gilt $\lambda_i \neq \mu_i$. Dann folgt

$$0 = \sum_{i \in I, \lambda_i \neq 0 \text{ oder } \mu_i \neq 0} (\lambda_i - \mu_i) v_i .$$

Da es ein ein i mit $\lambda_i \neq \mu_i$ gibt, ist dies eine nicht-triviale Linearkombination, welche den Nullvektor darstellt. Dies ist ein Widerspruch zur linearen Unabhängigkeit. Damit ist die Eindeutigkeit bewiesen. □

2.5.15 Beispiele

1. In $\mathbb{R}^3$ sind $(1,2,3), (0,-1,2)$ linear unabhängig.
2. Die Vektoren $e_1, \ldots, e_n$ sind linear unabhängig in K^n.
3. Im Polynomring $K[t]$ (aufgefasst als K-Vektorraum) ist die Familie $(t^n)_{n \in \mathbb{N} \cup \{0\}}$ linear unabhängig.

2.5.16 Bemerkung

In jedem K-Vektorraum V gilt:

1. Ein einzelner Vektor $v \in V$ ist genau dann linear unabhängig, wenn $v \neq 0$ gilt.
2. Enthält eine Familie von Vektoren den Nullvektor, so ist sie linear abhängig.
3. Ist ein Vektor mehrfach in einer Familie enthalten, so ist diese linear abhängig.
4. Die Vektoren $v_1, \ldots, v_r$ mit $r \geq 2$ sind genau dann linear abhängig, wenn einer von ihnen eine Linearkombination der anderen ist.

Beweis:

1. Ist v linear abhängig, so gibt es ein $\lambda \neq 0$ mit $\lambda v = 0$. Hieraus folgt $v = 0$. Umgekehrt gilt $1 \cdot 0 = 0$, somit ist der Nullvektor linear abhängig.
2. $1 \cdot 0 = 0$
3. Angenommen es ist (v_i) eine Familie mit $v_1 = v_2$, dann gilt $1 \cdot v_1 + (-1) \cdot v_2 = 0$. Somit ist die Familie linear abhängig.
4. Sind die Vektoren $v_1, \ldots, v_r$ linear abhängig so gibt es $\lambda_1, \ldots, \lambda_r \in K$ mit $\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$ und es gibt mindestens ein k mit $\lambda_k \neq 0$. Für jedes solche k gilt

   $$v_k = \frac{-1}{\lambda_k} (\lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_{k+1} v_{k+1} + \cdots + \lambda_r v_r) .$$

   Ist Umgekehrt $v_k = \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \cdots + \mu_r v_r$, so gilt

   $$0 = \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + (-1) v_k + \mu_{k+1} v_{k+1} + \cdots + \mu_r v_r .$$

□

2.6 Basis und Dimension

2.6.1 Definition

Eine Familie $(v_i)_{i\in I}$ in einem Vektorraum V heißt *Erzeugendensystem* von V, wenn $V = \operatorname{span}(v_i)_{i\in I}$ gilt. Ein Erzeugendensystem heißt *Basis*, wenn es linear unabhängig ist. V heißt endlich erzeugt, wenn V ein endliches Erzeugendensystem hat. In diesem Fall gilt $V = \operatorname{span}(v_1, \dots, v_n)$.

2.6.2 Beispiele

1. Die leere Familie ist eine Basis des Nullraums.
2. $S := (e_1, \dots, e_n)$ ist eine Basis des K^n. Sie heißt die Standardbasis.
3. $(1, i)$ ist eine Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$.
4. $(1, t, t^2, t^3, \dots)$ ist eine unendliche Basis des K-Vektorraums aller Polynome $K[t]$.

2.6.3 Satz

Es sein $\mathcal{B} = (v_1, \dots, v_n)$ eine Familie von Vektoren eines K-Vektorraums $V \neq \{0\}$. Dann sind äquivalent:

1. $\mathcal{B}$ ist eine Basis. D.h. ein linear unabhängiges Erzeugendensystem.
2. $\mathcal{B}$ ist ein minimales Erzeugendensystem. D.h. $(v_1, \dots, v_{r-1}, v_{r+1}, \dots, v_n)$ ist bei jeder Wahl von $r \in \{1, \dots, n\}$ kein Erzeugendensystem mehr.
3. Zu jedem $v \in V$ gibt es eindeutig bestimmte $\lambda_1, \dots, \lambda_n \in K$ mit
$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$
d.h. jeder Vektor kann auf nur eine Art als Linearkombination geschrieben werden.
4. $\mathcal{B}$ ist eine maximale linear unabhängige Familie von Vektoren. D.h. für jedes $v \in V$ ist $(v_1, \dots, v_n, v)$ linear abhängig.

Beweis:

$1 \implies 2$: Sei $v_1, \dots, v_n$ ein nicht minimales Erzeugendensystem. Sei ohne Einschränkung $r = 1$, d.h. $v_2, \dots, v_n$ ist ein Erzeugendensystem vom V. Dann gilt $v_1 = \lambda_2 v_2 + \cdots \lambda_n v_n$. Hieraus folgt
$$0 = (-1)v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \,.$$
Also ist $v_1, \dots, v_n$ linear abhängig.

$2 \implies 3$: Angenommen es kann ein Vektor $v \in V$ auf zwei verschiedene Arten als Linearkombination von $v_1, \dots, v_n$ geschrieben werden, dann gilt
$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n \,.$$

Ohne Einschränkung gelte $\lambda_1 \neq \mu_1$. Dann folgt

$$v_1 = \frac{1}{\mu_1 - \lambda_1}((\lambda_2 - \mu_2)v_2 + \cdots + (\lambda_n - \mu_n)v_n)\,.$$

Somit war das Erzeugendensystem nicht minimal.

$3 \Longrightarrow 4$: Sei $v_1, \ldots, v_n$ ein Erzeugendensystem, sodass jeder Vektor auf nur eine Art als Linearkombination geschrieben werden kann. Dann ist $v_1, \ldots, v_n$ linear unabhängig, da der Nullvektor nur als triviale Linearkombination dargestellt werden kann. Sei nun $v \in V$ beliebig. Dann gibt es $\lambda_1, \ldots, \lambda_n \in K$ mit

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n\,.$$

Hieraus folgt

$$0 = (-1)v + \lambda_1 v_1 + \cdots + \lambda_n v_n\,.$$

Somit ist $(v_1, \ldots, v_n, v)$ linear abhängig.

$4 \Longrightarrow 1$: Sei $(v_1, \ldots, v_n)$ eine maximale linear unabhängige Familie von Vektoren in V. Es ist zu zeigen, dass sie ein Erzeugendensystem ist. Sei hierzu $v \in V$ beliebig. Dann ist $(v_1, \ldots, v_n, v)$ linear abhängig. Es gibt daher $\lambda_1, \ldots, \lambda_n, \lambda \in K$ mit

$$0 = \lambda_1 v_1 + \cdots + \lambda_n v_n + \lambda v\,,$$

wobei nicht alle skalare Null sind. Es ist nun $\lambda \neq 0$, da sonst $(v_1, \ldots, v_n)$ linear abhängig wäre. Somit folgt

$$v = \frac{-\lambda_1}{\lambda} v_1 + \cdots + \frac{-\lambda_n}{\lambda} v_n\,.$$

Daher ist $(v_1, \ldots, v_n)$ ein Erzeugendensystem. □

2.6.4 Folgerung

Ist V ein nicht endlich erzeugter Vektorraum, so gibt es eine unendliche linear unabhängige Familie von Vektoren.

Beweis:

Angenommen es wäre $(v_1, \ldots, v_n)$ eine maximale linear unabhängige Familie von Vektoren, so wäre sie nach obigem Satz ein Erzeugendensystem. Somit wäre V endlich erzeugt. □

2.6.5 Satz

Ist ein endliches Erzeugendensystem des Vektorraums V gegeben, so hat er eine endliche Basis.

Beweis:

Falls das Erzeugendensystem minimal ist, so haben wir eine Basis. Falls nicht, so kann ein Element im Erzeugendensystem gestrichen werden. Nach endlich vielen Schritten erhalten wir eine Basis, da wir mit einer endlichen Familie begonnen haben. □

2.6.6 Bemerkung

Allgemein gilt, dass jeder Vektorraum eine Basis hat. Für Vektorräume ohne endlichem Erzeugendensystem beweisen wir dies nicht in dieser Vorlesung.

2.6.7 Lemma

Gegeben sei ein K-Vektorraum V mit Basis $\mathcal{B} = (v_1, \ldots, v_r)$ und $w = \lambda_1 v_1 + \cdots + \lambda_r v_r \in V$. Ist $k \in \{1, \ldots, r\}$ mit $\lambda_k \neq 0$, so ist

$$\mathcal{B}' := (v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_r)$$

wieder eine Basis von V. D.h. man kann v_k durch w austauschen.

Beweis:

Wir betrachten den Fall $k = 1$, der allgemeine Fall folgt dann durch Umnummerierung. Es ist zu zeigen, dass $(w, v_2, v_3, \ldots, v_r)$ eine Basis von V ist. Ist $v \in V$ beliebig, so existieren $\mu_1, \ldots, \mu_r \in K$ mit $v = \mu_1 v_1 + \cdots + \mu_r v_r$. Dann gilt

$$v = \frac{\mu_1}{\lambda_1} w + \left(\mu_2 - \frac{\lambda_2 \mu_1}{\lambda_1}\right) v_2 + \cdots + \left(\mu_r - \frac{\lambda_r \mu_1}{\lambda_1}\right) v_r$$

womit gezeigt ist, dass $(w, v_2, v_3, \ldots, v_r)$ ein Erzeugendensystem ist.

Zur linearen Unabhängigkeit: Angenommen es ist $\mu w + \mu_2 v_2 + \cdots + \mu_r v_r = 0$. Dann folgt

$$\mu \lambda_1 v_1 + (\mu \lambda_2 + \mu_2) v_2 + \cdots + (\mu \lambda_r + \mu_r) v_r = 0\,.$$

Da $(v_1, \ldots, v_r)$ linear unabhängig ist, folgt

$$\mu \lambda_1 = 0, (\mu \lambda_2 + \mu_2) = 0, \ldots, (\mu \lambda_r + \mu_r) = 0\,.$$

Da $\lambda_1 \neq 0$ gilt folgt $\mu = 0$ und somit $\mu_2 = \mu_3 = \cdots = \mu_r = 0$. □

2.6.8 Austauschsatz von Steinitz

Ist V ein K-Vektorraum mit Basis $\mathcal{B} = (v_1, \ldots, v_r)$ und $(w_1, \ldots, w_n)$ linear unabhängig. Dann gilt $n \leq r$. Weiterhin kann $(v_1, \ldots, v_r)$ durch Umnummerierung so geordnet werden, dass

$$(w_1, \ldots, w_n, v_{n+1}, \ldots, v_r)$$

eine Basis von V ist.

Beweis:

Induktion nach n. Im Fall $n = 0$ ist nichts zu zeigen. Sei also $n \geq 1$ und der Satz für $n - 1$ bereits bewiesen.

Da $(w_1, \ldots, w_{n-1})$ linear unabhängig ist, ergibt die Induktionsannahme, dass $n - 1 \leq r$ gilt und $(w_1, \ldots, w_{n-1}, v_n, \ldots, v_r)$ eine Basis von V ist. Angenommen es wäre $r = n - 1$, so wäre $(w_1, \ldots, w_{n-1})$ bereits eine Basis und somit ein maximales System linear unabhängiger Vektoren. Dies widerspricht der linearen Unabhängigkeit von $(w_1, \ldots, w_n)$. Somit ist $n \leq r$ bewiesen.

Wir schreiben w_n als Linearkombination dieser Basis:

$$w_n = \lambda_1 w_1 + \cdots + \lambda_{n-1} w_{n-1} + \lambda_n v_n + \cdots + \lambda_r v_r$$

mit $\lambda_1, \ldots, \lambda_r \in K$. Wären $\lambda_n = \cdots = \lambda_r = 0$, so hätten wir einen Widerspruch zur linearen Unabhängigkeit von $(w_1, \ldots, w_n)$. Nach Umnummerierung von $v_n, \ldots, v_r$ können wir $\lambda_n \neq 0$ voraussetzen. Nun liefert das Lemma, dass wir v_n durch w_n austauschen können und die Basis $(w_1, \ldots, w_n, v_{n+1}, \ldots, v_r)$ erhalten. $\square$

2.6.9 Folgerung

Hat ein K-Vektorraum V eine endliche Basis, so ist jede Basis von V endlich.

Beweis:

Ist $(v_1, \ldots, v_r)$ eine Basis, so liefert der Austauschsatz, dass es kein System linear unabhängiger Vektoren mit mehr als r Elementen gibt. Insbesondere hat jede Basis höchstens r Elemente. $\square$

2.6.10 Folgerung

Hat ein K-Vektorraum V eine endliche Basis, so hat jede Basis von V gleich viele Elemente.

Beweis:

Sei eine endliche Basis mit r Elementen und eine zweite Basis $\mathcal{B}$ gegeben. Dann hat $\mathcal{B}$ nach obiger Folgerung höchstens r Elemente. Sei s die Anzahl der Elemente in $\mathcal{B}$. Erneutes Anwenden der Folgerung liefert, dass die erste Basis höchstens s Elemente hat. Es gilt somit $s \leq r$ und $r \leq s$. Zusammen erhalten wir $r = s$. $\square$

2.6.11 Definition

Ist V ein K-Vektorraum, so definieren wir

$$\dim_K(V) := \begin{cases} \infty & \text{falls } V \text{ keine endliche Basis hat,} \\ r & \text{falls } V \text{ eine Basis mit } r \text{ Elementen besitzt} \end{cases}$$

$\dim_K(V)$ heißt die *Dimension* von V über K. Man schreibt auch $\dim(V)$.

2.6.12 Folgerung

Ist $W \subset V$ ein Unterraum eines endlich erzeugten K-Vektorraums V, so ist auch W endlich erzeugt und es gilt $\dim W \leq \dim V$. Aus $\dim V = \dim W$ folgt $V = W$.

Beweis:

Wäre W nicht endlich erzeugt, so gäbe es in W eine unendliche linear unabhängige Familie von Vektoren. Dann wäre dies aber auch eine unendliche linear unabhängig Familie von Vektoren in V. Dies ist ein Widerspruch dazu, dass V endlich erzeugt ist. Insbesondere zeigt dies auch $\dim W \leq \dim V$.

Angenommen es ist $n = \dim V = \dim W$ und $w_1, \ldots, w_n$ eine Basis von W. Nehmen wir zusätzlich $V \neq W$ an, so gibt es ein $v \in V \setminus W$. Es ist dann $v, w_1, \ldots, w_n$ linear unabhängig. Dies ist ein Widerspruch. $\square$

2.6.13 Basisergänzungssatz

Ist V ein endlich erzeugter Vektorraum und $w_1, \ldots, w_n \in V$ linear unabhängige Vektoren. Dann gibt es $w_{n+1}, \ldots, w_r$ so, dass

$$(w_1, \ldots, w_n, w_{n+1}, \ldots, w_r)$$

eine Basis von V ist.

Beweis:

Sei zunächst $(v_1, \ldots, v_m)$ ein Erzeugendensystem vom V. Dann kann man hieraus eine Basis Auswählen. O.E. sei diese $(v_1, \ldots, v_r)$. Nach dem Basisaustauschsatz kann nun $(w_1, \ldots, w_n)$ durch $w_{n+1}, \ldots, w_r \in \{v_1, \ldots, v_r\}$ zu einer Basis $(w_1, \ldots, w_r)$ ergänzt werden. $\square$

2.6.14 Beispiele

1. Es gilt $\dim K^n = n$, denn $e_1, \ldots, e_n$ ist eine Basis. Jede Basis von K^n hat genau n Elemente.

2. Eine Gerade (oder Ebene) durch den Nullpunkt in $\mathbb{R}^n$ ist ein Unterraum der Dimension 1 (oder 2).

3. $\dim_K K[t] = \infty$.

4. $\dim_{\mathbb{R}} \mathbb{C} = 2$ mit Basis $1, i$. $\dim_{\mathbb{C}} \mathbb{C} = 1$.

5. $\dim_{\mathbb{Q}} \mathbb{R} = \infty$.

2.6.15 Definition

Es sei K ein Körper. Eine *$m \times n$-Matrix über K* ist eine rechteckige Anordnung

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

von Elementen aus K. Wir Bezeichnen mit $\mathrm{Mat}(m \times n, K)$ die Menge aller $m \times n$-Matrizen über K. Weiterhin heißen $(a_{11}, \ldots, a_{1n}), \ldots, (a_{m1}, \ldots, a_{mn}) \in K^n$ die *Zeilen* von A.

Sind $a_1, \ldots, a_m \in K^n$ die Zeilen der Matrix A, so heißt $\mathrm{span}(a_1, \ldots, a_m)$ der *Zeilenraum* von A. Die Dimension des Zeilenraums einer Matrix heißt auch *Zeilenrang*.

2.6.16 Beispiel

Der Zeilenraum der Einheitsmatrix $E_n \in \mathrm{Mat}(n \times n, K)$ ist der K^n, da die Zeilen die Standardbasis $e_1, \dots, e_n$ von K^n bilden. Man schreibt auch $E_n = (\delta_{ij})_{i,j=1,\dots,n}$, hierbei ist δ_{ij} das Kronecker-Delta.

2.6.17 Definition

Die folgenden Manipulationen einer Matrix

1. Skalierung einer Zeile mit $\lambda \in K^\times$
2. Addition des λ-fachen der i-ten Zeile zur j-ten Zeile ($i \neq j$)
3. Vertauschung der i-ten und der j-ten Zeile

werden als *elementare Zeilenoperationen (oder Zeilenumformungen)* bezeichnet. Man sagt, eine Matrix B entsteht aus der Matrix A durch Zeilenoperationen, wenn sich A durch endlich viele elementare Zeilenoperationen in B überführen lässt.

2.6.18 Satz

Elementare Zeilenumformungen ändern den Zeilenraum einer Matrix nicht.

Beweis:

Wir prüfen die drei Fälle. Zum dritten Fall: Es ist offensichtlich, dass Vertauschung der Vektoren den Spann nicht ändert.

Zum ersten Fall: Seien $a_1, \dots, a_n$ die Zeilen der Matrix A und $\mu_1, \dots, \mu_n \in K^\times$ beliebig. Dann gilt

$$\begin{aligned}\mathrm{span}(a_1, \dots, a_n) &= \{\lambda_1 a_1 + \cdots + \lambda_n a_n : \lambda_1, \dots, \lambda_n \in K\} \\ &= \left\{\frac{\lambda_1}{\mu_1}\mu_1 a_1 + \cdots + \frac{\lambda_n}{\mu_n}\mu_n a_n : \lambda_1, \dots, \lambda_n \in K\right\} \\ &= \{\nu_1\mu_1 a_1 + \cdots + \nu_n \mu_n a_n : \nu_1, \dots, \nu_n \in K\} \\ &= \mathrm{span}(\mu_1 a_1, \dots, \mu_n a_n)\,.\end{aligned}$$

Somit ändert die erste Sorte Umformungen den Zeilenraum nicht.

Zum zweiten Fall: Durch Zeilenvertauschung kann hier immer $i = 1$ und $j = 2$ erreicht werden. In diesem Spezialfall entsteht aus der Matrix mit den Zeilen $v_1, \dots, v_n$ die Matrix mit den Zeilen $v_1, \lambda v_1 + v_2, v_3, \dots, v_n$. Nun erhalten wir

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = (\lambda_1 - \lambda\lambda_2)v_1 + \lambda_2(\lambda v_1 + v_2) + \lambda_3 v_3 + \cdots + \lambda_n v_n$$

und

$$\lambda_1 v_1 + \lambda_2(\lambda v_1 + v_2) + \lambda_3 v_3 + \cdots + \lambda_n v_n = (\lambda_1 + \lambda_2\lambda)v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n\,.$$

D.h. der alte Zeilenraum ist im neuen Zeilenraum enthalten und umgekehrt. Somit sind die Zeilenräume gleich. $\square$

2.6.19 Definition

Eine Matrix hat *Zeilenstufenform*, wenn in jeder Zeile der erste von Null verschiedene Eintrag eine 1 ist und die führenden Einsen von Zeile zu Zeile immer weiter nach rechts rutschen. Formal $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ hat Zeilenstufenform, wenn es Indizes $i_1, \ldots, i_k \in \{1, \ldots, n\}$ mit $k \leq m$ und $i_1 < i_2 < \cdots < i_k$ gibt, sodass

1. In der j-ten Zeile $(1 \leq j \leq k)$ ist der Eintrag in Spalte i_j eine 1.
2. Alle Einträge der j-ten Zeile $(1 \leq j \leq k)$ in den Spalten vor i_j sind Null.
3. Alle Zeilen mit Index $> k$ sind Nullzeilen.

gilt.

2.6.20 Satz

Jede Matrix $A \in \mathrm{Mat}(m \times n, K)$ kann durch elementare Zeilenumformungen auf Zeilenstufenform transformiert werden.

Beweis:

Ist A die Nullmatrix, so hat A bereits Zeilenstufenform. Andernfalls beweisen wir die Aussage durch Induktion nach der Anzahl m der Zeilen von A.

Sei $m = 1$: Es sei i_1 der Index des ersten von Null verschiedenen Eintrags. Wir multiplizieren die erste Zeile von A mit dem Kehrwerte des Eintrags an der Position $(1, i_1)$ und erhalten eine Matrix in Zeilenstufenform.

$m - 1 \to m$: Es sei i_1 der kleinste Index einer Spalte von A, die nicht die Nullspalte ist. Sei j ein Zeilenindex, sodass $a_{j,i_1} \neq 0$ gilt. Durch Tauschen der 1-ten mit der j-ten Zeile können wir $a_{1,i_1} \neq 0$ erreichen. Wir multiplizieren nun die 1-te Zeile mit $\frac{1}{a_{1,i_1}}$. Dann subtrahieren wir für jedes $k \in 2, \ldots, m$ das $-a_{k,i_1}$-fache der ersten Zeile von der k-ten Zeile. Hierbei entstehen in den Zeilen 2 bis m der j_1-ten Spalte Nullen.

Nun betrachten wird die Matrix A', die aus der 2-ten bis m-ten Zeile unserer Matrix besteht. Sie hat in den Spalten 1 bis i_1 nur Nullen. Nach Induktionsannahme kann diese Teilmatrix durch Zeilenoperation auf Zeilenstufenform transformiert werden. Hierbei bleiben die Nullen in den Spalten 1 bis i_1 erhalten.

Zusammen entsteht die Zeilenstufenform von A. $\square$

2.6.21 Beispiel

Sei $a_1 = (0, 0, 0, 2, -1)$, $a_2 = (0, 1, -2, 1, 0)$, $a_3 = (0, -1, 2, 1, -1)$, $a_4 = (0, 0, 0, 1, 2)$ in $\mathbb{R}^5$ gegeben. Dann ergibt der Beweis folgende Rechnung:

$$\begin{matrix} 0 & 0 & 0 & 2 & -1 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{matrix} \quad \to \quad \begin{matrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & -1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{matrix} \quad \to \quad \begin{matrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{matrix}$$

$$\to \quad \begin{matrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 & -1 \end{matrix} \quad \to \quad \begin{matrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{matrix} \quad \to \quad \begin{matrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

Hier ist $k = 3$ und $i_1 = 2, i_2 = 4, i_3 = 5$. Wir erhalten $\operatorname{span}(a_1, a_2, a_3, a_4) = \operatorname{span}(b_1, b_2, b_3)$ mit $b_1 = (0, 1, -2, 1, 0)$, $b_2 = (0, 0, 0, 1, 2)$, $b_3 = (0, 0, 0, 0, 1)$.

2.6.22 Bemerkung

Hat die Matrix $A \in \operatorname{Mat}(m \times n, K)$ Zeilenstufenform, so sind die von Null verschiedenen Zeilen linear unabhängig in K^n. Insbesondere bilden sie eine Basis des Zeilenraums.

Beweis:

Seien $i_1, i_2, \ldots, i_k$ die Positionen der führenden Einsen in den Zeilen $a_1, \ldots, a_k$. Aus $\lambda_1 a_1 + \cdots + \lambda_k a_k = 0$ folgt dann durch betrachten der i_1-ten Komponente $\lambda_1 = 0$. Dann folgt durch betrachten der i_2-ten Komponente $\lambda_2 = 0$, usw. $\square$

2.6.23 Definition

Sei $A = (a_{i,j}) \in \operatorname{Mat}(m \times n, K)$ eine Matrix, so ist $A^\top = (b_{ij}) \in \operatorname{Mat}(n \times m, K)$ die Matrix die durch *transponieren* aus A entsteht. Es gilt dann $a_{ij} = b_{ji}$.

2.6.24 Bemerkung

Es gelten

1. $(A + B)^\top = A^\top + B^\top$
2. $(\lambda A)^\top = \lambda A^\top$
3. $(A^\top)^\top = A$

Man bezeichnet den Zeilenraum von $A^\top$ als *Spaltenraum* von A. Man nenne die Dimension des Spaltenraums von A auch den *Spaltenrang* von A.

2.6.25 Ausblick

Wir werden zeigen, dass der Zeilenrang einer Matrix stets gleich ihrem Spaltenrang ist. Man spricht dann nur noch vom Rang.

Kapitel 3

Lineare Abbildungen

3.1 Definition und Beispiele

3.1.1 Definition

Es sei K ein Körper und V, W seien K-Vektorräume. Eine Abbildung $\varphi\colon V \to W$ heißt *linear Abbildung* (oder K-lineare Abbildung oder Homomorphismus von Vektorräumen), wenn gilt:

$$\begin{aligned}&\forall v_1, v_2 \in V: \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)\\&\forall \lambda \in K, v \in V: \varphi(\lambda v) = \lambda\varphi(v)\,.\end{aligned}$$

Zudem bezeichnet man mit

$$\mathrm{Hom}_K(V, W) := \{\varphi\colon V \to W \mid \varphi \text{ ist linear }\}$$

die Menge aller K-linearen Abbildungen $V \to W$. Man schreibt $\mathrm{Hom}(V, W)$, wenn sich K aus dem Kontext ergibt.

3.1.2 Bemerkung

Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen, so gilt $\varphi(0) = 0$.

Beweis:

$$0 + \varphi(0) = \varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0) \Longrightarrow 0 = \varphi(0)\,.$$

Man beachte, dass dieser Beweis nur die Additivität von φ benötigt. □

3.1.3 Bemerkung

Eine Abbildung $\varphi\colon V \to W$ ist genau dann linear, wenn

$$\varphi(\lambda x + y) = \lambda\varphi(x) + \varphi(y)$$

für alle $x, y \in V$, $\lambda \in K$ gilt.

Beweis:

Es ist klar, dass jede linear Abbildung obige Gleichung erfüllt. Somit bleibt die zweite Schlussrichtung zu zeigen.

Im Spezialfall $\lambda = 1$ erhalten wir die Additivität. Hieraus folgt $\varphi(0) = 0$. Somit liefert der Spezialfall $y = 0$ die Aussage $\varphi(\lambda x) = \lambda\varphi(x)$. $\square$

3.1.4 Beispiele

Gegeben seien die folgenden Abbildungen

$$\begin{aligned}
\varphi_1 &\colon \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + y, x - y)\\
\varphi_2 &\colon \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (1 + x, y)\\
\varphi_3 &\colon \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x^2, y^2)
\end{aligned}$$

φ_1 ist linear, φ_2, φ_3 sind nicht linear.

Zu φ_1: Für beliebige $r, s, t, u, \lambda \in \mathbb{R}$ gilt:

$$\begin{aligned}
\varphi_1((r, s) + (t, u)) &= \varphi_1((r + t, s + u)) = (r + s + t + u, r + t - s - u)\\
&= (r + s, r - s) + (t + u, u - u) = \varphi_1(r, s) + \varphi_1(t, u)\\
\varphi_1(\lambda(r, s)) &= \varphi_1(\lambda r, \lambda s) = (\lambda r + \lambda s, \lambda r - \lambda s)\\
&= \lambda(r + s, r - s) = \lambda\varphi(r, s)
\end{aligned}$$

Zu φ_2: Es ist $\varphi_2(0) = (1, 0)$.

Zu φ_3: Es ist $\varphi_3(1, 1) = (1, 1)$, $\varphi_3(2, 2) = (4, 4)$. Somit ist $\varphi_3(2 \cdot (1, 1)) \neq 2 \cdot \varphi_3(1, 1)$.

3.1.5 Beispiel

Sei $A = (a_{ij}) \in \operatorname{Mat}(m \times n, K)$ eine Matrix. Zu A bilden wir die Abbildung

$$F_A\colon K^n \to K^m, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Man schreibt auch $F_A(x) = Ax$ und nennt dies die Multiplikation der Matrix A mit dem Spaltenvektor x. Die Abbildung F_A ist linear, denn für beliebige

$x, y \in K^n, \lambda \in K$ gilt:

$$
\begin{aligned}
F_A(x+y) &= \begin{pmatrix} a_{11}(x_1+y_1)+\cdots+a_{1n}(x_n+y_n) \\ \vdots \\ a_{m1}(x_1+y_1)+\cdots+a_{mn}(x_n+y_n) \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1+\cdots+a_{1n}x_n+a_{11}y_1+\cdots+a_{1n}y_n \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n+a_{m1}y_1+\cdots+a_{mn}y_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1+\cdots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1+\cdots+a_{1n}y_n \\ \vdots \\ a_{m1}y_1+\cdots+a_{mn}y_n \end{pmatrix} \\
&= F_A(x)+F_A(y)
\end{aligned}
$$

$$
\begin{aligned}
F_A(\lambda x) &= \begin{pmatrix} a_{11}(\lambda x_1)+\cdots+a_{1n}(\lambda x_n) \\ \vdots \\ a_{m1}(\lambda x_1)+\cdots+a_{mn}(\lambda x_n) \end{pmatrix} \\
&= \lambda \begin{pmatrix} a_{11}x_1+\cdots+a_{1n}x_n \\ \vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n \end{pmatrix} = \lambda F_A(x)
\end{aligned}
$$

Weiterhin sieht man leicht: $F_A(e_1), \ldots, F_A(e_n)$ sind genau die Spalten der Matrix A.

3.1.6 Beispiele

1. Das Transponieren von Matrizen $\mathrm{Mat}(m \times n, K) \to \mathrm{Mat}(n \times m, K), A \mapsto A^\top$ ist eine lineare Abbildung.
2. Ist $K[t]$ der Ring der Polynome über K und $x_0 \in K$ ein festes Element, so sind die Abbildungen

$$
\begin{aligned}
&K[t] \mapsto K,\ f \mapsto f(x_0) \\
&K[t] \to K[t],\ f \mapsto f' \\
&K[t] \to K[t],\ f(t) \mapsto f(t+1) \\
&K[t] \to K[t],\ f(t) \mapsto f(2t)
\end{aligned}
$$

   linear.
3. Die Drehung mit festem Winkel φ um $(0,0)$ ist als Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ eine lineare Abbildung.
   Übung: Begründen Sie dieses elementargeometrisch.

3.1.7 Definition

Sei $\varphi\colon V \to W$ eine lineare Abbildung.

1. φ heißt *Isomorphismus* wenn φ bijektiv ist.

2. φ heißt *Endomorphismus* wenn $V = W$ gilt.

3. φ heißt *Automorphismus* wenn $V = W$ gilt und φ bijektiv ist.

3.1.8 Bemerkung

Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Dann gilt:

1. $\varphi(v - w) = \varphi(v) - \varphi(w)$
2. $\varphi(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1\varphi(v_1) + \cdots + \lambda_n\varphi(v_n)$.
3. Ist $(v_i)_{i\in I}$ eine linear abhängige Familie von Vektoren, so ist $(\varphi(v_i))_{i\in I}$ ebenfalls linear abhängig.
4. Sind $V' \subset V, W' \subset W$ Unterräume, so sind auch $\varphi(V')$ und $\varphi^{-1}(W')$ Unterräume.
5. $\dim\varphi(V) \leq \dim V$.
6. Ist φ ein Isomorphismus, so ist auch φ^{-1} linear.

Beweis:

1. $\varphi(v - w) = \varphi(v + (-1)w) = \varphi(v) + \varphi((-1)w) = \varphi(v) + (-1)\varphi(w) = \varphi(v) - \varphi(w)$.
2. $\varphi(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \varphi(\lambda_1 v_1) + \cdots + \varphi(\lambda_n v_n) = \lambda_1\varphi(v_1) + \cdots + \lambda_n\varphi(v_n)$
3. Angenommen, es gilt $\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$. Dann folgt $\lambda_1\varphi(v_1) + \cdots + \lambda_r\varphi(v_r) = 0$.
4. Aus $0 \in V'$ folgt $0 = \varphi(0) \in \varphi(V')$. Sind nun $w_1 = \varphi(v_1), w_2 = \varphi(v_2) \in \varphi(V')$ mit $v_1, v_2 \in V'$, $\lambda \in K$ gegeben, so folgt:
$$w_1 + w_2 = \varphi(v_1) + \varphi(v_2) = \varphi(v_1 + v_2) \in \varphi(V'),\ \text{da } v_1 + v_2 \in V'$$
und
$$\lambda w_1 = \lambda\varphi(v_1) = \varphi(\lambda v_1) \in \varphi(V'),\ \text{da } \lambda v_1 \in V'\,.$$
Sind nun $v_1, v_2 \in \varphi^{-1}(W')$, $\lambda \in K$ gegeben, so gilt $w_1 := \varphi(v_1), w_2 := \varphi(v_2) \in W'$. Es folgt $\varphi(v_1 + v_2) = w_1 + w_2 \in W'$ und $\varphi(\lambda v_1) = \lambda w_1 \in W'$. Dies zeigt $v_1 + v_2, \lambda v_1 \in \varphi^{-1}(W')$.
5. Sind $\varphi(v_1), \ldots, \varphi(v_r) \in \varphi(V)$ linear unabhängig, so auch $v_1, \ldots, v_r \in V$.
6. Sind $w_1, w_2 \in W$, $\lambda \in K$ gegeben, so gibt es eindeutige v_1, v_2 mit $\varphi(v_1) = w_1, \varphi(v_2) = w_2$. Hierfür gilt
$$w_1 + w_2 = \varphi(v_1) + \varphi(v_2) = \varphi(v_1 + v_2) \text{ und } \lambda w_1 = \lambda\varphi(v_1) = \varphi(\lambda v_1)\,.$$
Anwendung von φ^{-1} liefert
$$\varphi^{-1}(w_1 + w_2) = v_1 + v_2 = \varphi^{-1}(w_1) + \varphi^{-1}(w_2)$$
und
$$\varphi^{-1}(\lambda w_1) = \lambda v_1 = \lambda\varphi^{-1}(w_1)\,.$$

□

3.1.9 Beispiel

Es sei V ein K-Vektorraum mit Basis $\mathcal{B} = (b_1, \dots, b_n)$. Dann ist die Linearkombinationsabbildung

$$\Phi_{\mathcal{B}} \colon K^n \to V, (x_1, \dots, x_n) \mapsto x_1 b_1 + \cdots + x_n b_n$$

ein Isomorphismus von Vektorräumen.

Beweis:

Die Injektivität ist äquivalent zur linearen Unabhängigkeit von $\mathcal{B}$. Die Surjektivität ist die Aussage, dass $\mathcal{B}$ ein Erzeugendensystem ist.

Zu zeigen bleibt die Linearität

$$\begin{aligned}
\Phi_{\mathcal{B}}(\lambda x + y) &= (\lambda x_1 + y_1) b_1 + \cdots + (\lambda x_n + y_n) b_n \\
&= \lambda (x_1 b_1 + \cdots + x_n b_n) + (y_1 b_1 + \cdots + y_n b_n) \\
&= \lambda \Phi_{\mathcal{B}}(x) + \Phi_{\mathcal{B}}(y) \,.
\end{aligned}$$

□

3.1.10 Bemerkung

Sind U, V, W drei K-Vektorräume, und $\varphi \colon U \to V$, $\psi \colon V \to W$ lineare Abbildungen, so ist $\psi \circ \varphi \colon U \to W$ linear.

Beweis:

Für $u_1, u_2 \in U$, $\lambda \in K$ gilt:

$$(\psi \circ \varphi)(u_1 + u_2) = \psi(\varphi(u_1 + u_2)) = \psi(\varphi(u_1) + \varphi(u_2)) = \psi(\varphi(u_1)) + \psi(\varphi(u_2))$$

und

$$(\psi \circ \varphi)(\lambda u_1) = \psi(\varphi(\lambda u_1)) = \psi(\lambda \varphi(u_1)) = \lambda \psi(\varphi(u_1)) \,.$$

□

3.1.11 Bemerkung

Sind V, W zwei K-Vektorräume, so ist $\operatorname{Hom}(V, W)$ ein Untervektorraum vom $\operatorname{Abb}(V, W)$. Hierbei verwenden wir die folgenden Verknüpfungen

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x), \;\; (\lambda \cdot \varphi)(x) = \lambda \cdot \varphi(x) \,.$$

Beweis:

Es sei $\varphi, \psi \in \operatorname{Hom}(V, W)$ und $\lambda \in K$ gegeben. Wir haben zu zeigen, dass $\varphi + \psi$ und $\lambda \cdot \varphi$ linear sind. Seien hierfür $v, w \in V, \sigma \in K$, dann gilt

$$\begin{aligned}
(\varphi + \psi)(\sigma v + w) &= \varphi(\sigma v + w) + \psi(\sigma v + w) \\
&= \sigma \varphi(v) + \varphi(w) + \sigma \psi(v) + \psi(w) \\
&= \sigma \left(\varphi(v) + \psi(v) \right) + \left(\varphi(w) + \psi(w) \right) \\
&= \sigma (\varphi + \psi)(v) + (\varphi + \psi)(w)
\end{aligned}$$

und

$$\begin{aligned}(\lambda\varphi)(\sigma v + w) &= \lambda\varphi(\sigma v + w) = \lambda\left(\sigma\varphi(v) + \varphi(w)\right)\\ &= \sigma(\lambda\varphi(v)) + \lambda\varphi(w) = \sigma(\lambda\varphi)(v) + (\lambda\varphi)(w)\,.\end{aligned}$$

□

3.1.12 Bemerkung

Im Spezialfall $V = W$ erhalten wir den Vektorraum $\mathrm{Hom}(V,V)$. Dieser wird auch als $\mathrm{End}(V)$ bezeichnet. Auf $\mathrm{End}(V)$ gibt es die Verkettung von Abbildungen als Verknüpfung. Mit obiger Addition und der Verkettung als Multiplikation wird $\mathrm{End}(V)$ zu einem Ring. Er heißt auch *Endomorphismenring* von V.

Beweis:

Zu zeigen sind nur noch die Distributivgesetze. Exemplarisch zeigen wir eines:

$$\begin{aligned}(\varphi\circ(\psi_1+\psi_2))(v) &= \varphi((\psi_1+\psi_2)(v)) = \varphi(\psi_1(v)+\psi_2(v))\\ &= \varphi(\psi_1(v)) + \varphi(\psi_2(v)) = (\varphi\circ\psi_1)(v) + (\varphi\circ\psi_2)(v)\\ &= (\varphi\circ\psi_1 + \varphi\circ\psi_2)(v)\end{aligned}$$

□

3.2 Bild, Kern und Faser

3.2.1 Definition

Sei $\varphi\colon V\to W$ eine lineare Abbildung, so nennen wir

$$\begin{aligned}\mathrm{Im}(\varphi) &:= \varphi(V) \text{ das \textit{Bild} von } \varphi\,,\\ \varphi^{-1}(w) &:= \{v\in V\colon \varphi(v)=w\} \text{ die \textit{Faser} über } w\in W\,,\\ \mathrm{Ker}(\varphi) &:= \varphi^{-1}(0) \text{ den \textit{Kern} von } \varphi\,.\end{aligned}$$

3.2.2 Bemerkung

Ist $\varphi\colon V\to W$ linear, so gilt:

1. $\mathrm{Im}(\varphi)\subset W$, $\mathrm{Ker}(\varphi)\subset V$ sind Unterräume.
2. φ ist surjektiv g.d.w. $\mathrm{Im}(\varphi)=W$ gilt.
3. φ ist injektiv g.d.w. $\mathrm{Ker}(\varphi)=\{0\}$ gilt.
4. Ist φ injektiv und $v_1,\dots,v_n\in V$ linear unabhängig, so ist auch $\varphi(v_1),\dots,\varphi(v_n)\in W$ linear unabhängig.

Beweis:

Wir zeigen Teil 3 Schlussrichtung $\Longleftarrow$: Sei $v_1, v_2 \in V$ mit $\varphi(v_1) = \varphi(v_2)$ gegeben. Dann folgt $0 = \varphi(v_1) - \varphi(v_2) = \varphi(v_1 - v_2)$. D.h. $v_1 - v_2 \in \operatorname{Ker}\varphi$. Somit folgt $v_1 - v_2 = 0$ und $v_1 = v_2$.

Zu Teil 4: Seien $\lambda_1, \ldots, \lambda_n \in K$ mit $\lambda_1\varphi(v_1) + \cdots + \lambda_n\varphi(v_n) = 0$ gegeben. Dann folgt $\varphi(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0$. Die Injektivität liefert $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. Da $v_1, \ldots, v_n$ linear unabhängig ist, folgt $\lambda_1 = \cdots = \lambda_n = 0$. $\square$

3.2.3 Beispiel

Wir betrachten die lineare Abbildung $\varphi_A \colon K^n \to K^m, x \mapsto Ax$, die durch $A \in \operatorname{Mat}(m \times n, K)$ gegeben ist. Wir bezeichnen mit $s_1, \ldots, s_n \in K^m$ die Spalten von A. Dann gilt

$$\begin{aligned} \operatorname{Im}\varphi_A &= \{Ax \colon x \in K^n\} = \{x_1 s_1 + \cdots + x_n s_n \colon x_1, \ldots, x_n \in K\} \\ &= \operatorname{span}(s_1, \ldots, s_n)\,. \end{aligned}$$

Das Bild ist also genau der Spaltenraum der Matrix. Bezeichnet man als *Rang der linearen Abbildung* die Dimension des Bildes, so ist dies genau der Spaltenrang der Matrix A.

Der Kern der Abbildung φ_A ist die Lösungsmenge des Gleichungssystems:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ \vdots \qquad\qquad & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Man bezeichnet ein solches Gleichungssystem auch als *homogen*, da auf der rechten Seite nur Nullen stehen.

3.2.4 Satz

Sei V ein endlichdimensionaler und W ein beliebiger K-Vektorraum und $\varphi \colon V \to W$ eine lineare Abbildung. Sind $(v_1, \ldots, v_k)$ eine Basis von $\operatorname{Ker}\varphi$ und $(w_1, \ldots, w_r)$ eine Basis von $\operatorname{Im}\varphi$, sowie $u_1, \ldots, u_k \in V$ beliebig mit $\varphi(u_1) = w_1, \ldots, \varphi(u_r) = w_r$, so ist

$$(v_1, \ldots, v_k, u_1, \ldots, u_r)$$

eine Basis von V. Insbesondere gilt die Dimensionsformel

$$\dim(V) = \dim(\operatorname{Ker}\varphi) + \dim(\operatorname{Im}\varphi)\,.$$

Beweis:

Sei $v \in V$ beliebig. Dann gibt es $\mu_1, \ldots, \mu_r$ mit $\varphi(v) = \mu_1 w_1 + \cdots + \mu_r w_r \in \operatorname{Im}\varphi$. Setzen wir $v' := \mu_1 u_1 + \cdots + \mu_r u_r$, so gilt $\varphi(v') = \varphi(v)$ und $v - v' \in \operatorname{Ker}\varphi$. Daher existieren $\lambda_1, \ldots, \lambda_k$ mit $v - v' = \lambda_1 v_1 + \cdots + \lambda_k v_k$. Dies liefert

$$v = (v - v') + v' = \lambda_1 v_1 + \cdots + \lambda_k v_k + \mu_1 u_1 + \cdots + \mu_r u_r\,.$$

Wir haben also ein Erzeugendensystem von V.

Zur linearen Unabhängigkeit: Sei

$$0 = \lambda_1 v_1 + \cdots + \lambda_k v_k + \mu_1 u_1 + \cdots + \mu_r u_r$$

so folgt

$$0 = \varphi(0) = \varphi(\lambda_1 v_1 + \cdots + \lambda_k v_k) + \varphi(\mu_1 u_1 + \cdots + \mu_r u_r) = \mu_1 w_1 + \cdots + \mu_r w_r \,.$$

Die lineare Unabhängigkeit von $w_1, \ldots, w_r$ liefert $0 = \mu_1 = \cdots = \mu_r$. Es folgt

$$0 = \lambda_1 v_1 + \cdots + \lambda_k v_k \,.$$

Die lineare Unabhängigkeit von $v_1, \ldots, v_k$ liefert $0 = \lambda_1 = \cdots = \lambda_k$.

Somit ist gezeigt, dass

$$(v_1, \ldots, v_k, u_1, \ldots, u_r)$$

eine Basis von V ist. Es folgt $\dim V = k + r$, was die Dimensionsformel beweist. □

3.2.5 Folgerung

Seien V und W zwei K-Vektorräume gleicher endlicher Dimension und $\varphi\colon V \to W$ eine lineare Abbildung. Dann sind äquivalent:

1. φ ist injektiv.
2. φ ist surjektiv.
3. φ ist bijektiv.

3.2.6 Folgerung

Zwischen zwei endlichdimensionalen K-Vektorräumen gibt es genau dann einen Isomorphismus, wenn sie die gleiche Dimension haben.

Beweis:

Ist φ ein Isomorphismus, so liefert die Dimensionsformel die Gleichheit der Dimensionen. Sind V, W zwei n-dimensionale K-Vektorräume mit Basen $\mathcal{B}, \mathcal{C}$ so liefern die Linearkombinationsabbildungen $\Phi_{\mathcal{B}}, \Phi_{\mathcal{C}}$ Isomorphismen $K^n \to V$, $K^n \to W$. Durch Invertieren und Verketten entsteht hieraus der Isomorphismus $\Phi_{\mathcal{C}} \circ \Phi_{\mathcal{B}}^{-1}$ zwischen V und W. □

3.3 Lineare Abbildungen und Matrizen

3.3.1 Ziel

Eine explizite Beschreibung von lineare Abbildung zwischen endlich-dimensionalen Vektorräumen, sodass man mit diesen rechnen kann.

3.3.2 Satz

Es sei V ein endlich-dimensionaler K-Vektorraum und W ein beliebiger K-Vektorraum sowie Vektoren $v_1, \dots, v_r \in V$ und $w_1, \dots, w_r \in W$. Dann gilt

1. Ist $(v_1, \dots, v_r)$ linear unabhängig, so gibt es mindestens eine lineare Abbildung $\varphi\colon V \to W$ mit $\varphi(v_i) = w_i$ für $i = 1, \dots, r$.
2. Ist $(v_1, \dots, v_r)$ eine Basis von V, so gibt es genau eine lineare Abbildung $\varphi\colon V \to W$ mit $\varphi(v_i) = w_i$ für $i = 1, \dots, r$. Diese hat zudem die Eigenschaften:
   (a) $\operatorname{Im}\varphi = \operatorname{span}(w_1, \dots, w_r)$
   (b) φ ist genau dann injektiv, wenn $(w_1, \dots, w_r)$ linear unabhängig ist.

Beweis:

Wir beginnen mit Teil 2). Jedes $v \in V$ hat eine eindeutige Darstellung $v = \lambda_1 v_1 + \dots + \lambda_r v_r$. Aus $\varphi(v_i) = w_i$ folgt mit der Linearität von φ

$$\varphi(v) = \lambda_1 w_1 + \dots + \lambda_r w_r\,.$$

Es gibt als höchstens eine solche lineare Abbildung. Um die Existenz zu zeigen, zeigen wir, dass obige Vorschrift wirklich eine lineare Abbildung definiert:

$$\begin{aligned}\varphi(v+v') &= \varphi(\lambda_1 v_1 + \dots + \lambda_r v_r + \lambda'_1 v_1 + \dots + \lambda'_r v_r)\\ &= \varphi((\lambda_1 + \lambda'_1)v_1 + \dots + (\lambda_r + \lambda'_r)v_r)\\ &= (\lambda_1 + \lambda'_1)w_1 + \dots + (\lambda_r + \lambda'_r)w_r\\ &= \lambda_1 w_1 + \dots + \lambda_r w_r + \lambda'_r w_r + \dots + \lambda'_r w_r\\ &= \varphi(v) + \varphi(v')\end{aligned}$$

$$\begin{aligned}\varphi(\lambda v) &= \varphi(\lambda(\lambda_1 v_1 + \dots + \lambda_r v_r)) = \varphi(\lambda\lambda_1 v_1 + \dots + \lambda\lambda_r v_r)\\ &= \lambda\lambda_1 w_1 + \dots + \lambda\lambda_r w_r = \lambda(\lambda_1 w_1 + \dots + \lambda_r w_r) = \lambda\varphi(v)\,.\end{aligned}$$

Zudem gilt

$$\begin{aligned}\operatorname{Im}(\varphi) &= \{\varphi(\lambda_1 v_1 + \dots + \lambda_r v_r)\colon \lambda_1, \dots, \lambda_r \in K\}\\ &= \{\lambda_1 w_1 + \dots + \lambda_r w_r\colon \lambda_1, \dots, \lambda_r \in K\} = \operatorname{span}(w_1, \dots, w_r)\end{aligned}$$

und

$$\operatorname{Ker}(\varphi) = \{\lambda_1 v_1 + \dots + \lambda_r v_r \mid \lambda_1 w_1 + \dots + \lambda_r w_r = 0\}\,.$$

Die letzte Gleichung zeigt, dass die lineare Unabhängigkeit von $(w_1, \dots, w_r)$ die Injektivität von φ impliziert. Falls $(w_1, \dots, w_r)$ linear abhängig ist, so führt die lineare Unabhängigkeit von $(v_1, \dots, v_r)$ auf nicht-triviale Kernelemente und φ ist folglich nicht injektiv.

Zu Teil 1). Ist $(v_1, \dots, v_r)$ linear unabhängig, so ergänzen wir zu einer Basis $(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$. Zudem ergänzen wir $w_{r+1}, \dots, w_n$ beliebig und erhalten mit Teil 2) eine lineare Abbildung

$$\varphi\colon V \to W \text{ mit } \varphi(v_i) = w_i \text{ für } i = 1, \dots, n\,.$$

Die Konstruktion zeigt, dass φ in der Regel nicht eindeutig ist. $\square$

3.3.3 Folgerung

Ist V ein K-Vektorraum und $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V, so gibt es genau einen Isomorphismus

$$\Phi_{\mathcal{B}} \colon K^n \to V \text{ mit } \Phi_{\mathcal{B}}(e_i) = v_i \text{ für } i = 1, \ldots, n\,.$$

Hierbei ist $(e_1, \ldots, e_n)$ die Standardbasis von K^n. Wir hatten $\Phi_{\mathcal{B}}$ als Linearkombinationsabbildung bezeichnet. $\Phi_{\mathcal{B}}$ wird auch *Koordinatensystem* genannt.

3.3.4 Folgerung

Zu jeder linearen Abbildung $\varphi \colon K^n \to K^m$ gibt es eine eindeutige Matrix $A \in \operatorname{Mat}(m \times n, K)$ mit

$$\varphi(x) = Ax$$

wobei hier $x \in K^n$ ein Spaltenvektor ist.

Beweis:

Man schreibt $\varphi(e_1), \ldots, \varphi(e_n)$ als Spaltenvektoren in die Matrix A. □

3.3.5 Satz

Gegeben seien K-Vektorräume V mit Basis $\mathcal{A} = (v_1, \ldots, v_n)$ und W mit Basis $\mathcal{B} = (w_1, \ldots, w_m)$. Dann gibt es zu jeder linearen Abbildung $\varphi \colon V \to W$ genau eine Matrix $A = (a_{ij}) \in \operatorname{Mat}(m \times n, K)$, so dass

$$\varphi(v_j) = \sum_{i=1}^{m} a_{ij} w_i \text{ für } j = 1, \ldots, n$$

gilt. Dies definiert eine Abbildung

$$M_{\mathcal{B}}^{\mathcal{A}} \colon \operatorname{Hom}(V, W) \to \operatorname{Mat}(m \times n, K), \varphi \mapsto A = M_{\mathcal{B}}^{\mathcal{A}}(\varphi)\,.$$

$M_{\mathcal{B}}^{\mathcal{A}}$ ist ein Isomorphismus von K-Vektorräumen. Insbesondere gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(\varphi + \psi) = M_{\mathcal{B}}^{\mathcal{A}}(\varphi) + M_{\mathcal{B}}^{\mathcal{A}}(\psi) \text{ und } M_{\mathcal{B}}^{\mathcal{A}}(\lambda\varphi) = \lambda M_{\mathcal{B}}^{\mathcal{A}}(\varphi)\,.$$

$M_{\mathcal{B}}^{\mathcal{A}}(\varphi)$ heißt die *Darstellungsmatrix* der Abbildung φ bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$.

Kurz: Nach Wahl der Basen $\mathcal{A}, \mathcal{B}$ sind lineare Abbildungen $V \to W$ und $m \times n$-Matrizen über K das gleiche. Insbesondere geht das Rechnen mit linearen Abbildungen in das Rechnen mit Matrizen über.

Beweis:

Da $\mathcal{B}$ eine Basis ist, ist jede Spalte der Matrix A eindeutig bestimmt. Somit ist $M_{\mathcal{B}}^{\mathcal{A}}$ eine Abbildung.

Sei nun $B = (b_{ij})$ die Matrix zur Abbildung ψ, so ist

$$(\varphi + \psi)(v_j) = \varphi(v_j) + \psi(v_j) = \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i = \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i$$

und für $\lambda \in K$ ist

$$(\lambda\varphi)(v_j) = \lambda\varphi(v_j) = \lambda \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} (\lambda a_{ij}) w_i \,.$$

Daher ist die Abbildung $M_{\mathcal{B}}^{\mathcal{A}}$ linear.

Da $\mathcal{A}$ eine Basis ist, gibt es nach 3.3.2 zu jeder Matrix eine lineare Abbildung. Insbesondere ist $M_{\mathcal{B}}^{\mathcal{A}}$ surjektiv.

Die Injektivität von $M_{\mathcal{B}}^{\mathcal{A}}$ folgt aus der Tatsache, dass eine lineare Abbildung durch ihre Werte auf einer beliebigen Basis eindeutig bestimmt ist (s. 3.3.2). □

3.3.6 Zusatz

Betrachten wir die Vektorräume $V = K^m$ und $W = K^n$ mit den Standardbasen $\mathcal{A} \coloneqq (e_1, \ldots, e_m)$ und $\mathcal{B} \coloneqq (e_1, \ldots, e_n)$, so erhalten wir den Isomorphismus

$$M_{\mathcal{B}}^{\mathcal{A}} \colon \mathrm{Hom}(K^m, K^n) \to \mathrm{Mat}(n \times m, K), \varphi \mapsto M_{\mathcal{B}}^{\mathcal{A}}(\varphi) \,.$$

3.3.7 Bemerkung

Die Standardbasis des Matrizenraums $\mathrm{Mat}(n \times m, K)$ besteht aus den $n \cdot m$ Matrizen $E_{i,j}$, die in Zeile i und Spalte j eine Eins und sonst nur Nullen haben. Die lineare Abbildung φ_j^i mit $M_{\mathcal{B}}^{\mathcal{A}}(\varphi_j^i) = E_{j,i}$ ist dann durch $\varphi_j^i(a_i) = b_j$ und $\varphi_j^i(a_k) = 0$ für $k \neq i$ gegeben. Somit bilden die Abbildungen φ_j^i eine Basis von $\mathrm{Hom}(V, W)$. Die Skalare um eine konkrete lineare Abbildung als Linearkombination dieser Basis zu schreiben sind genau die Einträge ihrer Darstellungsmatrix.

3.3.8 Folgerung

Sei $\varphi \colon V \to W$ linear, $m = \dim V$, $n = \dim W$ und $r = \dim \mathrm{Im}\,\varphi$. Dann gibt es Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W, sodass

$$M_{\mathcal{B}}^{\mathcal{A}}(\varphi) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

gilt. Dabei ist E_r die $r \times r$-Einheitsmatrix. Obige Darstellung nennt man auch *Blockmatrix.*

Beweis:

Wir wiederholen die Konstruktion von Basen für V und W aus dem Beweis der Dimensionsformel (3.2.4).

D.h., wir wählen $a_{r+1}, \ldots, a_m$ als Basis von $\ker \varphi$ und ergänzen dies mit $a_1, \ldots, a_r$ zu einer Basis von V. Dann setzen wir $b_i \coloneqq \varphi(a_i)$ für $i = 1, \ldots, r$ und erhalten eine Basis von $\mathrm{Im}\,\varphi$. Abschließend ergänzen wir mit $b_{r+1}, \ldots, b_n$ zu einer Basis von W. Mit diesen Basen ergibt sich die obige Matrix. □

3.3.9 Bemerkung

Ist der Argumentraum gleich dem Bildraum, d.h. betrachtet man Endomorphismen $\varphi \colon V \to V$, so betrachtet man oft $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$. Zur Vereinfachung schreibt man $M_{\mathcal{B}}(\varphi)$.

Unabhängig von der Wahl der Basis $\mathcal{B}$ ist die Abbildung φ mit $M_{\mathcal{B}}(\varphi) = E_n$ immer die identische Abbildung id_V.

3.4 Matrixmultiplikation

3.4.1 Ziel

Seien φ, ψ linear. Frage: Wie hängt $M_{\mathcal{C}}^{\mathcal{A}}(\psi \circ \varphi)$ von $M_{\mathcal{B}}^{\mathcal{A}}(\varphi)$ und $M_{\mathcal{C}}^{\mathcal{B}}(\psi)$ ab?

3.4.2 Bemerkung

Seien $U = K^m, V = K^n, W = K^r$ und $\psi\colon U \to V$, $\varphi\colon V \to W$ linear. Seien A, B, C die Matrizen, sodass $\varphi(x) = Ax$, $\psi(y) = By$ und $(\varphi \circ \psi)(x) = Cx$ für alle Spaltenvektoren $x \in K^m, y \in K^n$ gilt. Dann gilt

$$c_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,n}b_{n,j}\,.$$

Beweis:

Wir bezeichnen mit $(e_1, \ldots, e_m), (f_1, \ldots, f_n), (g_1, \ldots, g_r)$ die Standardbasen von K^m, K^n, K^r. In der j-ten Spalte von C steht das Bild des j-ten Standardbasisvektors unter $(\varphi \circ \psi)$. Daher gilt

$$\begin{aligned}(\varphi \circ \psi)(e_j) &= \varphi(b_{1,j}f_1 + \cdots + b_{n,j}f_n) = b_{1,j}\varphi(f_1) + \cdots + b_{n,j}\varphi(f_n)\\ &= b_{1,j}(a_{1,1}g_1 + a_{2,1}g_2 + \cdots + a_{r,1}g_r) + \cdots\\ &\quad + b_{n,j}(a_{1,n}g_1 + a_{2,n}g_2 + \cdots + a_{r,n}g_r)\\ &= c_{1,j}g_1 + \cdots + c_{r,j}g_r\end{aligned}$$

Da $(g_1, \ldots, g_r)$ linear unabhängig ist, folgt

$$\begin{aligned}c_{1,j} &= a_{1,1}b_{1,j} + a_{1,2}b_{2,j} + \cdots + a_{1,n}b_{n,j}\\ &\vdots\\ c_{r,j} &= a_{r,1}b_{1,j} + a_{r,2}b_{2,j} + \cdots + a_{r,n}b_{n,j}\,.\end{aligned}$$

$\square$

3.4.3 Definition

In obiger Situation nennt man die Matrix C das Produkt $A \cdot B$ der Matrizen A und B.

3.4.4 Beispiele

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

Im Spezialfall einer einzeiligen und einer einspaltigen Matrix erhält man:

$$A \cdot B = (a_1 \, \cdots \, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \in K = \mathrm{Mat}(1 \times 1, K)\,,$$

multipliziert man in anderer Reihenfolge so entsteht

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} (b_1 \, \cdots \, b_n) = \begin{pmatrix} a_1 b_1 & \cdots & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_n b_1 & \cdots & a_n b_n \end{pmatrix} \in \mathrm{Mat}(n \times n, K).$$

Es gilt in der Regel $A \cdot B \neq B \cdot A$:

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

Zudem ist es möglich, dass ein Matrizenprodukt Null ist, auch wenn beide Faktoren nicht Null sind.

3.4.5 Beispiel: Drehmatrizen

Die Drehung im $\mathbb{R}^2$ um $(0,0)$ mit Winkel α ist eine lineare Abbildung. Sie hat die Darstellungsmatrix

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

Führt man zwei Drehungen mit Winkel α und β hintereinander aus, so erhält man

$$\begin{aligned} & \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \\ = & \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\sin\alpha\cos\beta - \cos\alpha\sin\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{pmatrix} \\ = & \begin{pmatrix} \cos\alpha + \beta & -\sin\alpha + \beta \\ \sin\alpha + \beta & \cos\alpha + \beta \end{pmatrix} \end{aligned}$$

Will man dieses direkt zeigen, so muss man die Additionstheoreme der Winkelfunktionen verwenden.

3.4.6 Satz: Rechenregel für Matrizen

Sind Matrizen $A, A_1, A_2 \in \mathrm{Mat}(m \times n, K)$, $B, B_1, B_2 \in \mathrm{Mat}(n \times r, K)$, $C \in \mathrm{Mat}(r \times s, K)$ und $\lambda \in K$ gegeben, so gilt:

$$\begin{aligned} A \cdot (B_1 + B_2) &= A \cdot B_1 + A \cdot B_2 \\ (A_1 + A_2) \cdot B &= A_1 \cdot B + A_2 \cdot B \\ A \cdot (\lambda B) &= (\lambda A) \cdot B = \lambda (A \cdot B) \\ (A \cdot B) \cdot C &= A \cdot (B \cdot C) \\ (A \cdot B)^\top &= B^\top \cdot A^\top \\ E_m \cdot A &= A \cdot E_n = A \end{aligned}$$

Beweis:

Wir zeigen nur die 5. und die 4. Gleichung:

An Position i, j von $(A \cdot B)^\top$ steht der Eintrag von $A \cdot B$ mit Index j, i

$$a_{j,1}b_{1,i} + \cdots + a_{j,n}b_{n,i} = b_{1,i}a_{j,1} + \cdots + b_{n,i}a_{j,n} .$$

An Position i, j von $B^\top A^\top$ steht

$$(B^\top)_{i,1}(A^\top)_{1,j} + \cdots + (B^\top)_{i,n}(A^\top)_{n,j} = b_{1,i}a_{j,1} + \cdots + b_{n,i}a_{j,n} .$$

Die Matrizen A, B, C beschreiben lineare Abbildungen

$$K^s \xrightarrow{\varphi_C} K^r \xrightarrow{\varphi_B} K^n \xrightarrow{\varphi_A} K^m ,$$

welche durch $\varphi_A \colon x \mapsto Ax$, $\varphi_B \colon x \mapsto Bx$ und $\varphi_C \colon x \mapsto Cx$ gegeben sind. Wir erhalten

$$(\varphi_A \circ \varphi_B) \circ \varphi_C = \varphi_A \circ (\varphi_B \circ \varphi_C)$$

aus der Assoziativität der Verkettung von Abbildungen. Da das Verketten lineare Abbildungen durch die Matrizenmultiplikation beschrieben wird, folgt die Gleichheit

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

von Matrixprodukten. □

3.4.7 Folgerung

Die Menge der $\mathrm{Mat}(n \times n, K)$ ist mit der Addition und der Matrizenmultiplikation ein Ring.

3.4.8 Definition

Eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ heißt invertierbar, wenn es eine Matrix $A' \in \mathrm{Mat}(n \times n, K)$ mit $AA' = A'A = E_n$ gibt.

Die Menge

$$\mathrm{Gl}_n(K) = \{A \in \mathrm{Mat}(n \times n, K) \colon A \text{ ist invertierbar } \}$$

bildet mit der Matrixmultiplikation als Verknüpfung eine Gruppe. Sie heißt *allgemeine lineare Gruppe* (general linear group). Man schreibt oft A^{-1} statt A' für die inverse Matrix.

Beweis:

Es ist zu zeigen, dass das Produkt zweier invertierbarer Matrizen wieder invertierbar ist:

$$\begin{aligned}(AB)(B'A') &= A(BB')A' = AA' = E_n \\ (B'A')(AB) &= B'(A'A)B = B'B = E_n\end{aligned}$$

Alle weiteren Gruppeneigenschaften wurden bereits gezeigt.

3.4.9 Anmerkung

Da die Verkettung von linearen Abbildungen durch die Matrixmultiplikation beschrieben wird, beschreiben die invertierbaren $n \times n$-Matrizen genau die bijektiven linearen Abbildungen $K^n \to K^n$. Die inverse Matrix beschreibt die Umkehrabbildung.

3.4.10 Bemerkung

Für $A \in \mathrm{Mat}(n \times n, K)$ sind äquivalent:

1. A ist invertierbar.
2. $A^\top$ ist invertierbar.
3. A hat Spaltenrang n.
4. A hat Zeilenrang n.

Anmerkung: Es gilt die Rechenregel $(A^{-1})^\top = (A^\top)^{-1}$.

Beweis:

Zu $1 \iff 2$:

$$(A^{-1})^\top A^\top = (AA^{-1})^\top = E_n^\top = E_n$$

$$(A((A^\top)^{-1})^\top) = ((A^\top)^\top((A^\top)^{-1})^\top) = ((A^\top)^{-1}A^\top)^\top = E_n^\top = E_n$$

Zu $1 \iff 3$:
Eine lineare Abbildung $K^n \to K^n$ ist genau dann invertierbar, wenn sie surjektiv ist. Folglich ist A genau dann invertierbar wenn A Spaltenrang n hat. Dies zeigt $1 \iff 3$.
Die Aussage $2 \iff 4$ entsteht aus $1 \iff 3$ durch Transponieren.

3.5 Koordinatentransformationen

3.5.1 Definition

Sei V ein K-Vektorraum mit Basis $\mathcal{B} = (v_1, \dots, v_n)$. Die Linearkombinationsabbildung

$$\Phi_\mathcal{B} \colon K^n \to V \text{ mit } \Phi_\mathcal{B}(e_i) = v_i \text{ für } i = 1, \dots, n$$

liefert das durch $\mathcal{B}$ gegebene Koordinatensystem in V. Es ist

$$\Phi_\mathcal{B}(x_1, \dots, x_n) = x_1v_1 + \cdots + x_nv_n\,.$$

Für $v \in V$ nennt man $x = (x_1, \dots, x_n) = \Phi_\mathcal{B}^{-1}(v) \in K^n$ die *Koordinaten* von $v = x_1v_1 + \cdots + x_nv_n$.

3.5.2 Bemerkung: Koordinatenwechsel

Es seien $\mathcal{A} = (a_1, \dots, a_n)$ und $\mathcal{B} = (b_1, \dots, b_n)$ zwei verschiedene Basen des K-Vektorraums V. Es gibt das Diagramm von Isomorphismen:

$$\begin{tikzcd} K^n \arrow[d, "T_{\mathcal{B}}^{\mathcal{A}} := \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{A}}"'] \arrow[dr, "\Phi_{\mathcal{A}}"] & \\ K^n \arrow[r, "\Phi_{\mathcal{B}}"'] & V \end{tikzcd}$$

Man versteht $T_{\mathcal{B}}^{\mathcal{A}}$ als Matrix in $\mathrm{Gl}_n(K)$ und nennt sie die *Transformationsmatrix* des Basiswechsels. Es ist

$$v = x_1 a_1 + \cdots + x_n a_n = y_1 b_1 + \cdots + y_n b_n \in V$$

mit

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_{\mathcal{B}}^{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Es gilt $T_{\mathcal{B}}^{\mathcal{A}} = M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id})$.

Im Spezialfall $V = K^n$ erhalten wir Matrizen A und B, deren Spalten die Vektoren der Basen $\mathcal{A}$ und $\mathcal{B}$ sind. Dann ergibt sich:

$$\begin{tikzcd} K^n \arrow[d, "T"'] \arrow[dr, "A"] & \\ K^n \arrow[r, "B"'] & K^n \end{tikzcd} \quad \text{mit } T = B^{-1}A.$$

Ist $\mathcal{B}$ die Standardbasis, so ist B die Einheitsmatrix und es gilt $T = A$.
Ist $\mathcal{A}$ die Standardbasis, so ist A die Einheitsmatrix und es gilt $T = B^{-1}$.

3.5.3 Beispiel

Im Fall $\mathbb{R}^2$ betrachten wir die Basis $\mathcal{A} = (e_1, e_2)$ und $\mathcal{B} = (b_1, b_2)$ mit

$$b_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ und } b_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Dann ergibt sich $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ und

$$B^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}.$$

Somit hat $x = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ die Darstellung $y = \begin{pmatrix} \frac{-2}{5} \\ \frac{-1}{5} \end{pmatrix}$ im durch $\mathcal{B}$ gegebenen Koordinatensystem. Man sagt gelegentlich y-Koordinaten des Punkts $(-1, -1)^\top$.

3.5.4 Bemerkung

Seien eine lineare Abbildung $\varphi\colon V \to W$ und Basen $\mathcal{A} = (a_1, \dots, a_n), \mathcal{B} = (b_1, \dots, b_m)$ von V und W gegeben. Dann hat man das folgende Diagramm von Abbildungen:

$$\begin{tikzcd} K^n \arrow[r, "\Phi_{\mathcal{A}}"] \arrow[d, "M_{\mathcal{B}}^{\mathcal{A}}(\varphi)"'] & V \arrow[d, "\varphi"] \\ K^m \arrow[r, "\Phi_{\mathcal{B}}"'] & W \end{tikzcd}$$

und es gelten

$$\Phi_{\mathcal{B}} \circ M_{\mathcal{B}}^{\mathcal{A}}(\varphi) = \varphi \circ \Phi_{\mathcal{A}}, \text{ und folglich } M_{\mathcal{B}}^{\mathcal{A}}(\varphi) = \Phi_{\mathcal{B}}^{-1} \circ \varphi \circ \Phi_{\mathcal{A}} \,.$$

Beweis:

Da alle Abbildungen linear sind, genügt es die Gleichheit auf einer beliebigen Basis von K^n zu bestätigen. Wir wählen die Standardbasis $(e_1, \dots, e_n)$.

Sei $M_{\mathcal{B}}^{\mathcal{A}}(\varphi) = A = (a_{i,j})$, so gilt

$$\Phi_{\mathcal{B}}(M_{\mathcal{B}}^{\mathcal{A}}(F)(e_j)) = \Phi_{\mathcal{B}}(a_{1,j}, \dots, a_{m,j}) = \sum a_{i,j} b_i$$

$$F(\Phi_{\mathcal{A}}(e_j)) = F(a_j) = \sum a_{i,j} b_i$$

Die zweite Behauptung folgt aus der ersten durch Verkettung mit $\Phi_{\mathcal{B}}^{-1}$.

3.5.5 Definition

Obiges Diagramm wird als kommutativ bezeichnet. Das beutet, es gibt verschiedene Wege um von K^n nach W abzubilden. Unabhängig von der Wahl des Wegs entsteht das gleiche Bild. Dies ist also eine Art die Gleichheit von Verkettungen von Abbildungen graphisch darzustellen.

3.5.6 Satz

Seien U, V, W Vektorräume der Dimensionen r, m, n mit Basen $\mathcal{A}, \mathcal{B}, \mathcal{C}$ und $\psi\colon U \to V$ sowie $\varphi\colon V \to W$ lineare Abbildungen. Dann gilt

$$M_{\mathcal{C}}^{\mathcal{A}}(\varphi \circ \psi) = M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot M_{\mathcal{B}}^{\mathcal{A}}(\psi) \,.$$

In anderen Worten: Man verkettet lineare Abbildungen, indem man ihre Darstellungsmatrizen multipliziert.

Beweis:

Mit $A = M_{\mathcal{C}}^{\mathcal{B}}(\varphi), B = M_{\mathcal{B}}^{\mathcal{A}}(\psi)$ behaupten wird die Kommutativität des äußeren Quadrats von:

$$\begin{tikzcd}
K^r \arrow[rrr, "\Phi_{\mathcal{A}}"] \arrow[dr, "B"] \arrow[dd, "A \cdot B"'] & & & U \arrow[dl, "\psi"'] \arrow[dd, "\varphi \circ \psi"] \\
& K^m \arrow[r, "\Phi_{\mathcal{B}}"] \arrow[dl, "A"] & V \arrow[dr, "\varphi"'] & \\
K^n \arrow[rrr, "\Phi_{\mathcal{C}}"] & & & W
\end{tikzcd}$$

Da die Kommutativität aller inneren Drei- und Vierecke bereits gezeigt ist (3.5.4 und 3.4.2), folgt auch die des äußeren Quadrats. □

3.5.7 Folgerung

Sei V ein n-dimensionaler Vektorraum mit Basis $\mathcal{B}$ sowie φ, ψ Endomorphismen von V. Dann gilt

$$M_{\mathcal{B}}(\varphi \circ \psi) = M_{\mathcal{B}}(\varphi) \cdot M_{\mathcal{B}}(\psi)\,.$$

Insbesondere ist $M_B \colon \mathrm{End}(V) \to \mathrm{Mat}(n \times n; K)$ ein Ringisomorphismus.

3.5.8 Transformationsformel

Sei $\varphi \colon V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Vektorräumen V und W. Weiterhin seien $\mathcal{A}, \mathcal{A}'$ Basen von V und $\mathcal{B}, \mathcal{B}'$ Basen von W. Dann gilt

$$M_{\mathcal{B}'}^{\mathcal{A}'}(\varphi) = T_{\mathcal{B}'}^{\mathcal{B}} M_{\mathcal{B}}^{\mathcal{A}}(\varphi)(T_{\mathcal{A}'}^{\mathcal{A}})^{-1}\,.$$

Beweis:

Es gilt

$$\begin{aligned}
M_{\mathcal{B}'}^{\mathcal{A}'}(\varphi) &= M_{\mathcal{B}'}^{\mathcal{A}'}(\mathrm{id} \circ \varphi \circ \mathrm{id}) = M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}) M_{\mathcal{B}}^{\mathcal{A}'}(\varphi \circ \mathrm{id}) \\
&= M_{\mathcal{B}'}^{\mathcal{B}}(\mathrm{id}) M_{\mathcal{B}}^{\mathcal{A}}(\varphi) M_{\mathcal{A}}^{\mathcal{A}'}(\mathrm{id}) = T_{\mathcal{B}'}^{\mathcal{B}} M_{\mathcal{B}}^{\mathcal{A}}(\varphi) T_{\mathcal{A}}^{\mathcal{A}'} = T_{\mathcal{B}'}^{\mathcal{B}} M_{\mathcal{B}}^{\mathcal{A}}(\varphi)(T_{\mathcal{A}'}^{\mathcal{A}})^{-1}\,.
\end{aligned}$$

□

3.5.9 Folgerung

Ist V ein endlich-dimensionaler Vektorraum mit Basen $\mathcal{A}$ und $\mathcal{B}$ sowie ein Endomorphismus φ gegeben, so ist

$$M_{\mathcal{B}}(\varphi) = T_{\mathcal{B}}^{\mathcal{A}} M_{\mathcal{A}}(\varphi) T_{\mathcal{A}}^{\mathcal{B}},$$

bzw.

$$B = SAS^{-1}$$

wenn $A = M_{\mathcal{A}}(\varphi), B = M_{\mathcal{B}}(\varphi), S = T_{\mathcal{B}}^{\mathcal{A}}$ gesetzt wird.

3.5.10 Satz

Für jede Matrix $A \in \mathrm{Mat}(m \times n, K)$ gilt die Gleichheit von Zeilenrang und Spaltenrang.

Beweis:

Wir betrachten die durch A gegebene lineare Abbildung $\varphi_A\colon K^n \to K^m, x \mapsto Ax$. Dann gibt es Basen $\mathcal{A}, \mathcal{B}$ von K^n und K^m, sodass

$$M_{\mathcal{B}}^{\mathcal{A}}(\varphi_A) = B = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

gilt (Folgerung 3.3.8). Der Zeilen- und Spaltenrang von B ist r. Somit gilt hier Gleichheit. Wir haben nun zu zeigen, dass dies auch die Gleichheit von Zeilen- und Spaltenrang für A impliziert. In obiger Situation erhalten wir invertierbare Basiswechselmatrizen Matrizen S und T mit $B = SAT$. Die Restbehauptung folgt dann aus dem nächsten Lemma. $\square$

3.5.11 Lemma

Für $A \in \mathrm{Mat}(m \times n, K)$, $S \in \mathrm{Gl}_m(K), T \in \mathrm{Gl}_n(K)$, gilt:

1. Die Spaltenränge von SAT und A sind gleich.
2. Die Zeilenränge von SAT und A sind gleich.

Beweis:

Wir erhalten das folgende kommutative Diagramm linearer Abbildungen:

$$\begin{tikzcd} K^n \arrow[r, "A"] & K^m \arrow[d, "S"] \\ K^n \arrow[u, "T"] \arrow[r, "SAT"'] & K^m \end{tikzcd}$$

Da S und T Isomorphismen definieren, haben die Bilder der durch A und SAT gegebenen linearen Abbildungen die gleiche Dimension. Somit haben die Matrizen den gleichen Spaltenrang.

Wendet man das gleiche Argument auf die transponierten Matrizen an, so entsteht die Gleichheit der Spaltenränge von $A^\top$ und $T^\top \cdot A^\top \cdot S^\top$. Dies zeigt die zweite Behauptung. $\square$

Anmerkung: A und AT haben den gleichen Spaltenraum, bei A und SA stimmt nur die Dimensionen der Spalträume überein.

3.5.12 Definition

Zwei Matrizen $A, B \in \mathrm{Mat}(m \times n, K)$ heißen *äquivalent*, wenn es $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = SAT^{-1}$ gibt. Zwei quadratische Matrizen $A, B \in \mathrm{Mat}(n \times n, K)$ heißen *ähnlich*, wenn es $S \in \mathrm{Gl}_n(K)$ mit $B = SAS^{-1}$ gibt.

3.5.13 Bemerkung

Zwei Matrizen sind genau dann äquivalent, wenn sie die gleiche lineare Abbildung bezüglich verschiedenen Paaren von Basen darstellen.

Zwei quadratische Matrizen sind genau dann ähnlich, wenn sie den gleichen Endomorphismus bezüglich verschiedener Basen darstellen.

3.5.14 Lemma

Zwei Matrizen sind genau dann äquivalent, wenn sie den gleichen Rang haben. Insbesondere ist jede Matrix mit Rang r äquivalent zu

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Beweis:

Nach Folgerung 3.3.8 ist jede Matrix äquivalent zu einer der oben gegebenen Matrizen. Aus Lemma 3.5.11 folgt, dass äquivalente Matrizen gleichen Rang haben. Somit ist eine Matrix zu nur einer Matrix der obigen Form äquivalent.

□

3.6 Elementarmatrizen und Matrixumformungen

3.6.1 Definition

Es ist $\mathrm{Mat}(m \times n, K)$ ein mn-dimensionaler K-Vektorraum. Eine Basis ist durch die Matrizen $E_{i,j}$ gegeben, welche an der Position (i, j) eine 1 und sonst nur Nullen hat.

Ist $\lambda \in K^\times, i, j \in \{1, \ldots, m\}, i \neq j$, so bezeichnen wir in $\mathrm{Mat}(m \times m, K)$ die folgenden Matrizen

$$\begin{aligned} &S_i(\lambda) := E_m + (\lambda - 1)E_{i,i},\ Q_{ij} := E_m + E_{ij},\ Q_{ij}(\lambda) := E_m + \lambda E_{ij}, \\ &P_{ij} := E_m - E_{ii} - E_{jj} + E_{ij} + E_{ji} \end{aligned}$$

als Elementarmatrizen.

3.6.2 Bemerkungen

Obige Elementarmatrizen sind invertierbar. Es gilt

$$S_i(\lambda)^{-1} = S_i(\lambda^{-1}),\ Q_{ij}(\lambda)^{-1} = Q_{ij}(-\lambda),\ P_{ij}^{-1} = P_{ij}\,.$$

Ist $A \in \mathrm{Mat}(m \times n, K)$, so beschreibt die Multiplikation von Links eine Zeilenumformung der Matrix A. Es ist

1. $S_i(\lambda) \cdot A$ die Matrix, die aus A durch Skalierung der i-ten Zeile mit λ entsteht.

2. $Q_{ij}(\lambda) \cdot A$ die Matrix, die aus A durch Addition des λ-fachen der j-ten Zeile zur i-ten Zeile entsteht.

3. $P_{ij} \cdot A$ die Matrix, die aus A durch Vertauschung der i-ten und der j-ten Zeile entsteht.

Ist $A \in \mathrm{Mat}(n \times m, K)$, so beschreibt die Multiplikation von Rechts eine Spaltenumformung der Matrix A. Es ist

1. $A \cdot S_i(\lambda)$ die Matrix, die aus A durch Skalierung der i-ten Spalte mit λ entsteht.

2. $A \cdot Q_{ij}(\lambda)$ die Matrix, die aus A durch Addition des λ-fachen der i-ten Spalte zur j-ten Spalte entsteht.

3. $A \cdot P_{ij}$ die Matrix, die aus A durch Vertauschung der i-ten und der j-ten Spalte entsteht.

Nachweis: Übung.

3.6.3 Satz

Jede invertierbare Matrix $A \in \mathrm{Gl}_n(K)$ ist das Produkt von Elementarmatrizen. Man sagt auch $\mathrm{Gl}_n(K)$ wird von Elementarmatrizen erzeugt.

Beweis:

Sei $A \in \mathrm{Gl}_n(K)$ beliebig. A kann durch endlich viele elementare Zeilenoperationen auf Zeilenstufenform transformiert werden. Schreibt man die einzelnen Schritte als Multiplikationen mit Elementarmatrizen $B_1, \ldots, B_r$, so erhält man

$$B_r \cdots B_1 A = B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ & \ddots & \vdots \\ 0 & & b_{nn} \end{pmatrix}.$$

Da A invertierbar ist, ist B ebenfalls invertierbar und hat somit den Zeilenrang n. Dies zeigt, dass B keine Nullzeile hat und alle Diagonaleinträge von B gleich 1 sind.

Durch weitere Zeilenumformungen kann man alle Einträge in B oberhalb der Diagonalen eliminieren. Dies kann durch die Linksmultiplikation mit den Elementarmatrizen $B_{r+1}, \ldots, B_s$ kodiert werden. Wir erhalten

$$E_n = B_s \cdots B_{r+1} B = B_s \cdots B_1 A\,.$$

Hieraus folgt

$$A^{-1} = B_s \cdots B_1 \text{ und } A = B_1^{-1} \cdots B_s^{-1}\,.$$

□

3.6.4 Beispiel

Der obige Beweis liefert ein Verfahren um die Inverse einer Matrix A zu bestimmen. Dies wird auch als Gauß-Algorithmus bezeichnet. Man beginnt hierbei mit der Matrix A und der Einheitsmatrix E_n. Dann führt man auf beiden Matrizen die gleichen Zeilenoperationen aus und transformiert dabei A zur Einheitsmatrix. Dabei entsteht:

A	E_n
$B_1 \cdot A$	$B_1 \cdot E_n$
$\vdots$	$\vdots$
$B_s \cdots B_1 A = E_n$	$B_2 \cdots B_1 E_n = A^{-1}$

Im Fall der Matrix $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 2 & 3 & 3 \end{pmatrix}$ verläuft die Rechnung wie folgt:

$$\begin{array}{|ccc|ccc|}
\hline
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 1 & 0 \\
2 & 3 & 3 & 0 & 0 & 1 \\
\hline
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 1 & 0 \\
0 & 1 & 1 & -2 & 0 & 1 \\
\hline
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 1 & 0 \\
0 & 0 & -1 & -2 & -1 & 1 \\
\hline
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 1 & 0 \\
0 & 0 & 1 & 2 & 1 & -1 \\
\hline
1 & 1 & 0 & -1 & -1 & 1 \\
0 & 1 & 0 & -4 & -1 & 2 \\
0 & 0 & 1 & 2 & 1 & -1 \\
\hline
1 & 0 & 0 & 3 & 0 & -1 \\
0 & 1 & 0 & -4 & -1 & 2 \\
0 & 0 & 1 & 2 & 1 & -1 \\
\hline
\end{array}$$

Insbesondere sehen wir beim erreichen einer Zeilenstufenform, dass A vollen Rang hat und somit invertierbar ist.

3.7 Lineare Gleichungssysteme

3.7.1 Definition

Ein Gleichungssystem der Form

$$\begin{array}{ccccccc}
a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
\vdots & & \ddots & & \vdots & & \vdots \\
a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

mit gegebenen $a_{ij}, b_i \in K$ und gesuchten $x_1, \ldots, x_n$ wird als lineares Gleichungssystem (LGS) über K mit m Gleichungen und n Unbekannten bezeichnet.

Schreibt man $A = (a_{ij})$ sowie $x = (x_1, \ldots, x_n)^\top$ und $b = (b_1, \ldots, b_m)^\top$, so kann man das LGS auch als $Ax = b$ schreiben.

Die Matrix

$$(A \mid b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

wird als erweiterte Matrix A_{erw} des LGS bezeichnet.

Ein lineares Gleichungssystem heißt homogen, wenn b der Nullvektor ist. Sonst heißt es inhomogen.

3.7.2 Satz

Das lineare Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn der Rang von A und der Rang der erweiterten Matrix $(A \mid b)$ gleich sind.

Beweis:

Bezeichnen $a_1, \dots, a_n$ die Spalten der Matrix A, so ist das LGS äquivalent zu $x_1 a_1 + \cdots + x_n a_n = b$. Dies ist genau dann lösbar, wenn $b \in \operatorname{span}(a_1, \dots, a_n)$ gilt. Dies ist äquivalent zu $\operatorname{span}(a_1, \dots, a_n) = \operatorname{span}(a_1, \dots, a_n, b)$. Die letzte Gleichheit gilt genau dann, wenn die Räume gleiche Dimension haben. Sie ist somit zu $\mathrm{rk}A = \mathrm{rk}A_{\text{erw}}$ äquivalent. $\square$

3.7.3 Satz

Sei A eine Matrix. Es ist die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ ein Untervektorraum U_0. Ist l_0 eine Lösung des inhomogenen lineare Gleichungssystems $Ax = b$, so sind alle Lösungen des inhomogenen Systems durch

$$\{l_0 + u \colon u \in U_0\}$$

gegeben.

Beweis:

Die Lösungsmenge des homogenen Systems ist der Kern der durch A gegebenen linearen Abbildung und somit ein Unterraum U_0. Ist l_0 eine Lösung des inhomogenen Systems und $u \in U_0$, so gilt $A(l_0 + u) = Al_0 + Au = Al_0 = b$. Es ist also jeder Vektor obiger Menge eine Lösung.

Sei nun l eine zweite Lösung des inhomogenen Systems. Dann gilt $A(l - l_0) = Al - Al_0 = b - b = 0$. Es ist also $l - l_0 = u \in U_0$. Somit gilt $l = l_0 + u$. Jede weitere Lösung hat somit obige Form. $\square$

3.7.4 Bemerkung

Die algorithmische Lösung linearer Gleichungssysteme erfolgt durch Zeilenumformungen der erweiterten Matrix. Man kann direkt zeigen, dass dies Äquivalenzumformungen des Gleichungssystems sind, man kann dies aber auch als Linksmultiplikation der Gleichung $Ax = b$ mit invertierbaren Elementarmatrizen interpretieren. In jedem Fall ist klar, dass dies die Lösungsmenge nicht verändert.

3.7.5 Beispiel

Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Was ist die Lösung von $Ax = (-3, 0, 3)^\top$ in $\mathbb{R}^3$?

$$\begin{array}{ccc|c}
1 & 2 & 3 & -3 \\
4 & 5 & 6 & 0 \\
7 & 8 & 9 & 3 \\
\hline
1 & 2 & 3 & -3 \\
0 & -3 & -6 & 12 \\
0 & -6 & -12 & 24 \\
\hline
1 & 2 & 3 & -3 \\
0 & 1 & 2 & -4 \\
0 & 0 & 0 & 0
\end{array}$$

Somit ist das System lösbar. Die Lösungsmenge ist

$$\{(-3 - 3t - 2(-4 - 2t), -4 - 2t, t) : t \in \mathbb{R}\} = \{(5 + t, -4 - 2t, t) : t \in \mathbb{R}\}.$$

Man nennt $(5, -4, 0)$ auch Partikularlösung. Weiterhin ist $\{(t, -2t, t) : t \in \mathbb{R}\}$ die allgemeine Lösung des zugehörigen homogenen Systems.

Kapitel 4

Determinanten

4.1 Permutationen

4.1.1 Definition

Eine bijektive Abbildung $\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\}$ wird als *Permutation* bezeichnet. Die Menge aller Permutationen von $\{1,\dots,n\}$ wird mit S_n bezeichnet. Sie heißt die *symmetrische Gruppe* auf n Elementen. Die Gruppenverknüpfung ist durch die Verkettung von Abbildungen $\circ$ gegeben.

Man schreibt

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 1 & 3 & 6 & 5 \end{bmatrix} \text{ oder } (1\,2\,4\,3)(5\,6)$$

für die Permutation in S_6, welche $1,2,3,4,5,6$ auf $2,4,1,3,6,5$ abbildet. Die zweite Schreibweise heißt auch *Zykelschreibweise*. Ein Zykel der Länge 2 (d.h., eine Permutation die genau zwei Elemente vertauscht) heißt auch *Transposition*.

Ist $\sigma \in S_n$, so nennt man ein Paar (i,j) mit $i < j$ und $\sigma(i) > \sigma(j)$ eine *Fehlstellung*.

$$\#\{(i,j) \mid i < j \text{ und } \sigma(i) > \sigma(j)\}$$

heißt die Anzahl der Fehlstellungen. Man nennt eine Permutation *ungerade*, wenn die Anzahl der Fehlstellungen ungerade ist. Andernfalls heißt sie *gerade*. Man setzt

$$\operatorname{sgn}(\sigma) \coloneqq \begin{cases} 1 & \text{falls } \sigma \text{ gerade ist} \\ -1 & \text{falls } \sigma \text{ ungerade ist} \end{cases}$$

und nennt es das *Signum* der Permutation.

4.1.2 Bemerkung

Die Gruppe S_n hat $n!$ Elemente. Im Fall $n > 2$ ist S_n nicht abelsch. Dies sieht man am folgenden Beispiel

$$(1\,2) \circ (2\,3) = (1\,2\,3) \text{ und } (2\,3) \circ (1\,2) = (1\,3\,2)\,.$$

4.1.3 Beispiel

Die Transposition $(5\,6) \in S_{10}$ hat genau einen Fehlstellung $(5,6)$. Ihr Signum ist -1.

Ist $i < j$, so hat die Transposition (i,j) genau die Fehlstellungen $(i,i+1),\dots,(i,j-1),(i,j)$ und $(i+1,j),\dots,(j-1,j)$. Dies sind $2(j-i)-1$. Insbesondere ist jede Transposition ungerade.

4.1.4 Satz

Ist $\sigma \in S_n$, so gilt

$$\operatorname{sgn}(\sigma) = \prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}\,.$$

Beweis:

Sei $\sigma \in S_n$, dann gilt

$$\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\prod_{i<j} \sigma(j)-\sigma(i)}{\prod_{i<j} j-i} = \operatorname{sgn}(\sigma) \cdot \frac{\prod_{i<j} |\sigma(j)-\sigma(i)|}{\prod_{i<j} j-i}\,.$$

Da σ bijektiv ist, durchläuft $\{\sigma(i),\sigma(j)\}$ alle Paare genau dann einmal, wenn $\{i,j\}$ alle Paare einmal durchläuft. Somit stimmen im letzten Bruch Zähler und Nenner überein. Es folgt

$$\prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i} = \operatorname{sgn}(\sigma)\,.$$

$\square$

Anmerkung: Da $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$ gilt, kann man obiges Produkt auch als Produkt über alle 2-elementigen Teilmengen von $\{1,\dots,n\}$ verstehen. D.h.

$$\operatorname{sgn}(\sigma) = \prod_{\{i,j\}\in P} \frac{\sigma(j)-\sigma(i)}{j-i} \quad \text{mit } P = \{\{k,l\} \subset \{1,\dots,n\} \colon k \neq l\}.$$

4.1.5 Satz

Sind $\sigma,\tau \in S_n$, so gilt $\operatorname{sgn}(\sigma\circ\tau) = \operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\tau)$. D.h. $\operatorname{sgn}\colon S_n \to \{\pm 1\}$ ist ein Gruppenhomomorphismus.

Beweis:

$$\begin{aligned}
\operatorname{sgn}(\sigma\circ\tau) &= \prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{j-i} = \prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \frac{\tau(j)-\tau(i)}{j-i} \\
&= \left(\prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)}\right) \cdot \left(\prod_{i<j} \frac{\tau(j)-\tau(i)}{j-i}\right) \\
&= \prod_{i<j} \frac{\sigma(\tau(j))-\sigma(\tau(i))}{\tau(j)-\tau(i)} \cdot \operatorname{sgn}(\tau)
\end{aligned}$$

Durchläuft $\{i, j\}$ alle zweielementigen Teilmengen von $\{1, \dots, n\}$, so auch $\{\tau(i), \tau(j)\}$. Somit ist das erste Produkt der letzten Zeile gleich $\operatorname{sgn}(\sigma)$ und es folgt $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau)$. □

4.1.6 Folgerung

Jede Permutation $\sigma \in S_n$ kann als Produkt endlich vieler Transpositionen geschrieben werden. Sie ist genau dann gerade, wenn die Anzahl der Faktoren in dieser Darstellung gerade ist.

Beweis:

Wir zeigen zuerst durch Induktion nach n, dass eine solche Darstellung existiert. Im Fall $n \in \{1, 2\}$ ist nichts zu zeigen. Sei nun $n \geq 3$ und $\sigma \in S_n$ beliebig.

Im Fall $\sigma(n) = n$ kann $\sigma_{\{1,\dots,n-1\}}$ als Element von S_{n-1} aufgefasst werden und somit nach Induktionsvoraussetzung als Produkt von Transpositionen geschrieben werden.

Im Fall $\sigma(n) = m \neq n$ erfüllt $\tau := (n, m) \circ \sigma \in S_n$ die Gleichung $\tau(n) = n$ und τ kann nach dem soeben gezeigten Fall als Produkt von Transpositionen geschrieben werden. Mit $(n, m) \circ \tau = \sigma$ ist dann auch eine Darstellung von σ als Produkt von Transpositionen gefunden.

Die Behauptung über das Signum folgt mit Satz 4.1.5, da alle Transpositionen ungerade sind. □

4.1.7 Folgerung

Die Abbildung $S_n \to S_n, \sigma \mapsto \sigma^{-1}$ ist bijektiv. Zudem gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ für alle $\sigma \in S_n$.

4.1.8 Definition

Es wird

$$A_n := \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$$

als *alternierende Gruppe* bezeichnet. Insbesondere ist das Produkt zweier gerade Permutationen wieder gerade.

4.1.9 Bemerkung

Ist $\tau \in S_n \setminus A_n$ eine beliebige ungerade Permutation, so gilt

$$S_n = A_n \cup \tau A_n = A_n \cup A_n \tau \text{ und } A_n \cap A_n \tau = \emptyset\,.$$

Hierbei steht τA_n für die Menge $\{\tau \circ \sigma : \sigma \in A_n\}$.

4.2 Definition der Determinante

4.2.1 Definition

Sei K ein Körper und $n \in \mathbb{N}$. Eine Abbildung

$$\det\colon \operatorname{Mat}(n \times n, K) \to K, A \mapsto \det A$$

heißt *Determinante*, falls folgendes gilt:

1. Sie ist linear in jeder Spalte. D.h. schreibt man $A = (a_1, \dots, a_n)$ mit Spaltenvektoren $a_1, \dots, a_n \in K^n$, so gilt

$$\begin{aligned} &\det(a_1, \dots, a_{k-1}, u + v, a_{k+1}, \dots, a_n) \\ = \; &\det(a_1, \dots, a_{k-1}, u, a_{k+1}, \dots, a_n) + \det(a_1, \dots, a_{k-1}, v, a_{k+1}, \dots, a_n) \end{aligned}$$

   für alle $u, v \in K^n$ und

$$\det(a_1, \dots, a_{k-1}, \lambda u, a_{k+1}, \dots, a_n) = \lambda \det(a_1, \dots, a_{k-1}, u, a_{k+1}, \dots, a_n)$$

   für alle $u \in K^n, \lambda \in K$.

2. Sie ist *alternierend*, d.h. hat A zwei gleiche Spalten so gilt $\det A = 0$.

3. Sie ist *normiert*, d.h. es ist $\det E_n = 1$.

Man schreibt auch

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} := \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} .$$

4.2.2 Satz

Sei $\det\colon \mathrm{Mat}(n \times n, K) \to K$ eine Determinante, so gilt

1. Für $\lambda \in K$ ist $\det(\lambda A) = \lambda^n \det(A)$.

2. Ist eine Spalte von A Null, so ist $\det A = 0$.

3. Entsteht B aus A durch Vertauschung zweier Spalten, so ist $\det B = -\det A$.

4. Entsteht B aus A durch Addition des λ-fachen der i-ten Spalte zu j-ten Spalte ($i \neq j$), so ist $\det B = \det A$.

Insbesondere versteht man, wie elementare Spaltenoperationen die Determinante beeinflussen.

Beweis:

1. Skaliert man A mit λ, so skaliert man jede Spalte mit λ. n-faches Anwenden der Linearität (einmal für jede Spalte) liefert

$$\det(\lambda a_1, \dots, \lambda a_n) = \lambda \det(a_1, \lambda a_2, \dots, \lambda a_n) = \cdots = \lambda^n \det(a_1, \dots, a_n) .$$

2. Ist die i-te Spalte von A die Nullspalte, so folgt

$$\begin{aligned} 0 = 0 \cdot \det(a_1, \dots, a_n) &= \det(a_1, \dots, a_{i-1}, 0 \cdot a_i, a_{i+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) . \end{aligned}$$

3. Seien $i < j$ zwei Spaltenindizes. Wir betrachten $\det(a_1, \ldots, a_{i-1}, a_i + a_j, a_{i+1}, \ldots, a_{j-1}, a_i + a_j, a_{j+1}, \ldots, a_n)$. Wir notieren ab jetzt nur noch die i-te und die j-te Spalte. Es folgt

$$\begin{aligned}
0 &= \det(\ldots, a_i + a_j, \ldots, a_i + a_j, \ldots) \\
&= \det(\ldots, a_i, \ldots, a_i + a_j, \ldots) + \det(\ldots, a_j, \ldots, a_i + a_j, \ldots) \\
&= \det(\ldots, a_i, \ldots, a_i, \ldots) + \det(\ldots, a_j, \ldots, a_i, \ldots) \\
&\quad + \det(\ldots, a_i, \ldots, a_j, \ldots) + \det(\ldots, a_j, \ldots, a_j, \ldots) \\
&= \det(\ldots, a_j, \ldots, a_i, \ldots) + \det(\ldots, a_i, \ldots, a_j, \ldots)
\end{aligned}$$

da eine Determinante Null ist, sobald zwei Spaltenvektoren gleich sind.

4. Seien $i \neq j$ zwei Spaltenindizes. Wir erhalten

$$\begin{aligned}
&\det(a_1, \ldots, a_{j-1}, \lambda a_i + a_j, a_{j+1}, \ldots, a_n) \\
= {}&\lambda \det(a_1, \ldots, a_{j-1}, a_i, a_{j+1}, \ldots, a_n) \\
&+ \det(a_1, \ldots, a_{j-1}, a_j, a_{j+1}, \ldots, a_n) \\
= {}&\det(a_1, \ldots, a_n)
\end{aligned}$$

□

4.2.3 Beispiel

$$\begin{aligned}
\begin{vmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} &= - \begin{vmatrix} 2 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 1 & 1 \end{vmatrix} = - \begin{vmatrix} 2 & 1 & 0 \\ 1 & -1 & 3 \\ 0 & 0 & 1 \end{vmatrix} \\
&= - \begin{vmatrix} 3 & 1 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \end{vmatrix} = 3 \begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{vmatrix} = 3 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 3
\end{aligned}$$

4.2.4 Beispiel

Die Determinante einer oberen Dreiecksmatrix ist das Produkt ihrer Diagonalelemente

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{vmatrix} = a_{11} a_{22} \cdots a_{nn} \,.$$

Beweis:

Angenommen, es ist i der kleinste Index mit $a_{ii} = 0$. Dann kann durch Spaltenoperationen auf den ersten i Spalten eine Nullspalte erzeugt werden ohne die Determinante zu verändern. Somit ist $\det A = 0$.

Ist kein Diagonaleintrag Null, so erhalten wir mit Spaltenoperationen

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{vmatrix} = a_{11} a_{22} \cdots a_{nn} \det E_n \,.$$

□

4.2.5 Satz

Für $A \in \mathrm{Mat}(n \times n, K)$ ist $\det A = 0$ gleichbedeutend mit $\mathrm{rk}A < n$.

Beweis:

Durch Spaltenoperationen kann A auf Dreiecksform gebracht werden. Dabei ändert sich der Rang nicht und die Determinante höchstens im Vorzeichen. D.h. die Behauptung ist nur für Dreiecksmatrizen zu zeigen. Dies wurde bereits erledigt. $\square$

4.2.6 Bemerkung

Sind A_1, A_2 quadratische Matrizen mit n und m Zeilen sowie C eine $n \times m$-Matrix. Dann ist

$$A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$$

eine quadratische Matrix und es gilt

$$\det A = \det A_1 \det A_2 \,.$$

Nachweis: Man überlegt sich leicht, dass $\mathrm{rk}A_1 < n \implies \mathrm{rk}A < n + m$ gilt. Im Fall $\mathrm{rk}A_1 < n$ ist die Aussage somit gezeigt. Andernfalls kann durch Spaltenumformungen

$$\begin{aligned} \det A &= \det A_1 \det \begin{pmatrix} E_n & C \\ 0 & A_2 \end{pmatrix} = \det A_1 \det \begin{pmatrix} E_n & 0 \\ 0 & A_2 \end{pmatrix} \\ &= \det A_1 \det A_2 \det \begin{pmatrix} E_n & 0 \\ 0 & E_m \end{pmatrix} = \det A_1 \det A_2 \end{aligned}$$

erreicht werden. $\square$

4.3 Existenz und Eindeutigkeit

4.3.1 Theorem (Leibnizformel)

Sei K ein Körper und $n \in \mathbb{N}$. Dann gibt es genau eine Determinante

$$\det\colon \mathrm{Mat}(n \times n, K) \to K \,.$$

Für $A = (a_{ij}) \in \mathrm{Mat}(n \times n, K)$ ist diese durch

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

gegeben. Diese Summe mit $n!$ Summanden heißt Leibnizformel.

Beweis:

Wir zeigen zuerst, dass die durch die Leibnizformel definierte Abbildung alle abstrakten Eigenschaften hat. Damit ist dann die Existenz der Determinante bestätigt.

Sei $\sigma \in S_n$ beliebig. Dann ist die Abbildung

$$L_\sigma \colon \mathrm{Mat}(n \times n, K) \to K, (a_{i,j}) \mapsto a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

eine in jedem Spaltenvektor lineare Abbildung. Sei hierzu i ein Spaltenindex und $j \in \{1, \dots, n\}$ mit $i = \sigma(j)$. Dann gilt für $u, v \in K^n$ und $\lambda \in K$

$$\begin{aligned} &L_\sigma(a_1, \dots, a_{i-1}, u + \lambda v, a_{i+1}, \dots, a_n) \\ &= a_{1,\sigma(1)} \cdots a_{j-1,\sigma(j-1)} \cdot (u_j + \lambda v_j) \cdot a_{j+1,\sigma(j+1)} \cdots a_{n,\sigma(n)} \\ &= L_\sigma(a_1, \dots, a_{i-1}, u, a_{i+1}, \dots, a_n) + \lambda L_\sigma(a_1, \dots, a_{i-1}, v, a_{i+1}, \dots, a_n)\,. \end{aligned}$$

Da die Leibnizformel eine Linearkombination von Abbildungen $\{L_\sigma \colon \sigma \in S_n\}$ ist, folgt die Linearität der Leibnizformel in jeder Spalte der Matrix.

Nehmen wir an, dass die i-te und die j-te Spalte $(i \neq j)$ von A gleich sind. Dann können wir die disjunkte Zerlegung $S_n = A_n \cup (i\,j)A_n$ betrachten. Nun gilt

$$\begin{aligned} &\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in A_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} - \sum_{\sigma \in A_n} a_{1,((i\,j)\circ\sigma)(1)} \cdots a_{n,((i\,j)\circ\sigma)(n)} \\ &= \sum_{\sigma \in A_n} \big(a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} - a_{1,((i\,j)\circ\sigma)(1)} \cdots a_{n,((i\,j)\circ\sigma)(n)}\big) \end{aligned}$$

Da die i-te und die j-te Spalte von A gleich sind, sind in der letzten Summe alle Summanden null. Dies zeigt, dass die Leibnizformel alternierend ist.

Einsetzen der Einheitsmatrix in die Leibnizformel führt auf

$$\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \delta_{1,\sigma(1)} \cdots \delta_{n,\sigma(n)}\,.$$

Gibt es ein i mit $i \neq \sigma(i)$, so ist $\delta_{i,\sigma(i)} = 0$. Folglich ist höchstens der Summand mit $\sigma = \mathrm{id}$ von null verschieden. Dieser hat den Wert 1. Somit ist gezeigt, dass die Leibnizformel normiert ist.

Zur Eindeutigkeit: Sei $A = (a_1, \dots, a_n) = (a_{i,j}) \in \mathrm{Mat}(n \times n, K)$. Dann folgt mittels der Linearität in jeder Spalte

$$\begin{aligned} \det A &= \det(a_{1,1}e_1 + \dots + a_{n,1}e_n, a_2, \dots, a_n) \\ &= \sum_{i_1=1}^{n} a_{i_1,1} \det(e_{i_1}, a_{1,2}e_2 + \dots + a_{n,2}e_n, a_3, \dots, a_n) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{i_1,1} a_{i_2,2} \det(e_{i_1}, e_{i_2}, a_3, \dots, a_n) \\ &= \cdots \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} \det(e_{i_1}, e_{i_2}, \dots, e_{i_n})\,. \end{aligned}$$

Nutzen wir nun, dass die Gleichheit zweier Spalten impliziert, dass die Determinante Null ist, sind hier nur paarweise verschiedene Indizes $i_1, \dots, i_n$ zu betrachten und es folgt

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \cdot \operatorname{sgn}(\sigma) \\
&= \sum_{\sigma \in S_n} a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} \cdots a_{n,\sigma^{-1}(n)} \cdot \operatorname{sgn}(\sigma^{-1}) \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \cdot a_{1,\tau(1)} a_{2,\tau(2)} \cdots a_{n,\tau(n)} \,.
\end{aligned}$$

□

4.3.2 Beispiel

Es gilt

$$\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}$$

$$\begin{aligned}
\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} &= a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\
&\quad - a_{2,1}a_{1,2}a_{3,3} - a_{1,3}a_{1,3}a_{2,2} - a_{1,1}a_{2,3}a_{3,2}
\end{aligned}$$

4.3.3 Folgerung

Für $A \in \operatorname{Mat}(n \times n, K)$ gilt $\det A = \det A^\top$.

Insbesondere gelten alle Aussagen welche die Determinante mit den Spaltenvektoren einer Matrix in Verbindung bringen in entsprechender Art auch für die Zeilenvektoren.

Beweis:

Mit $A = (a_{i,j})$ und $B = (b_{i,j}) = A^\top$ folgt

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) a_{\tau(1),1} \cdots a_{\tau(n),n} \\
&= \sum_{\tau \in S_n} \operatorname{sgn}(\tau) b_{1,\tau(1)} \cdots b_{n,\tau(n)} = \det B = \det A^\top \,.
\end{aligned}$$

□

4.3.4 Bemerkung

Oft benötigt man Determinanten von Matrizen deren Einträge nicht in einem Körper liegen sondern beispielsweise Polynome sind. Da alle Rechnungen zum Beweis der Leibnizformel ohne Divisionen durchgeführt wurden übertragen sich alle Ergebnisse auch auf quadratische Matrizen deren Einträge in einem beliebigen kommutativen Ring mit 1 liegen.

Insbesondere gibt es für jeden kommutativen Ring R mit 1 eine durch die Leibnizformel definierte Abbildung

$$\det\colon \mathrm{Mat}(n \times n, R) \to R,$$

welche linear in jeder Zeile und Spalte, alternierend und normiert ist.

4.3.5 Multiplikationssatz für Determinanten

Sind $A, B \in \mathrm{Mat}(n \times n, K)$, so gilt

$$\det(A \cdot B) = \det A \det B\,.$$

Beweis:

Angenommen, der Rang einer der beiden Matrizen ist kleiner als n. Dann steht auf beiden Seiten der Gleichung 0 und die Behauptung ist in diesem Spezialfall gezeigt.

Sei nun A eine invertierbare Matrix, so betrachten wir die Abbildung

$$d\colon \mathrm{Mat}(n \times n, K) \to K, M \mapsto \frac{\det(A \cdot M)}{\det A}\,.$$

Die Abbildung d erfüllt $d(E_n) = 1$.

Sind $m_1, \ldots, m_n$ die Spalten der Matrix M, so sind $Am_1, \ldots, Am_n$ die Spalten der Matrix AM. Hieraus folgt: Sind die Spalten mit den Indizes i und j der Matrix M gleich, so sind auch die Spalten mit Indizes i und j von $A \cdot M$ gleich. Es folgt $d(M) = 0$ und d ist alternierend.

Weiterhin hängt die i-te Spalte von AM linear von der i-ten Spalte von M ab und alle anderen Spalten von AM bleiben bei Änderung der i-ten Spalte von M unverändert. Somit ist die Zuordnung $M \mapsto d(M)$ als Verkettung linearer Abbildungen linear in der i-ten Spalte von M.

Es folgt $d(M) = \det M$, da wir alle Eigenschaften der Definition der Determinante gezeigt haben. Es folgt die Behauptung. □

4.3.6 Folgerung

Ist $A \in \mathrm{Gl}_n(K)$, so gilt $\det(A^{-1}) = (\det A)^{-1}$.

4.3.7 Satz und Definition

Sei V ein endlich-dimensionaler K-Vektorraum und $\varphi\colon V \to V$ eine lineare Abbildung. Die Determinante der Darstellungsmatrix

$$\det M^{\mathcal{B}}_{\mathcal{B}}(\varphi)$$

ist unabhängig von der Wahl der Basis $\mathcal{B}$ von V. Sie heißt die Determinante des Endomorphismus φ.

Beweis:

Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Basen von V. Dann gilt

$$M_{\mathcal{A}}^{\mathcal{A}}(\varphi) = M_{\mathcal{A}}^{\mathcal{B}}(\mathrm{id}) M_{\mathcal{B}}^{\mathcal{B}}(\varphi) M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}) = M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id})^{-1} M_{\mathcal{B}}^{\mathcal{B}}(\varphi) M_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id})\,.$$

Hieraus folgt mit dem Multiplikationssatz

$$\det M_{\mathcal{A}}^{\mathcal{A}}(\varphi) = \det M_{\mathcal{B}}^{\mathcal{B}}(\varphi)\,.$$

□

4.4 Minoren

4.4.1 Definition

Ein Minor einer Matrix ist die Determinante einer quadratischen Untermatrix. Im Fall einer $n \times n$-Matrix A bezeichnen wir mit $A'_{i,j} \in \mathrm{Mat}((n-1) \times (n-1), K)$ die Matrix die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

Die Matrix $A^{\#} \in \mathrm{Mat}(n \times n, K)$, welche an Position (i, j) den Eintrag $(-1)^{i+j} \det A'_{j,i}$ hat, wird als Adjunkte (oder komplementäre Matrix) zu A bezeichnet. (Man beachte die Vertauschung von Zeilen- und Spaltenindex.)

4.4.2 Lemma

Sind $a_1, \ldots, a_n$ die Spalten der Matrix A, so gilt

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

Beweis:

Durch Spaltenoperation können in der Matrix mit den Spalten

$$a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n$$

alle Einträge der i-ten Zeile mit Ausnahme des Eintrags in Spalte j eliminiert werden. Dabei bleiben alle Einträge außerhalb der i-ten Zeile unverändert.

Bringt man jetzt durch $(i-1)$-maliges Vertauschen benachbarter Zeilen und $(j-1)$-maliges Vertauschen benachbarter Spalten den Eintrag an Position (i, j) an Position $(1, 1)$, so entsteht die Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & A'_{i,j} \end{pmatrix}.$$

Die Determinante dieser Matrix ist $\det A'_{i,j}$. Zusammen folgt

$$\det(a_1, \ldots, a_{j-1}, e_i, a_{j+1}, \ldots, a_n) = (-1)^{i+j-2} \det A'_{i,j} = (-1)^{i+j} \det A'_{i,j}\,.$$

□

4.4.3 Satz

Ist $A \in \mathrm{Mat}(n \times n, K)$, so gilt

$$A^{\#} \cdot A = A \cdot A^{\#} = (\det A) \cdot E_n\,.$$

Beweis:

Wir berechnen den Eintrag an Position (i,k) von $A^{\#} \cdot A$:

$$\begin{aligned}
\sum_{j=1}^{n} a^{\#}_{i,j} a_{j,k} &= \sum_{j=1}^{n} (-1)^{i+j} \det A'_{j,i} a_{j,k} \\
&= \sum_{j=1}^{n} a_{j,k} \det(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_n) \\
&= \det(a_1, \ldots, a_{i-1}, \left(\sum_{j=1}^{n} a_{j,k} e_j\right), a_{i+1}, \ldots, a_n) \\
&= \det(a_1, \ldots, a_{i-1}, a_k, a_{i+1}, \ldots, a_n) = \delta_{i,k} \det A
\end{aligned}$$

In gleicher Art untersucht man $A \cdot A^{\#}$. □

4.4.4 Bemerkung

Verwendet man obigen Satz, um die Inverse einer Matrix zu bestimmen, so spricht man von der *Cramerschen Regel* für die inverse Matrix.

Betrachtet man das Element an Position (i,i) von $A^{\#} \cdot A = A \cdot A^{\#}$, so erhält man

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{j,i} \det A'_{j,i} = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A'_{i,j}$$

für jedes $i \in \{1, \ldots, n\}$. Dies wird als *Laplace-Entwicklung* der Determinante nach der i-ten Spalte oder Zeile bezeichnet.

4.4.5 Beispiele

Für eine invertierbare 2×2-Matrix gilt:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Berechnung einer 3×3 Determinante durch Laplace-Entwicklung bezüglich der ersten Zeile:

$$\begin{vmatrix} 0 & 1 & 3 \\ 3 & 2 & 5 \\ 1 & 1 & 0 \end{vmatrix} = 0 \cdot \begin{vmatrix} 2 & 5 \\ 1 & 0 \end{vmatrix} - 1 \cdot \begin{vmatrix} 3 & 5 \\ 1 & 0 \end{vmatrix} + 3 \begin{vmatrix} 3 & 2 \\ 1 & 1 \end{vmatrix} = -1 \cdot (-5) + 3 \cdot 1 = 8\,.$$

Die Vorzeichen $(-1)^{i+j}$ werden oft als Schachbrettregel bezeichnet.

4.4.6 Cramersche Regel für lineare Gleichungssysteme

Ist $A = (a_1, \ldots, a_n)$ eine invertierbare Matrix. Dann gilt für die Lösung $x = (x_1, \ldots, x_n)$ das LGS $Ax = b$

$$x_i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A}\,.$$

Beweis:

x_i ist die i-te Komponente von $A^{-1}b$. Nach der Cramerschen Regel für die inverse Matrix gilt

$$\begin{aligned}
x_i &= \frac{(A^{\#}b)_i}{\det A} \\
&= \frac{1}{\det A}\sum_{j=1}^{n}(-1)^{i+j}\det A'_{j,i}b_j \\
&= \frac{1}{\det A}\sum_{j=1}^{n} b_j \det(a_1,\ldots,a_{i-1},e_j,a_{i+1},\ldots,a_n) \\
&= \frac{1}{\det A}\det(a_1,\ldots,a_{i-1},\sum_{j=1}^{n} b_j e_j,a_{i+1},\ldots,a_n) \\
&= \frac{1}{\det A}\det(a_1,\ldots,a_{i-1},b,a_{i+1},\ldots,a_n)\,.
\end{aligned}$$

□

4.5 Beispiele und Anwendungen

4.5.1 Die Vandermonde-Determinante

Sind $x_1,\ldots,x_n$ beliebig, so gilt

$$V(x_1,\ldots,x_n) := \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{-1} \end{vmatrix} = \prod_{1\le i<j\le n}(x_j-x_i)\,.$$

Beweis:

Durch Laplace-Entwicklung nach der letzten Zeile sieht man, dass $V(x_1,\ldots,x_n)$ ein Polynom in x_n vom Grad höchstens n mit Leitkoeffizient $V(x_1,\ldots,x_{n-1})$ ist. Dieses Polynom hat die Nullstellen $x_n = x_1,\ldots,x_n = x_{n-1}$. Dies zeigt

$$V(x_1,\ldots,x_n) = V(x_1,\ldots,x_{n-1})(x_n-x_1)\cdots(x_n-x_{n-1})\,.$$

Wiederholt man dieses Argument für $V(x_1,\ldots,x_{n-1})$, so entsteht die behauptete Formel. □

4.5.2 Definition

Ist $b_1,\ldots,b_n \in \mathbb{R}^n$ beliebig, so bezeichnet man

$$\{\lambda_1 b_1 + \cdots + \lambda_n b_n \colon \lambda_1,\ldots,\lambda_n \in [0,1]\}$$

als den von $b_1,\ldots,b_n$ aufgespannten Parallelkörper. Mit $\mu(b_1,\ldots,b_n)$ bezeichnet man das n-dimensionale Volumen des von $b_1,\ldots,b_n$ aufgespannten Körpers.

4.5.3 Bemerkung

Der geometrischen Anschauung für das Volumen eines Parallelkörpers entnimmt man:

1. $\mu(b_1, \ldots, b_n)$ ist unabhängig von der Reihenfolge $b_1, \ldots, b_n$.
2. $\mu(b_1, \ldots, b_n) = 0$, falls $b_1 = b_2$ gilt.
3. $\mu(\lambda b_1, \ldots, b_n) = |\lambda| \mu(b_1, \ldots, b_n)$
4. Für $\lambda \in \mathbb{R}$ gilt $\mu(b_1 + \lambda b_2, b_2, \ldots, b_n) = \mu(b_1, \ldots, b_n)$. (Scherung ändert das Volumen nicht.)
5. $\mu(e_1, \ldots, e_n) = 1$.

Hieraus folgt

$$\mu(b_1, \ldots, b_n) = |\det(b_1, \ldots, b_n)|\,.$$

Man bezeichnet $\det(b_1, \ldots, b_n)$ auch als das orientierte Volumen das Parallelkörpers.

4.5.4 Folgerung

Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, so gilt $\mu(Ab_1, \ldots, Ab_n) = |\det A| \mu(b_1, \ldots, b_n)$. D.h. die durch A gegebene lineare Abbildung ändert Volumina um den Faktor $|\det A|$.

4.5.5 Definition

Man bezeichnet

$$\mathrm{Sl}_n(K) := \{A \in \mathrm{Gl}_n(K) \mid \det A = 1\}$$

als *spezielle lineare Gruppe*. $\mathrm{Sl}_n(\mathbb{R})$ besteht aus allen Matrizen, die das orientierte Volumen nicht verändern.

Kapitel 5

Weiteres zu Unterräumen

5.1 Der Summenraum

5.1.1 Definition

Sind $U_1, \ldots, U_r \subset V$ Unterräume des K-Vektorraums V, so wird

$$U_1 + \cdots + U_r := \{u_1 + \cdots + u_r \colon u_1 \in U_1, \ldots, u_r \in U_r\}$$

als die Summe von $U_1, \ldots, U_r$ bezeichnet.

5.1.2 Satz

Der Summenraum $U_1 + \cdots + U_r$ ist der kleinste Unterraum von V, der $U_1, \ldots, U_r$ enthält.

Beweis:

Sei $U \subset V$ ein Unterraum der $U_1, \ldots, U_r$ enthält, so ist auch $U_1 + \cdots + U_r \subset U$, da U als Unterraum unter der Addition abgeschlossen ist. Es bleibt zu zeigen, dass $U_1 + \cdots + U_r$ wirklich ein Unterraum ist. Sind $u_1 + \cdots + u_r, v_1 + \cdots + v_r \in U_1 + \cdots + U_r$ mit $u_i, v_i \in U_i$, so gilt

$$\lambda(u_1 + \cdots + u_r) = (\lambda u_1) + \cdots + (\lambda u_r) \in U_1 + \cdots + U_r$$

und

$$(u_1 + \cdots + u_r) + (v_1 + \cdots + v_r) = (u_1 + v_1) + \cdots + (u_r + v_r) \in U_1 + \cdots + U_r$$

da $U_1, \ldots, U_r$ als Unterräume unter Addition und Skalarmultiplikation abgeschlossen sind. $\square$

5.1.3 Satz

Sind $U_1, U_2 \subset V$ zwei endlich-dimensionale Unterräume, so gilt

$$\dim U_1 + \dim U_2 = \dim(U_1 + U_2) + \dim(U_1 \cap U_2)\,.$$

Beweis:

Sei $(a_1, \ldots, a_k)$ eine Basis von $U_1 \cap U_2$. Durch Basisergänzung finden wir $b_1, \ldots, b_m$ und $c_1, \ldots, c_n$ so, dass $(a_1, \ldots, a_k, b_1, \ldots, b_m)$ eine Basis von U_1 und $(a_1, \ldots, a_k, c_1, \ldots, c_n)$ eine Basis von U_2 ist. Die Behauptung ist gezeigt, wenn wir zeigen können, dass $\mathcal{B} \coloneqq (a_1, \ldots, a_k, b_1, \ldots, b_m, c_1, \ldots, c_n)$ eine Basis von $U_1 + U_2$ ist. Es ist klar, dass $\mathcal{B}$ ein Erzeugendensystem von $U_1 + U_2$ ist. Wir haben die lineare Unabhängigkeit zu zeigen. Angenommen es ist $\lambda_1 a_1 + \cdots + \lambda_k a_k + \mu_1 b_1 + \cdots + \mu_m b_m + \nu_1 c_1 + \cdots + \nu_n c_n = 0$. Dann folgt

$$-(\mu_1 b_1 + \cdots + \mu_m b_m) = \lambda_1 a_1 + \cdots + \lambda_k a_k + \nu_1 c_1 + \cdots + \nu_n c_n \in U_1 \cap U_2$$

Da $(a_1, \ldots, a_k)$ eine Basis von $U_1 \cap U_2$ und $(a_1, \ldots, a_k, c_1, \ldots, c_n)$ linear unabhängig sind, folgt $\nu_1 = \cdots = \nu_u = 0$. Es bleibt $\lambda_1 a_1 + \cdots + \lambda_k a_k + \mu_1 b_1 + \cdots + \mu_m b_m = 0$. Da $(a_1, \ldots, a_k, n_1, \ldots, b_m)$ eine Basis von U_1 ist, folgt $\lambda_1 = \cdots = \lambda_k = \mu_1 = \cdots = \mu_m = 0$. □

5.1.4 Definition

Sind $U_1, U_2 \subset V$ Unterräume. Gilt $U_1 \cap U_2 = \{0\}$, so nennt man die Summe $U_1 + U_2$ direkt. In Zeichen $U_1 \oplus U_2$.

Es gilt $\dim U_1 \oplus U_2 = \dim U_1 + \dim U_2$.

Allgemein nennt man $U_1 + \cdots + U_k$ direkt, wenn

$$U_i \cap \bigoplus_{j \neq i} U_j = \{0\}$$

für alle i gilt.

5.1.5 Satz

Für Unterräume $U_1, U_2 \subset V$ sind äquivalent:

1. $U_1 \cap U_2 = \{0\}$
2. Jedes $u \in U_1 + U_2$ ist eindeutig als $u_1 + u_2$ mit $u_1 \in U_1, u_2 \in U_2$ darstellbar.
3. Zwei von Null verschiedene Vektoren $u_1 \in U_1, u_2 \in U_2$ sind linear unabhängig.

Beweis:

$1 \implies 2$: Angenommen es ist $u_1 + u_2 = v_1 + v_2$ mit $u_1, v_1 \in U_1, u_2, v_2 \in U_2$. Dann folgt $u_1 - v_1 = v_2 - u_2 \in U_1 \cap U_2 = \{0\}$. Dies zeigt $u_1 - v_1 = v_2 - u_2 = 0$. Somit gilt $u_1 = v_1$ und $u_2 = v_2$. D.h., die Darstellung ist eindeutig.

$2 \implies 3$: Angenommen es ist $\lambda u_1 + \mu u_2 = 0$ mit $0 \neq u_1 \in U_1, 0 \neq u_2 \in U_2$. Dann folgt

$$\lambda u_1 + 0 = 0 + (-\mu) u_2 \in U_1 + U_2 \,.$$

Da die Darstellung als Summe eindeutig ist, folgt $\lambda u_1 = 0, 0 = -\mu u_2$. Dies zeigt $\lambda = \mu = 0$, da u_1 und u_2 beide nicht Null sind.

$3 \implies 1$: Sei $u_1 = u_2 \in U_1 \cap U_2$. Dann sind u_1, u_2 linear abhängig. Da wir 3 voraussetzen folgt $u_1 = u_2 = 0$. Somit ist $U_1 \cap U_2 = \{0\}$. □

5.1.6 Bemerkung

Sind $(b_1, \dots, b_{i_1}), (b_{i_1+1}, \dots, b_{i_2}) \dots, (b_{i_{k-1}+1}, \dots, b_{i_k})$ Basen der Unterräume $U_1, \dots, U_k$. Es ist $(b_1, \dots, b_{i_k})$ ein Erzeugendensystem des Summenraums $U_1 + \dots + U_k$. Die Summe ist genau dann direkt, wenn $(b_1, \dots, b_{i_k})$ linear unabhängig ist.

Beweis:

Übung. $\square$

5.1.7 Satz

Sei $U \subset V$ eine Unterraum eines endlichdimensionalen Vektorraums. Dann gibt es einen Unterraum $W \subset V$ mit $U \oplus W = V$.

Beweis:

Sei $a_1, \dots, a_k$ eine Basis von U. Basisergänzung liefert $a_{k+1}, \dots, a_n$, sodass $a_1, \dots, a_n$ eine Basis von V ist. Wir wählen $W \coloneqq \operatorname{span}(a_{k+1}, \dots, a_n)$. $\square$

5.1.8 Beispiel

Sei $S \coloneqq \{M \in \operatorname{Mat}(n \times n, \mathbb{R}) \mid M = M^\top\}$ und $A \coloneqq \{M \in \operatorname{Mat}(n \times n, \mathbb{R}) \mid M = -M^\top\}$. Dann gilt

$$\operatorname{Mat}(n \times n, \mathbb{R}) = S \oplus A\,.$$

Beweis:

Es ist klar, dass S und A Unterräume sind. Der Schnitt besteht aus allen Matrizen M mit $-M^\top = M = M^\top$. Dies ist der Nullraum.

Sei nun $M \in \operatorname{Mat}(n \times n, \mathbb{R})$ beliebig. Es ist $M = \frac{1}{2}(M + M^\top) + \frac{1}{2}(M - M^\top)$. Da $\frac{1}{2}(M + M^\top) \in S$ und $\frac{1}{2}(M - M^\top) \in A$ gilt, folgt die Behauptung. $\square$

5.2 Unterraumketten

5.2.1 Definition

Sei V ein Vektorraum. Seien $U_0, \dots, U_k$ Unterräume von V mit

$$U_0 \subset U_1 \subset \dots \subset U_k$$

heißt *Unterraumkette*. Gilt zudem $U_0 \neq U_1 \neq \dots \neq U_k$, so spricht man von einer *Flagge* oder einer *Fahne*.

Eine Fahne heißt *vollständig*, wenn $\dim U_i = i$ und $k = \dim V$ gilt. In anderen Worten, wenn man keinen weiteren Unterraum in die Flagge einschieben kann.

5.2.2 Bemerkung

Ist V endlich-dimensional, so kann jede Fahne zu einer vollständigen Fahne ergänzt werden.

5.2.3 Konstruktion

Ist $(b_1, \dots, b_n)$ eine Basis von V, so kann man ihr die Flagge $\{0\} \subset \operatorname{span}(b_1) \subset \operatorname{span}(b_1, b_2) \subset \cdots \operatorname{span}(b_1, \dots, b_n)$ zuordnen.

5.2.4 Bemerkung

Ist $U_0 \subset \cdots \subset U_n$ eine vollständige Flagge in V, so gibt es eine Basis $(b_1, \dots, b_n)$ mit $U_i = \operatorname{span}(b_1, \dots, b_i)$.

Beweis:

Dies zeigt man durch wiederholte Basisergänzung. □

5.2.5 Bemerkung

Ist $\varphi\colon V \to W$ eine lineare Abbildung und $U_1 \subset \cdots, \subset U_k$ eine Unterraumkette in V, so ist $\varphi(U_1) \subset \cdots \subset \varphi(U_k)$ eine Unterraumkette in W.

Beweis:

Klar. □

5.2.6 Beispiel

Ist φ ein Endomorphismus, so ist

$$0 \subset \ker \varphi \subset \ker \varphi^2 \subset \ker \varphi^3 \cdots$$

eine Unterraumkette.

5.2.7 Satz

Es sei $(e_1, \dots, e_n)$ die Standardbasis des K^n und $U_i := \operatorname{span}(e_1, \dots, e_i)$ die Unterräume der Standardflagge.

Dann ist $\{M \in \operatorname{Mat}(n \times n, K) \mid MU_i = U_i \text{ für alle } i = 1, \dots, n\}$ die Menge der invertierbaren oberen Dreiecksmatrizen.

Beweis:

Übung. □

5.2.8 Satz

Es gilt

$$\begin{aligned} &\{M \in \operatorname{Mat}(n \times n, K) \mid MU \subset U \text{ für jeden Unterraum } U \subset K^n\} \\ = &\{\lambda E_n \mid \lambda \in K\} \\ = &\{M \in \operatorname{Mat}(n \times n, K) \mid MA = AM \text{ für alle } A \in \operatorname{Mat}(n \times n, K)\} \end{aligned}$$

Beweis:

Sei $M \in \{M \in \mathrm{Mat}(n \times n, K) \mid MU \subset U$ für jeden Unterraum $U\}$, so gilt speziell $M\mathrm{span}(e_i) \subset \mathrm{span}(e_i)$. Hieraus folgt $Me_i \in \mathrm{span}(e_i)$. Dies zeigt, dass $M = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ eine Diagonalmatrix ist. In gleicher Art zeigt man $M(e_1 + e_i) = \lambda_1 e_1 + \lambda_i e_i \in \mathrm{span}(e_1 + e_i)$. Hieraus folgt $\lambda_1 = \lambda_i$.

Sei nun $M \in \{M \in \mathrm{Mat}(n \times n, K) \mid MA = AM$ für alle $A \in \mathrm{Mat}(n \times n, K)\}$. Wir wählen $A = E_{i,j}$ (die Matrix, die an der Position (i,j) eine Eins und sonst nur Nullen hat).

Es ist $ME_{i,j}$ die Matrix, die in der j-ten Spalte die i-te Spalte von M und sonst nur Nullen hat. Es ist $E_{i,j}M$ die Matrix, die in der i-ten Zeile die j-te Zeile von M und sonst nur Nullen hat. Die Gleichheit dieser Produkte zeigt, die i-te Spalte von M nur einen Eintrag in der i-ten Zeile und die j-te Zeile nur einen Eintrag in der j-te Spalte hat. Zudem sind die Einträge an den Positionen (i,i) und (j,j) gleich. Folglich haben alle Matrizen die Form λE_n.

Die jeweils fehlende zweite Inkusionsrichtung ist offensichtlich. $\square$

5.2.9 Definition

Man nennt $\{\lambda E_n \colon \lambda \in K\}$ das *Zentrum* des Matrizenrings $\mathrm{Mat}(n \times n, K)$, da dies die Matrizen sind, die mit allen anderen Matrizen kommutieren.

5.3 Der Quotientenraum

5.3.1 Satz

Sei $U \subset V$ ein Unterraum des Vektorraums V. Dann wird durch

$$u \sim_U v :\Longleftrightarrow u - v \in U$$

eine Äquivalenzrelation auf V definiert.

Beweis:

Da $0 \in U$ gilt, folgt $v \sim_U v$ für alle $v \in V$. Ist $u \sim_U v$, so folgt $u - v \in U$ und auch $v - u \in U$. Dies zeigt $v \sim_U u$. Sei $u \sim_U v$ und $v \sim_U w$, so gilt $u - v, v - w \in U$. Dies liefert $u - w \in U$ und folglich $u \sim_U w$. $\square$

5.3.2 Beispiel

Es gilt $v \sim_U 0$ genau dann, wenn $v \in U$ ist.

5.3.3 Satz

Ist $U \subset V$ ein Unterraum und $x \in V$, so ist

$$x + U := \{x + u \mid u \in U\} = \{y \in V \mid y \sim_U x\}.$$

D.h., die Äquivalenzklassen der Relation $\sim_U$ sind genau die Translate von U.

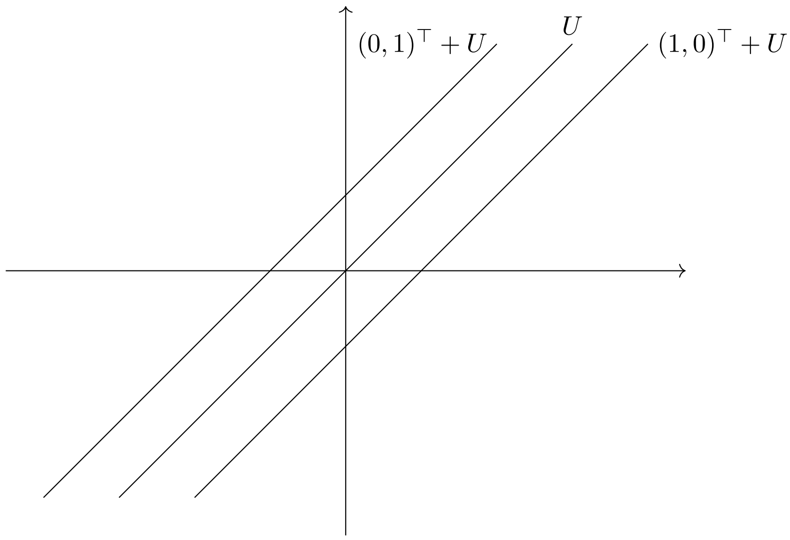
Beweis:

$$\{y \in V \mid y \sim_U x\} = \{y \in V \mid y - x = u \in U\} = \{x + u \colon u \in U\}$$

□

5.3.4 Zeichnung

Im Fall $U = \operatorname{span}((1,1)) \subset \mathbb{R}^2$ sehen die Klassen wie folgt aus:



Geometrisch gesprochen sind die Äquivalenzklassen genau die zu U parallelen Geraden.

5.3.5 Definition

Sei $U \subset V$ ein Unterraum, so setzen wir

$$V/U \coloneqq \{x + U \colon x \in V\} = \text{Menge der Äquivalenzklassen von } \sim_U \ .$$

5.3.6 Satz

Sei $U \subset V$ ein Unterraum des K-Vektorraums V. Auf V/U wird durch

$$(x+U) + (y+U) \coloneqq (x+y) + U, \lambda(x+U) \coloneqq (\lambda x) + U$$

eine Addition und eine Skalarmultiplikation erklärt. Sie gibt V/U die Struktur eines K-Vektorraums. Er wird als *Quotientenraum* bezeichnet.

Beweis:

Wir haben zu zeigen, dass die obigen Verknüpfungen wohldefiniert sind. Sei hierzu $x_1, x_2, y_1, y_2 \in V$ mit $x_1 + U = x_2 + U$ und $y_1 + U = y_2 + U$. Dann ist $x_1 - x_2 = u, y_1 - y_2 = u' \in U$ und es gilt

$$\begin{aligned}(x_1 + y_1) + U &= (x_2 + u + y_2 + u') + U = (x_2 + y_2) + (u + u') + U \\ \lambda x_1 + U &= \lambda(x_2 + u) + U = \lambda x_2 + \lambda u + U.\end{aligned}$$

Daher sind Addition und Skalarmultiplikation repräsentantenunabhängig und somit wohldefiniert. Alle Rechenregeln (Kommutativität, Assoziativität, etc.) gelten für V/U, da sie bereits in V gelten. $\square$

5.3.7 Satz

Ist $U \subset V$ ein Untervektorraum. Dann ist die Abbildung $\pi\colon V \to V/U, v \mapsto v+U$ surjektiv und linear. Ihr Kern ist U.

Beweis:

Die Linearität und die Surjektivität folgen sofort aus der Definition der Verknüpfungen auf V/U. Der Kern von π ist U, da $v \sim_U 0$ genau dann gilt, wenn $v \in U$ ist. $\square$

5.3.8 Folgerung

Ist $U \subset V$ ein Unterraum eines endlich-dimensionalen Vektorraums. Dann folgt

$$\dim(U) + \dim(V/U) = \dim(V)$$

aus der Dimensionsformel.

5.3.9 Folgerung

Ist $(b_1, \dots, b_k)$ eine Basis von U und $(b_1, \dots, b_k, b_{k+1}, \dots, b_n)$ eine Basis von V, so ist $(b_{k+1} + U, \dots, b_n + U)$ eine Basis von V/U.

5.3.10 Homomorphiesatz

Es sei $\varphi\colon V \to W$ eine lineare Abbildung und $U \subset \operatorname{Ker}\varphi$ ein Unterraum sowie $\pi\colon V \to V/U$ die Projektion. Dann gibt es eine eindeutig bestimmte lineare Abbildung $\psi\colon V/U \to W$ mit $\varphi = \psi \circ \pi$.

Beweis:

Wir zeigen Existenz und Eindeutigkeit der Abbildung ψ: Sei $v_1 + U = v_2 + U \in V/U$. Somit ist $u := v_1 - v_2 \in U$. Dann gilt

$$\varphi(v_1) = \varphi(v_2 + u) = \varphi(v_2) + \varphi(u) = \varphi(v_2)\,.$$

Die Setzung $\psi(v_1 + U) := \varphi(v_1)$ ist somit unabhängig von der Wahl des Repräsentanten und daher wohldefiniert. Die Linearität von ψ folgt aus der Linearität von φ:

$$\begin{aligned}
\psi(v_1 + U + v_2 + U) &= \psi((v_1 + v_2) + U) = \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \\
&= \psi(v_1 + U) + \psi(v_2 + U) \\
\psi(\lambda(v_1 + U)) &= \psi((\lambda v_1) + U) = \varphi(\lambda v_1) = \lambda\varphi(v_1) = \lambda\psi(v_1 + U)
\end{aligned}$$

$\square$

Kapitel 6

Eigenwerte

Motivation

Diagonalmatrizen sind viel einfacher als voll besetzte Matrizen. Kann man einen Endomorphismus durch Wahl einer passenden Basis mit einer Diagonalmatrix darstellen?

6.1 Definition und Beispiele

6.1.1 Definition

Es sei $\varphi\colon V \to V$ eine lineare Abbildung. Angenommen es gilt

$$\varphi(v) = \lambda v$$

für einen Vektor $v \in V \setminus \{0\}$ und $\lambda \in K$, so wird λ als *Eigenwert* und v als *Eigenvektor* von φ bezeichnet.

Die Menge

$$E(\varphi, \lambda) \coloneqq \{v \in V \mid \varphi(v) = \lambda v\} = \ker(\varphi - \lambda \mathrm{id})$$

wird als *Eigenraum* zum Eigenwert λ bezeichnet.

6.1.2 Beispiele

1. Der Kern einer linearen Abbildung ist ihr Eigenraum zum Eigenwert 0.

2. Sei $U \subset \mathbb{R}^2$ ein eindimensionaler Unterraum und $\pi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die senkrechte Projektion auf U. Dann ist U Eigenraum von φ zum Eigenwert 1.

3. Eine Drehung $\mathbb{R}^2 \to \mathbb{R}^2$ mit einem Drehwinkel $\alpha \in (0, \pi)$ hat keinen Eigenvektor und keinen Eigenwert.

4. Sei D die Differentiationsabbildung reeller Funktionen. Dann gilt $D(x \mapsto \exp(\lambda x)) = (x \mapsto \lambda \exp(\lambda x))$. Somit ist jedes $\lambda \in \mathbb{R}$ ein Eigenwert von D und $x \mapsto \exp(\lambda x)$ ein Eigenvektor von D.

5. Sei $M = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ eine Diagonalmatrix. Dann gilt für $\varphi_M \colon K^n \to K^n, x \mapsto Mx$
$$\varphi_M(e_i) = \lambda_i e_i \,.$$
Somit sind $\lambda_1, \dots, \lambda_n$ Eigenwerte und $e_1, \dots, e_n$ Eigenvektoren von φ_M.

6.1.3 Bemerkung

1. Es ist $E(\varphi, \lambda) \subset V$ ein Unterraum.
2. λ ist Eigenwert von φ g.d.w. $E(\varphi, \lambda) \neq \{0\}$.
3. $E(\varphi, \lambda) \setminus \{0\}$ ist die Menge der zu λ gehörenden Eigenvektoren.
4. Sind $\lambda_1, \lambda_2 \in K$ verschieden, so ist $E(\varphi, \lambda_1) \cap E(\varphi, \lambda_2) = \{0\}$.

6.1.4 Definition

Sei V ein endlichdimensionaler Vektorraum und $\varphi \colon V \to V$ ein Endomorphismus. Es φ heißt *diagonalisierbar*, wenn es eine Basis $\mathcal{B}$ von V gibt, sodass $M_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix ist.

6.1.5 Satz

Sei V ein endlichdimensionaler Vektorraum und $\varphi \colon V \to V$ ein Endomorphismus. Es φ ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von φ hat.

Beweis:

Sei $\mathcal{B}$ eine Basis von V, sodass $M_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix ist. Dann sind alle Vektoren von $\mathcal{B}$ Eigenvektoren.

Ist andererseits $\mathcal{B}$ eine Basis von $\mathcal{B}$ aus Eigenvektoren von φ, so ist $M_{\mathcal{B}}(\varphi)$ eine Diagonalmatrix. □

6.1.6 Definition

Eine Matrix $A \in \operatorname{Mat}(n \times n, K)$ heißt diagonalisierbar, wenn es ein $S \in \operatorname{Gl}_n(K)$ gibt, sodass SAS^{-1} eine Diagonalmatrix ist.

Dies ist dazu äquivalent, dass die durch A gegebene lineare Abbildung diagonalisierbar ist.

6.1.7 Lemma

Seien $\lambda_1, \dots, \lambda_n$ paarweise verschiedene Eigenwerte einer linearen Abbildung φ und $v_1, \dots, v_n$ zugehörige Eigenvektoren. Dann ist $(v_1, \dots, v_n)$ linear unabhängig.

Beweis:

Angenommen die Aussage wäre falsch, so gäbe es ein minimales Gegenbeispiel. D.h., es gibt ein minimales n, sodass $(v_1, \dots, v_n)$ linear abhängig ist und jedes echte Teilsystem dieser Vektoren linear unabhängig ist.

Dann gibt es $(\alpha_1, \dots, \alpha_n)$ mit

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0 \text{ und } \alpha_i \neq 0.$$

Wenden wir hierauf φ an, so folgt

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n = 0\,.$$

Zieht man das λ_n-fache der ersten Gleichung von der zweiten ab, so folgt

$$\alpha_1(\lambda_1 - \lambda_n)v_1 + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1} + \alpha_n(\lambda_n - \lambda_n)v_n = 0$$

Folglich ist $(v_1, \dots, v_{n-1})$ linear abhängig. Dies ist ein Widerspruch zur Minimalität des Gegenbeispiels. □

6.1.8 Folgerung

Ist V ein endlich-dimensionaler Vektorraum und $\varphi \in \mathrm{End}(V)$, so hat φ höchstens $\dim V$ verschiedene Eigenwerte.

6.1.9 Folgerung

Hat eine $(n \times n)$-Matrix n paarweise verschiedene Eigenwerte, so ist sie diagonalisierbar.

Beweis:

Zu jedem Eigenwert gibt es mindestens einen Eigenvektor. Nach obigem Lemma sind sie linear unabhängig und formen somit eine Basis des K^n. □

6.2 Das charakteristische Polynom

In diesem Abschnitt sind alle Vektorräume endlich-dimensional.

6.2.1 Bemerkung

Ist φ ein Endomorphismus des endlichdimensionalen Vektorraums V und $\lambda \in K$. Es ist λ genau dann ein Eigenwert von φ, wenn $\det(\varphi - \lambda \mathrm{id}) = 0$ gilt.

Beweis:

λ ist genau dann Eigenwert von φ, wenn $E(\varphi, \lambda) = \mathrm{Ker}(\varphi - \lambda \mathrm{id}) \neq \{0\}$ gilt. Dies ist zu $\det(\varphi - \lambda \mathrm{id}) = 0$ äquivalent. □

6.2.2 Definition

Ist A eine $(n \times n)$-Matrix, so heißt

$$\chi_A(X) := \det(A - X \cdot E_n) \in K[X]$$

das charakteristische Polynom von A.

6.2.3 Lemma

Sei $A \in \operatorname{Mat}(n \times n, K)$ und $S \in \operatorname{Gl}_n(K)$ sowie $B := SAS^{-1}$. Dann gilt

$$\chi_A(X) = \chi_B(X)\,.$$

D.h., ähnliche Matrizen haben gleiche charakteristische Polynome.

Beweis:

$$\begin{aligned}\chi_B(X) &= \det(SAS^{-1} - X \cdot E_n) = \det(SAS^{-1} - X \cdot SE_nS^{-1}) \\ &= \det(S(A - X \cdot E_n)S^{-1}) = \det(S)\det(A - X \cdot E_n)\det(S^{-1}) \\ &= \chi_A(X)\end{aligned}$$

□

6.2.4 Definition

Sei φ ein Endomorphismus eines endlichdimensionalen K-Vektorraums V, so heißt

$$\chi_\varphi(X) := \det(\varphi - X \cdot \mathrm{id}) \in K[X]$$

das charakteristische Polynom von φ.

6.2.5 Bemerkung

Die Eigenwerte sind genau die Nullstellen des charakteristischen Polynoms.

6.2.6 Beispiel

$$A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$$

Es ergibt sich

$$\begin{aligned}\chi_A(X) = \begin{vmatrix} -1-X & 6 \\ -1 & 4-X \end{vmatrix} &= (-1-X)(4-X) + 6 \\ = X^2 - 3X + 2 &= (X-1)(X-2)\,.\end{aligned}$$

Die Eigenwerte sind 1 und 2. Lösungen von $(A - E_n)x = 0$ und $(A - 2E_n)x = 0$ sind $b_1 := (3,1)^\top$ und $b_2 := (2,1)^\top$. Somit ist $\mathcal{B} = (b_1, b_2)$ eine Basis des $\mathbb{R}^2$ aus Eigenvektoren von A. Weiterhin ist A diagonalisierbar. Es gilt

$$SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \text{ mit } S^{-1} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix},\ S = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}.$$

6.2.7 Bemerkung

Ist $A = (a_{i,j})$ eine $(n \times n)$-Matrix, so ist

$$\begin{aligned}\chi_A(X) &= \begin{vmatrix} a_{1,1} - X & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} - X & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} - X \end{vmatrix} \\ &= (-1)^n \left(X^n - (a_{1,1} + a_{2,2} + \cdots + a_{n,n})X^{n-1} + \cdots + \det(A)X^0\right) .\end{aligned}$$

Man nennt $a_{1,1} + a_{2,2} + \cdots + a_{n,n} = \mathrm{Tr}(A)$ die *Spur* von A.

6.2.8 Vietascher Wurzelsatz

Sind $\lambda_1, \ldots, \lambda_n$ die Nullstellen des Grad-n-Polynoms $f = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$, so gilt

$$\begin{aligned} a_{n-1} &= -(\lambda_1 + \cdots + \lambda_n) \\ a_{n-2} &= (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \cdots + \lambda_{n-1}\lambda_n) \\ &\vdots \\ a_{n-k} &= (-1)^k \sum_{1 \le i_1 < \cdots < i_k \le n} \lambda_{i_1} \cdots \lambda_{i_k} \\ &\vdots \\ a_0 &= (-1)^n \lambda_1 \cdots \lambda_n \end{aligned}$$

Beweis:

Dieses erhält man durch Ausmultiplizieren von $(X - \lambda_1) \cdots (X - \lambda_n)$. $\square$

6.2.9 Folgerung

Die Spur einer quadratischen Matrix ist die Summe ihrer Eigenwerte (mit Vielfachheit).

6.2.10 Beispiel

Sei

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

dann ist $\chi_A(X) = (-1)^n(X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0)$. D.h. jedes Polynom vom Grad n mit Leitkoeffizient $(-1)^n$ ist charakteristische Polynom einer passend gewählten Matrix.

Nachweis: Übung.

6.3 Diagonalisierbarkeit

In diesem Abschnitt sind alle Vektorräume endlich-dimensional.

6.3.1 Definition

Es sei λ der Eigenwert eines Endomorphismus oder einer Matrix. Die Vielfachheit der Nullstelle λ des charakteristischen Polynoms wird als *algebraische Vielfachheit* von λ bezeichnet. Die Dimension des Eigenraums zum Eigenwert λ heißt *geometrische Vielfachheit* des Eigenwerts.

6.3.2 Bemerkung

Sei φ ein Endomorphismus eines n-dimensionalen Vektorraums. Dann hat das charakteristische Polynom von φ Grad n. Es zerfällt genau dann vollständig in Linearfaktoren, wenn die Summe der algebraischen Vielfachheiten aller Eigenwerte n ist. Andernfalls ist die Summe der Vielfachheiten kleiner.

6.3.3 Satz

Ist λ der Eigenwert eines Endomorphismus φ von V. Dann ist die geometrische Vielfachheit von λ höchstens so groß wie die algebraischen Vielfachheit.

Beweis:

Sei $b_1, \dots, b_k$ eine Basis des Eigenraums zum Eigenwert λ. Wir ergänzen diese zu einer Basis $\mathcal{B} = (b_1, \dots, b_n)$ von V. Dann erhalten wir

$$A \coloneqq M_{\mathcal{B}}(\varphi) = \begin{pmatrix} \lambda & \cdots & 0 & * & \cdots & * \\ & \ddots & & \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda & * & \cdots & * \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix}.$$

Entwickelt man $\chi_A(X) = \det(A - X \cdot E_n)$ nach den ersten k Spalten, so sieht man, dass λ mindestens k-fache Nullstelle des charakteristischen Polynoms von A ist. □

6.3.4 Satz

Sind $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte des Endomorphismus φ und $\mathcal{B}_1 \coloneqq (b_1, \dots, b_{i_1}), \mathcal{B}_2 \coloneqq (b_{i_1+1}, \dots, b_{i_2}), \dots, \mathcal{B}_k \coloneqq (b_{i_{k-1}+1}, \dots, b_{i_k})$ Basen der Eigenräume $E(\varphi, \lambda_1), \dots, E(\varphi, \lambda_k)$, so ist $(b_1, \dots, b_{i_k})$ linear unabhängig.

In anderen Worten: Die Summe $E(\varphi, \lambda_1) + \cdots + E(\varphi, \lambda_k)$ ist direkt.

Beweis:

Sei $0 = \mu_1 b_1 + \cdots + \mu_{i_k} b_{i_k}$ eine lineare Relation. Wir gruppieren die Summanden wie folgt

$$\begin{aligned} v_1 &:= \mu_1 b_1 + \cdots + \mu_{i_1} b_{i_1} \\ &\vdots \\ v_k &:= \mu_{i_{k-1}+1} b_{i_{k-1}+1} + \cdots + \mu_{i_k} b_{i_k} . \end{aligned}$$

Dann ist v_i im Eigenraum zum Eigenwert λ_i. Weiterhin ist $v_1 + \cdots + v_k = 0$. Da Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, folgt $0 = v_1 = \cdots = v_k$. Da $\mathcal{B}_1, \ldots, \mathcal{B}_k$ Basen sind, sind alle μ_j Null. □

6.3.5 Folgerung

Ein Endomorphismus $\varphi \colon V \to V$ ist genau dann diagonalisierbar, wenn sein charakteristisches Polynom vollständig in Linearfaktoren zerfällt und die algebraischen und geometrischen Vielfachheiten aller Eigenwerte übereinstimmen.

Beweis:

Es ist klar, dass die Aussage für jeden diagonalisierbaren Endomorphismus gilt.

Sei nun φ ein Endomorphismus dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt und die algebraische und geometrische Vielfachheit für jeden Eigenwert übereinstimmt. Seien $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von φ und $\mathcal{B}_1 := (b_1, \ldots, b_{i_1}), \mathcal{B}_2 := (b_{i_1+1}, \ldots, b_{i_2}), \ldots, \mathcal{B}_k := (b_{i_{k-1}+1}, \ldots, b_{i_k})$ Basen der zugehörigen Eigenräume. Dann ist $\dim V = \deg \chi_\varphi(X) = \#\mathcal{B}_1 + \cdots + \#\mathcal{B}_k$. Somit ist $(b_1, \ldots, b_{i_k})$ eine Basis von V. Folglich ist φ diagonalisierbar. □

6.3.6 Zusammenfassung

Sei $\varphi \colon V \to V$ ein Endomorphismus eines n-dimensionalen K-Vektorraums. Seien weiterhin $\lambda_1, \ldots, \lambda_k$ alle Eigenwerte von ϕ. Wir haben die Äquivalenz der folgenden Aussagen gezeigt:

1. φ ist diagonalisierbar.
2. χ_φ zerfällt vollständig in Linearfaktoren und für jeden Eigenwert stimmen die algebraische und die geometrische Vielfachheit überein.
3. $V = E(\varphi, \lambda_1) \oplus \cdots \oplus E(\varphi, \lambda_k)$
4. $V = E(\varphi, \lambda_1) + \cdots + E(\varphi, \lambda_k)$
5. $\dim V = \dim E(\varphi, \lambda_1) + \cdots + \dim E(\varphi, \lambda_k)$

Kapitel 7

Skalarprodukte

7.1 Das Standardskalarprodukt

7.1.1 Definition

Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $V = K^n$, so betrachtet man die Abbildung

$$\langle \ldots | \ldots \rangle \colon K^n \times K^n \to K, (x, y) \mapsto \langle x|y \rangle \coloneqq x_1\overline{y_1} + \cdots + x_n\overline{y_n}\,.$$

Sie wird als *Standardskalarprodukt* bezeichnet. Ist $K = \mathbb{R}$, so schreibt man auch $\langle x|y \rangle = x^\top \cdot y$.

Ist $K = \mathbb{R}$, so ist $t \mapsto \bar{t}$ die identische Abbildung.

7.1.2 Bemerkung

Für das Standardskalarprodukt gelten die Rechenregeln

1. $\langle \lambda x + y|z \rangle = \lambda\langle x|z \rangle + \langle y|z \rangle$
2. $\langle x|\lambda y + z \rangle = \overline{\lambda}\langle x|y \rangle + \langle x|z \rangle$
3. $\langle x|y \rangle = \overline{\langle y|x \rangle}$
4. $\langle x|x \rangle \geq 0$ und $\langle x|x \rangle = 0$ g.d.w. $x = 0$.

7.1.3 Definition

Für $x \in \mathbb{R}^n$ oder $x \in \mathbb{C}^n$ nennt man $\|x\| \coloneqq \sqrt{\langle x|x \rangle} = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$ die *Norm* oder Länge von x. Man nennt $d(x, y) \coloneqq \|x - y\|$ den *Abstand* zweier Punkte $x, y \in \mathbb{R}^n$.

7.1.4 Bemerkung

Norm und Abstand haben die folgenden Eigenschaften:

1. $\|x\| = 0$ g.d.w. $x = 0$
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$

3. $\|x+y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)
4. $d(x,y) = 0$ g.d.w. $x = y$
5. $d(x,y) = d(y,x)$ (Symmetrie)
6. $d(x,z) \leq d(x,y) + d(y,z)$ (Dreiecksungleichung)

Zudem gilt die *Cauchy-Schwarz Ungleichung* $|\langle x|y\rangle| \leq \|x\| \cdot \|y\|$. Hier gilt Gleichheit g.d.w. x, y linear abhängig ist.

7.1.5 Definition

Sind $x, y \in \mathbb{R}^n \setminus \{0\}$, so gilt $-1 \leq \frac{\langle x|y\rangle}{\|x\| \cdot \|y\|} \leq 1$. Man nennt den Winkel $\alpha \in [0, \pi]$ mit

$$\cos(\alpha) = \frac{\langle x|y\rangle}{\|x\| \cdot \|y\|}$$

den Winkel zwischen x und y.

Man nennt x *orthogonal* zu y, falls $\langle x|y\rangle = 0$ gilt.

7.2 Bilinearformen und Sesquilinearformen

7.2.1 Definition

Ist V ein K-Vektorraum, so nennt man eine Abbildung $\beta\colon V \times V \to K$ eine *Bilinearform*, wenn

$$\beta(x + \lambda y, z) = \beta(x,z) + \lambda\beta(y,z)$$
$$\beta(x, y + \lambda z) = \beta(x,y) + \lambda\beta(x,z)$$

gilt. β heißt *symmetrisch*, falls

$$\beta(x,y) = \beta(y,x)$$

gilt.

Ist $K = \mathbb{C}$, so nennt man eine Abbildung $\beta\colon V \times V \to K$ eine *Sesquilinearform*, wenn

$$\beta(x + \lambda y, z) = \beta(x,z) + \lambda\beta(y,z)$$
$$\beta(x, y + \lambda z) = \beta(x,y) + \overline{\lambda}\beta(x,z)$$

gilt. Weiterhin heißt eine Bilinearform β *hermitesch*, wenn

$$\beta(x,y) = \overline{\beta(y,x)}$$

gilt. Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$, so heißt β *positiv definit*, wenn

$$\beta(x,x) \geq 0 \text{ und } \beta(x,x) = 0 \Longleftrightarrow x = 0$$

gilt. **Anmerkung:** Ist β hermitesch, so gilt $\beta(x,x) \in \mathbb{R}$.

Auf einem reellen Vektorraum nennt man eine symmetrische, positiv definite Bilinearform ein *Skalarprodukt*. Auf einem komplexen Vektorraum nennt man eine hermitesche, positiv definite Sesquilinearform ein *Skalarprodukt*. Ein reeller/komplexer Vektorraum mit Skalarprodukt heißt auch euklidischer/unitärer Raum. Weiterhin nennt man Vektorräume mit Skalarprodukt auch *Skalarprodukträume*.

7.2.2 Beispiele

1. Ist $A = (a_{ij}) \in \mathrm{Mat}(n \times n, K)$, so ist

$$\beta_A \colon K^n \times K^n \to K, (x, y) \mapsto x^\top A y$$

eine Bilinearform. Es gilt $\beta_A(e_i, e_j) = a_{i,j}$. Folglich ist β_A genau dann symmetrisch, wenn $A = A^\top$ gilt. In diesem Fall heißt A symmetrisch.

2. Ist $A = (a_{ij}) \in \mathrm{Mat}(n \times n, \mathbb{C})$, so ist

$$\beta_A \colon \mathbb{C}^n \times \mathbb{C}^n \to K, (x, y) \mapsto x^\top A \overline{y}$$

eine Sesquilinearform.

3. Ist V die Menge der auf dem Intervall $[0, 1]$ stetigen reellen Funktionen, so ist

$$\langle f|g\rangle \coloneqq \int_0^1 f(x)g(x)dx$$

ein Skalarprodukt auf V.

7.2.3 Satz

Ist $\langle \ldots | \ldots \rangle$ ein Skalarprodukt auf V, so gilt die Cauchy-Schwarz-Ungleichung

$$|\langle x|y\rangle| \leq \|x\| \cdot \|y\| .$$

Es gilt Gleichheit genau dann, wenn x, y linear abhängig ist.

Beweis:

Ist $x = 0$ oder $y = 0$, so ist die Behauptung trivial. Im Fall $x \neq 0 \neq y$ setzen wir $z \coloneqq \langle x|x\rangle y - \langle y|x\rangle x$ und erhalten

$$\begin{aligned}
\langle z|z\rangle &= \Big\langle \langle x|x\rangle y - \langle y|x\rangle x \;\Big|\; \langle x|x\rangle y - \langle y|x\rangle x \Big\rangle \\
&= \langle x|x\rangle \overline{\langle x|x\rangle} \langle y|y\rangle - \langle x|x\rangle \overline{\langle y|x\rangle} \langle y|x\rangle \\
&\quad - \langle y|x\rangle \overline{\langle x|x\rangle} \langle x|y\rangle + \langle y|x\rangle \overline{\langle y|x\rangle} \langle x|x\rangle \quad \text{(2. und 4. Term gleich)} \\
&= \langle x|x\rangle \left(\langle x|x\rangle \langle y|y\rangle - \langle x|y\rangle \overline{\langle x|y\rangle} \right) \\
&= \langle x|x\rangle \left(\langle x|x\rangle \langle y|y\rangle - |\langle x|y\rangle|^2 \right) .
\end{aligned}$$

Hieraus folgt $|\langle x|y\rangle|^2 \leq \langle x|x\rangle\langle y|y\rangle$. Da die Quadratwurzel monoton ist folgt die Cauchy-Schwarz-Ungleichung. Gilt in der Cauchy-Schwarz-Ungleichung Gleichheit, so ist $z = 0$ und folglich x, y linear abhängig. Ist (x, y) linear abhängig, etwa $y = \lambda x$, so gilt

$$|\langle x|y\rangle|^2 = |\overline{\lambda}\langle x \mid x\rangle|^2 = \lambda\overline{\lambda}\langle x \mid x\rangle\langle x \mid x\rangle = \langle x \mid x\rangle\langle y \mid y\rangle ,$$

welches die Gleichheit in der Cauchy-Schwarz-Ungleichung zeigt. $\square$

7.2.4 Definition

Ist $\langle \ldots | \ldots \rangle$ ein Skalarprodukt auf V, so ist

$$\|v\| := \sqrt{\langle v|v\rangle}$$

die Norm (oder Länge) eines Vektors. Es gilt $\|v\| \geq 0$ und $\|v\| = 0$ g.d.w. $v = 0$.

7.2.5 Definition

Ist V ein $\mathbb{R}$-Vektorraum mit Skalarprodukt und x, y zwei von Null verschiedene Vektoren, so ist der Winkel $\alpha \in [0, \pi]$ zwischen x und y durch

$$\cos\alpha = \frac{\langle x|y\rangle}{\|x\| \cdot \|y\|}$$

definiert.

7.2.6 Satz

Aus der Cauchy-Schwarz-Ungleichung folgt die Dreiecksungleichung.

Beweis:

$$\begin{aligned}\|x+y\|^2 &= \langle x+y|x+y\rangle = \langle x|x\rangle + \langle x|y\rangle + \langle y|x\rangle + \langle y|y\rangle \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 .\end{aligned}$$

□

7.2.7 Satz

In jedem Vektorraum mit Skalarprodukt gelten die Parallelogrammgleichung und die Diagonalengleichung

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= 2\left(\|x\|^2 + \|y\|^2\right) \\ 2\langle x|y\rangle &= \|x+y\|^2 - \|x\|^2 - \|y\|^2 \\ &\quad + i\left(\|x+iy\|^2 - \|x\|^2 - \|y\|^2\right)\end{aligned}$$

Im Fall eines reellen Vektorraums entfällt der imaginäre Summand der Diagonalengleichung.

Beweis:

Beide Gleichungen zeigt man durch Ausmultiplizieren nachdem man die Definition der Norm eingesetzt hat. □

7.3 Orthogonalität

7.3.1 Definition

Sei V ein Skalarproduktraum. Zwei Vektoren $u, v \in V$ heißen *orthogonal* (in Zeichen $u \perp v$), wenn $\langle u|v\rangle = 0$ gilt. Weiterhin heißen zwei Teilmengen $A, B \subset V$ orthogonal (in Zeichen $A \perp B$), wenn $a \perp b$ für alle $a \in A, b \in B$ gilt.

7.3.2 Bemerkungen

1. Der Nullvektor ist zu jedem Vektor orthogonal. Insbesondere ist der Nullvektor zu sich selber orthogonal.

2. Ist $M \subset V$ eine Teilmenge, so ist $M^\perp = \{x \in V \mid \langle x \mid u\rangle = 0 \text{ für alle } u \in M\}$ ein Untervektorraum. Ist U ein Untervektorraum, so nennt man $U^\perp$ das *orthogonale Komplement* von U.

   Beweis: Übung.

3. Sind v_1, v_2 zwei Vektoren so, dass $\langle v_1|w\rangle = \langle v_2|w\rangle$ für jedes w gilt, so folgt $v_1 = v_2$.

   Beweis: $v_1 - v_2$ ist orthogonal zu jedem Vektor w. Insbesondere auch zu sich selber. Daher gilt $v_1 - v_2 = 0$.

7.3.3 Definition

Sei $(v_1, \dots, v_n)$ eine Familie von Vektoren in einem Skalarproduktraum.

1. $(v_1, \dots, v_n)$ heißt *Orthogonalsystem*, wenn $v_i \perp v_j$ für alle $i, j = 1, \dots, n$ mit $i \neq j$ gilt.

2. Ein Orthogonalsystem $(v_1, \dots, v_n)$ heißt *Orthonormalsystem*, wenn zusätzlich $\|v_i\| = 1$ für alle $i = 1, \dots, n$ gilt.

3. Ist eine Basis eines Vektorraums ein Orthonormalsystem, so nennt man sie *Orthonormalbasis*.

7.3.4 Bemerkung

Sei $(v_1, \dots, v_n)$ ein Orthogonalsystem und kein $v_i = 0$.

1. Es ist $(v_1, \dots, v_n)$ linear unabhängig.

2. Es ist $(\frac{1}{\|v_1\|}v_1, \dots, \frac{1}{\|v_n\|}v_n)$ ein Orthonormalsystem.

3. Ist $(v_1, \dots, v_n)$ eine Orthonormalbasis, so gilt für jeden Vektor v

$$v = \langle v \mid v_1\rangle v_1 + \cdots + \langle v \mid v_n\rangle v_n \,.$$

   Dies wird auch als *Fourier-Entwicklung* von v bezeichnet.

Beweis:

Zur ersten Aussage:
Sei $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ eine lineare Relation. Dann folgt für jedes $i \in \{1, \dots, n\}$

$$0 = \langle 0 \mid v_i\rangle = \langle \lambda_1 v_1 + \cdots + \lambda_n v_n \mid v_i\rangle = \lambda_1\langle v_1 \mid v_i\rangle + \cdots + \lambda_n\langle v_n \mid v_i\rangle = \lambda_i\langle v_i \mid v_i\rangle \,.$$

Da $v_i \neq 0$ gilt, folgt $\lambda_i = 0$.

Zur zweiten Aussage: Diese ist klar.

Zur dritten Aussage: Da $(v_1, \dots, v_n)$ eine Basis ist, existieren zu v eindeutige $\lambda_1, \dots, \lambda_n$ mit

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n\,.$$

Für das Skalarprodukt $\langle v \mid v_i \rangle$ erhält man

$$\begin{aligned}\langle v \mid v_i \rangle &= \langle \lambda_1 v_1 + \cdots + \lambda_n v_n \mid v_i \rangle \\ &= \lambda_1 \langle v_1 \mid v_i \rangle + \cdots + \lambda_n \langle v_n \mid v_i \rangle = \lambda_i \langle v_i \mid v_i \rangle = \lambda_i\,.\end{aligned}$$

□

7.3.5 Satz (Gram Schmidt Verfahren)

Es sei $(b_1, \dots, b_n)$ eine Basis des Skalarproduktraums V. Dann hat V eine Orthonomalbasis $(v_1, \dots, v_n)$ mit $\operatorname{span}(b_1, \dots, b_i) = \operatorname{span}(v_1, \dots, v_i)$ für jedes $i \in \{1, \dots, n\}$.

Beweis:

Wir setzen

$$\begin{aligned} w_1 &:= b_1,\ v_1 := \frac{w_1}{\|w_1\|} \\ w_2 &:= b_2 - \langle b_2 \mid v_1 \rangle v_1,\ v_2 := \frac{w_2}{\|w_2\|} \\ w_3 &:= b_3 - \langle b_3 \mid v_1 \rangle v_1 - \langle b_3 \mid v_2 \rangle v_2,\ v_3 := \frac{w_3}{\|w_3\|} \\ &\vdots \\ w_n &:= b_n - \langle b_n \mid v_1 \rangle v_1 - \cdots - \langle b_n \mid v_{n-1} \rangle v_{n-1},\ v_n := \frac{w_n}{\|w_n\|} \end{aligned}$$

Wir zeigen durch Induktion nach i dass $(w_1, \dots, w_i)$ linear unabhängig ist, $v_1, \dots, v_i$ wohldefiniert und eine Orthonormalbasis von $\operatorname{span}(b_1, \dots, b_i)$ bilden.

Induktionsanfang: Ist $i = 1$, so ist $w_1 = b_1$ und $v_1 = \frac{w_1}{\|w_1\|}$. Somit ist $w_1 \neq 0$. Dies zeigt die Wohldefiniertheit von v_1 sowie $\operatorname{span}(b_1) = \operatorname{span}(w_1) = \operatorname{span}(v_1)$. Aus $\|v_1\| = 1$ folgt die Orthonormalität.

Induktionsschritt: Ist $(v_1, \dots, v_i)$ eine Orthonormalbasis von $\operatorname{span}(b_1, \dots, b_i)$, so ist $b_{i+1} \notin \operatorname{span}(v_1, \dots, v_i)$, da $(b_1, \dots, b_n)$ linear unabhängig ist. Somit ist $w_{i+1} \neq 0$ und v_{i+1} wohldefiniert. Weiterhin gilt

$$\begin{aligned} &\operatorname{span}(v_1, \dots, v_{i+1}) = \operatorname{span}(v_1, \dots, v_i, w_{i+1}) \\ &= \operatorname{span}(v_1, \dots, v_i, b_{i+1} - \langle b_{i+1} \mid v_1 \rangle v_1 - \cdots - \langle b_{i+1} \mid v_i \rangle v_i) \\ &= \operatorname{span}(v_1, \dots, v_i, b_{i+1}) = \operatorname{span}(b_1, \dots, b_i, b_{i+1}) \end{aligned}$$

Da alle v_i nach Konstruktion Norm 1 haben bleibt noch zu zeigen, dass sie ein Orthogonalsystem bilden. Nach Induktionsvoraussetzung gilt $\langle v_k \mid v_l \rangle = \delta_{k,l}$ bereits für alle $k, l \in \{1, \dots, i\}$. Sei nun $i + 1 > k \geq 1$, dann gilt

$$\begin{aligned} &\langle w_{i+1} \mid v_k \rangle = \langle b_{i+1} - \langle b_{i+1} \mid v_1 \rangle v_1 - \cdots - \langle b_{i+1} \mid v_i \rangle v_i \mid v_k \rangle \\ &= \langle b_{i+1} \mid v_k \rangle - \langle b_{i+1} \mid v_1 \rangle \langle v_1 \mid v_k \rangle - \cdots - \langle b_{i+1} \mid v_i \rangle \langle v_i \mid v_k \rangle \\ &= \langle b_{i+1} \mid v_k \rangle - \langle b_{i+1} \mid v_k \rangle \langle v_k \mid v_k \rangle = \langle b_{i+1} \mid v_k \rangle - \langle b_{i+1} \mid v_k \rangle = 0 \end{aligned}$$

Dies zeigt die Orthogonalität. $\square$

7.3.6 Beispiel

Wir betrachten den Raum der reellen Polynome mit Skalarprodukt $\langle f \mid g\rangle := \int_{-1}^{1} f(x)g(x)dx$. Wir wenden das Gram-Schmidt Verfahren auf die Basis $b_1 := 1, b_2 := x, b_3 := x^2$ des Unterraums der Polynom vom Höchstgrad 2 an. Diese liefert:

$$
\begin{aligned}
w_1 &:= 1, v_1 := \frac{1}{\sqrt{\int_{-1}^{1} 1dx}} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \\
w_2 &:= x - \int_{-1}^{1} \frac{\sqrt{2}}{2} x\, dx \cdot \frac{\sqrt{2}}{2} = x, v_2 = \frac{x}{\sqrt{\int_{-1}^{1} x^2 dx}} = \frac{\sqrt{6}}{2} x \\
w_3 &:= x^2 - \int_{-1}^{1} \frac{\sqrt{2}}{2} \cdot x^2 dx \cdot \frac{\sqrt{2}}{2} - \int_{-1}^{1} \frac{\sqrt{6}}{3} x x^2 dx \frac{\sqrt{6}}{3} x = x^2 - \frac{2}{3} \cdot \frac{2}{4} = x^2 - \frac{1}{3} \\
v_3 &:= \frac{3\sqrt{10}}{4} \left(x^2 - \frac{1}{3} \right)
\end{aligned}
$$

Anmerkung: Die hier entstehenden Polynome werden auch als *Legendre-Polynome* bezeichnet.

7.3.7 Folgerung

Jeder endlichdimensionale Skalarproduktraum hat eine Orthonormalbasis.

7.3.8 Satz

Ist $U \subset V$ ein endlich-dimensionaler Unterraum eines Skalarproduktraums und $(v_1, \ldots, v_n)$ eine Orthonormalbasis von U. Wir setzen

$$\pi_U \colon V \to U, x \mapsto \langle x \mid v_1\rangle v_1 + \cdots + \langle x \mid v_n\rangle v_n \,.$$

Die Abbildung π_U hat die Eigenschaften

1. π ist linear und surjektiv.
2. $\pi \circ \pi = \pi$
3. $(x - \pi_U(x)) \in U^\perp$

π_U wird als *orthogonale Projektion* auf U bezeichnet.

Beweis:

Die Linearität von π_U folgt aus der Linearität das Skalarprodukts in der ersten Komponente. Da $\pi(x) = x$ für alle $x \in U$ gilt, folgt die Surjektivität und die zweite Aussage.

Zur dritten Aussage: Es gilt

$$
\begin{aligned}
&\langle x - \pi_U(x) \mid v_i\rangle = \langle x \mid v_i\rangle - \langle x \mid v_1\rangle\langle v_1 \mid v_i\rangle - \cdots - \langle x \mid v_n\rangle\langle v_n \mid v_i\rangle \\
&= \langle x \mid v_i\rangle - \langle x \mid v_i\rangle\langle v_i \mid v_i\rangle = \langle x \mid v_i\rangle - \langle x \mid v_i\rangle = 0 \,.
\end{aligned}
$$

Dies zeigt $x - \pi_U(x) \perp v_i$ und es folgt $x - \pi_U(x) \perp \operatorname{span}(v_i \colon i = 1, \ldots, n) = U$. $\square$

7.3.9 Folgerung

Ist $U \subset V$ ein endlich-dimensionaler Unterraum, so gilt $V = U \oplus U^\perp$.

Beweis:

$U \cap U^\perp = \{0\}$ ist klar. Für jedes $v \in V$ gilt $v = \pi_U(v) + (v - \pi_U(v)) \in U + U^\perp$. $\square$

7.3.10 Satz

Es sei $(v_1, \ldots, v_n)$ ein Orthonormalsystem in einem Skalarproduktraum V sowie $U = \operatorname{span}(v_1, \ldots, v_n)$. Dann gilt:

$$\begin{aligned}
&\|v\|^2 \geq |\langle v|v_1\rangle|^2 + \cdots + |\langle v|v_n\rangle|^2 \\
&\|v\|^2 = |\langle v|v_1\rangle|^2 + \cdots + |\langle v|v_n\rangle|^2 \text{ genau dann, wenn } v \in U
\end{aligned}$$

Dies wird als Bessel-Ungleichung und Parsevalsche Gleichung bezeichnet.

Beweis:

Ist $v \in U$, so gilt $v = \langle v|v_1\rangle v_1 + \cdots + \langle v|v_n\rangle v_n$. Hieraus folgt

$$\begin{aligned}
\|v\|^2 &= \langle \langle v|v_1\rangle v_1 + \cdots + \langle v|v_n\rangle v_n \mid \langle v|v_1\rangle v_1 + \cdots + \langle v|v_n\rangle v_n\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle \langle v|v_i\rangle v_i \mid \langle v|v_j\rangle v_j\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle v|v_i\rangle \overline{\langle v|v_j\rangle} \langle v_i \mid v_j\rangle \\
&= \sum_{i=1}^n \langle v|v_i\rangle \overline{\langle v|v_i\rangle} = \sum_{i=1}^n |\langle v|v_i\rangle|^2
\end{aligned}$$

Diese Rechnung zeigt, dass im Fall $v \in U$ die Parsevalsche Gleichung gilt. Weiterhin zeigt sie $\|\pi_U(v)\|^2 = |\langle v|v_1\rangle|^2 + \cdots + |\langle v|v_n\rangle|^2$. Es folgt

$$\begin{aligned}
\|v\|^2 &= \langle \pi_U(x) - (\pi_U(v) - v) \mid \pi_U(x) - (\pi_U(v) - v)\rangle \\
&= \langle \pi_U(v) \mid \pi_U(v)\rangle - \langle \pi_U(x) \mid -(\pi_U(v) - v)\rangle \\
&\quad - \langle -(\pi_U(v) - v) \mid \pi_U(v)\rangle + \langle -(\pi_U(v) - v) \mid -(\pi_U(v) - v)\rangle \\
&= \|\pi_U(v)\|^2 + \|(\pi_U(v) - v)\|^2 \\
&= |\langle v|v_1\rangle|^2 + \cdots + |\langle v|v_n\rangle|^2 + \|(\pi_U(v) - v)\|^2 .
\end{aligned}$$

Dieses zeigt die Bessel-Ungleichung. Es gilt Gleichheit genau dann, wenn $(\pi_U(v) - v) = 0$ gilt. Letzteres ist zu $v \in U$ äquivalent. $\square$

7.4 Bilinearformen und Matrizen

7.4.1 Satz

Es sei V ein n-dimensionaler Vektorraum über dem Körper K. Dann bildet die Menge der Bilinearformen auf V einen n^2 dimensionalen Vektorraum. Ist

$\mathcal{B} = (b_1, \dots, b_n)$ eine Basis von V, so ist jede Bilinearform $\beta\colon V \times V \to K$ durch die Werte $\beta(b_i, b_j)$ für $i, j = 1, \dots, n$ eindeutig bestimmt. Weiterhin gibt es zu jeder Wahl von $\beta(b_i, b_j) \in K$ für $i, j = 1, \dots, n$ genau eine solche Bilinearform.

Beweis:

Sind $v, w \in V$ beliebig. Dann gibt es $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in K$ mit $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$ und $w = \mu_1 b_1 + \cdots + \mu_n b_n$. Hiermit gilt:

$$\beta(v, w) = \beta(\lambda_1 b_1 + \cdots + \lambda_n b_n, \mu_1 b_1 + \cdots + \mu_n b_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \beta(b_i, b_j) \mu_j$$

Somit ist β durch $\beta(b_i, b_j)$ eindeutig bestimmt. Weiterhin ist

$$\beta(v, w) = (\lambda_1 \dots \lambda_n) \begin{pmatrix} \beta(b_1, b_1) & \cdots & \beta(b_1, b_n) \\ \vdots & \ddots & \vdots \\ \beta(b_n, b_1) & \cdots & \beta(b_n, b_n) \end{pmatrix} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Man nennt die obige Matrix $M_{\mathcal{B}}(\beta) := (\beta(b_i, b_j))_{i,j=1,\dots,n}$ die Darstellungsmatrix der Bilinearform β bezüglich der Basis $\mathcal{B}$.

Sind die Werte $\beta(b_i, b_j)$ für alle $i, j \in \{1, \dots, n\}$ beliebig vorgegeben, so kann man die Darstellungsmatrix bilden und erhält mit obiger Formel eine Bilinearform auf V. Wir erhalten mit der Zuordnung $\beta \mapsto M_{\mathcal{B}}(\beta)$ einen Isomorphismus zwischen dem Raum der Bilinearformen auf V und dem Raum der $n \times n$-Matrizen über K. $\square$

7.4.2 Beispiel

Wir betrachten die Standardbilinearform $\beta(x, y) = x_1 y_1 + x_2 y_2$ auf $\mathbb{R}^2$ und $\mathcal{B} = ((1, -1), (1, 2))$. Es ist

$$M_{\mathcal{B}}(\beta) = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}.$$

7.4.3 Bemerkung

Eine Bilinearform β ist genau dann symmetrisch, wenn die Darstellungsmatrix $M_{\mathcal{B}}(\beta)$ bezüglich einer beliebigen Basis $\mathcal{B}$ symmetrisch ist.

Beweis:

Ist β symmetrisch, so ist $\beta(b_i, b_j) = \beta(b_j, b_i)$. Folglich ist die Darstellungsmatrix symmetrisch. Ist andererseits die Matrix symmetrisch und sind $(\lambda_1, \dots, \lambda_n)$ und $(\mu_1, \dots, \mu_n)$ die Koordinaten von $v, w \in V$, so gilt

$$\begin{aligned} \beta(v, w) &= (\lambda_1 \dots \lambda_n) M_{\mathcal{B}}(\beta) \begin{pmatrix} \mu_1 \dots \mu_n \end{pmatrix}^\top \\ &= \Big((\lambda_1 \dots \lambda_n) M_{\mathcal{B}}(\beta) \begin{pmatrix} \mu_1 \dots \mu_n \end{pmatrix}^\top \Big)^\top \\ &= (\mu_1 \dots \mu_n) M_{\mathcal{B}}(\beta)^\top \begin{pmatrix} \lambda_1 \dots \lambda_n \end{pmatrix}^\top \\ &= (\mu_1 \dots \mu_n) M_{\mathcal{B}}(\beta) \begin{pmatrix} \lambda_1 \dots \lambda_n \end{pmatrix}^\top = \beta(w, v). \end{aligned}$$

$\square$

7.4.4 Satz

Es seien $\mathcal{B}, \mathcal{C}$ Basen von V, $S = T_{\mathcal{B}}^{\mathcal{C}}$ die Basiswechselmatrix und β eine Bilinearform so gilt

$$M_{\mathcal{C}}(\beta) = S^\top M_{\mathcal{B}}(\beta) S\,.$$

Beweis:

Wir bezeichnen mit $\Phi_{\mathcal{B}}$, $\Phi_{\mathcal{C}}\colon K^n \to V$ die Koordinatensysteme. Dann gilt

$$\begin{aligned}\beta(v,w) &= \Phi_{\mathcal{B}}^{-1}(v)^\top M_{\mathcal{B}}(\beta)\Phi_{\mathcal{B}}^{-1}(w) = (S\cdot\Phi_{\mathcal{C}}^{-1}(v))^\top M_{\mathcal{B}}(\beta)(S\cdot\Phi_{\mathcal{C}}^{-1}(w))\\ &= \Phi_{\mathcal{C}}^{-1}(v)^\top (S^\top M_{\mathcal{B}}(\beta)S)\Phi_{\mathcal{C}}^{-1}(w)\,.\end{aligned}$$

Somit ist $S^\top M_{\mathcal{B}}(\beta)S$ die Darstellungsmatrix $M_{\mathcal{C}}(\beta)$. □

7.4.5 Bemerkung

Sei β eine positiv definite Bilinearform auf $\mathbb{R}^n$. Es ist $\mathcal{B}$ genau dann eine Orthonormalbasis bezüglich des durch β gegebenen Skalarprodukts, wenn $M_{\mathcal{B}}(\beta) = E_n$ gilt.

Beweis:

Es ist $\mathcal{B} = (b_1, \ldots, b_n)$ genau dann eine Orthonormalbasis, wenn $\beta(b_i, b_j) = \delta_{i,j}$ für alle $i, j \in \{1, \ldots, n\}$ gilt. In Matrixschreibweise ist dies $M_{\mathcal{B}}(\beta) = E_n$. □

7.4.6 Bemerkung

Genau wie oben zeigt man, dass die Menge der Sesquilinearformen auf einem $\mathbb{C}$-Vektorraum V einen $\mathbb{C}$-Vektorraum bildet. Dieser hat die Dimension $\dim(V)^2$. Ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V und β eine Sesquilinearform auf V, so nennt man $M_{\mathcal{B}}(\beta) \coloneqq (\beta(b_i, b_j))_{i,j=1,\ldots,n}$ die Darstellungsmatrix von β bezüglich $\mathcal{B}$. Es gilt dann

$$\beta(u,v) = \Phi_{\mathcal{B}}^{-1}(u)^\top M_{\mathcal{B}}(\beta)\overline{\Phi_{\mathcal{B}}^{-1}(v)}\,.$$

Hierbei bezeichnet $\overline{}$ die komplexe Konjugation aller Einträge. Für den Basiswechsel erhält man die Formel

$$M_{\mathcal{C}}(\beta) = S^\top M_{\mathcal{B}}(\beta)\overline{S}$$

mit $S = T_{\mathcal{B}}^{\mathcal{C}}$. Die Form β ist genau dann hermitesch, wenn die Darstellungsmatrix $M^\top = \overline{M}$ erfüllt. Daher nennt man solche Matrizen auch hermitesch.

Ist β zudem ein Skalarprodukt (d.h. positiv definit und hermitesch), so ist die Darstellungsmatrix von β genau dann die Einheitsmatrix, wenn die gewählte Basis eine Orthonormalbasis ist.

Beweis:

Durch direktes Übertragen aller Argumente für Bilinearformen.

Kapitel 8

Normale Endomorphismen

8.1 Die adjungierte Abbildung

8.1.1 Definition

Es seien $(V, \beta), (W, \gamma)$ Skalarprodukträume über dem gleichen Körper und $\varphi\colon V \to W$ eine lineare Abbildung. Eine Abbildung $\varphi^*\colon W \to V$ heißt zu φ *adjungierte Abbildung* (bezüglich β und γ), falls

$$\forall v \in V, w \in W\colon \gamma(\varphi(v), w) = \beta(v, \varphi^*(w))$$

gilt. Da Skalarprodukte symmetrisch bzw. hermitesch sind, ist die Forderung zu

$$\forall v \in V, w \in W\colon \gamma(w, \varphi(v)) = \beta(\varphi^*(w), v)$$

äquivalent.

8.1.2 Satz

Zu jeder linearen Abbildung $\varphi\colon V \to W$ gibt es höchstens eine adjungierte Abbildung.

Beweis:

Wir bezeichnen die Skalarprodukte auf V und W mit β und γ.

Angenommen es sind $\psi, \varphi^*\colon W \to V$ zwei adjungierte Abbildungen von φ, dann gilt

$$\beta(v, \psi(w) - \varphi^*(w)) = \beta(v, \psi(w)) - \beta(v, \varphi^*(w)) = \gamma(\varphi(v), w) - \gamma(\varphi(v), w) = 0\,.$$

Setzen wir nun für v den Wert $\psi(w) - \varphi^*(w)$ ein, so folgt da β positiv definit ist $\psi(w) - \varphi^*(w) = 0$. $\square$

8.1.3 Satz

Sind $\varphi, \varphi_1, \varphi_2 \colon V \to W$ und $\psi \colon W \to X$ lineare Abbildungen zwischen Skalarprodukträumen. Falls adjungierte Abbildungen existieren, so gelten

$$\begin{aligned}(\varphi_1 + \varphi_2)^* &= \varphi_1^* + \varphi_2^* \,, \\ (\lambda\varphi)^* &= \overline{\lambda}\varphi^* \,, \\ (\varphi^*)^* &= \varphi \,, \\ (\psi \circ \varphi)^* &= \varphi^* \circ \psi^* \,.\end{aligned}$$

Beweis:

Wir bezeichnen die Skalarprodukte auf V, W und X mit β, γ und δ.

Es gilt

$$\begin{aligned}\beta(v, (\varphi_1 + \varphi_2)^*(w)) &= \gamma((\varphi_1 + \varphi_2)(v), w) = \gamma(\varphi_1(v) + \varphi_2(v), w) \\ &= \gamma(\varphi_1(v), w) + \gamma(\varphi_2(v), w) = \beta(v, \varphi_1^*(w)) + \beta(v, \varphi_2^*(w)) \\ &= \beta(v, \varphi_1^*(w) + \varphi_2^*(w)) = \beta(v, (\varphi_1^* + \varphi_2^*)(w)) \,.\end{aligned}$$

Dies zeigt $(\varphi_1 + \varphi_2)^* = (\varphi_1^* + \varphi_2^*)$.

Zur 2. Behauptung:

$$\gamma((\alpha\varphi)(v), w) = \gamma(\alpha \cdot \varphi(v), w) = \gamma(\varphi(\alpha v), w) = \beta(\alpha v, \varphi^*(w)) = \beta(v, \overline{\alpha}\varphi^*(w)) \,.$$

Dies zeigt $(\alpha\varphi)^* = \overline{\alpha}\varphi^*$.

Zur 3. Behauptung: Es ist $\gamma(\varphi(v), w) = \beta(v, \varphi^*(w)) = \gamma((\varphi^*)^* (v), w)$. Somit ist $\varphi = (\varphi^*)^*$.

Zur 4. Behauptung:

$$\begin{aligned}\delta((\psi \circ \varphi)(v), x) &= \delta(\psi(\varphi(v)), x) = \gamma(\varphi(v), \psi^*(x)) \\ &= \beta(v, \varphi^*(\psi^*(x))) = \beta(v, (\varphi^* \circ \psi^*)(x))\end{aligned}$$

Dies zeigt $(\varphi^* \circ \psi^*) = (\psi \circ \varphi)^*$. $\square$

8.1.4 Satz

Sind V, W endlich-dimensional und $\mathcal{B}, \mathcal{C}$ Orthonormalbasen von V und W, so gilt

$$M_{\mathcal{B}}^{\mathcal{C}}(\varphi^*) = \overline{M_{\mathcal{C}}^{\mathcal{B}}(\varphi)}^\top =: \left(M_{\mathcal{C}}^{\mathcal{B}}(\varphi)\right)^* \,.$$

Insbesondere existiert die adjungierte Abbildung. Zu einer Matrix A bezeichnen wir $A^* = \overline{A}^\top = \overline{A^\top}$ als *adjungierte Matrix*.

Beweis:

Da $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen sind, erhalten wir für $M_{\mathcal{B}}(\beta)$ und $M_{\mathcal{C}}(\gamma)$ jeweils die Einheitsmatrix. Somit gilt für $v \in V, w \in W$

$$\begin{aligned}\gamma(\varphi(v), w) &= \Phi_{\mathcal{C}}^{-1}(\varphi(v))^\top M_{\mathcal{C}}(\beta)\overline{\Phi_{\mathcal{C}}^{-1}(w)} = \left(M_{\mathcal{C}}^{\mathcal{B}}(\varphi)\Phi_{\mathcal{B}}^{-1}(v)\right)^\top \overline{\Phi_{\mathcal{C}}^{-1}(w)} \\ &= \Phi_{\mathcal{B}}^{-1}(v)^\top \left(M_{\mathcal{C}}^{\mathcal{B}}(\varphi)\right)^\top \overline{\Phi_{\mathcal{C}}^{-1}(w)} = \Phi_{\mathcal{B}}^{-1}(v)^\top \overline{\left(M_{\mathcal{C}}^{\mathcal{B}}(\varphi)\right)^* \Phi_{\mathcal{C}}^{-1}(w)} \\ &= \beta(v, \psi(w)) \,,\end{aligned}$$

wobei die Abbildung $\psi \colon W \to V$ durch $M_{\mathcal{B}}^{\mathcal{C}}(\psi) := M_{\mathcal{C}}^{\mathcal{B}}(\varphi)^*$ gegeben ist. Somit ist ψ eine adjungierte Abbildung von φ. $\square$

8.1.5 Folgerung

Ist φ ein Endomorphismus des endlich dimensionalen Skalarproduktraums V mit Orthonormalbasis $\mathcal{B}$, so gilt

1. $M_{\mathcal{B}}(\varphi^*) = M_{\mathcal{B}}(\varphi)^*$
2. $\mathrm{rk}(\varphi) = \mathrm{rk}(\varphi^*)$
3. $\det M_{\mathcal{B}}(\varphi^*) = \det \overline{M_{\mathcal{B}}(\varphi)} = \overline{\det M_{\mathcal{B}}(\varphi)}$
4. $(\varphi(V))^{\perp} = \ker \varphi^*$

Beweis:

Wir zeigen nur die letzte Aussage:

$$\begin{aligned}(\varphi(V))^{\perp} &= \{v \in V \mid \beta(v, \varphi(w)) = 0 \,\forall w \in V\} \\ &= \{v \in V \mid \beta(\varphi^*(v), w) = 0 \,\forall w \in V\} \\ &= \{v \in V \mid \varphi^*(v) = 0\} = \ker(\varphi^*)\end{aligned}$$

□

8.1.6 Definition

Sei φ ein Endomomorphismus des Skalarproduktraums V. Wir nennen φ normal falls φ^* existiert und $\varphi \circ \varphi^* = \varphi^* \circ \varphi$ gilt.

8.1.7 Satz

Ist φ ein Endomorphismus eines Skalarproduktraums V, so gilt:

1. φ ist genau dann normal, wenn $\beta(\varphi v, \varphi w) = \beta(\varphi^* v, \varphi^* w)$ für alle $v, w \in V$ gilt.
2. Ist φ normal so gilt $\ker \varphi = \ker \varphi^*$.

Beweis:

Sei φ normal, so gilt

$$\beta(\varphi v, \varphi w) = \beta(v, \varphi^*(\varphi w)) = \beta(v, \varphi(\varphi^* w)) = \beta(\varphi^* v, \varphi^* w)\,.$$

Angenommen es gilt $\beta(\varphi v, \varphi w) = \beta(\varphi^* v, \varphi^* w)$ für alle $v, w \in V$, so folgt

$$\beta(v, \varphi^*(\varphi(w))) = \beta(v, \varphi(\varphi^*(w)))$$

Dies zeigt $\varphi^*(\varphi(w)) = \varphi(\varphi^*(w))$. Weiterhin folgt $\|\varphi v\| = \|\varphi^* v\|$ für jeden normalen Endomorphismus.

Zum zweiten Teil: Mit dem ersten Teil erhalten wir für den Kern eines normalen Endomorphismus φ

$$\ker \varphi = \{x \in V \mid \beta(\varphi x, \varphi x) = 0\} = \{x \in V \mid \beta(\varphi^* x, \varphi^* x) = 0\} = \ker \varphi^*\,.$$

□

8.2 Beispiele normaler Abbildungen

8.2.1 Definition

Der Endomorphismus φ des Skalarproduktraums V heißt *selbstadjungiert*, wenn $\varphi^* = \varphi$ gilt. Der Endomorphismus φ des Skalarproduktraums V heißt *anti-selbstadjungiert*, wenn $\varphi^* = -\varphi$ gilt.

8.2.2 Bemerkung

Ist ein Endomorphismus selbstadjungiert oder anti-selbstadjungiert so ist er normal.

8.2.3 Satz

Sei φ ein Endomorphismus des endlich-dimensionalen euklidischen Raums V mit Orthonormalbasis $\mathcal{B}$ so gilt

1. φ ist genau dann selbstadjungiert, wenn $M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(\varphi)^\top$ gilt. Man nennt φ und $M_{\mathcal{B}}(\varphi)$ dann auch symmetrisch.

2. φ ist genau dann anti-selbstadjungiert, wenn $M_{\mathcal{B}}(\varphi) = -M_{\mathcal{B}}(\varphi)^\top$ gilt. Man nennt φ und $M_{\mathcal{B}}(\varphi)$ dann auch schiefsymmetrisch (oder antisymmetrisch).

Beweis:

Zu 1: φ selbstadjungiert $\iff \varphi = \varphi^* \iff M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(\varphi^*) \iff M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(\varphi)^\top$
Zu 2: φ anti-selbstadjungiert $\iff \varphi = -\varphi^* \iff M_{\mathcal{B}}(\varphi) = -M_{\mathcal{B}}(\varphi^*) \iff M_{\mathcal{B}}(\varphi) = -M_{\mathcal{B}}(\varphi)^\top$ $\square$

8.2.4 Satz

Sei φ ein Endomorphismus des endlich-dimensionalen unitären Raums V mit Orthonormalbasis $\mathcal{B}$ so gilt

1. φ ist genau dann selbstadjungiert, wenn $M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(\varphi)^*$ gilt. Man nennt φ und $M_{\mathcal{B}}(\varphi)$ dann auch hermitesch.

2. φ ist genau dann anti-selbstadjungiert, wenn $M_{\mathcal{B}}(\varphi) = -M_{\mathcal{B}}(\varphi)^*$ gilt. Man nennt φ und $M_{\mathcal{B}}(\varphi)$ dann auch schiefhermitesch (oder antihermitesch).

Beweis:

Zu 1: φ selbstadjungiert $\iff \varphi = \varphi^* \iff M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(\varphi^*) \iff M_{\mathcal{B}}(\varphi) = M_{\mathcal{B}}(\varphi)^*$.
Zu 2: φ anti-selbstadjungiert $\iff \varphi = -\varphi^* \iff M_{\mathcal{B}}(\varphi) = -M_{\mathcal{B}}(\varphi^*) \iff M_{\mathcal{B}}(\varphi) = -M_{\mathcal{B}}(\varphi)^*$. $\square$

8.2.5 Beispiel

Sei $U \subset V$ ein Unterraum eines Skalarproduktraums. Es bezeichnet $\pi_U : V \to U$ die Orthogonalprojektion auf U. D.h. $\pi_U \circ \pi_U = \pi_U$ und $x - \pi_U(x) \in U^\perp$ für alle $x \in V$. Es ist π_U selbstadjungiert.

Beweis:

$$\begin{aligned}
\langle \pi_U(x)|w\rangle &= \langle \pi_U(x)|\pi_U(w) + (w - \pi_U(w))\rangle \\
&= \langle \pi_U(x)|\pi_U(w)\rangle + \langle \pi_U(x)|(w - \pi_U(w))\rangle \\
&= \langle \pi_U(x)|\pi_U(w)\rangle + 0 \\
&= \langle \pi_U(x)|\pi_U(w)\rangle + \langle (x - \pi_U(x))|\pi_U(w)\rangle \\
&= \langle \pi_U(x) + (x - \pi_U(x))|\pi_U(w)\rangle \\
&= \langle x|\pi_U(w)\rangle
\end{aligned}$$

□

8.2.6 Übungsaufgabe

Ist $\varphi\colon V \to V$ selbstadjungiert und V endlich-dimensional, so gilt $\det \varphi \in \mathbb{R}$.

8.2.7 Satz

Sei $\varphi\colon V \to V$ ein Endomorphismus so, dass φ^* existiert. Dann gibt es eindeutig bestimmte Endomorphismen φ_1, φ_2 mit $\varphi = \varphi_1 + \varphi_2$ mit φ_1 selbstadjungiert, φ_2 anti-selbstadjungiert.

Beweis:

Ist $\varphi = \varphi_1 + \varphi_2$ mit $\varphi_1 = \varphi_1^*$ und $\varphi_2 = -\varphi_2^*$. Es folgt $\varphi^* = \varphi_1 - \varphi_2$. Somit setzen wir $\varphi_1 := \frac{1}{2}(\varphi + \varphi^*)$ und $\varphi_2 := \frac{1}{2}(\varphi - \varphi^*)$. Wir erhalten $\varphi_1 = \varphi_1^*$, $\varphi_2 = -\varphi_2^*$ und $\varphi = \varphi_1 + \varphi_2$. □

8.2.8 Definition

Es seien (V, β) und (W, γ) zwei Skalarprodukträume. Eine lineare Abbildung $\varphi\colon V \to W$ heißt *Isometrie* oder *isometrisch* falls $\beta(v_1, v_2) = \gamma(\varphi(v_1), \varphi(v_2))$ für alle $v_1, v_2 \in V$ gilt. Ist der Grundkörper $\mathbb{R}$ so nennt man diese Abbildungen auch *orthogonal*. Im Fall des Grundkörpers $\mathbb{C}$ heißen sie *unitär*.

8.2.9 Bemerkung

Sei $\varphi\colon V \to W$ eine Isometrie. Dann gilt

1. $\|v\| = \|\varphi(v)\|$ für alle $v \in V$.
2. $v \perp w \Longrightarrow \varphi(v) \perp \varphi(w)$.
3. Ist φ surjektiv, so ist φ^{-1} ebenfalls eine Isometrie von V.
4. Ist φ zudem ein Endomorphismus und λ ein Eigenwert von φ, so ist $|\lambda| = 1$.

5. Sei (X, δ) eine weiterer Skalarproduktraum und $\psi\colon W \to X$ eine Isometrie, so ist $\psi \circ \varphi$ ebenfalls eine Isometrie.

Beweis:

Die ersten beiden Aussagen folgen direkt aus der Definition. Sei nun φ surjektiv. Dann gibt es zu $w_1, w_2 \in W$ passende $v_1, v_2 \in V$ mit $\varphi(v_i) = w_i$. Wir erhalten

$$\beta(\varphi^{-1}(w_1), \varphi^{-1}(w_2)) = \beta(v_1, v_2) = \gamma(\varphi(v_1), \varphi(v_2)) = \gamma(w_1, w_2)\,.$$

Zum vierten Punkt: Sei λ ein Eigenwert zum Eigenvektor v von φ und $V = W$. Dann gilt

$$|\lambda| \cdot \|v\| = \|\lambda v\| = \|\varphi(v)\| = \|v\|\,.$$

Somit ist $|\lambda| = 1$.

Zum letzten Punkt: Dies folgt aus

$$\delta(\psi(\varphi(v)), \psi(\varphi(w))) = \gamma(\varphi(v), \varphi(w)) = \beta(v, w)\,.$$

□

8.2.10 Satz

Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine Matrix. Die durch A gegebene Abbildung $\mathbb{R}^n \to \mathbb{R}^n$ ist genau dann bezüglich des Standard-Skalarprodukts orthogonal, wenn $A^{-1} = A^\top$ gilt.

Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ eine Matrix, die durch A gegebene Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ ist genau dann bezüglich des Standard-Skalarprodukts unitär, wenn $A^{-1} = A^*$ gilt.

Beweis:

Für eine beliebige Matrix A gilt

$$\langle Ax \mid Ay \rangle = (Ax)^\top \overline{Ay} = x^\top A^\top \overline{A}\, \overline{y} = x^\top \overline{A^* A}\, \overline{y}\,.$$

Die durch A gegebene Abbildung ist somit genau dann unitär, wenn A^*A die Einheitsmatrix ist. □

8.2.11 Definition

Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ heißt *orthogonal*, falls $A^\top = A^{-1}$ gilt. Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ heißt *unitär*, falls $A^* = A^{-1}$ gilt.

8.2.12 Bemerkung

Die Determinante einer orthogonalen Matrix ist ± 1. Die Determinante eine unitären Matrix hat Betrag 1.

Nachweis: Sei A orthogonal, dann gilt

$$(\det A)^2 = \det A \det A^\top = \det A \det A^{-1} = 1\,.$$

Sei A unitär, dann gilt

$$|\det A|^2 = \det A \cdot \overline{\det A} = \det A \cdot \det \overline{A} = \det A \cdot \det A^* = \det A \cdot \det A^{-1} = 1\,.$$

□

8.2.13 Definition

Wir definieren

$$\begin{aligned} O(n) &:= \{A \in \mathrm{Gl}_n(\mathbb{R}) \mid A^{-1} = A^\top\} \\ SO(n) &:= \{A \in O(n) \mid \det A = 1\} \\ U(n) &:= \{A \in \mathrm{Gl}_n(\mathbb{C}) \mid A^{-1} = A^*\} \end{aligned}$$

als *orthogonale*, *spezielle orthogonale* und als *unitäre Gruppe*.

Dieses sind Gruppen da Verkettungen und Inverse von Isometrien nach 8.2.9 wieder Isometrien sind.

8.2.14 Beispiel

Wir bestimmen $O(2)$. Sei

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

eine orthogonale Matrix, so gilt $A^\top A = E_2$. Diese liefert die Gleichungen

$$a^2 + c^2 = 1, ab + cd = 0, b^2 + d^2 = 1\,.$$

Es gibt somit ein $\alpha, \beta \in [0, 2\pi[$ mit $a = \cos\alpha, c = \sin\alpha$ und $b = \sin\beta, d = \cos\beta$. Die mittlere Gleichung liefert

$$0 = \cos\alpha \sin\beta + \sin\alpha\cos\beta = \sin(\alpha + \beta)$$

Somit ist $\alpha+\beta \in \{0, \pi, 2\pi, 3\pi\}$. Dies führt auf $b = c, a = -d$ oder $b = -c, a = d$. Somit hat eine Matrix in $O(2)$ die Form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ oder } \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

Im ersten Fall beschreibt die Matrix eine Drehung um α, im zweiten Fall eine Spiegelung.

8.2.15 Satz

Ist A eine reelle oder komplexe $n \times n$-Matrix, so sind äquivalent:

1. A ist orthogonal bzw. unitär.
2. Die Spalten von A bilden eine Orthonormalbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.
3. Die Zeilen von A bilden eine Orthonormalbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$.

Beweis:

Die zweite Aussage bedeutet $A^\top \overline{A} = E_n$, die dritte bedeutet $A\overline{A^\top} = E_n$. Beides ist zu $A^* = A^{-1}$ äquivalent. □

8.2.16 Satz

Sei V ein euklidischer oder unitärer Raum mit Orthonormalbasis $\mathcal{B}$ und $\varphi\colon V \to V$ ein Endomorphismus. Dann ist φ genau dann orthogonal bzw. unitär, wenn $M_{\mathcal{B}}(\varphi)$ es ist.

Insbesondere sind orthogonale und unitäre Abbildungen normal.

Beweis:

Sei $A = M_{\mathcal{B}}(\varphi)$ und $v = \Phi_{\mathcal{B}}(x), w = \Phi_{\mathcal{B}}(y)$ zwei beliebige Vektoren. Dann gilt

$$\langle v \mid w \rangle = x^\top \overline{y}$$

und

$$\langle \varphi(v) \mid \varphi(w) \rangle = (Ax)^\top \overline{Ay} = x^\top \left(A^\top \overline{A} \right) \overline{y}\,.$$

Somit ist φ genau dann orthogonal bzw. unitär wenn $A^\top \overline{A} = E_n$ gilt. □

8.3 Der Spektralsatz für normale Endomorphismen

8.3.1 Definition

Die Menge der Eigenwerte eines Endomorphismus eines endlich-dimensionalen Vektorraums wird auch als sein *Spektrum* bezeichnet. Spektraltheorie ist ein anderes Wort für die Theorie der Eigenwerte und Eigenvektoren.

8.3.2 Satz

Es sei V ein endlich-dimensionaler unitärer Vektorraum, φ ein normaler Endomorphismus von V und λ ein Eigenwert von φ mit Eigenvektor y. Dann gilt

1. $\varphi - \lambda \mathrm{id}$ ist normal.
2. $\varphi^*(y) = \overline{\lambda} y$.
3. Sind λ, μ zwei verschiedene Eigenwerte so ist $E(\lambda, \varphi) \perp E(\mu, \varphi)$. D.h. Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.

Beweis:

1. Es gilt

   $$(\varphi - \lambda \mathrm{id})^* \circ (\varphi - \lambda \mathrm{id}) = (\varphi^* - \overline{\lambda} \mathrm{id}) \circ (\varphi - \lambda \mathrm{id}) = \varphi^* \circ \varphi - \overline{\lambda}\varphi - \lambda\varphi^* + \lambda\overline{\lambda}\mathrm{id}\,.$$

   In gleicher Art zeigt man

   $$(\varphi - \lambda \mathrm{id}) \circ (\varphi - \lambda \mathrm{id})^* = \varphi \circ \varphi^* - \overline{\lambda}\varphi - \lambda\varphi^* + \lambda\overline{\lambda}\mathrm{id}\,.$$

   Da φ normal ist, sind die rechten Seiten gleich und somit ist $\varphi - \lambda \mathrm{id}$ normal.

2. Aus $(\varphi - \lambda\mathrm{id})(y) = 0$ folgt

$$\begin{aligned}0 &= \langle (\varphi - \lambda\mathrm{id})(y) \mid (\varphi - \lambda\mathrm{id})(y)\rangle \\ &= \langle y \mid (\varphi - \lambda\mathrm{id})^* \circ (\varphi - \lambda\mathrm{id})(y)\rangle \\ &= \langle y \mid (\varphi - \lambda\mathrm{id}) \circ (\varphi - \lambda\mathrm{id})^*(y)\rangle \\ &= \langle (\varphi - \lambda\mathrm{id})^*(y) \mid (\varphi - \lambda\mathrm{id})^*(y)\rangle\end{aligned}$$

   Dies liefert $(\varphi - \lambda\mathrm{id})^*(y) = 0$ und somit $\varphi^*(y) = \overline{\lambda} y$.

3. Sei z ein Eigenvektor zum Eigenwert μ, so gilt

$$\lambda\langle y \mid z\rangle = \langle \lambda y \mid z\rangle = \langle \varphi(y) \mid z\rangle = \langle y \mid \varphi^*(z)\rangle = \langle y \mid \overline{\mu} z\rangle = \mu\langle y \mid z\rangle\,.$$

   Aus $\mu \neq \lambda$ folgt $\langle y \mid z\rangle = 0$.

8.3.3 Bemerkung

Sei φ ein beliebiger Endomorphismus eines Skalarproduktraums mit Eigenwert λ und A seine Darstellungsmatrix bezüglich einer beliebigen Orthonormalbasis. Dann gilt

$$\begin{aligned} 0 &= \det(\varphi - \lambda\mathrm{id}) = \det(A - \lambda E_n) \\ \Longrightarrow 0 &= \det(\overline{A - \lambda E_n}^{\top}) = \det(A^* - \overline{\lambda} E_n) = \det(\varphi^* - \overline{\lambda}\mathrm{id})\,.\end{aligned}$$

D.h. adjungieren eines Endomorphismus konjugiert die Eigenwerte. Im Spezialfall eines normalen Endomorphismus bleiben die Eigenvektoren erhalten.

8.3.4 Definition

Es sei $\varphi\colon V \to V$ eine lineare Abbildung. Ein Unterraum $U \subset V$ ist *φ-invariant*, wenn $\varphi(U) \subset U$ gilt.

8.3.5 Spektralsatz

Es sei φ ein Endomorphismus des endlich-dimensionalen unitären Vektorraums V und $\lambda_1, \dots, \lambda_k$ die Eigenwerte von φ. Dann sind äquivalent:

1. φ ist normal.

2. $E(\lambda_i, \varphi) \perp E(\lambda_j, \varphi)$ für alle $i \neq j$ und für jeden Eigenwert stimmen die algebraische und die geometrische Vielfachheit überein.

3. $V = E(\lambda_1, \varphi) \oplus \dots \oplus E(\lambda_k, \varphi)$ und die Summanden sind zueinander paarweise orthogonal.

4. Es existiert eine Orthonormalbasis von V aus Eigenvektoren von φ.

5. Es existieren Orthogonalprojektionen $\pi_1, \dots, \pi_k$ mit $\pi_1 + \dots + \pi_k = \mathrm{id}$ und $\pi_i \circ \pi_j = 0$ für $i \neq j$ sowie

$$\varphi = \lambda_1 \pi_1 + \dots + \lambda_k \pi_k\,.$$

   Dies wird als *Spektraldarstellung* von φ bezeichnet.

Beweis:

$1 \Longrightarrow 2$**:** Es ist noch zu zeigen, dass die algebraische und geometrische Vielfachheit eines jeden Eigenwerts gleich sind. Die Unterräume $E(\lambda_i, \varphi)$ sind φ- und φ^*-invariant, da $\varphi(v) = \lambda v$ und $\varphi^*(v) = \overline{\lambda} v$ für alle $v \in E(\lambda_i, \varphi)$ gilt.

Es ist aber auch $E(\lambda_i, \varphi)^\perp$ ein φ-invarianter Unterraum, denn es gilt

$$\begin{aligned}
& v \in E(\lambda_i, \varphi)^\perp \\
\Longleftrightarrow\ & \forall w \in E(\lambda_i, \varphi) \colon \langle v \mid w \rangle = 0 \\
\Longrightarrow\ & \forall w \in E(\lambda_i, \varphi) \colon \langle v \mid \varphi^*(w) \rangle = 0 \text{ da } \varphi^* E(\lambda_i, \varphi) \subset E(\lambda_i, \varphi) \\
\Longleftrightarrow\ & \forall w \in E(\lambda_i, \varphi) \colon \langle \varphi(v) \mid w \rangle = 0 \\
\Longleftrightarrow\ & \varphi(v) \in E(\lambda_i, \varphi)^\perp \,.
\end{aligned}$$

Wählen wir eine Orthonormalbasis $(b_1, \ldots, b_\ell)$ des Eigenraums $E(\lambda_i, \varphi)$, so können wir diese zu einer Orthonormalbasis $\mathcal{B} = (b_1, \ldots, b_n)$ ergänzen. Es ist dann $E(\lambda_i, \varphi)^\perp = \operatorname{span}(b_{\ell+1}, \ldots, b_n)$. Wir erhalten

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix}
\lambda_i & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & \lambda_i & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_i & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & c_{1,1} & \cdots & c_{1,n-\ell} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & c_{n-\ell,1} & \cdots & c_{n-k,n-\ell}
\end{pmatrix}$$

$$\chi_\varphi(x) = (\lambda_i - x)^\ell \chi_C(x) \text{ mit}$$

$$C = \begin{pmatrix}
c_{1,1} & \cdots & c_{1,n-\ell} \\
\vdots & \ddots & \vdots \\
c_{n-\ell,1} & \cdots & c_{n-k,n-\ell}
\end{pmatrix}$$

Die Matrix C ist die Darstellungsmatrix von $\varphi_{E(\lambda_i,\varphi)^\perp}$ bezüglich der Basis $(b_{\ell+1}, \ldots, b_n)$. Weiterhin ist ℓ die geometrische Vielfachheit des Eigenwerts λ_i. Angenommen, die algebraische Vielfachheit von λ_i wäre noch größer als ℓ, so wäre $\chi_C(\lambda_i) = 0$. Dann wäre $\operatorname{Ker}(C - \lambda_i \mathrm{id}) \neq \{0\}$ und $\varphi_E(\lambda_i, \varphi)^\perp$ hätte den Eigenwert λ_i. Dies ist ein Widerspruch, da ein zugehöriger Eigenvektor in $E(\lambda_i, \varphi)^\perp \cap E(\lambda_i, \varphi) = \{0\}$ liegen würde.

Somit stimmen die algebraischen und die geometrischen Vielfachheiten aller Eigenwerte überein.

$2 \Longrightarrow 3$**:** Da im Körper der komplexen Zahlen jedes Polynom vollständig in Linearfaktoren zerfällt und für jeden Eigenwert die geometrische und die algebraische Vielfachheit übereinstimmen ist φ diagonalisierbar. Somit bilden die Eigenräume in direkter Summe den gesamten Raum. Die Orthogonalität der Summanden ist bereits vorausgesetzt.

$3 \Longrightarrow 4$**:** Wir wählen zu jedem Eigenraum $E(\lambda_i, \varphi)$ eine Orthonormalbasis. Da die verschiedenen Eigenräume zueinander orthogonal sind, entsteht so ein Orthonormalsystem. Da die Summe der Eigenräume den gesamten Raum V ergibt entsteht so eine Basis von V.

$4 \Longrightarrow 5$**:** Es sei $(b_1, \dots, b_n)$ eine Orthonormalbasis von V aus Eigenvektoren von φ geordnet nach Eigenwerten. D.h., $b_1, \dots, b_{i_1}$ sind Eigenvektoren zum Eigenwert λ_1, $b_{i_1+1}, \dots, b_{i_2}$ sind Eigenvektoren zum Eigenwert λ_2, usw.

Wir setzen

$$\pi_j(x) := \pi_{\mathrm{span}(b_{i_{j-1}+1}, \dots, b_{i_j})}(x) = \langle x \mid b_{i_{j-1}+1} \rangle b_{i_{j-1}+1} + \cdots + \langle x \mid b_{i_j} \rangle b_{i_j} \, .$$

Dieses sind Orthogonalprojektionen. Nach 7.3.4 gilt $\pi_1(v) + \cdots + \pi_k(v) = v$ und

$$\begin{aligned} \varphi(v) &= \varphi(\pi_1(v) + \cdots + \pi_k(v)) = \varphi(\pi_1(v)) + \cdots + \varphi(\pi_k(v)) \\ &= \lambda_1 \pi_1(v) + \cdots + \lambda_k \pi_k(v) \end{aligned}$$

Die Gleichheit $\pi_i \circ \pi_j = 0$ für $i \neq j$ folgt aus der Orthogonalität von $(b_1, \dots, b_n)$.

$5 \Longrightarrow 1$**:** Wir haben $\pi_i = \pi_i^*$ bereits gezeigt. Somit gilt

$$\begin{aligned} \varphi \circ \varphi^* &= (\lambda_1 \pi_1 + \cdots + \lambda_k \pi_k) \circ (\lambda_1 \pi_1 + \cdots + \lambda_k \pi_k)^* \\ &= (\lambda_1 \pi_1 + \cdots + \lambda_k \pi_k) \circ (\overline{\lambda_1} \pi_1 + \cdots + \overline{\lambda_k} \pi_k) \\ &= \lambda_1 \overline{\lambda_1} \pi_1 + \cdots + \lambda_k \overline{\lambda_k} \pi_k \end{aligned}$$

und

$$\varphi^* \circ \varphi = \overline{\lambda_1} \lambda_1 \pi_1 + \cdots + \overline{\lambda_k} \lambda_k \pi_k$$

welches die Normalität von φ zeigt. □

8.3.6 Beispiel

Sei $\varphi \in \mathrm{End}(\mathbb{R}^3)$ durch die Matrix

$$\begin{bmatrix} 2 & -4 & -2 \\ -4 & 2 & -2 \\ -2 & -2 & 5 \end{bmatrix}$$

gegeben. Da die Standardbasis des $\mathbb{R}^3$ eine Orthonormalbasis ist, ist φ selbstadjungiert und insbesondere normal. Es gilt $\langle x \mid \varphi(y) \rangle = \langle \varphi(x) \mid y \rangle$.

Berechnung der Eigenwerte:

$$\begin{aligned} &\begin{vmatrix} 2-\lambda & -4 & -2 \\ -4 & 2-\lambda & -2 \\ -2 & -2 & 5-\lambda \end{vmatrix} = \\ &= (2-\lambda)^2 \cdot (5-\lambda) + (2-\lambda) \cdot (-8) + (5-\lambda)(-16) - 32 \\ &= -\lambda^3 + 9\lambda^2 - 24\lambda + 20 - 16 + 8\lambda - 80 + 16\lambda - 32 \\ &= -\lambda^3 + 9\lambda^2 - 108 = (\lambda+3)(-\lambda^2 + 12\lambda - 36) = -(\lambda+3)(\lambda-6)^2 \end{aligned}$$

φ hat die Eingenwerte -3 und 6 mit algebraischer und geometrischer Vielfachheit 1 und 2.

Berechnung der Eigenräume:

$$E(-3,\varphi)=\operatorname{Ker}\begin{pmatrix}5&-4&-2\\-4&5&-2\\-2&-2&8\end{pmatrix}=\operatorname{span}(2,2,1)^\top$$

$$\begin{aligned}E(6,\varphi)&=\operatorname{Ker}\begin{pmatrix}-4&-4&-2\\-4&-4&-2\\-2&-2&-1\end{pmatrix}=\operatorname{span}\left((1,0,-2)^\top,(0,1,-2)^\top\right)\\&=\operatorname{span}\left((1,0,-2)^\top,(4,-5,2)^\top\right)\end{aligned}$$

Es ist $\mathcal{B}=\left(\frac{1}{3}(2,2,1)^\top,\frac{1}{\sqrt{5}}(1,0,-2)^\top,\frac{1}{3\sqrt{5}}(4,-5,2)^\top\right)$ eine Orthonormalbasis des $\mathbb{R}^3$ aus Eigenvektoren von φ und

$$M_{\mathcal{B}}(\varphi)=\begin{pmatrix}-3&0&0\\0&6&0\\0&0&6\end{pmatrix},\quad T_{\mathcal{E}}^{\mathcal{B}}=\begin{pmatrix}\frac{2}{3}&\frac{1}{\sqrt{5}}&\frac{4}{3\sqrt{5}}\\\frac{2}{3}&0&\frac{-5}{3\sqrt{5}}\\\frac{1}{3}&\frac{-2}{\sqrt{5}}&\frac{2}{3\sqrt{5}}\end{pmatrix},$$

$$T_{\mathcal{B}}^{\mathcal{E}}=(T_{\mathcal{E}}^{\mathcal{B}})^{-1}=\begin{pmatrix}\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\\frac{1}{\sqrt{5}}&0&\frac{-2}{\sqrt{5}}\\\frac{4}{3\sqrt{5}}&\frac{-5}{3\sqrt{5}}&\frac{2}{3\sqrt{5}}\end{pmatrix}$$

Spektraldarstellung:

$$\varphi(v)=-3\pi_{E(-3,\varphi)}(v)+6\pi_{E(6,\varphi)}(v)$$

mit

$$M_{\mathcal{B}}(\pi_{E(-3,\varphi)})=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}\text{ und }M_{\mathcal{B}}(\pi_{E(6,\varphi)})=\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

Endomorphismenpolynome: Allgemein gilt

$$M_{\mathcal{B}}(\varphi^k)=M_{\mathcal{B}}(\varphi)^k\,.$$

Wir betrachten

$$\chi_\varphi(X)=-(X+3)(X-6)^2$$

und

$$M_{\mathcal{B}}(\chi_\varphi(\varphi))=\chi_\varphi(M_{\mathcal{B}}(\varphi))=\begin{pmatrix}\chi_\varphi(-3)&0&0\\0&\chi_\varphi(6)&0\\0&0&\chi_\varphi(6)\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}.$$

8.3.7 Definition

Ein Endomorphismus φ eines Skalarproduktraums heißt *idempotent (Projektion)* wenn $\varphi=\varphi\circ\varphi$ gilt. Er heißt *strikt positiv*, wenn $\langle\varphi(v)\mid v\rangle>0$ für alle $v\in V, v\neq 0$ gilt.

8.3.8 Satz (Charakterisierung von Abbildungen durch ihre Eigenwerte)

Sei φ ein Endomorphismus eines endlich-dimensionalen Vektorraums und $\sigma(\varphi)$ sei die Menge aller Eigenwerte. Dann ist φ genau dann invertierbar, wenn $0 \notin \sigma(\varphi)$ gilt. Ist V zudem ein unitärer Raum und φ normal, so gilt:

1. φ ist hermitesch $\iff \sigma(\varphi) \subset \mathbb{R}$.

2. φ ist unitär $\iff \sigma(\varphi) \subset \{z \in \mathbb{C} \mid |z| = 1\}$.

3. φ ist idempotent $\iff \sigma(\varphi) \subset \{0, 1\}$.

4. φ ist strikt positiv $\iff \sigma(\varphi) \subset \,]0, \infty[$.

Beweis:

Es ist φ genau dann invertierbar, wenn $\det(\varphi) = \chi_\varphi(0) \neq 0$ gilt. Für die weiteren Aussagen ist φ als normal vorausgesetzt und der Grundkörper ist $\mathbb{C}$. Nach dem Spektralsatz gibt es eine Orthonormalbasis $\mathcal{B} = (b_1, \ldots, b_n)$ von V aus Eigenvektoren von φ. Dann ist $M_\mathcal{B}(\varphi) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ mit $\{\lambda_1, \ldots, \lambda_n\} = \sigma(\varphi)$. Es folgt

1. φ ist hermitesch $\iff M_\mathcal{B}(\varphi)$ ist hermitesch $\iff M_\mathcal{B}(\varphi) = \overline{M_\mathcal{B}(\varphi)^\top}$. Somit sind alle λ_i reell. D.h. alle Eigenwerte sind reell.

2. φ ist unitär $\iff M_\mathcal{B}(\varphi)$ ist unitär $\iff M_\mathcal{B}(\varphi)^{-1} = \overline{M_\mathcal{B}(\varphi)^\top}$. D.h. es gilt $\lambda_i^{-1} = \overline{\lambda_i}$. Dies ist zu $|\lambda_i| = 1$ äquivalent.

3. φ ist idempotent $\iff M_\mathcal{B}(\varphi) = M_\mathcal{B}(\varphi)^2$. Dies ist zu $\lambda_i = \lambda_i^2$ äquivalent. Dies impliziert $\lambda_i \in \{0, 1\}$.

4. Wir schreiben v als $\sum_i \mu_i b_i$. Dann gilt

$$\begin{aligned}\langle \varphi(v) \mid v \rangle &= \langle \varphi(\sum_i \mu_i b_i) \mid \sum_j \mu_j b_j \rangle = \langle \sum_i \mu_i \lambda_i b_i \mid \sum_j \mu_j b_j \rangle \\ &= \sum_i \mu_i \lambda_i \overline{\mu_i} \langle b_i \mid b_i \rangle = \sum_i \lambda_i |\mu_i|^2 \,.\end{aligned}$$

   Somit ist die strikte Positivität von φ zu $\lambda_i > 0$ für alle $i = 1, \ldots, n$ äquivalent.

□

8.4 Anwendung auf symmetrische und hermitesche Matrizen

8.4.1 Motivation und Erinnerung

Symmetrische und hermitesche Matrizen können spezielle Endomorphismen oder spezielle Bi- und Sesquilinearformen repräsentieren. Wir formulieren die bisherigen Ergebnisse für diese Situation im Detail.

8.4.2 Satz

Sei V ein endlich-dimensionaler unitärer (euklidischer) Raum und φ ein selbstadjungierter (symmetrischer) Endomorphismus. Dann existiert eine Orthonormalbasis $\mathcal{B}$ von V, sodass

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

gilt. Alle Eigenwerte sind reell und mit ihrer Vielfachheit vertreten. Insbesondere hat eine reelle symmetrische $n \times n$-Matrix immer n reelle Eigenwerte (mit Vielfachheit).

Beweis:

Im unitären Fall ist φ normal. Nach dem Spektralsatz gibt es eine Orthonormalbasis aus Eigenvektoren. Insbesondere ist φ diagonalisierbar und die algebraische und geometrische Vielfachheit stimmt für jeden Eigenwert überein. Nach 8.3.8 sind alle Eigenwerte reell.

Zum euklidischen Fall: Sei $\mathcal{A}$ eine beliebige Orthonormalbasis von V. Dann ist die Matrix $M := M_{\mathcal{A}}(\varphi)$ eine reelle symmetrische $n \times n$-Matrix. Diese Matrix definiert eine lineare Abbildung $\varphi_{\mathbb{C}} \colon \mathbb{C}^n \to \mathbb{C}^n$ die selbstadjungiert bezüglich des Standardskalarprodukts auf $\mathbb{C}^n$ ist. Anwendung obiger Überlegungen auf $\varphi_{\mathbb{C}}$ liefert, dass $\varphi_{\mathbb{C}}$ nur reelle Eigenwerte hat. Insbesondere zerfällt das charakteristische Polynom von M bereits im reellen vollständig in Linearfaktoren.

Es bleibt zu zeigen, dass die algebraische und die geometrische Vielfachheit für jeden Eigenwert von φ übereinstimmt. Weiterhin ist zu zeigen, dass Eigenvektoren zu verschiedenen Eigenwerten zueinander orthogonal sind.

Es sei $U = E(\lambda_1, \varphi) + \cdots + E(\lambda_n, \varphi)$ dann ist U als Erzeugnis aller Eigenräume ein φ-invarianter Unterraum und $\varphi|U \colon U \to U$ ist diagonalisierbar.

Nehmen wir an, dass $U \neq V$ gilt, so gilt für $u \in U$ und $v \in U^\perp$

$$\langle u \mid \varphi(v) \rangle = \langle \varphi^*(u) \mid v \rangle = \langle \varphi(u) \mid v \rangle = 0\,.$$

Somit ist $\varphi(v) \in U^\perp$ und man kann $\varphi_{U^\perp} \colon U^\perp \to U^\perp$ bilden. Das charakteristische Polynom von $\varphi_{U^\perp}$ ist ein Teiler des charakteristischen Polynoms von φ. Somit hat auch $\varphi_{U^\perp}$ nur reelle Eigenwerte. Da es zu jedem Eigenwert λ mindestens einen Eigenvektor gibt, hat $\varphi_{U^\perp}$ mindestens einen Eigenvektor in $U^\perp \cap E(\lambda, \varphi)$. Dies widerspricht $U^\perp \perp E(\lambda_1, \varphi) + \cdots + E(\lambda_n, \varphi)$. Folglich ist U der Nullraum.

Seien nun v_1, v_2 Eigenvektoren zu verschiedenen Eigenwerten λ_1, λ_2. Dann gilt

$$\begin{aligned} \lambda_1 \langle v_1 | v_2 \rangle &= \langle \lambda_1 v_1 | v_2 \rangle = \langle \varphi(v_1) | v_2 \rangle = \langle v_1 | \varphi^*(v_2) \rangle \\ &= \langle v_1 | \varphi(v_2) \rangle = \langle v_1 | \lambda_2 v_2 \rangle = \lambda_2 \langle v_1 | v_2 \rangle \end{aligned}$$

Da $\lambda_1 \neq \lambda_2$ gilt, folgt $\langle v_1 | v_2 \rangle = 0$.

Dies zeigt: Wählt man für jeden Eigenraum eine Orthonormalbasis, so bilden diese Vektoren zusammengenommen eine Orthonormalbasis von V. $\square$

8.4.3 Satz

Sei V ein endlich-dimensionaler unitärer (euklidischer) Raum.

1. Ist φ eine lineare Abbildung, so ist $(v, w) \mapsto \langle \varphi(v)|w\rangle$ eine Sesquilinearform (Bilinearform).
2. Es sei β eine Sesquilinearform (Bilinearform). Dann existiert eine eindeutig bestimmte lineare Abbildung $\varphi\colon V \to V$ so, dass $\beta(v, w) = \langle \varphi(v)|w\rangle$ gilt. Notation: φ_β

Insbesondere sind der Raum der Endomorphismen und der Raum der Sesquilinearformen (Bilinearformen) isomorph.

Beweis:

1. Die Sesqui- bzw. Bilinearität von $\langle \ldots | \ldots\rangle$ impliziert die gleiche Eigenschaft für $(v, w) \mapsto \langle \varphi(v)|w\rangle$.
2. Sei $(b_1, \ldots, b_n)$ eine Orthonormalbasis von V. Falls ein solches φ existiert, so gilt

$$\begin{aligned}\varphi(b_i) &= \langle \varphi(b_i)|b_1\rangle b_1 + \cdots + \langle \varphi(b_i)|b_n\rangle b_n \\ &= \beta(b_i, b_1)b_1 + \cdots + \beta(b_i, b_n)b_n\end{aligned}$$

   Somit gibt es höchstens ein solches φ.

   Ist φ wie oben gewählt, so gilt

$$\begin{aligned}\langle \varphi(b_i)|b_j\rangle &= \langle \beta(b_i, b_1)b_1 + \cdots + \beta(b_i, b_n)b_n|b_j\rangle \\ &= \langle \beta(b_i, b_1)b_1|b_j\rangle + \cdots + \langle \beta(b_i, b_n)b_n|b_j\rangle = \beta(b_i, b_j)\,.\end{aligned}$$

   Somit stimmen die beiden Formen für jedes Paar von Basisvektoren überein. Somit sind sie gleich.

Die Linearität der Zuordnung *Lineare Abbildungen* $\to$ *Sesquilinearformen (Bilinearformen)* ist klar. □

8.4.4 Bemerkung

Betrachten wir $\mathbb{C}^n$ mit dem Standard-Skalarprodukt und die Form $\beta(x, y) = x^\top A\overline{y}$, so gilt

$$\beta(x, y) = x^\top A\overline{y} = (A^\top x)^\top \overline{y} = \langle A^\top x|y\rangle\,.$$

D.h. in dieser Situation entsteht die Darstellungsmatrix der obigen linearen Abbildung durch transponieren aus der Darstellungsmatrix von β.

8.4.5 Satz

Sei β eine hermitesche (symmetrische) Sesquilinearform (Bilinearform) auf einem endlich-dimensionalen unitären (euklidischen) Raum. Dann ist φ_β selbstadjungiert.

Es existiert eine Orthonormalbasis $\mathcal{B}$, sodass $M_\mathcal{B}(\beta)$ eine Diagonalmatrix ist. Die Basis $\mathcal{B}$ besteht aus Eigenvektoren von φ_β. Die Diagonaleinträge von $M_\mathcal{B}(\beta)$ sind die Eigenwerte von φ_β. D.h., es gilt

$$M_\mathcal{B}(\beta) = M_\mathcal{B}(\varphi_\beta)\,.$$

Beweis:

Sei β hermitesch (symmetrisch), so gilt

$$\langle \varphi_\beta(v)|w\rangle = \beta(v,w) = \overline{\beta(w,v)} = \overline{\langle \varphi_\beta(w)|v\rangle} = \langle v|\varphi_\beta(w)\rangle\,.$$

Somit ist φ_β selbstadjungiert (symmetrisch). Nach 8.4.2 existiert somit eine Orthonormalbasis $\mathcal{B} = (b_1,\ldots,b_n)$ aus Eigenvektoren von φ_β. Sei $M_\mathcal{B}(\varphi_\beta) = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$, so folgt

$$\beta(b_i,b_j) = \langle \varphi_\beta(b_i)|b_j\rangle = \langle \lambda_i b_i|b_j\rangle = \lambda_i\delta_{i,j}\,.$$

Dies zeigt $M_\mathcal{B}(\beta) = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. □

8.4.6 Satz

Sei $M \in \operatorname{Mat}(n\times n,\mathbb{C})$ hermitesch. Dann existiert eine unitäre Matrix S mit $S^*MS = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Die λ_i sind die Eigenwerte von M. Sie sind reell und treten entsprechend ihrer Vielfachheit auf.

Sei $M \in \operatorname{Mat}(n\times n,\mathbb{R})$ symmetrisch. Dann existiert eine orthogonale Matrix S mit $S^\top MS = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Die λ_i sind die Eigenwerte von M. Sie treten entsprechend ihren Vielfachheiten auf.

Beweis:

Das Standardskalarprodukt gibt $V = \mathbb{C}^n$ (bzw. $\mathbb{R}^n$) die Stuktur eines Skalarproduktraums. Die Standardbasis $\mathcal{E} = (e_1,\ldots,e_n)$ ist dann eine Orthonormalbasis. Die lineare Abbildung $\varphi\colon v\mapsto Mv$ ist dann hermitesch (bzw. symmetrisch). Nach 8.4.2 existiert eine Orthonormalbasis $\mathcal{B}$, sodass $M_\mathcal{B}(\varphi) = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$ gilt. Zudem sind alle Eigenwerte reell.

Die Basiswechselmatrix $S \coloneqq T_\mathcal{E}^\mathcal{B}$ ist unitär (orthogonal), da beides Orthonormalbasen sind. Somit haben wir

$$\operatorname{diag}(\lambda_1,\ldots,\lambda_n) = M_\mathcal{B}(\varphi) = S^{-1}M_\mathcal{E}(\varphi)S = S^{-1}MS$$

mit $S^* = S^{-1}$ (bzw. $S^\top = S^{-1}$). Es folgt die Behauptung. □

8.5 Unitäre und orthogonale Endomorphismen

8.5.1 Satz

Sei V ein endlich-dimensionaler unitärer Raum und φ unitär. Dann existiert eine Orthonormalbasis $\mathcal{B}$ von V, sodass $M_\mathcal{B}(\varphi) = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$ gilt. Die λ_i sind die Eigenwerte (mit Vielfachheit gezählt) von φ. Zudem gilt $|\lambda_i| = 1$ für $i = 1,\ldots,n$.

Beweis:

Dies folgt direkt aus dem Spektralsatz und 8.3.8.

8.5.2 Lemma

Es sei φ eine orthogonaler Endomorphismus des endlich-dimensionalen Raums V mit $\dim(V) > 0$.

1. Ist $U \subset V$ ein φ-invarianter Unterraum, so ist auch $U^\perp$ ein φ-invarianter Raum.

2. In V gibt es einen 1- oder 2-dimensionalen φ-invarianten Unterraum U. Gibt es keinen 1-dimensionalen φ-invarianten Unterraum, so hat U eine Orthonormalbasis $\mathcal{A}$, sodass

$$M_\mathcal{A}(\varphi|_U) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \text{ mit } \theta \in (0, 2\pi)$$

gilt.

Beweis:

1. φ ist als orthogonaler Endomorphismus bijektiv. Aus $\varphi(U) \subset U$ folgt dann $U = \varphi(U)$. Sei nun $u \in U$ und $v \in U^\perp$, so gibt es ein $u_0 \in U$ mit $\varphi(u_0) = u$ und es folgt
$$\langle u|\varphi(v)\rangle = \langle \varphi(u_0)|\varphi(v)\rangle = \langle u_0|v\rangle = 0\,.$$
Somit ist $\varphi(v) \in U^\perp$.

2. Angenommen φ hat einen reellen Eigenwert, so gibt es zu diesem mindestens einen Eigenvektor. Der von einem Eigenvektor aufgespannte Unterraum ist dann φ-invariant.

   Angenommen φ hat keinen reellen Eigenwert, so muss man die Eigenwerte im Komplexen suchen. Sei hierzu $\mathcal{B}$ eine Orthonormalbasis von V dann ist $M = M_\mathcal{B}(\varphi)$ eine reelle orthogonale $n \times n$-Matrix.

   Sei $\tilde{\varphi}\colon \mathbb{C}^n \to \mathbb{C}^n$ die durch $x \mapsto Mx$ gegebene Abbildung. Sie ist bezüglich des Standardskalarprodukts unitär.

   Sei $\lambda = \cos(\theta) + i\sin(\theta)$ mit $\theta \in (0, 2\pi)$ ein echt komplexer Eigenwert mit Eigenvektor $v = s + it \in \mathbb{C}^n$ (hier $s, t \in \mathbb{R}^n$). Dann gilt

$$\begin{aligned} &\cos(\theta)s - \sin(\theta)t + i\,(\sin(\theta))s + \cos(\theta)t) \\ &= (\cos(\theta) + i\sin(\theta))(s + it) = Ms + iMt \end{aligned}$$

   Dieses zeigt $Ms = \cos(\theta)s - \sin(\theta)t$ und $Mt = \sin(\theta))s + \cos(\theta)t$.

   Weiterhin zeigt

$$\lambda(s + it) = M(s + it) \Longrightarrow \overline{\lambda}(s - it) = M(s - it)\,,$$

   dass das komplex Konjugierte eines Eigenwerts ebenfalls Eigenwert mit komplex konjugiertem Eigenvektor ist. Nach dem Spektralsatz gilt $(s + it)\perp(s - it)$ und somit

$$\begin{aligned} 0 &= \langle s + it|s - it\rangle = \langle s|s\rangle + \langle it|s\rangle + \langle s| - it\rangle + \langle it| - it\rangle \\ &= \langle s|s\rangle + 2i\langle t|s\rangle - \langle t|t\rangle \\ \Longrightarrow \quad & \|s\| = \|t\| \text{ und } s\perp t\,. \end{aligned}$$

Wählen wir $\mathcal{A} = (\frac{1}{\|s\|}s, \frac{1}{\|t\|}t)$ und $U = \operatorname{span}_{\mathbb{R}}(\mathcal{A}) \subset \mathbb{R}^n$ so erhalten wir einen φ-invarianten Unterraum mit

$$M_{\mathcal{A}}(\varphi|_U) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \text{ mit } \theta \in (0, 2\pi)\,.$$

□

8.5.3 Satz

Sei V ein endlich-dimensionaler euklidischer Raum und $\varphi\colon V \to V$ eine orthogonale Abbildung. Dann existiert eine Orthonormalbasis

$$\mathcal{B} = (a_1, \ldots, a_r, b_1, c_1, \ldots, b_s, c_s)$$

von V, sodass

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} \alpha_1 &&&&& \\ & \ddots &&&& \\ && \alpha_r &&& \\ &&& \Delta_1 && \\ &&&& \ddots & \\ &&&&& \Delta_s \end{pmatrix}$$

mit $\alpha_1, \ldots, \alpha_r \in \{-1, 1\}$ und

$$\Delta_i = \begin{pmatrix} \cos(\theta_i) & \sin(\theta_i) \\ -\sin(\theta_i) & \cos(\theta_i) \end{pmatrix} \text{ mit } \sin(\theta_i) \neq 0$$

gilt. Hierbei ist $n = \dim V = r + 2s$.

Beweis:

Sei $\mathcal{B}$ eine beliebige Orthonormalbasis von V. Wir betrachten zunächst ein paar Spezialfälle:

$\dim(V) = 0$: trivial.

$\dim(V) = 1$: Es ist $M_{\mathcal{B}}(\varphi)$ eine orthogonale 1×1-Matrix. Diese ist $[-1]$ oder $[1]$.

$\dim(V) = 2$: Hat φ keinen 1-dimensionalen invarianten Unterraum, so sind alle Eigenwerte von φ echt komplex. Nach obigem Lemma oder 8.2.14 hat $M_{\mathcal{B}}(\varphi)$ obige Form.

Den allgemeinen Fall beweisen wir durch Induktion nach $\dim V$. Der Induktionsanfang ist durch obige Spezialfälle abgedeckt.

Sei nun $\dim V > 1$ und die Behauptung für alle Räume kleinerer Dimension bereits bewiesen. Sei U ein φ-invarianter Unterraum kleinster positiver Dimension. Dann ist $\dim(V) > \dim(U) \in \{1, 2\}$.

Nach obigem Lemma gibt $\varphi(U) \subset U$ und $\varphi(U^\perp) \subset U^\perp$. Es sind $\varphi|_U : U \to U$ und $\varphi|_{U^\perp} : U^\perp \to U^\perp$ orthogonale Endomorphismen. Somit existiert nach Induktionsvoraussetzung für U und $U^\perp$ Orthonormalbasen

$$\begin{aligned} \mathcal{B}' &= (a_1, \ldots, a_k, b_1, c_1, \ldots, b_l, c_l) \\ \mathcal{B}'' &= (a_{k+1}, \ldots, a_r, b_{l+1}, c_{l+1}, \ldots, b_s, c_s) \end{aligned}$$

so, dass die Darstellungsmatrizen $M_{\mathcal{B}'}(\varphi|_U)$ und $M_{\mathcal{B}''}(\varphi|_{U^\perp})$ die obige Form haben.

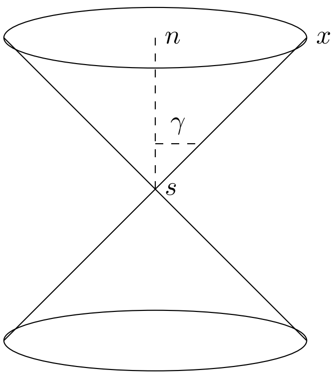
Da $V = U \oplus U^\perp$ gilt, ist $\mathcal{B} = (a_1, \ldots, a_r, b_1, c_1, \ldots, b_s, c_s)$ eine Orthonormalbasis von V, sodass $M_{\mathcal{B}}(\varphi)$ die gewünschte Form hat. $\square$

Kapitel 9

Quadriken im $\mathbb{R}^n$

9.1 Kegelschnitte und Quadriken

9.1.1 Bemerkung



Der Kreiskegel K mit Spitze s und Achse n und Öffnungswinkel γ ist durch

$$x \in K \Longleftrightarrow x = s \text{ oder } \frac{\langle x - s | n \rangle}{\|n\| \cdot \|x - s\|} = \cos\gamma$$

gegeben. Durch Quadrieren entsteht hieraus die Polynomgleichung

$$\langle x - s | n \rangle^2 = \|n\|^2 \|x - s\|^2 (\cos\gamma)^2 ,$$

welche ebenfalls für $x = s$ gültig ist. Der Schnitt mit der $x_3 = 0$ Ebene liefert die Gleichung

$$\begin{aligned} KS &= \{(x_1, x_2) \in \mathbb{R}^2 \mid ax_1^2 + 2bx_1x_2 + cx_2^2 + 2dx_1 + 2ex_2 + f = 0\} \\ &= \{x \in \mathbb{R}^2 \mid x^\top \begin{pmatrix} a & b \\ b & c \end{pmatrix} x + 2\,(d\ e)\,x + f = 0\} \end{aligned}$$

Jeder Kegelschnitt ist von dieser Form, jedoch sind die Koeffizienten $a, \ldots, f$ noch gewissen Einschränkungen unterlegen.

9.1.2 Definition

Sei $A \neq 0$ eine symmetrische reelle $n \times n$-Matrix und $b \in \mathbb{R}^n$ sowie $c \in \mathbb{R}$. Dann ist eine Quadrik im $\mathbb{R}^n$ die Menge

$$\{x \in \mathbb{R}^n \mid x^\top Ax + 2b^\top x + c = 0\}.$$

Beispiel: $x^2 + y^2 + 1 = 0$ ist eine Quadrik aber kein Kegelschnitt.

9.1.3 Bemerkung

Interpretiert man $x \mapsto x^\top Ax$ als quadratische Form und $x \mapsto 2b^\top x$ als Linearform, so ist eine Quadrik die Nullstellenmenge eines inhomogenen quadratischen Polynoms.

Betrachtet man einen abstrakten $\mathbb{R}$-Vektorraum V, so heißt die Menge

$$Q := \{v \in V \mid \beta(v, v) + l(v) + c = 0\}$$

eine Quadrik. Hierbei ist $\beta\colon V \times V \to \mathbb{R}$ eine Bilinearform und $l \in \operatorname{Hom}(V, \mathbb{R})$ eine Linearform sowie $c \in \mathbb{R}$.

9.1.4 Bemerkung: Homogene Darstellung

Ist

$$Q := \{x \in \mathbb{R}^n \mid x^\top Ax + 2b^\top x + c = 0\}$$

eine Quadrik, so bezeichnet man mit

$$A_{\text{erw}} := \begin{pmatrix} c & b^\top \\ b & A \end{pmatrix} \text{ und } x_{\text{erw}} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

die erweiterte Koeffizientenmatrix und den erweiterten Vektor, so gilt:

$$\begin{aligned} x_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} &= (1, x_1, \ldots, x_n) \begin{pmatrix} c & b^\top \\ b & A \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= (1, x_1, \ldots, x_n) \begin{pmatrix} c + b^\top x \\ b + Ax \end{pmatrix} \\ &= c + b^\top x + x^\top b + x^\top Ax \\ &= x^\top Ax + 2b^\top x + c \end{aligned}$$

Somit ist die Q definierende Gleichung zu $x_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} = 0$ äquivalent. Dies wird als *homogene Darstellung* bezeichnet.

9.2 Klassifikation von Quadriken

9.2.1 Definition

Die Menge

$$\mathrm{Aff}(K^n) := \{x \mapsto Mx + s \mid M \in \mathrm{Gl}_n(K), s \in K^n\}$$

wird als Menge der *affinen Abbildungen* des K^n bezeichnet. Abbildungen der Form $x \mapsto Mx + s$ mit $M \in O(n)$, $s \in \mathbb{R}^n$ werden als *affin-orthogonale Abbildungen* bezeichnet.

9.2.2 Übungsaufgaben

1. Zeigen Sie, dass die Menge der affinen Abbildungen eine Gruppe bildet.
2. Zeigen Sie, dass die Menge der affin-orthogonalen Abbildungen eine Untergruppe von $\mathrm{Aff}(\mathbb{R}^n)$ ist.

9.2.3 Satz

Sei Q eine Quadrik im $\mathbb{R}^n$ und φ eine affine Abbildung auf $\mathbb{R}^n$. Dann ist $\varphi(Q)$ ebenfalls eine Quadrik. Man sagt auch: Die Gruppe der affinen Abbildungen operiert auf der Menge der Quadriken.

Beweis:

Es sei $\varphi^{-1} : x \mapsto Mx + s$ die Inverse von φ. Dann gilt

$$\begin{aligned} Q &= \{x \in \mathbb{R}^n \mid x^\top Ax + 2b^\top x + c = 0\} \\ \Longrightarrow \varphi(Q) &= \{\varphi(x) \mid x \in \mathbb{R}^n,\ x^\top Ax + 2b^\top x + c = 0\} \\ &= \{y \in \mathbb{R}^n \mid \varphi^{-1}(y)^\top A\varphi^{-1}(y) + 2b^\top \varphi^{-1}(y) + c = 0\} \\ &= \{y \in \mathbb{R}^n \mid (My + s)^\top A(My + s) + 2b^\top (My + s) + c = 0\} \\ &= \{y \in \mathbb{R}^n \mid y^\top A'y + 2b'^\top y + c' = 0\} \end{aligned}$$

mit

$$A' = M^\top AM,\ b' = M^\top As + M^\top b = M^\top (As + b),\ c' = s^\top As + 2b^\top s + c\,.$$

Da M invertierbar und $A \neq 0$ ist, folgt $A' \neq 0$. Zudem gilt $A'^\top = M^\top A^\top M = M^\top AM = A'$. Somit ist das affine Bild einer Quadrik eine Quadrik. □

9.2.4 Definition

Zwei Quadriken $Q_1, Q_2 \subset \mathbb{R}^n$ heißen *affin äquivalent*, wenn es eine affine Abbildung gibt, welche die definierende Gleichung der einen in die definierende Gleichung der anderen transformiert. Die einfachste so erreichbare Form wird als *affine Normalform* der Quadrik bezeichnet.

Die einfachste Form, in die eine Quadrik mit affin-orthogonalen Transformationen überführt werden kann wird als *euklidische Normalform* bezeichnet.

Die einfachste Form, in die eine Quadrik mit orthogonalen Transformationen überführt werden kann wird als *Hauptachsenform* bezeichnet.

9.2.5 Bemerkung

Wir können die gleiche Art der Normierung für eine Hyperebene E, die durch die Gleichung

$$a^\top x = c \text{ bzw. } \langle a \mid x \rangle = c$$

gegeben ist, studieren. Ohne Einschränkung können wir $\|a\| = 1$ annehmen. Ist dann φ eine orthogonale Abbildung, so erhalten wir

$$\begin{aligned} \varphi(E) &= \{\varphi(x) \mid \langle a \mid x \rangle = c\} = \{y \mid \langle a \mid \varphi^{-1}(y) \rangle = c\} \\ &= \{y \mid \langle (\varphi^{-1})^*(a) \mid y \rangle = c\} = \{y \mid \langle \varphi(a) \mid y \rangle = c\}\,. \end{aligned}$$

Wählen wir φ als eine orthogonale Abbildung, welche a auf e_1 abbildet, so entsteht als „Hauptachsenform" der Ebene die Gleichung

$$y_1 = c\,.$$

Erlauben wir statt orthogonaler auch affin-orthogonale Abbildungen, so kann die Ebene noch mit $(-c, 0, \ldots, 0)$ verschoben werden. Somit ergibt sich als „euklidische Normalform" der Ebene die Gleichung

$$y_1 = 0\,.$$

Eine weitere Normierung mit affinen Abbildungen ist nicht mehr möglich.

9.2.6 Bemerkung

1. Ist Q durch die Gleichung $x^\top Ax + 2b^\top x + c = 0$ mit einer symmetrischen Matrix A gegeben, so gibt es nach dem Spektralsatz eine orthogonale Matrix S so, dass $S^{-1}AS = S^\top AS = D$ eine Diagonalmatrix ist. Diese Transformation S liefert die Hauptachsenform von Q. Man nennt $y^\top Dy + 2b^\top y + c = 0$ auch Q im y-Koordinatensystem.

2. Sei Q durch die Gleichung $y^\top Dy + 2b^\top y + c = 0$ mit einer Diagonalmatrix D gegeben. Die Gleichung hat dann die Form

   $$d_1 y_1^2 + 2b_1 y_1 + d_2 y_2^2 + 2b_2 y_2 + \cdots + d_n y_n^2 + 2b_n y_n + c = 0\,.$$

   Seien die Koordinaten so sortiert, dass $d_1, \ldots, d_i \neq 0$ und $d_{i+1} = \cdots = d_n = 0$ gilt. Die Identität

   $$d_j y_j^2 + 2b_j y_j = d_j \left(y_j + \frac{b_j}{d_j} \right)^2 - \frac{b_j^2}{d_j}$$

   zeigt, dass es eine Verschiebung gibt, welche in der Gleichung von Q die Normierung $b_1 = \cdots = b_i = 0$ erreicht und dabei nur den konstanten Term c stört.

   Die obigen Überlegungen zur Normierung einer Ebenengleichung zeigen, dass es eine affin-orthogonale Abbildung gibt, welche nur die Koordinaten $y_{i+1}, \ldots, y_n$ verändert und

   $$2b_{i+1} y_{i+1} + \cdots + 2b_n y_n + c$$

   entweder zu $\tilde{b}_{i+1} y_{i+1}$ oder zu $\tilde{c}$ normiert. Dies ist die euklidische Normalform.

3. Sei Q durch die Gleichung
$$d_1 y_1^2 + \cdots + d_i y_i^2 + 2b_{i+1} y_{i+1} + c = 0$$
mit $b_{i+1} = 0, c = 1$ oder $c = 0$ gegeben. So kann durch Skalieren der Variablen mit $\frac{1}{\sqrt{d_i}}$ und $\frac{1}{2b_{i+1}}$ sowie einem sortieren der Variablen nach Vorzeichen die Form
$$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_{p+q}^2 + \left\{ \begin{array}{c} 0 \\ 1 \\ z_{p+q+1} \end{array} \right\} = 0$$
erreicht werden. Dies ist die affine Normalform.

Wir werden später untersuchen, in wieweit die so beschriebenen Normalformen eindeutig sind.

9.2.7 Beispiel

Es sei die Quadrik $Q \subset \mathbb{R}^3$ durch $-x_1^2 - x_2^2 + x_3^2 + 2x_1x_2 - 2x_3 = 0$ gegeben. Die Gleichung von Q hat auch die Darstellung
$$x^\top Ax - 2[0\ 0\ 1]x = 0 \text{ mit } A = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Bestimmung von Eigenwerten und Eigenvektoren:
$$\begin{vmatrix} -1-\lambda & 1 & 0 \\ 1 & -1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)\big((-1-\lambda)^2 - 1\big)$$
$$= (1-\lambda)(\lambda^2 + 2\lambda) = (1-\lambda)\lambda(\lambda+2)$$

Die Eigenwerte sind $1, -2, 0$.
$$\ker \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \operatorname{span}(0\ 0\ 1)^\top$$
$$\ker \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \operatorname{span}(1\ -1\ 0)^\top$$
$$\ker \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \operatorname{span}(1\ 1\ 0)^\top$$

Wir erhalten die Orthonormalbasis und die Transformationsmatrix
$$\mathcal{B} = \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right), \ V = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix}.$$

Es ist

$$V^{-1} = V^\top \text{ und } D = V^\top A V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

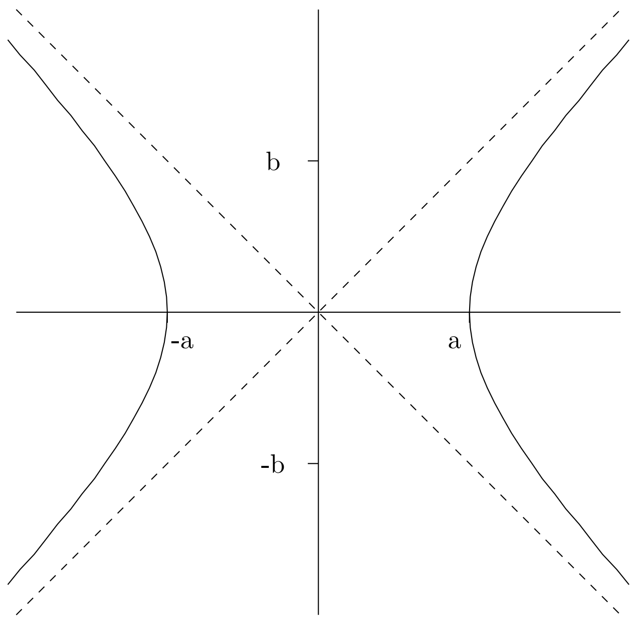
Mit $x = Vy$ bzw. $y = V^\top x$ erhalten wir

$$\begin{aligned} Q &= \{x \in \mathbb{R}^3 \mid y^\top \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} y - 2[0\ 0\ 1] \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} y = 0\} \\ &= \{x \in \mathbb{R}^3 \mid y^\top \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} y - 2[1\ 0\ 0]y = 0\} \\ &= \{x \in \mathbb{R}^3 \mid (y_1 - 1)^2 - 2y_2^2 - 1 = 0\} \\ &= \{x \in \mathbb{R}^3 \mid \frac{(y_1-1)^2}{1} - \frac{y_2^2}{\left(\frac{1}{\sqrt{2}}\right)^2} = 1\}. \end{aligned}$$

Die Quadrik ist ein Zylinder über einer Hyperbel.

9.2.8 Die Hyperbel

Die Gleichung $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ beschreibt eine Hyperbel. Es ist $y = \pm \frac{b}{a}\sqrt{x^2 - a^2}$. Somit sind $y = \pm \frac{b}{a} x$ Asymptoten.

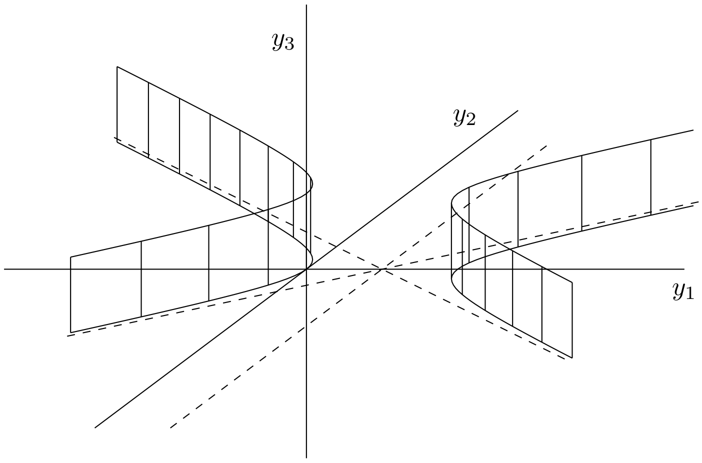


9.2.9 Zeichnung der Quadrik

Zeichnung der durch

$$\{x \in \mathbb{R}^3 \mid \frac{(y_1-1)^2}{1} - \frac{y_2^2}{\left(\frac{1}{\sqrt{2}}\right)^2} = 1\}$$

gegebenen Quadrik Q.



9.2.10 Satz

Ist $Q\colon x^\top Ax + 2b^\top x + c = 0$ eine Quadrik im $\mathbb{R}^n$, so sind $\mathrm{rk}(A)$ und $\mathrm{rk}(A_{\text{erw}})$ affine Invarianten.

Beweis:

Wir betrachten die affine Abbildung $x = Ry + s$ mit einer invertierbaren $n \times n$-Matrix R und $s \in \mathbb{R}^n$. Wenden wir diese auf die Q definierende Gleichung an, so geht die Matrix A in die Matrix $R^\top AR$ über. Da R invertierbar ist, bleibt somit der Rang der Matrix A unter affinen Abbildungen erhalten.

Für den 2. Teil definieren wir R_{erw} als

$$R_{\text{erw}} := \begin{pmatrix} 1 & 0 \\ s & R \end{pmatrix} \text{ und benutzen } x_{\text{erw}} := \begin{pmatrix} 1 \\ x \end{pmatrix},\ A_{\text{erw}} := \begin{pmatrix} c & b^\top \\ b & A \end{pmatrix}.$$

Dann gilt

$$x_{\text{erw}} = R_{\text{erw}} y_{\text{erw}}$$

und wir erhalten

$$x_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} = 0 \Longleftrightarrow y_{\text{erw}}^\top \left(R_{\text{erw}}^\top A_{\text{erw}} R_{\text{erw}}\right) y_{\text{erw}} = 0\,.$$

Somit ist $R_{\text{erw}}^\top A_{\text{erw}} R_{\text{erw}}$ die erweiterte Matrix der Bildquadrik. Da R_{erw} regulär ist, wird der Rang der erweiterten Matrix durch affine Abbildungen nicht verändert. $\square$

9.2.11 Definition

Man bezeichnet eine Quadrik $Q\colon x^\top Ax + 2b^\top x + c = 0$ mit $A \neq 0$ im Fall

$\mathrm{rk}(A) = \mathrm{rk}(A_{\mathbf{erw}})$ als kegelige Quadrik.

$1 + \mathrm{rk}(A) = \mathrm{rk}(A_{\mathbf{erw}})$ als Mittelpunktsquadrik.

$2 + \mathrm{rk}(A) = \mathrm{rk}(A_{\mathbf{erw}})$ als parabolische Quadrik.

9.2.12 Übungsaufgabe

Ist $Q\colon x^\top Ax + 2b^\top x + c = 0$ mit $A \neq 0$ eine Quadrik im $\mathbb{R}^n$, so sind genau die folgenden Fälle möglich:

1. $\mathrm{rk}(A) = \mathrm{rk}(A_{\mathrm{erw}}) \in \{1, \ldots, n\}$
2. $1 + \mathrm{rk}(A) = \mathrm{rk}(A_{\mathrm{erw}})$ mit $\mathrm{rk}(A) \in \{1, \ldots, n\}$
3. $2 + \mathrm{rk}(A) = \mathrm{rk}(A_{\mathrm{erw}})$ mit $\mathrm{rk}(A) \in \{1, \ldots, n-1\}$

9.2.13 Satz (Sylvesterscher Trägheitssatz)

Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine symmetrische Matrix. Dann gibt es eine invertierbare Matrix $R \in \mathrm{Gl}_n(\mathbb{R})$, sodass $R^\top AR$ die Form $\mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$ hat. Die Anzahl p der $+1$ und die Anzahl q der -1 Einträge hängt nicht von der konkreten Matrix R ab.

Beweis:

Da A symmetrisch ist, existiert einen Orthonormalbasis aus Eigenvektoren. Dies liefert eine orthogonale Matrix S mit $S^\top AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Bezeichnet D die Diagonalmatrix mit den Diagonaleinträgen $1/\sqrt{|\lambda_i|}$ im Fall $\lambda_i \neq 0$ und 1 sonst, so gilt

$$D^\top S^\top ASD = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)\,.$$

Es bleibt zu zeigen, dass p und q eindeutig sind. Es ist $p+q = \mathrm{rk}(A)$. Nehmen wir an, es gäbe eine Matrix R', sodass eine Diagonalmatrix mit p' und q' Einträgen $+1$ und -1 entsteht. Es gilt $p + q = p' + q'$. Wir setzen

$$U_+ := \mathrm{span}\{Re_i : i = 1, \ldots, p\}$$

und

$$U_- := \mathrm{span}\{R'e_i : i = p' + 1, \ldots, n\}\,.$$

Für jeden Vektor $x \in U_+ \setminus \{0\}$ gilt $x^\top Ax > 0$. Für jeden Vektor $y \in U_-$ gilt $y^\top Ay \leq 0$. Dies beweist $U_+ \cap U_- = \{0\}$. Somit gilt $n \geq \dim U_+ + \dim U_- = p + (n - p') = n + (p - p')$. Es folgt $p \leq p'$. In gleicher Art kann man $p' \leq p$ zeigen. Somit gilt $p = p'$ und $q = q'$. □

9.2.14 Bemerkung (Effektive Berechnung von p und q)

Im Sylvesterschen Trägheitssatz sind p und q die Anzahl der positiven und negativen Eigenwerte. Will man nicht zuerst die Eigenwerte berechnen, so kann man die durch R gegebenen Transformation auch direkt mittels quadratischer Ergänzung ermitteln. Wir zeigen dies an einem Beispiel:

$$\begin{aligned}
&x_1^2 + 4x_1x_2 + 2x_1x_3 + 4x_2x_3 + 2x_3x_4 + x_4^2\\
&= (x_1 + 2x_2 + x_3)^2 - (2x_2 + x_3)^2 + 4x_2x_3 + 2x_3x_4 + x_4^2\\
&= (x_1 + 2x_2 + x_3)^2 - 4x_2^2 - 4x_2x_3 - x_3^2 + 4x_2x_3 + 2x_3x_4 + x_4^2\\
&= (x_1 + 2x_2 + x_3)^2 - 4x_2^2 - x_3^2 + 2x_3x_4 + x_4^2\\
&= (x_1 + 2x_2 + x_3)^2 - 4x_2^2 - (x_3 - x_4)^2 + 2x_4^2
\end{aligned}$$

In Matrixschreibweise erhalten wir

$$\begin{pmatrix} 1 & 2 & 1 & 0\\ 2 & 0 & 2 & 0\\ 1 & 2 & 0 & 1\\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 1 \end{pmatrix}^\top \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & -4 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In dieser Form kann $p = q = 2$ direkt abgelesen werden. Die Ermittlung der Eigenwerte der ursprünglichen Matrix zu -2.3319394832804485724, $-0.7220637129691247565$3, 1.2484852535126999161, 3.8055179427368734128 ist hingegen wesentlich aufwendiger.

9.2.15 Klassifikation der Quadriken im $\mathbb{R}^2$

Mit den in 9.2.6 angegebenen Methoden kann eine Quadrik immer auf affine Normalform transformiert werden. Dabei bleiben nach obigen Sätzen $\mathrm{rk}(A)$ und $\mathrm{rk}(A_{\mathrm{erw}})$ unverändert. Skaliert man die Gleichung mit -1, so vertauscht man die Rolle von p und q im Sylvesterschen Trägheitssatz. Somit erhalten wir die folgende Liste von Quadriken:

Rang A	Rang A_{erw}	(p, q)	Gleichung	Name
1	1	(1,0)	$x_1^2 = 0$	Doppelgerade
1	2	(1,0)	$x_1^2 + 1 = 0$	leere Menge
1	2	(0,1)	$-x_1^2 + 1 = 0$	parallele Geraden
1	3	(1,0)	$x_1^2 + x_2 = 0$	Parabel
2	2	(2,0)	$x_1^2 + x_2^2 = 0$	Punkt
2	2	(1,1)	$x_1^2 - x_2^2 = 0$	sich schneidende Geraden
2	3	(2,0)	$x_1^2 + x_2^2 + 1 = 0$	leere Menge
2	3	(1,1)	$-x_1^2 + x_2^2 + 1 = 0$	Hyperbel
2	3	(0,2)	$-x_1^2 - x_2^2 + 1 = 0$	Ellipse

Man bezeichnet $p - q$ auch als Signatur der Quadrik.

9.2.16 Einige Quadriken im $\mathbb{R}^3$

Genau wie oben kann man auch Quadriken im $\mathbb{R}^3$ durch den Rang, den Rang der erweiterten Matrix und ihre Signatur klassifizieren. Wir geben nur ein paar Beispiele an:

1. $x^2 + y^2 = z^2$: Kegel
2. $x^2 - y^2 = 0$: Ebenenpaar
3. $z^2 = x^2 + y^2 + 1$: Zweischaliges Hyperboloid
4. $z^2 = x^2 + y^2 - 1$: Einschaliges Hyperboloid
5. $x^2 + y^2 + z^2 = 1$: Kugel
6. $z = \frac{-1}{2}(x^2 + y^2)$: Paraboloid
7. $x^2 - y^2 + 2z = 0$: hyperbolisches Paraboloid.

9.3 Pol und Polare

9.3.1 Definition

Es sei $Q\colon x^\top Ax + 2b^\top x + c = 0$ eine Quadrik im $\mathbb{R}^n$. Zu jedem $x \in \mathbb{R}^n$ wird die Menge

$$P_{Q,x} \coloneqq \{y \in \mathbb{R}^n \mid x_{\text{erw}}^\top A_{\text{erw}} y_{\text{erw}} = 0\}$$

als Polare zum Punkt x bezüglich Q bezeichnet. Der Punkt x heißt Pol der Polare.

9.3.2 Bemerkung

In gewöhnlichen Koordinaten geschrieben ist die Gleichung der Polare:

$$x^\top Ay + b^\top x + b^\top y + c = 0\,.$$

Dies ist zu

$$(x^\top A + b^\top)y + b^\top x + c = 0$$

äquivalent. Bei festem x ist dies eine lineare Gleichung in y. Somit gibt es 3 Fälle

1. Ist $(x^\top A + b^\top)$ nicht Null, so ist die Polare eine Hyperebene.
2. Ist $(x^\top A + b^\top)$ Null und $b^\top x + c$ ungleich Null, so ist die Polare die leere Menge.
3. Sind $(x^\top A + b^\top)$ und $b^\top x + c$ Null, so ist die Polare der gesamte $\mathbb{R}^n$.

9.3.3 Satz

Sei $Q\colon x^\top Ax + 2b^\top x + c = 0$ eine Quadrik. Dann gilt:

1. Ist $x \in \mathbb{R}^n$, so gilt $x \in Q \iff x \in P_{Q,x}$.
2. Sind $x, y \in \mathbb{R}^n$, so gilt $x \in P_{Q,y} \iff y \in P_{Q,x}$.
3. Sei $x \in Q$ und $y \in P_{Q,x}$ ein von x verschiedener Punkt. Ist y ein Punkt von Q, so ist die Verbindungsgerade von x und y bereits in Q enthalten.

Beweis:

1. $x \in P_{Q,x} \iff x_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} = 0 \iff x \in Q$.
2. $x \in P_{Q,y} \iff y_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} = 0 \iff x_{\text{erw}}^\top A_{\text{erw}} y_{\text{erw}} = 0 \iff y \in P_{Q,x}$.
3. Ein Punkt auf der Verbindungsgerade von x und y hat die Form $\lambda x + (1-\lambda)y$. Wir erhalten $(\lambda x + (1-\lambda)y)_{\text{erw}} = \lambda x_{\text{erw}} + (1-\lambda)y_{\text{erw}}$. Es folgt

$$\begin{aligned}
&(\lambda x_{\text{erw}} + (1-\lambda)y_{\text{erw}})^\top A_{\text{erw}} \lambda(x_{\text{erw}} + (1-\lambda)y_{\text{erw}}) \\
&= \lambda x_{\text{erw}}^\top A_{\text{erw}} \lambda x_{\text{erw}} + (1-\lambda) y_{\text{erw}}^\top A_{\text{erw}} \lambda x_{\text{erw}} \\
&\quad + \lambda x_{\text{erw}}^\top A_{\text{erw}} (1-\lambda) y_{\text{erw}} + (1-\lambda)^2 y_{\text{erw}}^\top A_{\text{erw}} y_{\text{erw}} = 0
\end{aligned}$$

9.3.4 Definition

Sei Q eine Quadrik im $\mathbb{R}^n$ und $x \in Q$. Im Fall $P_{Q,x} = \mathbb{R}^n$ heißt der Punkt x singulär. Ist x nicht singulär, so wird $P_{Q,x}$ auch als Tangentialhyperebene (Tangente) von Q in x bezeichnet.

9.3.5 Übungsaufgabe

Ist $Q\colon x^\top Ax + 2b^\top x + c = 0$ eine Quadrik im $\mathbb{R}^n$ und $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung. Zeigen Sie, dass

$$\varphi(P_{Q,x}) = P_{\varphi(Q),\varphi(x)}$$

gilt.

9.3.6 Satz

Es sei $Q\colon x^\top Ax + 2b^\top x + c = 0$ eine Quadrik im $\mathbb{R}^n$ und $\mathrm{rk} A_{\text{erw}} = n+1$, dann gilt for $y, z \in \mathbb{R}^n$

$$P_{Q,y} = P_{Q,z} \Longrightarrow y = z\,.$$

D.h., verschiedene Punkte haben immer verschiedene Polaren und man kann eine Funktion

Menge der Polaren $\to$ Menge der Pole

definieren.

Beweis:

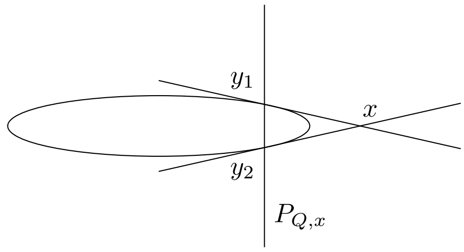
Seien $y, z \in \mathbb{R}^n$ mit $P_{Q,y} = P_{Q,z}$, so sind die Gleichungen

$$y_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} = 0 \iff z_{\text{erw}}^\top A_{\text{erw}} x_{\text{erw}} = 0\,.$$

D.h. die Zeilenvektoren $y_{\text{erw}}^\top A_{\text{erw}}$ und $z_{\text{erw}}^\top A_{\text{erw}}$ sind skalare Vielfache von einander. Da A_{erw} maximalen Rang hat, sind bereits y_{erw} und z_{erw} skalare Vielfache von einander. Die erste Komponente von y_{erw} und z_{erw} ist jeweils 1. Somit gilt $y_{\text{erw}} = z_{\text{erw}}$, welches $y = z$ impliziert. □

9.3.7 Bemerkung (Geometrische Konstruktion der Polaren)

Sei Q eine Quadrik im $\mathbb{R}^2$ und $x \in \mathbb{R}^2$. Wie konstruiert man $P_{Q,x}$ geometrisch? Im Fall $x \in Q$ ist $P_{Q,x}$ die Tangente im Punkt x. Andernfalls ist $P_{Q,x} \cap Q$ die Menge der Punkte von $y \in Q$, sodass $x \in P_{Q,y}$ gilt. D.h. der Schnitt sind die Punkte von Q auf deren Tangente x liegt. Somit ergibt sich folgendes Bild:



Kapitel 10

Normalformen und allgemeine Spektraltheorie

10.1 Äquivalenz und Ähnlichkeit von Matrizen, Jordan-Normalform und Spektralsatz

10.1.1 Zur Erinnerung: Äquivalenz

(Siehe 3.5.12 – 3.5.14)

1. Es seien $A, B \in \mathrm{Mat}(m \times n, K)$ zwei Matrizen. Wir nennen A, B äquivalent, wenn es invertierbare Matrizen S, T mit $SAT^{-1} = B$ gibt.

2. Äquivalenz ist eine Äquivalenzrelation auf der Menge $\mathrm{Mat}(m \times n, K)$.

3. Wir hatten gezeigt, dass zwei Matrizen gleicher Größe genau dann äquivalent sind, wenn ihre Ränge gleich sind. Weiterhin hatten wir zu jeder Matrix eine „einfachste" Matrix konstruiert, zu der sie äquivalent ist. Sie hatte die Form einer Blockmatrix
$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

4. Zwei Matrizen sind genau dann äquivalent, wenn sie die gleiche lineare Abbildung bezüglich verschiedenen Basen darstellen.

10.1.2 Zur Erinnerung: Ähnlichkeit

(Siehe 3.5.12, 6.2.3)

1. Es seien $A, B \in \mathrm{Mat}(n \times n, K)$ zwei Matrizen. Wir nennen A, B ähnlich, wenn es eine invertierbare Matrix S mit $SAS^{-1} = B$ gibt.

2. Ähnlichkeit ist eine Äquivalenzrelation auf der Menge $\mathrm{Mat}(n \times n, K)$.

3. Zwei Matrizen A, B sind genau dann ähnlich, wenn es einen Endomorphismus φ und Basen $\mathcal{A}, \mathcal{B}$ mit $A = M_{\mathcal{A}}(\varphi)$ und $B = M_{\mathcal{B}}(\varphi)$ gibt.

4. Wir hatten gezeigt, dass ähnliche Matrizen das gleiche charakteristische Polynom haben.

5. Offen sind noch: Wie entscheidet man Ähnlichkeit? Wie sieht eine „einfachste“ Matrix aus, zu der eine gegebene Matrix ähnlich ist.

10.1.3 Satz

Sind $A, B \in \mathrm{Mat}(n \times n, K)$ zwei diagonalisierbare Matrizen, so sind sie genau dann ähnlich, wenn ihre Eigenwerte (mit Vielfachheit) übereinstimmen.

Beweis:

Sind A, B ähnlich, so sind ihre charakteristischen Polynome gleich. Folglich stimmen ihre Eigenwerte mit Vielfachheit überein.

Sind A und B diagonalisierbar und stimmen ihre Eigenwerte samt Vielfachheiten überein, so gibt es invertierbare Matrizen S, T mit

$$SAS^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = TBT^{-1}\,.$$

Da Ähnlichkeit eine Äquivalenzrelation ist, sind A und B ähnlich. □

10.1.4 Bemerkung

1. Obiger Satz löst das Ähnlichkeitsproblem für diagonalisierbare Matrizen. Ferner zeigt er, dass die „einfachste“ Matrix zu der eine solche Matrix äquivalent ist, eine Diagonalmatrix ist.

2. Für Matrizen deren charakteristisches Polynom vollständig in Linearfaktoren zerfällt ist die einfachste ähnliche Matrix durch die sogenannte Jordan-Normalform gegeben. Sind $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte einer Matrix A, so hat die zugehörige Jordan-Normalform die Form

$$J(A) = \begin{pmatrix} J(\lambda_1) & & \\ & \ddots & \\ & & J(\lambda_m) \end{pmatrix}$$

Dabei heißen $J(\lambda_1), \ldots, J(\lambda_m)$ *Jordanblöcke.* Ein Jordanblock hat die Form

$$J(\lambda) = \begin{pmatrix} J_{i_1}(\lambda) & & \\ & \ddots & \\ & & J_{i_k}(\lambda) \end{pmatrix}.$$

Die Matrizen $J_{i_j}(\lambda)$ heißen *Jordankästchen.* Ein Jordankästchen $J_\ell(\lambda)$ ist eine $\ell \times \ell$-Matrix der Form

$$L_\ell(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ 0 & \lambda & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & \lambda & 1 \\ & & & 0 & \lambda \end{pmatrix}.$$

D.h., die Diagonaleinträge einer Jordanmatrix sind die Eigenwerte, abgesehen von den Einträgen direkt oberhalb der Hauptdiagonale (die möglicherweise 1 sind) sind alle weiteren Einträge Null. Sortiert man die Jordanblöcke nach den Eigenwerten und die Jordankästchen nach ihrer Größe, so ist die Jordan-Normalform eindeutig.

10.1.5 Ankündigung: Spektralsatz

Um obige Aussagen über die Jordan-Normalform zu beweisen, beweisen wir zunächst den folgenden Spektralsatz: Sei V ein endlich-dimensionaler Vektorraum und φ ein Endomorphismus von V, dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt mit Eigenwerten $\lambda_1, \ldots, \lambda_m$. Dann gibt es Exponenten $i(\lambda_1), \ldots, i(\lambda_m)$ so, dass für $H(\lambda_j) := \ker(\varphi - \lambda_j \mathrm{id})^{i(\lambda_j)}$ gilt:

1. $V = H(\lambda_1) \oplus \cdots \oplus H(\lambda_m)$
2. $\varphi(H(\lambda_j)) \subset H(\lambda_j)$
3. $\varphi|_{H(\lambda_j)} = \lambda_j \mathrm{id}_{H(\lambda_j)} + (\varphi - \lambda_j \mathrm{id})|_{H(\lambda_j)}$. Dabei ist $(\varphi - \lambda_j \mathrm{id})|_{H(\lambda_j)}$ nilpotent vom Index $i(\lambda_k)$.

10.2 Nilpotente Endomorphismen

10.2.1 Definition

Ein Endomorphismus $\psi\colon V \to V$ heißt *nilpontent vom Index k*, wenn $\psi^k = 0$ und $\psi^e \neq 0$ für $e = 0, \ldots, k-1$ gilt. Hierbei ist $\psi^0 = \mathrm{id}_V$. Entsprechend heißt eine quadratische Matrix A *nilpotent vom Index k*, wenn $A^k = 0$ und $A^e \neq 0$ für $e = 0, \ldots, k-1$ gilt.

10.2.2 Bemerkung

Es sei $\varphi\colon V \to V$ ein Endomorphismus und λ ein Eigenwert von φ. Ist φ nilpotent vom Index k, so ist $\lambda = 0$.

Beweis:

Sei v ein Eigenvektor zum Eigenwert λ. Dann gilt

$$\varphi(v) = \lambda v,\ \varphi^2(v) = \lambda^2 v, \ldots, \varphi^k(v) = \lambda^k v\,.$$

Aus $\varphi^k(v) = 0$ folgt $\lambda^k = 0$ und somit $\lambda = 0$. □

10.2.3 Satz

Es seien A, B zwei ähnliche Matrizen. Ist A nilpotent vom Index k, so ist auch B nilpotent vom Index k.

Beweis:

Es sei $B = S^{-1}AS$. Dann gilt

$$B^j = S^{-1}AS \cdot S^{-1}AS \cdots S^{-1}AS = S^{-1}A^jS\,.$$

Dies zeigt $B^j = 0$ g.d.w. $A^j = 0$. $\square$

Es gilt nicht: Sind A und B nilpotent vom Index k, so sind A und B ähnlich.

10.2.4 Beispiele nilpotenter Matrizen

1. Wir betrachten die Matrix $V_d \in \mathrm{Mat}(n \times n, K)$

$$V_d \coloneqq \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.$$

Es ist $V_d\, e_1 = 0$ und $V_d\, e_i = e_{i-1}$ für $i = 2, \ldots, d$. Es gilt

$$V_d^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}$$

und

$$V_d^{d-1} = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

sowie $V_d^d = 0$.

2. Nun betrachten wir die Block-Diagonalmatrix

$$W = \mathrm{diag}(V_{d_1}, \ldots, V_{d_k})\,.$$

Dies ist eine quadratische Matrix der Größe $n \coloneqq d_1 + \cdots + d_k$. Sie definiert einen Endomorphismus $\varphi\colon K^n \to K^n$. Wir gruppieren die Vektoren der Standardbasis gemäß der Struktur von W:

$$\begin{aligned} &\mathcal{A}_1 \coloneqq (e_1, \ldots, e_{d_1}), \mathcal{A}_2 \coloneqq (e_{d_1+1}, \ldots, e_{d_1+d_2}), \ldots \\ &\mathcal{A}_k \coloneqq (e_{d_1+\cdots+d_{k-1}+1}, \ldots, e_n) \end{aligned}$$

und setzen

$$U_i \coloneqq \mathrm{span}(\mathcal{A}_i) \text{ für } i = 1, \ldots, k\,.$$

Es gilt

(a) $K^n = U_1 \oplus \cdots \oplus U_k$

(b) $\varphi U_i \subset U_i$. D.h, die U_i sind φ invariante Unterräume. $\varphi|_{U_i}$ ist ein Endomorphismus von U_i.

(c) $M_{\mathcal{A}_i}(\varphi|_{U_i}) = V_{d_i}$

(d) φ ist nilpotent vom Index $\max(d_1, \dots, d_k)$.

(e) Es gilt

$$\begin{aligned}\ker W &= \operatorname{span}(e_1, e_{d_1+1}, e_{d_1+d_2+1}, \dots, e_{d_1+\dots+d_{k-1}+1})\\ \dim\ker W^m &= \#\{j \in \{1,\dots,k\} \mid d_j \geq 1\} + \#\{j \in \{1,\dots,k\} \mid d_j \geq 2\}\\ &\quad + \dots + \#\{j \in \{1,\dots,k\} \mid d_j \geq m\}\,.\end{aligned}$$

D.h. aus $\dim\ker W, \dim\ker W^2, \dots$ können k und $d_1, \dots, d_k$ rekonstruiert werden.

3. Sei K ein Körper mit $2 \neq 0$ und $\varphi\colon K^4 \to K^4$ durch

$$M_{\mathcal{E}}(\varphi) = \begin{pmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 \end{pmatrix}$$

gegeben. Dann ist $\mathcal{B} := (e_1 + e_3, e_2 + e_4, e_1 - e_3, e_2 - e_4)$ ebenfalls eine Basis von K^4. Es gilt

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

D.h., es gibt viele wesentlich verschiedene Basen bezüglich deren die Darstellungsmatrix von φ obige Form hat.

10.2.5 Satz

Es sei φ ein nilpotenter Endomorphismus eines n-dimensionalen Vektorraums V. Dann gibt es eine Basis $\mathcal{B}$ so, dass $M_{\mathcal{B}}(\varphi)$ eine obere Dreiecksmatrix ist, die nur Nullen auf der Diagonale hat.

Beweis:

Da φ nilpotent ist, gibt es ein k mit $0 = \varphi^k$. Nun betrachten wir die Unterraumkette

$$\ker\varphi \subset \ker\varphi^2 \subset \dots \subset \ker\varphi^k = V$$

und wählen eine angepasste Basis $\mathcal{B} = (b_1, \dots, b_{j_1}, b_{j_1+1}, \dots, b_{j_k})$ von V mit

$$\ker\varphi^e = \operatorname{span}(b_1, \dots, b_{j_e})\,.$$

Da ein Vektor aus $\ker\varphi^e \setminus \ker\varphi^{e-1}$ durch φ in $\ker\varphi^{e-1}$ abgebildet wird, hat die Darstellungsmatrix $M_{\mathcal{B}}(\varphi)$ die Gestalt

$$\begin{pmatrix} 0 & A_{1,2} & A_{1,3} & & \cdots & A_{1,k}\\ & 0 & A_{2,3} & A_{2,4} & & \\ & & \ddots & \ddots & \ddots & \vdots\\ & & & 0 & A_{k-2,k-1} & A_{k-2,k}\\ & & & & 0 & A_{k-1,k}\\ & & & & & 0 \end{pmatrix}.$$

Hierbei sind die $A_{i,j}$ Blockmatrizen vom Format $(j_i - j_{i-1}) \times (j_j - j_{j-1})$. $\square$

10.3 Der Spektralsatz

10.3.1 Definition und Bemerkung

1. Es sei V ein Vektorraum der Dimension n und φ ein Endomorphismus von V mit Eigenwert λ. Wir setzen

$$\begin{aligned} K(\lambda, 0) &:= \ker(\varphi - \lambda\mathrm{id})^0 = \ker \mathrm{id} = \{0\} \\ K(\lambda, 1) &:= \ker(\varphi - \lambda\mathrm{id})^1 \\ K(\lambda, 2) &:= \ker(\varphi - \lambda\mathrm{id})^2, \text{ usw.} \end{aligned}$$

Dies ist eine Kette von Unterräumen

$$K(\lambda, 0) \subset K(\lambda, 1) \subset K(\lambda, 2) \cdots .$$

Es ist $K(\lambda, 1)$ der Eigenraum zum Eigenwert λ. Seine Dimension ist die geometrische Vielfachheit des Eigenwerts.

2. Angenommen, es gibt einen Index $i(\lambda)$ mit

$$K(\lambda, 0) \subsetneq K(\lambda, 1) \subsetneq \cdots \subsetneq K(\lambda, i(\lambda)) = K(\lambda, i(\lambda) + 1)$$

so gilt $K(\lambda, i(\lambda)) = K(\lambda, i(\lambda) + j)$ für alle $j \in \mathbb{N}$. D.h., nach der ersten Wiederholung eines Unterraums tritt nur noch dieser Unterraum auf. Man bezeichnet $i(\lambda)$ als den *Index* des Eigenwerts λ.

Anmerkung: Da V endliche Dimension hat, gibt es eine solche Wiederholung. Zudem gilt die Ungleichung $i(\lambda) \leq \dim K(\lambda, i(\lambda))$.

Beweis: Die Inklusion $K(\lambda, i(\lambda)) \subset K(\lambda, i(\lambda) + j + 1)$ ist klar. Angenommen es ist $x \in K(\lambda, i(\lambda) + j + 1)$, so gilt

$$\begin{aligned} 0 &= (\varphi - \lambda\mathrm{id})^{i(\lambda)+j+1}(x) = (\varphi - \lambda\mathrm{id})^{i(\lambda)+1}\big((\varphi - \lambda\mathrm{id})^j(x)\big) \\ &= (\varphi - \lambda\mathrm{id})^{i(\lambda)}\big((\varphi - \lambda\mathrm{id})^j(x)\big) = (\varphi - \lambda\mathrm{id})^{i(\lambda)+j}(x) . \end{aligned}$$

Dies zeigt $K(\lambda, i(\lambda) + j + 1) = K(\lambda, i(\lambda) + j)$. Somit folgt die Behauptung induktiv.

3. Wir Definieren den Raum $\mathrm{H}(\varphi, \lambda)$ als $K(\lambda, i(\lambda))$. Er wird als *Hauptraum* (oder verallgemeinerter Eigenraum) zum Eigenwert λ bezeichnet.

4. Alle obigen Räume $K(\lambda, j)$ sind φ-invariant.

Beweis: Für $x \in V$ gilt

$$\begin{aligned} (\varphi - \lambda\mathrm{id})^j(\varphi(x)) &= (\varphi - \lambda\mathrm{id})^j((\varphi - \lambda\mathrm{id})(x) + \lambda x) \\ &= (\varphi - \lambda\mathrm{id})^{j+1}(x) + \lambda(\varphi - \lambda\mathrm{id})^j(x) \\ &= (\varphi - \lambda\mathrm{id})((\varphi - \lambda\mathrm{id})^j(x)) + \lambda(\varphi - \lambda\mathrm{id})^j(x) \\ &= \varphi((\varphi - \lambda\mathrm{id})^j(x)) . \end{aligned}$$

Dieses zeigt: $x \in K(\lambda, j) \Longrightarrow \varphi(x) \in K(\lambda, j)$.

5. Es ist $(\varphi - \lambda\mathrm{id})|_{\mathrm{H}(\varphi,\lambda)}$ nilpotent vom Index $i(\lambda)$.

Nachweis: Es ist

$$\ker(\varphi - \lambda\mathrm{id})^{i(\lambda)-1} \subsetneq \ker(\varphi - \lambda\mathrm{id})^{i(\lambda)} = \mathrm{H}(\varphi, \lambda) .$$

10.3.2 Satz

Sei V ein Vektorraum der Dimension n und φ ein Endomorphismus von V mit Eigenwert λ und Index $i(\lambda)$. Dann gilt

$$V = \mathrm{H}(\varphi, \lambda) \oplus \mathrm{Im}(\varphi - \lambda \mathrm{id})^{i(\lambda)} .$$

Beweis:

Sei $x \in \mathrm{H}(\varphi, \lambda) \cap \mathrm{Im}(\varphi - \lambda \mathrm{id})^{i(\lambda)}$, so gibt es ein $y \in V$ mit $x = (\varphi - \lambda \mathrm{id})^{i(\lambda)}(y)$ und $(\varphi - \lambda \mathrm{id})^{i(\lambda)}(x) = 0$. Hiermit folgt

$$0 = (\varphi - \lambda \mathrm{id})^{i(\lambda)}(x) = (\varphi - \lambda \mathrm{id})^{i(\lambda)} \Big((\varphi - \lambda \mathrm{id})^{i(\lambda)}(y) \Big) = (\varphi - \lambda \mathrm{id})^{2i(\lambda)}(y)$$

D.h., es ist $y \in K(\lambda, 2i(\lambda)) = K(\lambda, i(\lambda))$. Dieses zeigt

$$0 = (\varphi - \lambda \mathrm{id})^{i(\lambda)}(y) = x .$$

Aus $\dim(V) = \dim \ker(\varphi - \lambda \mathrm{id})^{i(\lambda)} + \dim \mathrm{Im}(\varphi - \lambda \mathrm{id})^{i(\lambda)}$ folgt nun

$$V = \mathrm{H}(\varphi, \lambda) \oplus \mathrm{Im}(\varphi - \lambda \mathrm{id})^{i(\lambda)} .$$

□

10.3.3 Satz

Sei V ein Vektorraum der Dimension n und φ ein Endomorphismus von V mit Eigenwert λ und Index $i(\lambda)$ und $k = \dim \mathrm{H}(\varphi, \lambda)$ sowie $U := \mathrm{Im}(\varphi - \lambda \mathrm{id})^{i(\lambda)}$. Dann gilt:

1. Das charakteristische Polynom von $\varphi|_{\mathrm{H}(\varphi,\lambda)}$ ist $(X - \lambda)^k$.
2. U ist ein φ-invarianter Unterraum.
3. λ ist kein Eigenwert von $\varphi|_U \colon U \to U$.
4. Die Dimension von $\mathrm{H}(\varphi, \lambda)$ ist die algebraische Vielfachheit des Eigenwerts λ von φ.

Beweis:

1. Da $(\varphi - \lambda \mathrm{id})|_{\mathrm{H}(\varphi,\lambda)}$ nilpotent ist, gibt es nach 10.2.5 eine Basis B, so dass die zugehörigen Darstellungsmatrix eine obere Dreiecksmatrix ist, die nur Nullen auf der Diagonale hat. Somit ist $M_{\mathcal{B}}(\varphi)$ eine obere Dreiecksmatrix, bei der nur λ auf der Diagonale steht. Somit hat das charakteristische Polynom die obige Form.
2. Ist $x \in U$, so gibt es ein $y \in V$ mit $x = (\varphi - \lambda \mathrm{id})^{i(\lambda)}(y)$. Es folgt

$$\begin{aligned}
\varphi(x) &= \varphi(x) - \lambda x + \lambda x \\
&= (\varphi - \lambda \mathrm{id})((\varphi - \lambda \mathrm{id})^{i(\lambda)}(y)) + \lambda(\varphi - \lambda \mathrm{id})^{i(\lambda)}(y) \\
&= (\varphi - \lambda \mathrm{id})^{i(\lambda)+1}(y) + (\varphi - \lambda \mathrm{id})^{i(\lambda)}(\lambda y) \\
&= (\varphi - \lambda \mathrm{id})^{i(\lambda)}((\varphi - \lambda \mathrm{id})y) + (\varphi - \lambda \mathrm{id})^{i(\lambda)}(\lambda y) \\
&= (\varphi - \lambda \mathrm{id})^{i(\lambda)}((\varphi - \lambda \mathrm{id})y + \lambda y) \\
&= (\varphi - \lambda \mathrm{id})^{i(\lambda)}(\varphi(y)) \in U .
\end{aligned}$$

3. Angenommen λ wäre ein Eigenwert von $\varphi|_U$, so gäbe es einen zugehörigen Eigenvektor $x \neq 0$ mit $(\varphi - \lambda)x = 0$. Somit gilt $x \in \mathrm{H}(\varphi, \lambda)$. Dies ist ein Widerspruch, da $\mathrm{H}(\varphi, \lambda) \cap U = \{0\}$ nach 10.3.2 gilt.

4. Nach 10.3.2 gilt $V = \mathrm{H}(\varphi, \lambda) \oplus U$. Somit gilt für Basen $\mathcal{B}_1 := (b_1, \dots, b_k)$ von $\mathrm{H}(\varphi, \lambda)$ und $\mathcal{B}_2 := (b_{k+1}, \dots, b_n)$ von U, dass $\mathcal{B} = (b_1, \dots, b_n)$ eine Basis von V ist. Da $\mathrm{H}(\varphi, \lambda)$ und U beide φ-invariante Unterräume sind, erhalten wir die Darstellungsmatrix
$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} M_{\mathcal{B}_1}(\varphi|_{\mathrm{H}(\varphi,\lambda)}) & 0 \\ 0 & M_{\mathcal{B}_2}(\varphi|_U) \end{pmatrix}.$$
Somit gilt für die charakteristischen Polynome
$$\chi(\varphi) = \chi(\varphi|_{\mathrm{H}(\varphi,\lambda)}) \cdot \chi(\varphi|_U)\,.$$
Nach 1 ist der erste Faktor $(X - \lambda)^k$, der zweite Faktor ist in λ ungleich Null. $\square$

10.3.4 Satz

Sei V ein Vektorraum der Dimension n und φ ein Endomorphismus von V mit Eigenwert λ und Index $i(\lambda)$ und $\mu \neq \lambda$ ein zweiter Eigenwert. Dann gilt
$$\mathrm{H}(\varphi, \mu) \subset \mathrm{Im}(\varphi - \lambda \mathrm{id})^{i(\lambda)}\,.$$

Beweis:

Wie oben setzen wir $U := \mathrm{Im}(\varphi - \lambda\mathrm{id})^{i(\lambda)}$. Es gilt $\mathrm{H}(\varphi|_U, \mu) = \mathrm{H}(\varphi, \mu) \cap U \subset \mathrm{H}(\varphi, \mu)$. Nach 10.3.3 sind die Dimensionen dieser Räume die algebraischen Vielfachheiten des Eigenwerts μ bezüglich der Abbildungen $\varphi|_U$ und φ. Nach 10.3.3 sind diese Vielfachheiten gleich, da sich die charakteristischen Polynome nur um den Faktor $(X - \lambda)^k$ unterscheiden. Dies zeigt die Behauptung. $\square$

10.3.5 Satz

Sei V ein Vektorraum der Dimension n und φ ein Endomorphismus von V mit charakteristischem Polynom
$$\chi(\varphi) = (X - \lambda_1)^{\mu(\lambda_1)} \cdots (X - \lambda_m)^{\mu(\lambda_m)} \text{ mit } \lambda_i \neq \lambda_j \text{ für } i \neq j\,,$$
welches vollständig in Linearfaktoren zerfällt. Sind $\mathcal{B}_1 = (b_1, \dots, b_{i_1}), \dots, \mathcal{B}_m = (b_{i_{m-1}+1}, \dots, b_{i_m})$ Basen der Haupträume $H(\varphi, \lambda_1), \dots, H(\varphi, \lambda_m)$, so ist $\mathcal{B} := (b_1, \dots, b_{i_m})$ eine Basis von V. In anderen Worten: Die Summe aller Haupträume ist direkt und ergibt V.

Beweis:

Es sind $\mu(\lambda_1), \dots, \mu(\lambda_m)$ die algebraischen Vielfachheiten der Eigenwerte $\lambda_1, \dots, \lambda_m$ und somit auch die Dimensionen der Haupträume.

Wir beweisen nun durch Induktion nach m, dass $\mathcal{B}$ eine Basis ist.

Induktionsanfang $m = 1$: In diesem Fall ist $\mu(\lambda_1) = \dim V$ und die Aussage ist trivial.

Induktionsschritt $m-1 \rightsquigarrow m$**:** Sei die Aussage bereits für alle Endomorphismen mit genau $m-1$ Eigenwerten bewiesen. Wir setzen $U := \operatorname{Im}((\varphi - \lambda_m)^{i(\lambda_m)})$. Es ist U ein φ-invarianter Raum mit

$$\chi_{\varphi|_U}(X) = (X-\lambda_1)^{\mu(\lambda_1)} \cdots (X-\lambda_{m-1})^{\mu(\lambda_{m-1})}\,.$$

Es gilt $\mathrm{H}(\varphi, \lambda_i) = \mathrm{H}(\varphi|_U, \lambda_i)$ für $i = 1, \ldots, m-1$, und somit ist $(b_1, \ldots, b_{i_{m-1}})$ nach Induktionsvoraussetzung eine Basis von U. Aus

$$V = \mathrm{H}(\varphi, \lambda_m) \oplus U$$

folgt dann, dass $\mathcal{B}$ eine Basis von V ist.

10.3.6 Zusammenfassung: Spektralsatz

Sei V ein Vektorraum der Dimension n und φ ein Endomorphismus von V dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Sind dann $\{\lambda_1, \ldots, \lambda_m\}$ alle Eigenwerte von φ, so gilt:

1. V ist direkte Summe der Haupträume $\mathrm{H}(\varphi, \lambda_i)$.
2. Die Dimensionen der Haupträume sind die algebraischen Vielfachheiten der Eigenwerte.
3. Für alle Haupträume gilt $\varphi(\mathrm{H}(\varphi, \lambda_i)) \subset \mathrm{H}(\varphi, \lambda_i)$.
4. $\varphi|_{\mathrm{H}(\varphi,\lambda_i)} = \lambda_i \mathrm{id}_{\mathrm{H}(\varphi,\lambda_i)} + (\varphi - \lambda_i \mathrm{id})|_{\mathrm{H}(\varphi,\lambda_i)}$. Hierbei ist der zweite Summand eine nilpotente Abbildung.
5. $1 \leq i(\lambda) \leq \dim \mathrm{H}(\varphi, \lambda)$ gilt für jeden Eigenwert λ.

Beweis:

Alle Aussagen wurden bereits gezeigt. □

10.4 Die Jordan-Normalform

10.4.1 Definition

Es sei V eine endlich-dimensionaler Vektorraum und φ ein nilpotenter Endomorphismus von V. Ein Vektor z heißt *φ-zyklischer Erzeuger* von V, falls $\varphi^i(z) = 0$ gilt und $(\varphi^{i-1}(z), \ldots, \varphi(z), z)$ eine Basis von V ist.

Ein System von Vektoren $(z_1, \ldots, z_k)$ heißt *φ-zyklisches Erzeugendensystem* von V, falls es Exponenten $i_1, \ldots, i_k \in \mathbb{N}$ gibt, so dass $\varphi^{i_j}(z_j) = 0$ und $\varphi^{i_j - 1}(z_j) \neq 0$ gilt und zudem

$$(\varphi^{i_1-1}(z_1), \ldots, \varphi(z_1), z_1, \ldots, \varphi^{i_k-1}(z_k), \ldots, \varphi(z_k), z_k)$$

eine Basis von V ist.

10.4.2 Bemerkung

Sei φ ein nilpotenter Endomorphismus des n-dimensionalen Vektorraums V.

1. Falls es einen φ-zyklischen Erzeuger z von V gibt, so gilt

$$M_{\mathcal{B}}(\varphi) = V_n$$

   mit der in 10.2.4 definierten Matrix V_n und der Basis $\mathcal{B} = (\varphi^{n-1}(z), \ldots, \varphi(z), z)$.

2. Falls es ein φ-zyklisches Erzeugendensystem $(z_1, \ldots, z_k)$ von V mit Exponenten $i_1, \ldots, i_k$ gibt, so gilt

$$M_{\mathcal{B}}(\varphi) = \operatorname{diag}(V_{i_1}, \ldots, V_{i_k})$$

   mit der Basis

$$\mathcal{B} = (\varphi^{i_1-1}(z_1), \ldots, \varphi(z_1), z_1, \ldots, \varphi^{i_k-1}(z_k), \ldots, \varphi(z_k), z_k)\,.$$

10.4.3 Satz

Es sei φ ein nilpotenter Endomorphismus des endlich-dimensionalen Vektorraums V, so existiert ein φ-zyklisches Erzeugendensystem von V.

Beweis:

Wir beweisen die Aussage durch Induktion nach dem Nilpotenzindex von φ. Angenommen der Nilpotenzindex von φ ist 1, so ist φ die Nullabbildung. In diesem Fall ist jede Basis $(z_1, \ldots, z_n)$ von V ein φ-zyklisches Erzeugendensystem mit den Exponenten $1, \ldots, 1$.

Sei nun die Behauptung für alle nilpotenten Endomorphismen mit Nilpotenzindex kleiner als φ bewiesen, so ist

$$\varphi|_{\operatorname{Im}(\varphi)} \colon \operatorname{Im}(\varphi) \to \operatorname{Im}(\varphi)$$

eine nilpotente Abbildung mit

$$i(\varphi|_{\operatorname{Im}(\varphi)}) = i(\varphi) - 1\,.$$

Somit gibt es nach Induktionsvoraussetzung ein $\varphi|_{\operatorname{Im}(\varphi)}$-zyklisches Erzeugendensystem $(u_1, \ldots, u_k)$ von $\operatorname{Im}(\varphi)$ mit Exponenten $(i_1, \ldots, i_k)$. Wir wählen $z_1, \ldots, z_k \in V$ mit $\varphi(z_j) = u_j$.

Behauptung:

$$\mathcal{A} = (\varphi^{i_1}(z_1), \ldots, \varphi(z_1), z_1, \ldots, \varphi^{i_k}(z_k), \ldots, \varphi(z_k), z_k)$$

ist linear unabhängig. Angenommen es gäbe eine lineare Relation

$$0 = \sum_{j=1}^{k} \sum_{e=0}^{i_j} a_{j,e} \varphi^e(z_j)\,,$$

so wenden wir auf diese Relation die Abbildung φ an. Es entsteht die Relation

$$0 = \sum_{j=1}^{k} \sum_{e=0}^{i_j-1} a_{j,e} \varphi^e(u_j)\,.$$

Da $(u_1, \dots, u_k)$ ein φ-zyklisches Erzeugendensystem von $\mathrm{Im}(\varphi)$ ist, sind alle in obiger Relation auftretenden Koeffizienten Null.

Setzen wir dieses in die Ausgangsrelation ein, so entsteht

$$0 = \sum_{j=1}^{k} a_{j,i_j} \varphi^{i_j - 1}(u_j)\,.$$

Da $(u_1, \dots, u_k)$ ein φ-zyklisches Erzeugendensystem von $\mathrm{Im}(\varphi)$ ist, sind auch die verbleibenden Koeffizienten Null.

Nun wählen wir $(z_{k+1}, \dots, z_m)$ so, dass $(\varphi^{i_1}(z_1), \dots, \varphi^{i_k}(z_k))$ zu einer Basis von $\ker(\varphi)$ ergänzt wird.

Wir behaupten, dass so ein φ-zyklisches Erzeugendensystem $(z_1, \dots, z_m)$ mit den Exponenten $(i_1 + 1, \dots, i_k + 1, 1, \dots, 1)$ entsteht. D.h.,

$$\mathcal{B} = (\varphi^{i_1}(z_1), \dots, \varphi(z_1), z_1, \dots, \varphi^{i_k}(z_k), \dots, \varphi(z_k), z_k, z_{k+1}, \dots, z_m)$$

ist eine Basis von V.

Nachweis der linearen Unabhängigkeit: Wenden wir wie oben zunächst φ auf eine mögliche lineare Relation an, so entsteht eine lineare Relation zwischen Null-Vektoren und den Vektoren einer Basis von $\mathrm{Im}(\varphi)$. Somit treten in der Relation nur Vektoren aus $\ker(\varphi)$ auf. Diese Vektoren bilden jedoch eine Basis von $\ker\varphi$ und sind somit linear unabhängig.

Nachweise „Erzeugendensystem“: Sei $v \in V$ beliebig, so kann mit den Vektoren aus $\mathcal{B}$ ein Vektor v_1 mit $\varphi(v) = \varphi(v_1)$ linearkombiniert werden, da die Bilder der Vektoren von $\mathcal{B}$ eine Basis von $\mathrm{Im}(\varphi)$ enthalten. Somit ist noch ein zweiter Vektor $v_2 = v - v_1 \in \ker(\varphi)$ als Linearkombination der Vektoren aus $\mathcal{B}$ darzustellen. Da $\mathcal{B}$ eine Basis von $\ker(\varphi)$ enthält, folgt $v_2 \in \mathrm{span}(\mathcal{B})$. Somit ist $v \in \mathrm{span}(\mathcal{B})$ gezeigt. □

10.4.4 Bemerkung

Obiger Beweis ist konstruktiv. Im Fall von kleinem Nilpotenzindex erfolgt die Berechnung der Basis wie folgt:

Index 1: Wähle eine Basis von $\ker\varphi$.

Index 2: Berechne eine Basis $(\varphi(z_1), \dots, \varphi(z_k))$ von $\mathrm{Im}\,\varphi$. Ergänze $(\varphi(z_1), \dots, \varphi(z_k))$ mit $z_{k+1}, \dots, z_m$ zu einer Basis von $\ker\varphi$. Es ist

$$\mathcal{B} = (\varphi(z_1), z_1, \dots, \varphi(z_k), z_k, z_{k+1}, \dots, z_m)\,.$$

Index 3: Berechne eine Basis $(\varphi^2(z_1), \dots, \varphi^2(z_k))$ von $\mathrm{Im}\,\varphi^2$. Ergänze diese Vektoren mit $\varphi(z_{k+1}), \dots, \varphi(z_\ell)$ zu einer Basis von $\ker\varphi|_{\mathrm{Im}\,\varphi}$. Ergänze erneut mit Vektoren $z_{\ell+1}, \dots, z_m$ zu einer Basis von $\ker\varphi$. Es ist

$$\begin{aligned}\mathcal{B} = (&\varphi^2(z_1), \varphi(z_1), z_1, \dots, \varphi^2(z_k), \varphi(z_k), z_k,\\ &\varphi(z_{k+1}), z_{k+1}, \dots, \varphi(z_\ell), z_\ell, z_{\ell+1}, \dots, z_m)\end{aligned}$$

die gesuchte Basis und

$$M_{\mathcal{B}}(\varphi) = \mathrm{diag}(V_3, \dots, V_3, V_2, \dots, V_2, 0, \dots, 0)\,.$$

10.4.5 Übungsaufgabe

Beschreiben Sie die Berechnung im Fall einer Abbildung vom Nilpotenzindex 4.

10.4.6 Satz

Es sei V ein endlich-dimensionaler Vektorraum und φ ein Endomorphismus von V dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Wir bezeichnen mit $\lambda_1, \dots, \lambda_m$ alle Eigenwerte von φ. Dann gibt es eine Basis $\mathcal{B}$ so, dass $M_{\mathcal{B}}(\varphi)$ eine Blockdiagonalmatrix $\operatorname{diag}(J(\lambda_1), \dots, J(\lambda_k))$ ist. Hierbei sind die Jordanblöcke $J(\lambda_i)$ Blockdiagonalmatrizen $\operatorname{diag}(J_{k_1(\lambda_i)}(\lambda_i), \dots, J_{k_{j(\lambda_i)}(\lambda_i)}(\lambda_i))$. Es ist $J_{k_i(\lambda)}(\lambda)$ eine $k_i(\lambda) \times k_i(\lambda)$-Matrix der Form

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}.$$

Es ist $k_1(\lambda) + \cdots + k_{j(\lambda)}(\lambda)$ die algebraische Vielfachheit des Eigenwerts λ. Im Fall $k_1(\lambda) \geq \cdots \geq k_j(\lambda)$ sind $k_1(\lambda), \cdots, k_j(\lambda)$ eindeutig bestimmt.

Beweis:

Nach dem Spektralsatz 10.3.6 ist V die direkte Summe der Haupträume $\mathrm{H}(\varphi, \lambda_i)$. Auf dem Hauptraum $\mathrm{H}(\varphi, \lambda_i)$ ist $\varphi - \lambda_i \mathrm{id}$ nilpotent vom Index $i(\lambda_i)$. Wählen wir im Hauptraum nach 10.4.3 eine Basis passend zur nilpotenten Abbildung $(\varphi - \lambda_{\mathrm{id}})|_{\mathrm{H}(\varphi,\lambda_i)}$ so erhalten wir für $\varphi|_{\mathrm{H}(\varphi,\lambda_i)}$ die Darstellungsmatrix $J(\lambda_i) = \operatorname{diag}(J_{k_1(\lambda)}(\lambda), \dots, J_{k_{j(\lambda)}(\lambda)})$. Fassen wir alle so gefundenen Hauptraumbasen zu einer Basis $\mathcal{B}$ zusammen, so hat $M_{\mathcal{B}}(\varphi)$ genau die gesuchte Jordan-Form. $\square$

10.4.7 Satz

Es seien A, B zwei quadratische Matrizen gleicher Größe deren charakteristischen Polynome vollständig in Linearfaktoren zerfallen. Dann sind A und B genau dann ähnlich, wenn ihre Jordan-Formen die selben Jordankästchen besitzen.

Beweis:

Dies folgt aus der Eindeutigkeit der Jordan-Form. $\square$

10.5 Endomorphismenpolynome

10.5.1 Definition

Es sei R ein kommutativer Ring. Ein Polynom mit Koeffizienten in R ist eine Abbildung

$$a\colon \mathbb{N} \cup \{0\} \to R,$$

die an allen bis auf endlich vielen Stellen den Wert Null hat. Ist n die größte Zahl mit $a(n) \neq 0$, so schreibt man das Polynom auch als

$$a(0) + a(1)t + \cdots + a(n)t^n$$

oder

$$a_0 + a_1 t + \cdots + a_n t^n \, .$$

Zwei Polynome sind genau dann gleich, wenn alle Koeffizienten gleich sind.

10.5.2 Bemerkung

1. Ist R ein kommutativer Ring, so gibt es eine Abbildung

$$R[t] \to \mathrm{Abb}(R, R), a_0 + \cdots + a_n t^n \mapsto (t \mapsto a_0 + \cdots + a_n t^n)$$

   die einem Polynom die zugehörige Polynomfunktion zuordnet.

2. Formale Polynome tragen die Struktur eines kommutativen Rings. Man addiert und multipliziert sie mit den üblichen Rechenregeln.

3. Ist K ein Körper und φ ein Endomorphismus eines Vektorraums V so kann man die Abbildung

$$K[t] \to \mathrm{End}(V), a_0 + \cdots + a_n t^n \mapsto a_0 \mathrm{id} + a_1 \varphi + \cdots + a_n \varphi^n$$

   betrachten. Man schreibt hierfür auch $P(\varphi)$ im Fall eines Polynoms $P \in K[t]$.

4. Ist φ ein Endomorphismus und P_1, P_2 zwei Polynome, so gilt $P_1(\varphi) \circ P_2(\varphi) = (P_1 \cdot P_2)(\varphi)$ und $P_1(\varphi) \circ P_2(\varphi) = P_2(\varphi) \circ P_1(\varphi)$.

5. Sind φ, ψ zwei Endomorphismen des gleichen Raums, so ist in der Regel

$$P_1(\varphi) \circ P_2(\psi) \neq P_2(\psi) \circ P_1(\varphi) \, .$$

6. Ist V ein endlich-dimensionaler K-Vektorraum mit Basis $\mathcal{B}$ und φ ein Endomorphismus sowie $P \in K[t]$ ein Polynom, so gilt

$$P(M_{\mathcal{B}}(\varphi)) = M_{\mathcal{B}}(P(\varphi)) \, .$$

   Hierbei setzen wir $P(M) = a_0 E_n + a_1 M + a_2 M \cdot M + \cdots + a_n M^n$ im Fall einer quadratischen Matrix M und eines Polynoms $P = a_0 + a_1 t + \cdots + a_n t^n$.

10.5.3 Satz

Sei K ein Körper, $M \in \mathrm{Mat}(n \times n, K)$ und $S \in \mathrm{Gl}_n(K)$ sowie $P \in K[t]$, so gilt

$$P(M) = S^{-1} P(SMS^{-1}) S \, .$$

Beweis:

Zunächst gilt

$$(SMS^{-1})^n = SMS^{-1}SMS^{-1} \cdots SMS^{-1} = SM^nS^{-1}\,.$$

Sei nun $P = a_0 + a_1 t + \cdots + a_n t^n$, so gilt

$$\begin{aligned} S^{-1}P(SMS^{-1})S &= S^{-1}(a_0E_n + a_1SMS^{-1} + \cdots + a_n(SMS^{-1})^n)S \\ &= S^{-1}(a_0SS^{-1} + a_1SMS^{-1} + \cdots + a_nSM^nS^{-1})S \\ &= S^{-1}S(a_0E_n + a_1M + \cdots + a_nM^n)S^{-1}S \\ &= a_0E_n + a_1M + \cdots + a_nM^n = P(M) \end{aligned}$$

□

10.5.4 Satz

Es sei $M \in \mathrm{Mat}(n \times n, K)$ eine quadratische Matrix, dann gibt es ein eindeutig bestimmtes Polynom kleinsten Grades P mit Leitkoeffizient 1 so, dass $P(M) = 0$ gilt.

Ist m_M obiges Polynom, so gilt für jedes Polynom Q

$$Q(M) = 0 \iff m_M \mid Q\,.$$

Anmerkung: Es gelten die entsprechenden Aussagen für Endomorphismen endlich-dimensionaler Vektorräume.

Beweis:

Es ist $\mathrm{Mat}(n \times n, K)$ ein n^2-dimensionaler Vektorraum. Somit ist das System $(E_n, M, M^2, \ldots, M^{(n^2)})$ linear abhängig. Folglich existiert ein solches Polynom.

Zur Eindeutigkeit: Angenommen es gäbe zwei solche Polynome P_1, P_2, so wäre $P_3 := P_1 - P_2 \neq 0$ ein Polynom kleineren Grades mit $P_3(M) = 0$.

Zur letzten Aussage: Wir teilen Q durch m_M mit Rest und erhalten $Q = F \cdot m_M + R$ mit Polynomen F, R sowie $\deg R < \deg m_M$.

Nun gilt

$$0 = Q(M) = F(M) \cdot m_M(M) + R(M) = R(M)\,.$$

Somit ist $Q(M) = 0$ g.d.w. $R(M) = 0$ gilt. Wäre $R \neq 0$, so erhalten wir einen Widerspruch zur Minimalität des Grades von m_P. □

10.5.5 Definition

Obiges Polynom m_M heißt *Minimalpolynom.* Das Minimalpolynom eines Endomorphismus definiert man in gleicher Art.

10.5.6 Satz

Es sei φ ein Endomorphismus eines endlich-dimensionalen Vektorraums V, dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Sind $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von φ, so gilt

$$m_\varphi = (t - \lambda_1)^{i(\lambda_1)} \cdots (t - \lambda_k)^{i(\lambda_k)}\,.$$

Beweis:

Aus dem Spektralsatz folgt, dass V die direkte Summe der Haupträume von φ ist. Da $P(\varphi)$ eine lineare Abbildung ist, ist $P(\varphi) = 0$ äquivalent zu $P(\varphi)(v) = 0$ für $v \in \mathrm{H}(\varphi, \lambda_1), \ldots, \mathrm{H}(\varphi, \lambda_k)$.

Da $\psi_i := (\varphi - \lambda_i \mathrm{id})|_{\mathrm{H}(\varphi,\lambda_i)}$ nilpotent vom Index $i(\lambda_i)$ ist, ist $t^{i(\lambda_i)}$ das Minimalpolynom von ψ_i.

Folglich ist $(t - \lambda_i)^{i(\lambda)}$ das Minimalpolynom von $\varphi|_{\mathrm{H}(\varphi,\lambda_i)}$. Wir erhalten

$$P(\varphi) = 0 \Longleftrightarrow (t - \lambda_i)^{i(\lambda_i)} \mid P \text{ für } i = 1, \ldots, k\,.$$

Hieraus folgt

$$m_\varphi = (t - \lambda_1)^{i(\lambda_1)} \cdots (t - \lambda_k)^{i(\lambda_k)}\,.$$

□

10.5.7 Korollar: Satz von Cayley-Hamilton

Es sei M eine quadratische Matrix mit Einträgen in $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ oder $\mathbb{C}$ mit charakteristischem Polynom χ_M. Dann gilt $\chi_M(M) = 0$.

Beweis:

Das charakteristische Polynom vom M zerfällt in $\mathbb{C}$ vollständig in Linearfaktoren und es gilt

$$\chi_m = (\pm 1)(t - \lambda_1)^{\mu(\lambda_1)} \cdots (t - \lambda_k)^{\mu(\lambda_k)}\,.$$

Hierbei ist $\mu(\lambda_i)$ die algebraische Vielfachheit des Eigenwerts λ_i. Es gilt $i(\lambda_i) \leq \mu(\lambda_i)$. Hieraus folgt

$$m_M \mid \chi_M\,.$$

□

10.6 Beispiel

10.6.1 Matrix

Wir betrachten die Matrix

$$M := \begin{pmatrix} 1 & 2 & -1 \\ -7 & 4 & -3 \\ -7 & -1 & 0 \end{pmatrix}.$$

10.6.2 Berechnung des charakteristischen Polynoms

$$
\begin{aligned}
&\det\begin{pmatrix} 1-X & 2 & -1 \\ -7 & 4-X & -3 \\ -7 & -1 & 0-X \end{pmatrix}\\
&= (1-X)(-X(4-X)-3) - 2(7X-21) - 1(7+7(4-X)))\\
&= (1-X)(X^2-4X-3) - 14X + 42 - 7 - 28 + 7X\\
&= -X^3 + X^2 + 4X^2 - 4X + 3X - 3 - 7X + 7\\
&= -X^3 + 5X^2 - 8X + 4\\
&= -(X-1)(X^2-4X+4) = -(X-1)(X-2)^2
\end{aligned}
$$

10.6.3 Interpretation des charakteristischen Polynoms

Das charakteristische Polynom zerfällt vollständig in Linearfaktoren. Somit hat M bereits über $\mathbb{Q}$ eine Jordan-Normalform. Aufgrund der Vielfachheiten der Eigenwerte ist die Jordan-Form

$$
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ oder } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.
$$

Im ersten Fall ist das Minimalpolynom $(X-1)(X-2)$ und der Index des Eigenwertes 2 ist 1. Im zweiten Fall ist das Minimalpolynom $(X-1)(X-2)^2$ und der Index des Eigenwertes 2 ist 2.

10.6.4 Berechnung der Indizes und Eigenräume

$$
\begin{aligned}
E(M,1) = H(M,1) = \ker M - E_3 &= \ker\begin{pmatrix} 0 & 2 & -1 \\ -7 & 3 & -3 \\ -7 & -1 & -1 \end{pmatrix}\\
&= \ker\begin{pmatrix} -7 & -1 & -1 \\ 0 & 2 & -1 \end{pmatrix} = \operatorname{span}((-3,7,14)^\top).
\end{aligned}
$$

$$
\begin{aligned}
E(M,2) = \ker M - 2E_3 &= \ker\begin{pmatrix} -1 & 2 & -1 \\ -7 & 2 & -3 \\ -7 & -1 & -2 \end{pmatrix}\\
&= \ker\begin{pmatrix} -1 & 2 & -1 \\ -4 & -4 & 0 \\ -5 & -5 & 0 \end{pmatrix}\\
&= \ker\begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \operatorname{span}((1,-1,-3)^\top)
\end{aligned}
$$

An dieser Stelle sehen wir $1 = \dim E(M,2) \neq \dim H(M,2) = 2$. Somit gilt $i(2) = 2$. M ist nicht diagonalisierbar. Die Jordan-Form ist

$$
J := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.
$$

10.6.5 Berechnung der Jordan-Basis

Gesucht wird eine Basis $\mathcal{B} = (b_1, b_2, b_3)$ von $\mathbb{R}^3$ mit

$$M_{\mathcal{B}}(x \mapsto Mx) = J\,.$$

An J lesen wir ab:

1. b_1 ein Basisvektor von $E(M, 1)$.
2. b_2 ist Basisvektor von $E(M, 2)$.
3. b_3 erfüllt $Mb_3 = 2b_3 + b_2$. D.h., b_3 löst das LGS $(M - 2E_3)b_3 = b_2$.

Wir wählen somit

$$b_1 = (-3, 7, 14)^\top, b_2 = (1, -1, -3)^\top\,.$$

Berechnung von b_3:

$$\begin{array}{ccc|c}
-1 & 2 & -1 & 1 \\
-7 & 2 & -3 & -1 \\
-7 & -1 & -2 & -3 \\
\hline
-15 & 0 & -5 & -5 \\
-21 & 0 & -7 & -7 \\
-7 & -1 & -2 & -3 \\
\hline
-7 & -1 & -2 & -3 \\
3 & 0 & 1 & 1 \\
\hline
-1 & -1 & 0 & -1 \\
3 & 0 & 1 & 1
\end{array}$$

Somit ist $b_3 = (0, 1, 1)^\top$ eine Lösung. Man kann b_3 um beliebige Vielfache von b_2 abändern. Wir erhalten $H(M, 2) = \operatorname{span}(b_2, b_3)$. Mit $S = (b_1, b_2, b_3)$ gilt $SJS^{-1} = M$ bzw. $J = S^{-1}MS$.

Literaturverzeichnis

[1] Artin, Michael: Algebra, Birkhäuser, 1993

[2] Ebbinghaus, Heinz-Dieter et al.: Zahlen, Springer, 1992

[3] Fischer, Gerd: Lineare Algebra, Vieweg, 2005

[4] Kowalsky, Hans-Joachim: Lineare Algebra, de Gruyter, 2013