Vorlesungsmanuskript zu

Lineare Algebra I

Werner Balser
Institut für Angewandte Analysis

Wintersemester 2007/08

Inhaltsverzeichnis

Kapitel 1

Vektorräume

1.1 Lineare Räume

Im Folgenden sei $\mathbb{K}$ ein Körper. Wir können uns unter $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$ vorstellen, außer wenn wir ausdrücklich einen anderen Körper voraussetzen. Die meisten der folgenden Resultate gelten aber genauso für einen beliebigen Körper $\mathbb{K}$. Wir nennen $\mathbb{K}$ auch den *Skalarenkörper*, und ein Element $\lambda \in \mathbb{K}$ heißt manchmal auch ein *Skalar*. Die genaue Definition eines Körpers, und speziell von $\mathbb{R}$ und $\mathbb{C}$, wird in der Vorlesung *Analysis I* gegeben. Allgemeine Körper werden in der Vorlesung *Algebra* behandelt.

Definition 1.1.1 *Eine Menge V, zusammen mit zwei Abbildungen*

$$+ \ : \ V \times V \longrightarrow V, \qquad (v,w) \longmapsto v+w,$$

$$\cdot \ : \ \mathbb{K} \times V \longrightarrow V, \qquad (\lambda, v) \longmapsto \lambda \cdot v \quad (= \lambda\, v),$$

heißt ein linearer Raum *oder* ein Vektorraum *über $\mathbb{K}$, wenn folgende Axiome alle gelten:*

(V1) $\forall\, u,v,w \in V \ : \quad u + (v+w) \ = \ (u+v)+w$ (Assoz.-Ges. der Addit.)

(V2) $\exists\, 0 \in V \quad \forall\, v \in V \ : \quad v \ + \ 0 \ = \ v$ (Existenz eines Nullvektors)

(V3) $\forall\, v \in V \quad \exists\, \tilde{v} \in V \ : \quad v \ + \ \tilde{v} \ = \ 0$ (Exist. eines additiven Inversen)

(V4) $\forall\, v_1, v_2 \in V \ : \quad v_1 \ + \ v_2 \ = \ v_2 \ + \ v_1$ (Kommutativges. der Addition)

(V5) $\forall\, v \in V \quad \forall\, \lambda, \mu \in \mathbb{K} \ : \quad \lambda\,(\mu\, v) \ = \ (\lambda\, \mu)\, v$ (Assoz.-Ges. der Multiplik.)

(V6) $\forall\, v,w \in V, \quad \forall\, \lambda \in \mathbb{K} \ : \quad \lambda\,(v+w) \ = \ \lambda\, v + \lambda\, w$ (1. Distrib.-Ges.)

(V7) $\forall\, v \in V, \quad \forall\, \lambda, \mu \in \mathbb{K} \ : \quad (\lambda + \mu)\, v \ = \ \lambda\, v + \mu\, v$ (2. Distrib.-Ges.)

(V8) $\forall\, v \in V \ : \quad 1\, v \ = \ v$

Ist dies der Fall, so heißen die Elemente von V auch Vektoren. *Für den Fall $\mathbb{K} = \mathbb{R}$ spricht man auch von einem* reellen, *für $\mathbb{K} = \mathbb{C}$ von einem* komplexen *Vektorraum.*

Beispiel 1.1.2 *Beachte für dieses Beispiel sowie die ganze Vorlesung, dass $\mathbb{N} = \{1, 2, 3, \ldots\}$ die Menge der natürlichen Zahlen bezeichnet. Insbesondere ist $0 \notin \mathbb{N}$, was in manchen Büchern und Vorlesungen anders sein kann, und wir schreiben $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.*

Die folgenden Mengen sind alle Vektorräume über $\mathbb{K}$:

(a) Die Menge $\mathbb{K}^n$ aller *Spaltenvektoren der Länge* $n \in \mathbb{N}$ mit Elementen in $\mathbb{K}$, d. h., die Menge aller

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \ldots, x_n)^T ,$$

mit Zahlen $x_j \in \mathbb{K}$. Dabei ist für x wie oben und einen zweiten Spaltenvektor

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (y_1, \ldots, y_n)^T$$

sowie $\lambda \in \mathbb{K}$ Addition und Skalarmultiplikation definiert durch

$$x + y = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \qquad \lambda\, x = \begin{pmatrix} \lambda\, x_1 \\ \lambda\, x_2 \\ \vdots \\ \lambda\, x_n \end{pmatrix}.$$

Für $n = 1$ ist $\mathbb{K}^n$ „praktisch gleich" $\mathbb{K}$. Für $n = 2$ bzw. $n = 3$ und $\mathbb{K} = \mathbb{R}$ kann man x als Punkt der Ebene bzw. des Raumes auffassen und nennt deshalb auch allgemein die Zahlen x_k die Koordinaten des Vektors x. Beachte auch, dass die Schreibweise $x = (x_1, \ldots, x_n)^T$ ein Spezialfall der sogenannten Transposition einer Matrix ist, welche in Definition 2.3.1 noch genauer behandelt wird.

(b) Die Menge $C[a, b]$ aller stetigen Funktionen auf einem festen abgeschlossenen Intervall $[a, b]$, $a \leq b$, mit Werten in $\mathbb{K}$ und der üblichen Addition und Multiplikation.

(c) Die Menge $\mathbb{K}[t]$ aller *Polynome in einer Variablen t mit Koeffizienten* in $\mathbb{K}$. Man liest das Symbol $\mathbb{K}[t]$ als „$\mathbb{K}$ adjungiert t".

(d) Die Menge $\mathbb{K}_n[t]$ aller Polynome vom Grad $\leq n$ mit Koeffizienten in $\mathbb{K}$.

(e) Die Menge aller Funktionen auf einem festen, nicht-leeren Definitionsbereich D mit Werten in $\mathbb{K}$, oder allgemeiner, mit Werten in einem festen aber beliebigen Vektorraum V über $\mathbb{K}$.

(f) Kartesische Produkte von beliebig vielen Vektorräumen über $\mathbb{K}$.

(g) Die Menge $\mathbb{K}_J$ aller Abbildungen f einer nicht-leeren Menge J in $\mathbb{K}$ mit $f(j) \neq 0$ höchstens für endlich viele $j \in J$. Vergleiche hierzu auch Aufgabe 5.1.6.

Falls $\mathbb{K} = \mathbb{R}$ oder $= \mathbb{C}$ ist, hat ein Vektorraum V entweder nur ein Element, welches dann der Nullvektor sein muss, oder unendlich viele Elemente, da ja dann $\mathbb{K}$ selber bereits eine unendliche Menge ist; vergleiche dazu auch Aufgabe 1.2.5. *Es gibt aber auch Körper mit endlich vielen, z. B. zwei, Elementen, und für solche Körper ist dann z. B. auch $\mathbb{K}^n$ eine endliche Menge. Solche Fälle spielen aber in dieser Vorlesung keine Rolle.*

Aufgabe 1.1.3 *Zeige, dass $\mathbb{K}^n$ ein Vektorraum ist, d. h., zeige dass für die oben definierten Verknüpfungen in $\mathbb{K}^n$ alle Axiome eines Vektorraumes erfüllt sind.*

Aufgabe 1.1.4 *Begründe, warum die Menge der Polynome vom Grad n, bei festem $n \in \mathbb{N}$, keinen Vektorraum bildet.*

Aufgabe 1.1.5 *Sei $\mathbb{K}$ ein Körper mit $p \geq 2$ Elementen. Berechne die Anzahl der Elemente von $\mathbb{K}^n$ bzw. $\mathbb{K}_n[t]$, für $n \in \mathbb{N}$, sowie von $\mathbb{K}[t]$.*

Aufgabe 1.1.6 (Vektoren und MAPLE) *Begleitend zur Vorlesung werden wir das Computeralgebra-Paket* MAPLE *benutzen, das, wie auch viele andere solche Systeme, mit Vektoren aus* $\mathbb{K}^n$*, aber auch mit Polynomen symbolisch, d. h. „formelmäßig“, rechnen kann. Die Kommandozeilen*

```
> restart;
> with(LinearAlgebra);
> x := Vector([1, 2, 3]);
> y := Vector([3, 2, 1]);
> z := VectorAdd(x,y);
```

berechnen die Summe der zuvor eingegebenen Vektoren x *und* y*. Finde einen Rechner, auf dem* MAPLE *in der Version* 8 *oder höher installiert ist, starte das Paket mit* xmaple8 *oder einem ähnlichen Kommando, gib dann obige Zeilen ein und beobachte, was passiert. Dabei ist das erste Kommando „*> restart*“ nicht unbedingt notwendig, aber immer angeraten, da es „Reste“ von etwaigen früheren Befehlen beseitigt und das Programm sozusagen in den Ausgangszustand versetzt. Der zweite Befehl ruft ein Programmpaket zum Thema lineare Algebra auf. Es gibt auch noch ein zweites Paket mit dem Namen „linalg“, welches sich in etlichen Punkten von diesem hier aufgerufenen Paket unterscheidet.*

1.2 Einfache Folgerungen aus den Axiomen

Im Folgenden bezeichnet V immer einen Vektorraum über $\mathbb{K}$. Allein mit Hilfe der Axiome kann man weitere Rechenregeln beweisen, die in jedem Vektorraum gelten müssen:

Behauptung 1.2.1 *In jedem Vektorraum* V *über* $\mathbb{K}$ *gelten folgende Aussagen:*

(a) Es gibt nur einen Nullvektor in V.

(b) Für alle $v \in V$ gilt $0\,v = 0$; dabei steht links die Zahl 0, rechts der Nullvektor.

(c) Zu jedem $v \in V$ gibt es nur ein additives Inverses $\tilde{v}$, das wir im Folgenden auch mit $-v$ bezeichnen, denn es gilt auch $-v = (-1)\,v$ für alle $v \in V$.

(d) $\lambda\, v \;=\; 0 \quad \Longleftrightarrow \quad \lambda = 0 \quad \text{oder} \quad v = 0 \quad \text{(oder beides)}.$

(e) Die Gleichung $v + x = w$, mit $v, w \in V$, besitzt genau eine Lösung $x \in V$, nämlich $x = w + (-v)$.

(f) Die Gleichung $\alpha\, x = v$, mit $\alpha \in \mathbb{K} \setminus \{0\}$ und $v \in V$, besitzt genau eine Lösung, nämlich $x = \alpha^{-1}\, v$.

Beweis: Zu (a): Vorausgesetzt sei $0 + v = \tilde{0} + v = v$ für alle $v \in V$; zu zeigen ist $0 = \tilde{0}$. Durch Einsetzen von 0 bzw. $\tilde{0}$ für v folgt $0 = \tilde{0} + 0 = 0 + \tilde{0} = \tilde{0}$, und das war zu zeigen. Zu (b): Sei $w = 0\,v$ gesetzt. Mit Hilfe des zweiten Distributivgesetzes folgt $w = (0 + 0)\,v = w + w$. Durch Addition eines additiven Inversen von w zu beiden Seiten der Gleichung folgt $w = 0$. Zu (c): Es gilt für jedes $v \in V$: $v + (-1)\,v = (1 - 1)\,v = 0\,v = 0$ nach (b). Also ist $(-1)\,v$ *ein* additives Inverses zu v. Sei $\tilde{v}$ ein weiteres solches. Dann folgt unter Verwendung von Kommutativ- und Assoziativgesetz der Vektoraddition: $\tilde{v} = \tilde{v} + 0 = \tilde{v} + v + (-1)\,v = v + \tilde{v} + (-1)\,v = 0 + (-1)\,v = (-1)\,v$. Zu (d): Genauso wie bei (b) zeigt man, dass $\lambda\,0 = 0$ für jedes $\lambda \in \mathbb{K}$. Sei jetzt $\lambda\, v = 0$, und sei o. B. d. A. angenommen, dass $\lambda \neq 0$ ist. Dann folgt $0 = \lambda^{-1}\,0 = (\lambda^{-1}\,\lambda)\,v = v$, wobei die Axiome (V5) und (V8) verwendet wurden. Die Beweise von (e) und (f) werden als Übungsaufgabe gestellt. □

Aufgabe 1.2.2 *Beweise die Aussagen* (e) *und* (f) *der obigen Behauptung.*

Bemerkung 1.2.3 *Wir schreiben in Zukunft also immer $-v$ für das additive Inverse eines Vektors v. Statt $w + (-v)$, für $v, w \in V$, schreiben wir dann auch kurz $w - v$ und sprechen von der* Differenz *der Vektoren w und v. In einem Vektorraum V gibt es also neben der Addition und der Multiplikation mit Skalaren eine dritte, abgeleitete Operation, die Subtraktion. Beachte aber, dass im Allgemeinen weder ein Produkt von Vektoren noch etwa gar ein Quotient definiert ist.*

Aufgabe 1.2.4 (Allgemeine Distributivgesetze) *Zeige: Für alle $n \in \mathbb{N}$, $\lambda, \lambda_1, \ldots, \lambda_n \in \mathbb{K}$ und $v, v_1, \ldots, v_n \in V$ gilt immer*

$$\lambda \sum_{k=1}^{n} v_k \;=\; \sum_{k=1}^{n} \lambda\, v_k\,, \qquad \Big(\sum_{k=1}^{n} \lambda_k \Big)\, v \;=\; \sum_{k=1}^{n} \lambda_k\, v\,.$$

Aufgabe 1.2.5 *Zeige: In einem Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ gibt es entweder nur einen oder unendlich viele Vektoren. Warum ist dies nicht richtig für allgemeine Körper $\mathbb{K}$?*

1.3 Unterräume

Definition 1.3.1 *Eine nichtleere Teilmenge $U \subset V$ heißt* ein Unterraum *oder* Teilraum, *falls U abgeschlossen bezüglich der Addition und der Multiplikation mit Skalaren ist, d. h., falls gilt:*

(U1) $\forall\, v, w \in U \;:\; \quad v + w \in U,$

(U2) $\forall\, v \in U \quad \forall\, \lambda \in \mathbb{K} \;:\; \quad \lambda\, v \in U.$

Abstrakt ausgedrückt heißt das, dass die Restriktionen der Abbildungen „$+$“ und „$\cdot$“ von $U \times U$ nach U abbilden.

Satz 1.3.2 *Sei U ein Unterraum eines Vektorraums V über $\mathbb{K}$. Dann ist U selber wieder ein Vektorraum über $\mathbb{K}$; d. h. also, in U sind alle Vektorraumaxiome* (V1) – (V8) *erfüllt.*

Beweis: Die Gültigkeit von (V1) und (V4) – (V8) ist sofort klar, da nur die Menge der Vektoren verkleinert wurde. Da aus (U2) für $\lambda = 0$ folgt, dass $0 \in U$ ist, gilt auch (V2), und für $\lambda = -1$ folgt aus (U2) auch (V3). $\square$

Aufgabe 1.3.3 *Zeige: V selber sowie die Teilmenge $\{0\}$, welche also nur aus dem Nullvektor besteht, sind immer Unterräume von V. Man nennt diese beiden auch die* trivialen Unterräume *von V. Zeige weiter, dass für ein $v \in V$ die Menge aller Vielfachen $\{\lambda\, v : \lambda \in \mathbb{K}\}$ ebenfalls ein Unterraum ist.*

Aufgabe 1.3.4 *Zeige: Die Menge aller Polynome ist ein Unterraum des Vektorraums aller Funktionen mit Definitionsbereich $\mathbb{K}$.*

Aufgabe 1.3.5 *Die Vektoren aus $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ kann man sich in natürlicher Weise als Punkte einer Ebene bzw. des Raumes veranschaulichen. Untersuche, welche Geraden in der Ebene bzw. im Raum Unterräume sind.*

Lemma 1.3.6 *Seien U_j, $j \in J$, Unterräume von V. Dann ist auch ihr Durchschnitt*

$$U \;=\; \bigcap_{j \in J} U_j$$

ein Unterraum von V.

Beweis: Die Gültigkeit von (U1) und (U2) für U ist unmittelbar klar auf Grund der Definition des Durchschnitts, aber wir müssen noch zeigen, dass U nicht leer ist. Dies folgt aber, da nach Satz 1.3.2 jedes U_j selber ein Vektorraum ist und somit den Nullvektor enthalten muss, und dieser ist dann auch Element des Durchschnitts U. $\square$

Aufgabe 1.3.7 *Untersuche, ob die Vereinigung zweier Unterräume wieder ein Unterraum ist.*

1.4 Lineare Hülle, lineare Unabhängigkeit

Definition 1.4.1 *Gegeben seien eine beliebige Anzahl von Vektoren $v_j \in V$, für $j \in J \neq \emptyset$. Dabei soll erlaubt sein, dass auch einige der Vektoren v_j gleich sind, und deshalb sprechen wir statt von der Menge besser von dem* System *der Vektoren $(v_j, j \in J)$. Ist $J_1 \subset J$, so heißt $(v_j, j \in J_1)$* Teilsystem *von $(v_j, j \in J)$. Gelegentlich werden wir auch Teilmengen von V als Systeme auffassen, was ohne weiteres möglich ist, während umgekehrt ein System im allgemeinen keine Teilmenge von V ist. Für endlich viele $j_1, \ldots, j_n \in J$ und Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ heißt die Summe*

$$v \;=\; \sum_{k=1}^{n} \lambda_k \, v_{j_k} \tag{1.4.1}$$

eine Linearkombination *der $(v_j, j \in J)$. Der Nullvektor ist also immer eine Linearkombination, da ja alle λ_k gleich 0 sein können. Die Menge aller Linearkombinationen von $(v_j, j \in J)$ heißt* die lineare Hülle *dieser Vektoren, und wir schreiben für diese Menge in Zeichen $\mathcal{L}(v_j, j \in J)$. Die Vektoren $(v_j, j \in J)$ heißen* linear unabhängig*, falls für beliebige Indizes $j_1, \ldots, j_n \in J$ und Zahlen $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ die Gleichung*

$$\sum_{k=1}^{n} \lambda_k \, v_{j_k} \;=\; 0 \tag{1.4.2}$$

nur dann gilt, wenn alle $\lambda_k = 0$ sind. Ist dies nicht so, d. h., gilt (1.4.2) für mindestens eine Wahl von $j_1, \ldots, j_n \in J$ und $\lambda_1, \ldots, \lambda_n$, die nicht alle gleich 0 sind, so heißen die Vektoren linear abhängig.

Bemerkung 1.4.2 *Es ist bequem, statt (1.4.1) kürzer*

$$v = \sum_{j \in J} \lambda_j \, v_j \tag{1.4.3}$$

zu schreiben. Dabei sollen die λ_j alle aus dem Körper $\mathbb{K}$ sein, und es ist wichtig zu beachten, dass höchstens endlich viele dieser Skalare von 0 verschieden sein dürfen. *Wir stellen uns dabei vor, dass die rechts stehende Summe zwar formal unendlich viele Elemente enthalten kann, dass wir aber alle Terme mit $\lambda_j = 0$ ignorieren können, sodass nur eine endliche Summe zu berechnen ist oder sogar nur die leere Summe übrig bleibt, welche per Definition immer den Nullvektor ergibt. Mit dieser Kurzschreibweise zeigt sich dann die lineare Unabhängigkeit des Systems $(v_j, j \in J)$ dadurch, dass die Gleichung (1.4.3) für $v = 0$ nur dann bestehen kann, wenn* alle *$\lambda_j = 0$ sind. Es ist sinnvoll, auch ein* leeres System *zuzulassen, für welches also $J = \emptyset$ ist. Dieses soll als linear unabhängig angesehen werden, und entsprechend der Konvention, dass eine leere Summe den Wert 0 haben soll, sehen wir den Nullvektor als die einzige Linearkombination des leeren Systems an.*

Beispiel 1.4.3 *Die Menge $\mathbb{K}[t]$ aller Polynome mit Koeffizienten in $\mathbb{K}$ ist ein Vektorraum über $\mathbb{K}$. Jedes Polynom $p(t) = \sum_{j=0}^{n} a_j \, t^j$ ist offenbar eine Linearkombination des Systems der* Monome *$(t^j, j \in \mathbb{N}_0)$. Nach dem sogenannten* Identitätssatz für Polynome*, der in der Vorlesung* Analysis *behandelt wird, ist $p(t) = 0$ für alle $t \in \mathbb{K}$ gleichwertig mit $a_0 = \ldots = a_n = 0$. Dies bedeutet dasselbe wie die Aussage, dass die Monome linear unabhängig sind.*

Aufgabe 1.4.4 *Formuliere die obigen Begriffe der Linearkombination und der linearen Unabhängigkeit noch einmal für den wichtigsten Spezialfall eines endlichen Systems $(v_1, \ldots, v_n)$ von Vektoren in V.*

Lösung: Da bei einer Linearkombination auch einige der Zahlen $\lambda_k = 0$ sein dürfen, kann man o. B. d. A. so sagen:

Für beliebige Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ heißt $\sum_{k=1}^{n} \lambda_k \, v_k$ eine *Linearkombination* von $(v_1, \ldots, v_n)$. Das System $(v_1, \ldots, v_n)$ heißt *linear unabhängig*, falls die Gleichung

$$\sum_{k=1}^{n} \lambda_k \, v_k \;=\; 0 \tag{1.4.4}$$

nur gilt, wenn alle $\lambda_k = 0$ sind. Ist dies nicht so, d. h., gilt (1.4.2) für mindestens eine Wahl von $\lambda_1, \ldots, \lambda_n$, die nicht alle gleich 0 sind, so heißt das System *linear abhängig*. Statt „das System $(v_1, \ldots, v_n)$ ist linear (un)abhängig" soll es auch erlaubt sein zu sagen „die Vektoren $v_1, \ldots, v_n$ sind linear (un)abhängig". □

Aufgabe 1.4.5 *Zeige: Ein System von unendlich vielen Vektoren ist genau dann linear unabhängig, wenn jedes endliche Teilsystem linear unabhängig ist.*

Aufgabe 1.4.6 *Zeige: Ein System aus einem Vektor $v \in V$ ist genau dann linear unabhängig, wenn v nicht der Nullvektor ist. Zwei Vektoren v_1, v_2 sind genau dann linear abhängig, wenn einer der Vektoren ein skalares Vielfaches des anderen ist.*

Behauptung 1.4.7 *Für jedes System $(v_j, j \in J)$ aus V gelten immer folgende Aussagen:*

(a) Sind die Vektoren $(v_j, j \in J)$ linear unabhängig, und ist $J_1 \subset J$, so ist auch $(v_j, j \in J_1)$ linear unabhängig.

(b) Gilt $v_j = 0$ für ein $j \in J$, oder $v_j = v_k$ für $j, k \in J$ mit $j \neq k$, so ist $(v_j, j \in J)$ linear abhängig.

(c) Ist $J \neq \emptyset$, und sind die Vektoren $(v_j, j \in J)$ linear unabhängig, so gibt es zu jedem $v \in \mathcal{L}(v_j, j \in J)$ eindeutig bestimmte Skalare $\lambda_j \in \mathbb{K}$, $j \in J$, von denen höchstens endlich viele $\neq 0$ sind, sodass (1.4.3) gilt.

(d) Wenn J mindestens zwei Elemente enthält, ist $(v_j, j \in J)$ genau dann linear abhängig, wenn es ein $j_0 \in J$ gibt, für welches v_{j_0} als Linearkombination der übrigen v_j geschrieben werden kann.

Beweis: Zu (a): Folgt sofort aus der Definition der linearen Unabhängigkeit. Zu (b): Setze $J_1 = \{j\}$ im ersten bzw. $J_1 = \{j, k\}$ im zweiten Fall und wende Teil (a) sowie Aufgabe 1.4.6 an. Zu (c): Jedes $v \in \mathcal{L}(v_j, j \in J)$ kann per Definition als Linearkombination (1.4.1) dargestellt werden, und dies kann kurz in der Form (1.4.3) geschrieben werden, wobei von den möglicherweise unendlich vielen Skalaren λ_j höchstens endlich viele von 0 verschieden sind. Wenn auch eine zweite solche Darstellung für v mit Koeffizienten α_j gilt, so folgt offenbar $0 = \sum_{j \in J} (\lambda_j - \alpha_j) \, v_j$, was auf Grund der linearen Unabhängigkeit nur sein kann, wenn für alle $j \in J$ gilt $\alpha_j = \lambda_j$. Zu (d): Genau dann ist $(v_j, j \in J)$ linear abhängig, wenn $\sum_{j \in J} \lambda_j \, v_j = 0$ gilt für eine Wahl von $\lambda_j \in \mathbb{K}$, wobei für mindestens ein $j_0 \in J$ gilt $\lambda_{j_0} \neq 0$. Diese Gleichung ist aber äquivalent zu

$$v_{j_0} \;=\; - \sum_{j \in J \setminus \{j_0\}} \frac{\lambda_j}{\lambda_{j_0}} \, v_j \,,$$

was zu zeigen war. □

Satz 1.4.8 *Für ein System $(v_j, j \in J)$ ist die lineare Hülle $\mathcal{L}(v_j, j \in J)$ ein Unterraum von V, und zwar genau der Durchschnitt aller Unterräume von V, welche alle v_j enthalten.*

Beweis: Sei U der Durchschnitt aller Unterräume, welche alle v_j enthalten. Nach Lemma 1.3.6 ist U selber ein Unterraum, und nach Definition folgt, dass U alle Linearkombinationen der v_j enthalten muss. Also ist $\mathcal{L}(v_j, j \in J) \subset U$. Umgekehrt zeigt man schnell, dass $\mathcal{L}(v_j, j \in J)$ selber ein Unterraum ist, und deshalb folgt $\mathcal{L}(v_j, j \in J) = U$. □

Aufgabe 1.4.9 *Zeige: Drei Vektoren in $\mathbb{R}^2$ sind immer linear abhängig. Finde zwei Vektoren in $\mathbb{R}^2$, welche linear unabhängig sind.*

1.5 Erzeugendensystem und Basis

Definition 1.5.1 *Vektoren $(v_j, j \in J)$ aus V heißen* ein Erzeugendensystem *für V, falls jeder Vektor aus V eine Linearkombination der v_j ist, d. h., falls $\mathcal{L}(v_j, j \in J) = V$ ist. Wenn die $(v_j, j \in J)$ zugleich ein Erzeugendensystem und linear unabhängig sind, heißen sie auch* eine Basis *von V.*

Offenbar ist das leere System Basis von V genau dann, wenn V nur aus dem Nullvektor besteht. Ob es in einem allgemeinen Vektorraum eine Basis gibt, ist zunächst nicht klar und soll noch genauer untersucht werden. In manchen Vektorräumen gibt es aber eine *natürliche Basis*:

Behauptung 1.5.2 *Im Vektorraum der Polynome sind die Monome eine Basis. In $\mathbb{K}^n$ entsprechen diesen Monomen die Spaltenvektoren*

$$e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k\,, \qquad 1 \le k \le n\,, \tag{1.5.1}$$

wobei der Pfeil andeuten soll, dass die einzige 1 an der Stelle Nr. k steht. Diese sind ebenfalls eine Basis, und wir nennen diese auch die kanonische Basis von $\mathbb{K}^n$. *Allgemeiner sind in $\mathbb{K}_J$ wie in* Beispiel 1.1.2 (g) *diejenigen Funktionen, welche an einer Stelle $j \in J$ den Wert 1 annehmen und sonst gleich 0 sind, eine Basis. Beachte aber, dass dies nicht so ist, wenn wir die Menge aller Funktionen auf J mit Werten in $\mathbb{K}$ betrachten.*

Aufgabe 1.5.3 *Zeige die Richtigkeit der vorstehenden Behauptung.*

Aufgabe 1.5.4 (Einheitsvektoren und MAPLE) MAPLE *kennt die oben definierten* Einheitsvektoren *in $\mathbb{K}^n$. Führe die Kommandos*

```
> restart;
> with(LinearAlgebra);
> x := UnitVector(1,5);
> y := UnitVector(4,5);
> z := VectorAdd(x,y,4,-1);
```

aus und finde insbesondere heraus, was die beiden Parameter 4 *und* -1 *im letzten Kommando bedeuten.*

Aufgabe 1.5.5 *Sei* $(v_j, j \in J)$ *ein Erzeugendensystem für* V*. Zeige: Falls ein* $j_0 \in J$ *existiert, für welches* v_{j_0} *eine Linearkombination der übrigen Vektoren* v_j *ist, so ist auch* $(v_j, j \in J \setminus \{j_0\})$ *ein Erzeugendensystem von* V*. Wir sagen dann auch, dass das Erzeugendensystem* $(v_j, j \in J)$ verkürzbar ist.

Eine Basis eines Vektorraums kann also aus endlich, aber auch aus unendlich vielen Vektoren bestehen; dies ist wichtig für den nächsten Abschnitt. Hier zeigen wir noch folgendes Resultat:

Proposition 1.5.6 *Für ein nicht-leeres Systems* $(v_j, j \in J)$ *von Vektoren in einem Vektorraum* V *über* $\mathbb{K}$ *sind folgende Aussagen äquivalent:*

(a) *Die* $(v_j, j \in J)$ *bilden eine Basis von* V*.*

(b) *Die* $(v_j, j \in J)$ *sind ein unverkürzbares Erzeugendensystem von* V*; d. h., sie sind ein Erzeugendensystem von* V*, und für jede Teilmenge* $J_1 \subset J$ *gilt: Wenn* $(v_j, j \in J_1)$ *ebenfalls eine Erzeugendensystem von* V *ist, so folgt* $J_1 = J$*.*

(c) *Die* $(v_j, j \in J)$ *sind ein unverlängerbares linear unabhängiges System; d. h., sie sind linear unabhängig, und für jede Obermenge* $J_1 \supset J$ *gilt: Falls* $(w_j, j \in J_1)$ *ein linear unabhängiges System ist, und falls* $w_j = v_j$ *ist für alle* $j \in J$*, so folgt* $J_1 = J$*.*

(d) *Jeder Vektor* $v \in V$ *lässt sich auf genau eine Weise als eine Linearkombination der* $(v_j, j \in J)$ *schreiben; d. h., es gibt eindeutig bestimmte* $\lambda_j \in \mathbb{K}$*,* $j \in J$*, nur endlich viele* $\neq 0$*, für welche* (1.4.3) *gilt.*

Beweis: Wir zeigen: (a)$\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (a).

Zu (a)$\Longrightarrow$ (b): Falls es ein $j_0 \in J \setminus J_1$ gäbe, dann müsste v_{j_0} eine Linearkombination der $(v_j, j \in J_1)$ sein, also gäbe es $\lambda_j \in \mathbb{K}$, $j \in J_1$ mit
$$v_{j_0} = \sum_{j \in J_1} \lambda_j \, v_j \,.$$
Setzt man $\lambda_{j_0} = -1$ und $\lambda_j = 0$ für $j \in J \setminus (J_1 \cup \{j_0\})$, so folgt
$$\sum_{j \in J} \lambda_j \, v_j = 0 \,,$$
was der linearen Unabhängigkeit von $(v_j, j \in J)$ widerspricht.

Zu (b) $\Longrightarrow$ (c): Da das System $(v_j, j \in J)$ ein Erzeugendensystem für $V \neq \{0\}$ ist, kann es nicht nur aus dem Nullvektor bestehen. Falls es linear abhängig wäre, wäre nach Behauptung 1.4.7 (d) ein v_{j_0} eine Linearkombination der übrigen Vektoren v_j, und nach Aufgabe 1.5.5 ergäbe sich ein Widerspruch zu (b). Also ist $(v_j, j \in J)$ linear unabhängig. Wenn das System verlängerbar wäre, dann gäbe es einen Vektor $v \in V$, der nicht als Linearkombination von $(v_j, j \in J)$ geschrieben werden könnte, was aber ebenfalls (b) widerspräche.

Zu (c) $\Longrightarrow$ (d): Falls ein $v \in V$ keine Linearkombination von $(v_j, j \in J)$ wäre, könnte man diesen Vektor zum System hinzufügen, ohne die lineare Unabhängigkeit zu verlieren. Also ist jedes v eine solche Linearkombination, und aus Behauptung 1.4.7 (c) folgt, dass diese Darstellung eindeutig ist.

Zu (d) $\Longrightarrow$ (a): Klar ist, dass $(v_j, j \in J)$ ein Erzeugendensystem ist. Da jeder Vektor, also auch der Nullvektor, nur auf eine Weise als Linearkombination darstellbar ist, folgt auch die lineare Unabhängigkeit. $\square$

1.6 Existenz von Basen in endlich-dimensionalen Räumen

Definition 1.6.1 *Ein Vektorraum V über $\mathbb{K}$ heißt* endlich-dimensional*, falls er ein Erzeugendensystem aus endlich vielen Vektoren besitzt; falls nicht, nennen wir ihn* unendlich-dimensional*. Da das leere System aus endlich vielen Vektoren besteht und Erzeugendensystem des Vektorraums ist, welcher nur aus dem Nullvektor besteht, folgt dass dieser endlich-dimensional ist.*

Aufgabe 1.6.2 *Zeige: Der Vektorraum $\mathbb{K}[x]$ aller Polynome ist unendlich-dimensional.*

Die Existenz einer Basis in einem endlich-dimensionalen Raum folgt aus dem nächsten Satz:

Satz 1.6.3 (Basisauswahlsatz) *Sei $V \neq \{0\}$ ein Vektorraum über $\mathbb{K}$, und sei $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V, also insbesondere ist V endlich-dimensional. Dann gibt es eine Teilmenge $\{j_1, \ldots, j_n\} \subset \{1, \ldots, m\}$ so, dass $(v_{j_1}, \ldots, v_{j_n})$ eine Basis von V ist.*

Beweis: Falls das System $(v_1, \ldots, v_m)$ linear unabhängig ist, dann ist es bereits eine Basis, und die Behauptung gilt für $\{j_1, \ldots, j_n\} = \{1, \ldots, m\}$. Falls nicht, ist einer der Vektoren v_j eine Linearkombination der übrigen. Durch Änderung der Numerierung können wir erreichen, dass dies für $j = n$ zutrifft, und mit Aufgabe 1.5.5 folgt, dass dann $(v_1, \ldots, v_{m-1})$ ebenfalls Erzeugendensystem von V ist. Durch Wiederholung dieses Schlusses ergibt sich die Behauptung in endlich vielen Schritten. □

Aufgabe 1.6.4 *Zeige, dass die Polynome $(t^2 - t + 1, t^2 + t, t - 1, t)$ ein Erzeugendensystem für $\mathbb{R}_2[t]$ sind, und wähle daraus ein Basis aus.*

Bemerkung 1.6.5 (Basen in unendlich-dimensionalen Räumen) *Auch in unendlich-dimensionalen Räumen gibt es immer eine Basis. Der Beweis beruht auf dem sogenannten* Zornschen Lemma*, was zum* Auswahlaxiom *der Mengenlehre äquivalent ist. In dieser und späteren Vorlesungen spielt die Existenz von Basen in unendlicher Dimension keine große Rolle, und wir lassen deshalb den Beweis hier aus.*

Aufgabe 1.6.6 (Basen und MAPLE) *Finde heraus, welche* MAPLE*-Kommandos geeignet sind, um Basen von Unterräumen von $\mathbb{K}^n$ zu finden.*

1.7 Dimension

In jedem Vektorraum $V \neq \{0\}$ gibt es unendlich viele verschiedene Basen, wie sich aus dem folgenden Lemma ergibt:

Lemma 1.7.1 (Austauschlemma) *Sei $V \neq \{0\}$ Vektorraum über $\mathbb{K}$, und sei $(v_j, j \in J)$ eine Basis von V. Sei weiter $v \in V \setminus \{0\}$; also kann v in der Form (1.4.3) geschrieben werden, wobei nicht alle $\lambda_j = 0$ sein können. Wähle ein $j_0 \in J$ mit $\lambda_{j_0} \neq 0$. Dann ist auch das System, welches aus der Basis durch Ersetzen von v_{j_0} durch v entsteht, wieder eine Basis von V.*

Beweis: Für ein beliebiges $w \in V$ wollen wir zeigen, dass die Gleichung

$$w = \alpha v + \sum_{j \in J \setminus \{j_0\}} \alpha_j v_j$$

durch eindeutig bestimmte Koeffizienten $\alpha, \alpha_j \in \mathbb{K}$ erfüllbar ist, von denen höchstens endlich viele nicht verschwinden. Durch Einsetzen von (1.4.3) kann diese Gleichung umgeschrieben werden in die Form

$$w \;=\; \alpha\,\lambda_{j_0}\,v_{j_0} \;+\; \sum_{j\in J\setminus\{j_0\}} \big(\alpha_j + \alpha\,\lambda_j\big)\,v_j\,.$$

Da $(v_j, j \in J)$ eine Basis von V ist, gibt es eindeutig bestimmte $\beta_j \in \mathbb{K}$ mit

$$w \;=\; \sum_{j\in j} \beta_j\,v_j\,,$$

und da nach Voraussetzung $\lambda_{j_0} \neq 0$ ist, können die Beziehungen $\beta_{j_0} = \alpha\,\lambda_{j_0}$ und $\beta_j = \alpha_j + \alpha\,\lambda_j$ für $j \in J \setminus \{j_0\}$ eindeutig nach α und α_j aufgelöst werden. Also ist das neue System ebenfalls eine Basis des Raumes V. □

Aufgabe 1.7.2 *Sei U ein Unterraum von $\mathbb{C}^n$, $n \geq 1$, mit folgender Eigenschaft: Für jeden Vektor $u = (u_1, \ldots, u_n)^T \in U$ ist auch $\overline{u} = (\overline{u}_1, \ldots, \overline{u}_n)^T$ in U enthalten. Sei weiter $(u_1, \ldots, u_m)$ eine Basis von U. Zeige mit Hilfe des Austauschlemmas: Für jedes $j = 1, \ldots, m$ kann man den Vektor u_j entweder durch $u_1 + \overline{u}_j$ oder durch $u_1 - \overline{u}_j$ ersetzen, ohne die lineare Unabhängigkeit des Systems zu verlieren. Schließe hieraus, dass es eine Basis von U gibt, welche nur aus reellen Vektoren u_j besteht.*

Aufgabe 1.7.3 *Die Monome sind bekanntlich eine Basis von $\mathbb{R}[t]$. Untersuche, welches Monom durch das Polynom $t^2 - 1$ ersetzt werden kann, so dass wir weiterhin eine Basis von $\mathbb{R}[t]$ haben.*

Satz 1.7.4 (Austauschsatz von Steinitz) *Sei $V \neq \{0\}$ Vektorraum über $\mathbb{K}$, und sei $(v_j, j \in J)$ eine Basis von V. Sei weiter $(w_1, \ldots, w_m)$ ein endliches linear unabhängiges System von Vektoren aus V. Dann gibt es $j_1, \ldots, j_m \in J$ derart, dass das System, welches aus $(v_j, j \in J)$ durch Ersetzen von v_{j_k} durch w_k, für $1 \leq k \leq m$, entsteht, ebenfalls Basis von V ist. Insbesondere sind die j_k alle voneinander verschieden, sodass J mindestens m Elemente haben muss.*

Beweis: Folgt aus dem Austauschlemma mit vollständiger Induktion über m: Für $m = 1$ ist nichts mehr zu zeigen. Also sei angenommen, dass der Satz für ein $m \geq 1$ richtig ist. Für ein linear unabhängiges System $(w_1, \ldots, w_m, w_{m+1})$ wenden wir dann die Induktionshypothese auf $(w_1, \ldots, w_m)$ an und können somit annehmen, dass es $j_1, \ldots, j_m \in J$ gibt, für welche (nach der Ersetzung) $v_{j_k} = w_k$ gilt für $k = 1, \ldots, m$. Der Vektor w_{m+1} kann dann in der Form

$$w_{m+1} \;=\; \sum_{j\in J} \lambda_j\,v_j$$

geschrieben werden. Wäre $\lambda_j = 0$ für alle $j \in J \setminus \{j_1, \ldots, j_m\}$, so wäre w_{m+1} eine Linearkombination der $v_{j_k} = w_k$, was der linearen Unabhängigkeit des Systems $(w_1, \ldots, w_m, w_{m+1})$ widerspräche. Also gibt es ein $j_{m+1} \in J \setminus \{j_1, \ldots, j_m\}$ mit $\lambda_{j_{m+1}} \neq 0$. Anwendung des Austauschlemmas auf dieses j_{m+1} zeigt dann die Gültigkeit der Behauptung auch für $m + 1$. □

Korollar zu Satz 1.7.4 *Sei V endlich-dimensional. Dann haben zwei Basen von V immer die gleiche Anzahl von Elementen.*

Beweis: Sei $(v_1, \ldots, v_n)$ Basis von V, und sei $(w_1, \ldots, w_m)$ linear unabhängig. Dann folgt $m \leq n$ aus dem Austauschsatz. Wenn $(w_1, \ldots, w_m)$ ebenfalls eine Basis von V ist, kann man die Bezeichnungen vertauschen, und daraus ergibt sich die Behauptung. □

Aufgabe 1.7.5 *Zeige: Genau dann gilt* $\dim V = \infty$, *wenn es ein abzählbar-unendliches System* $(v_j, j \in \mathbb{N})$ *in* V *gibt, welches linear unabhängig ist.*

Definition 1.7.6 *Als* Dimension eines Vektorraumes V *über* $\mathbb{K}$ *definieren wir*

$$\dim V \;=\; \begin{cases} \infty & \text{falls } V \text{ unendlich-dimensional ist,} \\ n \in \mathbb{N}_0 & \text{falls } V \text{ eine Basis aus } n \text{ Vektoren besitzt.} \end{cases}$$

Insbesondere ist $\dim V = 0$ *genau dann, wenn* V *nur aus dem Nullvektor besteht.*

Aufgabe 1.7.7 (Dimension und MAPLE) *Finde selber die Funktion der Kommandos* > Dimension *bzw.* > Dimensions *heraus.*

Aufgabe 1.7.8 *Finde die Dimension der Vektorräume* $\mathbb{K}^n$, $\mathbb{K}_n[t]$ *und* $\mathbb{K}[t]$.

Korollar zu Satz 1.7.4 *Ist* U *ein Unterraum von* V, *so folgt* $\dim U \leq \dim V$. *Ist sogar* $\dim U = \dim V < \infty$, *so folgt* $U = V$.

Beweis: Falls U unendlich-dimensional ist, gibt es wegen Aufgabe 1.7.5 ein abzählbar-unendliches System $(u_j, j \in \mathbb{N})$ in U, welches linear unabhängig ist, und dies ist selbstverständlich auch in V linear unabhängig. Also gilt die Behauptung in diesem Fall. Im anderen Fall sei $(u_1, \ldots, u_m)$ eine Basis von U, also insbesondere $\dim U = m$. Für $\dim V = \infty$ ist nichts mehr zu zeigen, also sei $n = \dim V < \infty$ vorausgesetzt. Mit dem Austauschsatz von Steinitz folgt dann $m \leq n$, und bei Gleichheit muss $(u_j, j \in \mathbb{N})$ bereits Basis von V sein, woraus $U = V$ folgt. □

Aufgabe 1.7.9 *Finde ein Beispiel eines Vektorraumes* V *über* $\mathbb{K}$, *welcher einen echten Unterraum* $U \neq V$ *besitzt, so dass* $\dim U = \dim V$ *ist.*

Aufgabe 1.7.10 *Sei* $n \in \mathbb{N}$, *und sei* V *ein* n*-dimensionaler Vektorraum über* $\mathbb{K}$. *Zeige:*

(a) Für $m > n$ sind m Vektoren in V immer linear abhängig.

(b) Für $m < n$ sind m Vektoren in V kein Erzeugendensystem von V.

(c) Falls n Vektoren in V linear unabhängig *oder* ein Erzeugendensystem sind, so bilden sie bereits eine Basis von V.

Sozusagen komplementär zum Basisauswahlsatz ist der folgende Ergänzungssatz, der auch oft angewandt werden kann:

Satz 1.7.11 (Basisergänzungssatz) *Sei* V *ein* n*-dimensionaler Vektorraum über* $\mathbb{K}$, *sei* $0 \leq m < n$, *und sei* $(v_1, \ldots, v_m)$ *ein linear unabhängiges System in* V. *Dann kann man Vektoren* $v_{m+1}, \ldots, v_n$ *so wählen, dass* $(v_1, \ldots, v_n)$ *eine Basis von* V *ist.*

Beweis: Wähle irgendeine Basis $(w_1, \ldots, w_n)$ von V und wende den Austauschsatz von Steinitz an! □

Aufgabe 1.7.12 *Ergänze den Vektor*

$$v \;=\; \begin{pmatrix} 1 \\ 1 \end{pmatrix} \;\in\; \mathbb{R}^2$$

zu einer Basis von $\mathbb{R}^2$.

Aufgabe 1.7.13 *Seien $n \in \mathbb{N}$ und D eine Menge mit n Elementen. Zeige: Der Vektorraum aller Abbildungen von D nach $\mathbb{K}$ hat die Dimension n.*

1.8 Wechsel des Skalarenkörpers

Jeder Vektorraum über $\mathbb{C}$ ist offenbar auch Vektorraum über $\mathbb{R}$, denn dazu muss die Multiplikation mit Skalaren nur auf $\mathbb{R}$ eingeschränkt werden. Hinsichtlich der Dimension gilt folgendes:

Satz 1.8.1 *Sei V ein n-dimensionaler Vektorraum über $\mathbb{C}$, wobei auch $n = \infty$ zugelassen sei. Fasst man V als Vektorraum über $\mathbb{R}$ auf, so hat V die Dimension $2\,n$, wobei $2\,\infty = \infty$ sein soll.*

Beweis: Sei $\dim V = \infty$ (über $\mathbb{C}$). Dann gibt es ein linear unabhängiges System $(v_j, j \in \mathbb{N})$. Dieses System bleibt linear unabhängig, wenn wir V über $\mathbb{R}$ betrachten, und deshalb gilt die Behauptung in diesem Fall. Im anderen Fall sei $(v_1, \ldots, v_n)$ eine Basis von V (über $\mathbb{C}$). Wir zeigen, dass das System $(v_1, \ldots, v_n, i\,v_1, \ldots, i\,v_n)$ Basis von V über $\mathbb{R}$ ist. Dazu sei ein $v \in V$ betrachtet. Mit $\lambda_j = \alpha_j + i\,\beta_j$, $\alpha_j, \beta_j \in \mathbb{R}$, gilt die Gleichung $v = \sum_{j=1}^{n} \lambda_j\, v_j$ genau dann, wenn

$$v \;=\; \Big(\sum_{j=1}^{n} \alpha_j\, v_j \Big) \;+\; \Big(\sum_{j=1}^{n} \beta_j\, i\, v_j \Big),$$

und das war zu zeigen. □

Wenn V Vektorraum über $\mathbb{R}$ ist, kann man zunächst ein $v \in V$ nicht mit einer komplexen Zahl multiplizieren. Man kann aber V „in kanonischer Weise vergrößern" und so zu einem Vektorraum über $\mathbb{C}$ machen; dies geschieht analog zur Konstruktion der komplexen Zahlen:

Satz 1.8.2 (Komplexifizierung) *Sei V ein Vektorraum über $\mathbb{R}$. Dann ist auch $V \times V = \{(v_1, v_2) : v_1, v_2 \in V\}$ ein Vektorraum über $\mathbb{R}$, und wir können jedes $v \in V$ mit $(v, 0) \in V \times V$ identifizieren, so dass V ein Unterraum von $V \times V$ wird. Für $\alpha, \beta \in \mathbb{R}$, und $(v_1, v_2) \in V \times V$ sei*

$$(\alpha + i\,\beta)\;(v_1, v_2) \;=\; (\alpha\, v_1 - \beta\, v_2, \beta\, v_1 + \alpha\, v_2)$$

gesetzt. Dann wird $V \times V$ zu einem Vektorraum über $\mathbb{C}$, wobei die Paare $(0, v)$ gerade den Vektoren $i\,v = i\,(v, 0)$ entsprechen.

Aufgabe 1.8.3 *Beweise den obigen Satz!*

1.9 Summe von Unterräumen

Definition 1.9.1 *Sei V ein Vektorraum über $\mathbb{K}$,und seien A und B beliebige Teilmengen von V. Wir definieren*

$$A \;+\; B \;=\; \{a + b : \; a \in A,\; b \in B\} \tag{1.9.1}$$

als die Summe *von A und B. Besteht A nur aus einem einzigen Vektor a, so schreiben wir statt $A + B$ auch $a + B$. Beachte, dass man sich die Menge $a + B$ als die* Translation *oder* Verschiebung *von B um den Vektor a vorstellen kann. Ist U ein Unterraum von V sowie $v \in V$, so heißt die Menge $v + U$ auch eine* lineare Mannigfaltigkeit *in V, und wir sagen, dass die Dimension dieser linearen Mannigfaltigkeit gleich der von U ist.*

Satz 1.9.2 *Seien U_1 und U_2 Unterräume eines Vektorraumes V über $\mathbb{K}$. Dann sind auch $U_1 + U_2$ und $U_1 \cap U_2$ Unterräume von V, und es gilt*

$$\dim(U_1 + U_2) \ + \ \dim(U_1 \cap U_2) \ = \ \dim U_1 \ + \ \dim U_2\,,$$

auch falls einige der Räume unendlich-dimensional sind.

Beweis: Dass der Durchschnitt von Unterräumen ebenfalls Unterraum ist, ist in Lemma 1.3.6 gezeigt worden. Seien jetzt $u, \tilde{u} \in U_1 + U_2$. Dann gibt es nach Definition Vektoren $u_1, \tilde{u}_1 \in U_1$ und $u_2, \tilde{u}_2 \in U_2$ mit $u = u_1 + u_2$ und $\tilde{u} = \tilde{u}_1 + \tilde{u}_2$. Also folgt $u + \tilde{u} = (u_1 + \tilde{u}_1) + (u_2 + \tilde{u}_2) = \hat{u}_1 + \hat{u}_2$, mit $\hat{u}_1 \in U_1$, $\hat{u}_2 \in U_2$, und somit ist $u + \tilde{u} \in U_1 + U_2$. Für $\alpha \in \mathbb{K}$ ist $\alpha\, u = \alpha\, u_1 + \alpha\, u_2$, und daraus folgt genauso $\alpha\, u \in U_1 + U_2$. Um die Dimensionsgleichung zu erhalten, bemerken wir zunächst, dass $(U_1 \cup U_2) \subset (U_1 + U_2)$ gilt. Wenn also einer der beiden Unterräume U_j unendlich-dimensional ist, dann gibt es dort eine Folge von linear unabhängigen Vektoren, und somit ist auch $U_1 + U_2$ unendlich-dimensional, und dann gilt die Gleichung. Seien jetzt beide Unterräume endlich-dimensional. Dann ist auch $U_1 \cap U_2$ endlich-dimensional, und wir wählen eine Basis $(v_1, \ldots, v_\nu)$ von $U_1 \cap U_2$ (beachte, dass der Durchschnitt auch nur aus dem Nullvektor bestehen kann, sodass diese Basis auch leer sein kann, was $\nu = 0$ entspricht. Mit dem Basisergänzungssatz können wir weitere Vektoren $u_{11}, \ldots, u_{1\mu_1}$ und $u_{21}, \ldots, u_{2\mu_2}$ so wählen, dass $(v_1, \ldots, v_\nu, u_{11}, \ldots, u_{1\mu_1})$ bzw. $(v_1, \ldots, v_\nu, u_{21}, \ldots, u_{2\mu_2})$ Basis von U_1 bzw. U_2 ist. Wir zeigen nun, dass $(v_1, \ldots, v_\nu, u_{11}, \ldots, u_{1\mu_1}, u_{21}, \ldots, u_{2\mu_2})$ Basis von $U_1 + U_2$ ist (woraus die Behauptung durch Abzählen der Basisvektoren folgt): Jeder Vektor aus $U_1 + U_2$ ist von der Form $u = u_1 + u_2$, und u_j ist Linearkombination der gewählten Basis von U_j. Also ist das angegebene System ein Erzeugendensystem von $U_1 + U_2$. Für die lineare Unabhängigkeit setzen wir an

$$0 \ = \ \sum_{j=1}^{\nu} \alpha_j\, v_j \ + \ \sum_{k=1}^{\mu_1} \beta_k\, u_{1k} \ + \ \sum_{\ell=1}^{\mu_2} \gamma_\ell\, u_{2\ell} \ = \ v + u_1 + u_2\,.$$

Dann ist $v + u_1 \in U_1$ und $u_2 \in U_2$. Wegen $-u_2 = v + u_1$ folgt aber, dass $v + u_1 \in U_2$, also sogar im Durchschnitt $U_1 \cap U_2$ ist. Daher muss gelten $\beta_k = 0$ für $k = 1, \ldots, \mu_1$. Wegen der linearen Unabhängigkeit von $(v_1, \ldots, v_\nu, u_{21}, \ldots, u_{2\mu_2})$ folgt dann aber, dass auch alle α_j und γ_ℓ verschwinden müssen. □

Definition 1.9.3 *Seien U_1 und U_2 Unterräume eines Vektorraumes V über $\mathbb{K}$, und sei $U_1 \cap U_2 = \{0\}$. Dann nennen wir die Summe $U_1 + U_2$ auch eine* direkte Summe *und schreiben $U_1 \oplus U_2$. In diesem Fall gilt also die Gleichung*

$$\dim(U_1 \oplus U_2) \ = \ \dim U_1 \ + \ \dim U_2\,,$$

und zu jedem $u \in U_1 \oplus U_2$ gibt es genau ein Paar $u_1 \in U_1$ und $u_2 \in U_2$ mit $u = u_1 + u_2$. Entsprechend definieren wir die Summe $U_1 + \ldots + U_m$ endlich vieler Unterräume $U_1, \ldots, U_m$ von V als die Menge aller $u = u_1 + \ldots + u_m$ mit $u_j \in U_j$, $1 \le j \le m$, und nennen die Summe direkt und schreiben dann $U_1 \oplus \ldots \oplus U_m$, wenn für $u_j \in U_j$ die Gleichung

$$0 \ = \ u_1 + \ldots + u_m$$

nur dann bestehen kann, wenn $u_1 = \ldots = u_m = 0$ ist. Beachte auch die Analogie zur linearen Unabhängigkeit.

Aufgabe 1.9.4 *Untersuche, wann die Summe zweier eindimensionaler Unterräume von V eine direkte Summe ist.*

Aufgabe 1.9.5 *Sei $m \in \mathbb{N}$, seien $U_1, \ldots, U_m$ Unterräume eines endlich-dimensionalen Vektorraumes V über $\mathbb{K}$, und sei $U = U_1 + \ldots + U_m$. Zeige: Genau dann ist diese Summe direkt, wenn folgendes gilt:*

- *Sind $(u_1^{(j)}, \ldots, u_{s_j}^{(j)})$ Basen für U_j, für jedes $j = 1, \ldots, m$, so ist das System*

$$(u_1^{(1)}, \ldots, u_{s_1}^{(1)}, u_1^{(2)}, \ldots, u_{s_2}^{(2)}, \ldots, u_1^{(m)}, \ldots, u_{s_m}^{(m)})\,,$$

also sozusagen die Vereinigung dieser Basen, eine Basis für U.

Aufgabe 1.9.6 *Sei* $m \in \mathbb{N}$, *seien* $U_1, \ldots, U_m$ *Unterräume eines Vektorraumes* V *über* $\mathbb{K}$, *und sei* $U = U_1 + \ldots + U_m$. *Zeige: Genau dann ist diese Summe direkt, wenn es zu jedem* $u \in U$ eindeutig bestimmte $u_j \in U_j$ *gibt, für welche* $u = u_1 + \ldots + u_m$ *ist.*

Aufgabe 1.9.7 *Sei* $m \in \mathbb{N}$, *seien* $U_1, \ldots, U_m$ *Unterräume eines Vektorraumes* V *über* $\mathbb{K}$, *und sei* $U = U_1 + \ldots + U_m$. *Zeige, dass* U *die lineare Hülle der Vereinigung aller* U_j *ist.*

Aufgabe 1.9.8 *Finde selbst eine mögliche Definition für die Summe, bzw. die direkte Summe, einer beliebigen Anzahl von Unterräumen eines Vektorraums.*

Kapitel 2

Matrizen

2.1 Definition und elementare Eigenschaften

Definition 2.1.1 *Seien $n, m \in \mathbb{N}$. Eine Abbildung*

$$\{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow \mathbb{K}\,, \qquad (j,k) \longmapsto a_{jk} \tag{2.1.1}$$

kann man sich als ein rechteckiges Zahlenschema

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

veranschaulichen. Dieses Schema heißt eine $m \times n$-Matrix, *oder eine* Matrix vom Typ $m \times n$. *Die Zahlen a_{jk} heißen* die Elemente *von A, und wir schreiben manchmal auch kurz $A = \left[a_{jk}\right]$. Die $m \times 1$-Matrizen*

$$s_k \;=\; \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}, \qquad 1 \;\le k\; \le n\,,$$

heißen die Spalten *von A, die $1 \times n$-Matrizen*

$$z_j \;=\; \left[a_{j1}, a_{j2}, \ldots, a_{jn}\right], \quad 1 \;\le\; j \;\le\; m\,,$$

werden die Zeilen *von A genannt. Also hat eine Matrix vom Typ $m \times n$ gerade m Zeilen und n Spalten. Im Fall $n = m$ sprechen wir von einer* quadratischen Matrix. *Die Menge der $m \times n$-Matrizen wird auch mit $\mathbb{K}^{m \times n}$ bezeichnet. Zwei Matrizen A und B* vom gleichen Typ $m \times n$ *werden* addiert, *indem man zu jedem Element von A das an der gleichen Stelle stehende Element von B hinzuzählt. Anders ausgedrückt heißt das:*

$$A = \left[a_{jk}\right], \quad B = \left[b_{jk}\right] \quad \Longrightarrow \quad A \;+\; B = \left[a_{jk} + b_{jk}\right].$$

Eine Matrix A wird mit einem Faktor $\lambda \in \mathbb{C}$ multipiziert, indem man jedes ihrer Elemente mit λ malnimmt.

Satz 2.1.2 *Für $n, m \in \mathbb{K}$ ist $\mathbb{K}^{m \times n}$ mit den oben eingeführten Verknüpfungen ein Vektorraum über $\mathbb{K}$ der Dimension $n\,m$.*

Beweis: Da diese Matrixmenge eigentlich Abbildungen der Form (2.1.1) sind, ist die Vektorraumeigenschaft klar wegen Beispiel 1.1.2. Da der Definitionsbereich der Abbildungen $n\,m$ Elemente umfasst, folgt die Behauptung mit Aufgabe 1.7.13. □

Aufgabe 2.1.3 *Seien $n, m \in \mathbb{N}$. Gib eine Basis in $\mathbb{K}^{m\times n}$ an und vergleiche mit der kanonischen Basis in $\mathbb{K}^n$. Wie sieht der Nullvektor in $\mathbb{K}^{m\times n}$ aus? Wir bezeichnen diesen auch als Nullmatrix.*

Aufgabe 2.1.4 (Eingabe von Matrizen in MAPLE) MAPLE *kennt auch Matrizen. Die Befehlsfolge*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3, 3, [ [1, 0, 1], [0, 1, 1], [2, -1, 1] ]);
> B := Matrix(3, 3, [ [1, 1, 2], [0, 1, 1], [0, 1, 1] ]);
> C := Matrix(2, 3, [[1,1,1], [1, -1, 0]]);
```

definiert z. B. drei Matrizen A, B, C, mit welchen dann auch gerechnet werden kann; siehe dazu die noch folgenden Aufgaben. Untersuche selber, ob man die Angabe der Zeilen- und/oder Spaltenzahl auch weglassen kann, und was geschieht, wenn die eingegebenen Zeilen zu kurz oder zu lang, zu wenige oder zu viele sind.

Aufgabe 2.1.5 *Entscheide, welche der folgenden Matrizen addiert werden können, und berechne gegebenenfalls ihre Summe.*

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix},$$

$$D = \begin{bmatrix} -1 & 2 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad F = \begin{bmatrix} -1 & -2 \end{bmatrix}.$$

Multipliziere alle diese Matrizen mit dem Faktor $\lambda = 2$.

Aufgabe 2.1.6 (Matrixaddition mit MAPLE) *Die folgende Sequenz von* MAPLE*-Kommandos addiert die entsprechenden Matrizen:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3, 3, [ [1, 0, 1], [0, 1, 1], [2, -1, 1] ]);
> B := Matrix(3, 3, [ [1, 1, 2], [0, 1, 1], [0, 1, 1] ]);
> C := Matrix(2, 3, [[1,1,1], [1, -1, 0]]);
> A1 := Add(A,B);
> A2 := Add(A,C);
> A3 := Add(A,B,2,3);
```

Finde heraus, welche Bedeutung die optionalen Parameter 2 und 3 im letzten Kommando haben. Finde auch einen Weg, wie man mit Hilfe von MAPLE *eine Matrix mit einer Zahl multipliziert.*

2.2 Das Produkt von Matrizen

Definition 2.2.1 *Wir setzen für* $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{C}$

$$\begin{bmatrix} a_1, \ldots, a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_{j=1}^{n} a_j\, b_j\,.$$

Das heißt: Für eine Zeile und eine Spalte gleicher Länge ist jeweils ein Produkt definiert, und das Ergebnis dieser Operation ist ein Skalar. Allgemein definieren wir: Wenn A *und* B *Matrizen sind, und wenn die Zeilen von* A *dieselbe Länge wie die Spalten von* B *haben, dann ist das Produkt* $A\,B$ *die Matrix* $C = \begin{bmatrix} c_{jk} \end{bmatrix}$*, für die* c_{jk} *gerade das Produkt der* j*-ten Zeile von* A *mit der* k*-ten Spalte von* B *ist. Mit anderen Worten: Ist*

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1s} \\ b_{21} & b_{22} & \ldots & b_{2s} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{ns} \end{bmatrix},$$

so ist $A\,B = \begin{bmatrix} c_{jk} \end{bmatrix}$ *mit*

$$c_{jk} = \sum_{\nu=1}^{n} a_{j\nu}\, b_{\nu k}\,, \qquad 1 \le j \le m\,, 1 \le k \le s\,.$$

Aufgabe 2.2.2 *Zeige für* A, B, C *wie oben: Die* k*-te Spalte von* C *ist gleich dem Produkt von* A *mit der* k*-ten Spalte von* B*, die* j*-te Zeile von* C *ist das Produkt der* j*-ten Zeile von* A *mit* B*.*

Aufgabe 2.2.3 *Berechne die Produkte folgender Matrizen:*

$$\begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a & -b \\ 2 & a+b \end{bmatrix}.$$

Lösung: Das Ergebnis ist

$$\begin{bmatrix} 0 \\ -3 \end{bmatrix} \qquad \text{bzw.} \qquad \begin{bmatrix} -a+2 & 2\,b+a \\ a+6 & 3\,a+2\,b \end{bmatrix}.$$

□

Aufgabe 2.2.4 (Matrixmultiplikation mit MAPLE) MAPLE *kann auch Matrizen multiplizieren:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3, 3, [ [1, 0, 1], [0, 1, 1], [2, -1, 1] ]);
> B := Matrix(3, 3, [ [1, 1, 2], [0, 1, 1], [0, 1, 1] ]);
> C := Matrix(2, 3, [[1,1,1], [1, -1, 0]]);
> V1 := VandermondeMatrix([1,2,3,4]);
```

> V2 := VandermondeMatrix([a,b,c,d]);

> A3:=MatrixMatrixMultiply(A,B);

> A4:=MatrixMatrixMultiply(A,C);

> A4:=MatrixMatrixMultiply(C,A);

> A5:=MatrixMatrixMultiply(V1,V2);

Wie steht es mit der Gültigkeit des Kommutativgesetzes bei der Matrixmultiplikation? Beachte, dass MAPLE *auch mit Matrizen umgehen kann, deren Elemente nicht Zahlen sondern Variablen sind, und dass das System auch sogenannte Vandermondesche Matrizen kennt, die später noch eine Rolle spielen werden.*

Beispiel 2.2.5 *Für*

$$A \;=\; \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad X \;=\; \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

gilt offenbar

$$A\,X \;=\; \begin{bmatrix} 0 & 0 \\ x_{11} & x_{12} \end{bmatrix}, \qquad X\,A \;=\; \begin{bmatrix} x_{12} & 0 \\ x_{22} & 0 \end{bmatrix},$$

und daraus folgt $A^2 = A\,A = 0$*. Das bedeutet, dass eine Matrixgleichung* $A\,X = B$ *im Allgemeinen keine Lösung haben wird, und wenn doch, braucht die Lösung nicht eindeutig bestimmt zu sein. Weiter sieht man, dass für Matrizen* A *und* X *die Produkte* $A\,X$ *und* $X\,A$ *beide definiert sein können, aber im Allgemeinen verschieden ausfallen. Mit anderen Worten heißt das, dass für die Matrixmultiplikation* im Allgemeinen kein Kommutativgesetz gilt.

Aufgabe 2.2.6 *Berechne und vergleiche* $A\,B$ *und* $B\,A$ *für*

$$A \;=\; \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad B \;=\; \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

2.3 Rechenregeln

Definition 2.3.1 *Für eine beliebige* $m \times n$*-Matrix*

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

heißt

$$\overline{A} \;=\; \begin{bmatrix} \overline{a}_{11} & \overline{a}_{12} & \dots & \overline{a}_{1n} \\ \overline{a}_{21} & \overline{a}_{22} & \dots & \overline{a}_{2n} \\ \vdots & \vdots & & \vdots \\ \overline{a}_{m1} & \overline{a}_{m2} & \dots & \overline{a}_{mn} \end{bmatrix}$$

die zu A konjugiert komplexe Matrix. *Weiter nennen wir*

$$A^T \;=\; \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix},$$

welche die Zeilen von A als Spalten enthält und umgekehrt, die transponierte Matrix *oder kurz die* Transponierte *zu A. Für eine quadratische Matrix*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

heißen die Zahlen $a_{11}, \dots, a_{nn}$ die Diagonalelemente *von A. Falls $a_{jk} = 0$ ist für alle $k < j$, d. h., wenn alle Elemente unterhalb der Diagonalen verschwinden, dann heißt A eine* obere Dreiecksmatrix. *Falls dagegen $a_{jk} = 0$ ist für alle $k > j$, dann sprechen wir von einer* unteren Dreiecksmatrix. *Falls sogar beides gilt, d. h., falls alle Elemente außer evtl. den Diagonalelementen verschwinden, dann heißt A eine* Diagonalmatrix. *Wenn A eine Diagonalmatrix mit den Diagonalelementen $\lambda_1, \dots, \lambda_n$ ist, schreiben wir manchmal*

$$A = \operatorname{diag}[\lambda_1, \dots, \lambda_n].$$

Die $n \times n$-Matrix

$$I = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

also die Diagonalmatrix, deren Diagonalelemente alle gleich 1 sind, heißt Einheitsmatrix. *Wenn man das sogenannte* Kronecker-Delta

$$\delta_{jk} = \begin{cases} 0 & \text{für} \quad j \neq k \\ 1 & \text{für} \quad j = k \end{cases}$$

einführt, kann man sagen, dass I die Elemente δ_{jk} hat.

Aufgabe 2.3.2 (Matrizen und MAPLE) *Finde* MAPLE-*Kommandos zur Berechnung der transponierten Matrix sowie zur einfachen Eingabe von Diagonalmatrizen und der Einheitsmatrix, welche in Englisch „identity matrix" heißt und deshalb in dieser Vorlesung auch mit I bezeichnet wird. In der Literatur ist aber auch das Symbol E üblich.*

Satz 2.3.3 (Rechenregeln für Matrizen) *Für feste Zahlen $n, m, r, s \in \mathbb{N}$ gelten die folgenden Aussagen:*

(a) $A \in \mathbb{K}^{m\times n},\ B \in \mathbb{K}^{n\times r},\ \lambda \in \mathbb{K} \implies \lambda\,(A\,B) = (\lambda A)\,B = A\,(\lambda B)$.

(b) $A \in \mathbb{K}^{m\times n},\ B, C \in \mathbb{K}^{n\times r} \implies A\,(B + C) = A\,B + A\,C$.

(c) $A, B \in \mathbb{K}^{m\times n},\ C \in \mathbb{K}^{n\times r} \implies (A + B)\,C = A\,C + B\,C$.

(d) $A \in \mathbb{K}^{m\times n},\ B \in \mathbb{K}^{n\times r},\ C \in \mathbb{K}^{r\times s} \implies (A\,B)\,C = A\,(B\,C)$.

(e) $A \in \mathbb{K}^{m\times n},\ I = [\delta_{jk}] \in \mathbb{K}^{n\times n} \implies A\ I = A$.

(f) $I = [\delta_{jk}] \in \mathbb{K}^{n\times n},\ A \in \mathbb{K}^{n\times r} \implies I\ A = A$.

(g) $A, B \in \mathbb{K}^{m\times n} \implies (A + B)^T = A^T + B^T$.

(h) $A \in \mathbb{K}^{m\times n},\ B \in \mathbb{K}^{n\times r} \implies (A\,B)^T = B^T\,A^T$.

Aufgabe 2.3.4 *Beweise den oben stehenden Satz.*

Aufgabe 2.3.5 *Zeige: Wenn man eine Matrix A passender Größe von rechts bzw. links mit einer Diagonalmatrix $\Lambda = \operatorname{diag}[\lambda_1, \dots, \lambda_n]$ multipliziert, so erhält man das Produkt, indem man die k-te Spalte bzw. Zeile von A mit dem Faktor λ_k malnimmt.*

2.4 Zeilen- und Spaltenrang

Definition 2.4.1 *Sei $A \in \mathbb{K}^{m\times n}$. Die lineare Hülle der Zeilen bzw. Spalten von A heißt* Zeilenraum *bzw.* Spaltenraum *von A; die Dimension dieser Räume heißt* Zeilenrang *bzw.* Spaltenrang *von A.*

Beispiel 2.4.2 *Für eine Matrix, deren erste Spalten die ersten s Vektoren der kanonischen Basis von $\mathbb{K}^m$ sind, während evtl. weitere Spalten nur noch Nullen enthalten, liest man ab dass ihr Zeilen- und Spaltenrang beide gleich s sind. Im Folgenden soll gezeigt werden, dass die Werte von Zeilen- und Spaltenrang für beliebige Matrizen immer übereinstimmen.*

Aufgabe 2.4.3 *Zeige: Für jede Matrix $A \in \mathbb{K}^{m\times n}$ ist der Spaltenrang von A höchstens gleich n, der Zeilenrang höchstens gleich m. Finde einfache Beispiele dafür, dass diese Abschätzungen scharf sind.*

Aufgabe 2.4.4 *Finde den Zeilen- und Spaltenrang der Einheitsmatrix.*

Aufgabe 2.4.5 *Zeige: Für $A \in \mathbb{K}^{m\times n}$ und $x \in \mathbb{K}^n$ ist $A\,x$ ein Vektor des Spaltenraums, also eine Linearkombination der Spalten von A. Finde selber eine analoge Aussage für die Zeilen von A.*

2.5 Elementare Operationen

Definition 2.5.1 *Sei $n \in \mathbb{N}$, $n \geq 2$. Jede der folgenden drei Arten von quadratischen n-reihigen Matrizen heißt eine* Elementarmatrix der Größe n:

(a) *Für $\lambda \in \mathbb{K} \setminus \{0\}$ und $1 \leq j \leq n$ sei $E_j(\lambda)$ die Diagonalmatrix, deren j-tes Diagonalelement gleich λ ist, während die übrigen Diagonalelemente gleich 1 sind.*

(b) *Für $1 \leq j < k \leq n$ sei $P_{jk} = P_{kj}$ die Matrix, welche aus der Einheitsmatrix durch Vertauschen der j-ten und k-ten Spalte entsteht.*

(c) *Für $1 \leq j, k \leq n$, $j \neq k$, sei E_{jk} die Matrix mit Einsen überall auf der Diagonalen und einer zusätzlichen Eins in der Position (j, k), sowie Nullen in allen übrigen Positionen.*

Aufgabe 2.5.2 *Seien $n \geq 2$ und $A \in \mathbb{K}^{n\times m}$. Zeige, dass die Multiplikation von links mit Elementarmatrizen der Größe n folgenden Zeilenoperationen für A entspricht:*

(a) *$E_j(\lambda)\,A$ entsteht aus A durch Multiplikation der Elemente in der j-ten Zeile mit dem Faktor λ.*

(b) *$P_{jk}\,A$ entsteht aus A durch Vertauschen der Zeilen Nr. j und k.*

(c) *$E_{jk}\,A$ entsteht aus A durch Addition der Zeile Nr. k zur Zeile Nr. j.*

Benutze Regel (h) *aus* Satz 2.3.3*, um herauszufinden, welche Wirkung eine Multiplikation von rechts mit Elementarmatrizen der Größe m hat.*

Definition 2.5.3 *Für eine Matrix A heißt jede der Operationen, welche nach* Aufgabe 2.5.2 *der Multiplikation von links bzw. rechts mit einer Elementarmatrix passender Größe entspricht, eine* elementare Zeilen- bzw. Spaltenoperation *für A. Beachte, dass jede solche Operation umkehrbar ist. Wenn eine Matrix B aus A durch endlich viele dieser elementaren Zeilen- und Spaltenoperationen entsteht, dann heißt B* äquivalent zu A.

Aufgabe 2.5.4 *Wir wollen im Folgenden jede Hintereinanderausführung von endlich vielen elementaren Zeilen- oder Spaltenoperationen als erlaubt bezeichnen. Überprüfe durch eine Kombination von Operationen des Typs* (a),(c)*, dass auch die Addition eines Vielfachen einer Zeile bzw. Spalte zu einer anderen Zeile bzw. Spalte eine erlaubte Operation ist.*

Aufgabe 2.5.5 (Elementare Operationen mit MAPLE) *Finde selber die Befehle in* MAPLE*, welche den oben eingeführten elementaren Zeilen- und Spaltenoperationen entsprechen.*

Definition 2.5.6 (Äquivalenzrelationen) *Sei eine nicht-leere Menge X gegeben. Eine Teilmenge $R \subset X \times X$ heißt* eine Relation *auf X. Wir sagen dass ein $x_1 \in X$ zu einem $x_2 \in X$* in Relation steht*, wenn das Paar (x_1, x_2) zu R gehört. Eine solche Relation auf X heißt eine* Äquivalenzrelation*, falls für beliebige $x, x_1, x_2, x_3 \in X$ gilt:*

$$
\begin{array}{lll}
(\mathrm{R}) & (x,x) \in R & (\text{Reflexivität}) \\
(\mathrm{S}) & (x_1,x_2) \in R \implies (x_2,x_1) \in R & (\text{Symmetrie}) \\
(\mathrm{T}) & (x_1,x_2) \in R \text{ und } (x_2,x_3) \in R \implies (x_1,x_3) \in R & (\text{Transitivität})
\end{array}
$$

Statt $(x_1, x_2) \in R$ schreiben wir auch $x_1 \sim x_2$ und sagen in Worten: x_1 ist äquivalent zu *x_2. Für $x \in X$ sei $A_x = \{\tilde{x} \in X : \quad x \sim \tilde{x} \}$. Wir nennen ein solches A_x eine* Äquivalenzklasse*. Ein beliebiges Element einer Äquivalenzklasse heißt auch* ein Repräsentant *dieser Äquivalenzklasse.*

Aufgabe 2.5.7 *Zeige: Ist auf $X \neq \emptyset$ eine Äquivalenzrelation gegeben, so bilden die Äquivalenzklassen A_x eine* Zerlegung von *X; d. h., aus $A_x \cap A_y \neq \emptyset$ folgt $A_x = A_y$, und $\cup_{x \in X} A_x = X$.*

Aufgabe 2.5.8 *Zeige: Auf der Menge $\mathbb{K}^{m \times n}$ aller Matrizen fester Größe hat der oben eingeführte Begriff der Äquivalenz die drei Eigenschaften einer Äquivalenzrelation.*

Satz 2.5.9 *Der Spaltenrang einer Matrix ist invariant gegenüber elementaren Spaltenoperationen; der Zeilenrang einer Matrix ist invariant gegenüber elementaren Zeilenoperationen.*

Beweis: Es genügt, die Aussage für Spalten zu beweisen, da dann durch Übergang zur transponierten Matrix auch die für Zeilen folgt. Sei A eine Matrix mit den Spalten $a_1, \ldots, a_n \in \mathbb{K}^m$, und sei $U = \mathcal{L}(a_1, \ldots, a_n)$ der Spaltenraum von A. Das Vertauschen von Vektoren in $(a_1, \ldots, a_n)$ ändert nichts daran, dass diese Spaltenvektoren ein Erzeugendensystem für U sind, und gleiches gilt für Multiplikation einer Spalte mit einem von 0 verschiedenen Faktor. Wir wollen nun dasselbe zeigen für den Fall der Ersetzung der Spalte a_k durch die Summe $a_k + a_j$, für $j \neq k$: Jeder Vektor $u \in U$ kann dargestellt werden als $u = \sum_{\nu=1}^n \alpha_\nu \, a_\nu$. Setzt man

$$
\beta_\nu \; = \; \begin{cases} \alpha_\nu & (\nu \neq j) \\[1ex] \alpha_j - \alpha_k & (\nu = j) \end{cases}
$$

so folgt

$$
\beta_k \, (a_k + a_j) \; + \; \sum_{\nu \neq k} \beta_\nu \, a_\nu \; = \; \alpha_k \, (a_k + a_j) + (\alpha_j - \alpha_k) \, a_j + \sum_{\nu \neq j,k} \alpha_\nu \, a_\nu \; = \; u \, .
$$

Dies war zu zeigen. □

Lemma 2.5.10 *Seien A, B so, dass das Matrixprodukt $B\,A$ definiert ist. Dann ist der Spaltenrang von $B\,A$ nicht größer als der von A.*

Beweis: Sei $C = B\,A$. Sind $a_1, \dots, a_n$ bzw. $c_1, \dots, c_n$ die Spalten von A bzw. C, so gilt nach Aufgabe 2.2.2 $c_k = B\,a_k$ für $1 \le k \le n$. Nach den Rechenregeln für Matrizen folgt aus $0 = \sum_{k=1}^n \lambda_k\, a_k$ die Gleichung $0 = \sum_{k=1}^n \lambda_k\, c_k$. Daraus schließen wir: *Wenn ein Teilsystem der Spalten von A linear abhängig ist, so ist dasselbe Teilsystem der Spalten von C ebenfalls linear abhängig.* Also kann der Spaltenrang von C nicht größer sein als der von A. □

Aufgabe 2.5.11 *Gib ein Beispiel für Matrizen A, B, für die der Spaltenrang von $B\,A$ echt kleiner als der von A ist.*

Satz 2.5.12 *Der Spaltenrang einer Matrix ist invariant gegenüber elementaren Zeilenoperationen; der Zeilenrang einer Matrix ist invariant gegenüber elementaren Spaltenoperationen.*

Beweis: Zeilenoperationen entsprechen der Multiplikation von links mit Elementarmatrizen. Also kann der Spaltenrang wegen Lemma 2.5.10 bei Zeilenoperationen nicht zunehmen. Da Zeilenoperationen reversibel sind, folgt sogar, dass der Spaltenrang bei elementaren Zeilenoperationen gleich bleibt. Die Aussage für den Zeilenrang folgt wieder durch die Betrachtung der transponierten Matrizen. □

2.6 Normalform und Rang einer Matrix

Satz 2.6.1 (Normalform unter Äquivalenz) *Jede Matrix $A \in \mathbb{K}^{m\times n}$ ist äquivalent zu einer Matrix der Form*

$$B \;=\; \left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right],$$

wobei I eine Einheitsmatrix der Größe s mit $0 \le s \le \min\{m, n\}$ ist, während die Nullen für Nullmatrizen entsprechender Größe stehen; dabei ist zu beachten, dass $s = 0$ so zu interpretieren ist, dass B die Nullmatrix ist, während für $s = n$ bzw. $s = m$ einige der Nullmatrizen in B als leer anzusehen sind.

Beweis: Die Umformung von A auf die Form B geschieht durch Anwenden des folgenden Algorithmus, in dem die nach jeder Umformung enstandene Matrix der Einfachheit halber weiterhin mit A bezeichnet sein soll:

(a) Falls A die Nullmatrix ist, ist $B = A$, also $s = 0$, und der Algorithmus ist beendet.

(b) Falls A nicht die Nullmatrix ist, so können wir durch evtl. Vertauschen von Zeilen und/oder Spalten erreichen, dass anschließend $a_{11} \ne 0$ ist. Danach können wir die erste Zeile oder Spalte mit einem Faktor multiplizieren, so dass sogar $a_{11} = 1$ gilt.

(c) Im Fall $a_{11} = 1$ können wir Vielfache der ersten Zeile von allen folgenden subtrahieren, so dass danach $a_{j1} = 0$ ist für alle $j = 2, \dots, m$. Danach können wir Vielfache der ersten Spalte von allen folgenden subtrahieren, so dass danach $a_{1k} = 0$ ist für alle $k = 2, \dots, n$. Beachte, dass dabei die Elemente der ersten Spalte nicht mehr verändert werden.

(d) Falls $m = 1$ oder $n = 1$ ist, ist der Algorithus beendet. Sonst wenden wir dieselben Schritte sinngemäß auf die kleinere Matrix

$$\begin{bmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m2} & \dots & a_{mn} \end{bmatrix}$$

an; diese Schritte verändern die erreichte Form der ersten Zeile und Spalte von A nicht.

Beachte für das Folgende, dass dieser Algorithmus auch die Berechnung von Zeilen- und Spaltenrang einer Matrix gestattet. □

Korollar zu Satz 2.6.1 *Der Zeilenrang und der Spaltenrang einer Matrix A stimmen immer überein.*

Beweis: Zeilen- und Spaltenrang sind beide invariant unter Zeilen-, aber auch unter Spaltenoperationen, und durch diese kann A auf die Form B wie im Satz 2.6.1 gebracht werden. Nach Beispiel 2.4.2 stimmen Zeilen- und Spaltenrang für B überein, und somit gilt dasselbe auch für A. □

Definition 2.6.2 (Rang einer Matrix) *Der gemeinsame Wert von Zeilen- und Spaltenrang einer Matrix wird im Folgenden als* Rang *der Matrix bezeichnet, und wir schreiben auch* $\operatorname{rang} A$ *für den Rang einer Matrix A.*

Proposition 2.6.3 (Rechenregeln für den Rang)

(a) $\forall\, A \in \mathbb{K}^{m\times n}: \quad \operatorname{rang} A \ \le \min\{n, m\}.$

(b) $\forall\, A \in \mathbb{K}^{m\times n}: \quad \operatorname{rang} A \ = \ \operatorname{rang} A^T.$

(c) $\forall\, A \in \mathbb{K}^{m\times n},\ B \in \mathbb{K}^{n,r}: \quad \operatorname{rang}(A\,B) \ \le\ \min\{\operatorname{rang} A, \operatorname{rang} B\}.$

Beweis: Zu (a): Klar nach Definition von Rang bzw. Zeilen- und Spaltenrang. Zu (b): Der Spaltenrang von A ist gleich dem Zeilenrang von A^T und umgekehrt, und daraus folgt die Behauptung. Zu (c): In Lemma 2.5.10 wurde gezeigt, dass der Spaltenrang einer Matrix bei Multiplikation von links mit einer beliebigen anderen Matrix nicht zunehmen kann, und daher folgt $\operatorname{rang}(A\,B) \le \operatorname{rang} B$. Durch Betrachten der Transponierten und Anwenden von (b) folgt dann $\operatorname{rang}(A\,B) = \operatorname{rang}(A\,B)^T = \operatorname{rang}(B^T A^T) \le \operatorname{rang} A^T = \operatorname{rang} A$, und daraus folgt die Behauptung. □

2.7 Zeilenstufenform und Rangberechnung

Definition 2.7.1 *Sei A eine beliebige Matrix. Wir sagen: Die j-te Zeile von A* hat k führende Nullen, *wenn die ersten k Einträge in dieser Zeile gleich 0 sind und, im Falle dass die Länge der Zeile größer als k ist, der folgende Eintrag von Null verschieden ist. Dabei darf $k = 0$ sein, was dann bedeutet, dass das erste Element der j-ten Zeile nicht verschwindet. Die Matrix heißt* in Zeilenstufenform, *falls jede ihrer Zeilen entweder nur Nullen enthält oder* mehr *führende Nullen als die Vorgängerzeile hat.*

Bei einer Matrix in Zeilenstufenform kann es vorkommen, dass die letzten Zeilen nur noch Nullen enthalten. Streicht man diese, so sind die übrigen Zeilen linear unabhängig. Mit anderen Worten: Der Rang einer Matrix in Zeilenstufenform ist gleich der Anzahl der nicht-trivialen Zeilen der Matrix. Daher genügt es zur Berechnung des Ranges von A die Matrix durch Zeilenoperationen so lange umzuformen, bis eine Zeilenstufenform erreicht ist. Dies kann folgendermaßen geschehen:

1. Setze $j = 1$.

2. Finde das kleinste k für welches ein $\ell \ge j$ existiert mit $a_{\ell k} \ne 0$; falls kein solches k existiert, so brich ab. Sonst vertausche die Zeilen j und ℓ, sodass danach $a_{jk} \ne 0$ ist. Dies bedeutet, dass die Zeile j genau, und die Zeilen mit höherer Nummer mindestens $k - 1$ führende Nullen enthalten.

3. Subtrahiere Vielfache der j-ten von allen folgenden Zeilen, um zu erreichen, dass diese Folgezeilen alle mindestens eine führende Null mehr besitzen als die Zeile Nr. j.

4. Erhöhe j um 1 und fahre fort mit Schritt 2.

Aufgabe 2.7.2 *Berechne die Ränge der Matrizen*

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \\ 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Aufgabe 2.7.3 (Rangberechnung mit MAPLE) *Führe die nachfolgenden* MAPLE-*Befehle aus:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3,5, [[1,0,0,1,-1], [1,1,-1,2,0], [2,-1,-1,0,0]]);
> Rank(A);
```

Interpretiere das Resultat, und variiere die Sequenz so, dass sie die Ränge der Matrizen aus der vorangegangenen Aufgabe löst.

2.8 Invertierbare Matrizen

Definition 2.8.1 *Eine quadratische n-reihige Matrix A heißt* invertierbar*, falls es eine quadratische n-reihige Matrix B gibt, für welche $B\,A = I$ ist. Die Menge aller invertierbaren n-reihigen Matrizen wird mit $GL(n, \mathbb{K})$ bezeichnet. Diese Bezeichnung ist die Abkürzung für die englische Bezeichnung „general linear group“.*

Satz 2.8.2 *Für jede quadratische n-reihige Matrix A gilt:*

(a) *A ist invertierbar genau dann, wenn $\operatorname{rang} A = n$ ist, also wenn alle Spalten, oder alle Zeilen, von A linear unabhängig sind.*

(b) *Sei B so, dass $B\,A = I$ gilt. Dann gilt auch $A\,B = I$.*

(c) *Aus $B\,A = I$ und $C\,A = I$ folgt $B = C$.*

Beweis: Zu (a): Falls $I = B\,A$ gilt, folgt mit Satz 2.6.3 (c) und der Definition des Rangs, dass $n = \operatorname{rang} I \le \min\{\operatorname{rang} B, \operatorname{rang} A\} \le n$, und somit ist $n = \operatorname{rang} A = \operatorname{rang} B$. Umgekehrt: Ist $\operatorname{rang} A = n$, so sind die Zeilen von A linear unabhängig und deshalb eine Basis im Raum aller Zeilenvektoren der Länge n (denn dieser Raum ist nach Satz 2.1.2 ein n-dimensionaler Vektorraum über $\mathbb{K}$). Also ist jeder Zeilenvektor eine Linearkombination der Zeilen von A, und deshalb gibt es Zeilenvektoren z_j (der Länge N) mit $e_j^T = z_j\,A$, $j = 1, \ldots, n$. Fasst man diese Zeilenvektoren zu einer Matrix B zusammen, so folgt $I = B\,A$. Zu (b): Im Beweis von (a) wurde bereits gezeigt, dass aus $I = B\,A$ folgt $\operatorname{rang} B = n$. Daher ist auch B nach (a) invertierbar, und somit gibt es ein C mit $C\,B = I$. Durch Multiplikation von rechts mit A folgt hieraus $A = C\,B\,A = C$. Zu (c): Aus $I = B\,A = C\,A$ folgt durch Multiplikation von rechts mit B unter Beachtung von $A\,B = I$, dass $B = C\,(A\,B) = C$ ist. □

Aufgabe 2.8.3 *Aus* Satz 2.8.2 *und* Satz 2.6.1 *folgt: Eine Matrix A ist genau dann invertierbar, wenn sie äquivalent zur Einheitsmatrix ist. Schließe hieraus, dass eine invertierbare Matrix ein Produkt von endlich vielen Elementarmatrizen ist.*

Definition 2.8.4 *Nach* Satz 2.8.2 *gibt es zu jedem $A \in GL(n, \mathbb{K})$ eine* eindeutige *Matrix B mit $B\,A = I$, und wir nennen dieses B die* inverse Matrix *zu A oder kürzer* Inverse *von A und schreiben auch A^{-1} für diese Matrix.*

Satz 2.8.5 (Rechenregeln für Inverse)

(a) $\forall\ A \in GL(n, \mathbb{K}): \quad A\,A^{-1} = A^{-1}\,A = I.$

(b) $A \in GL(n, \mathbb{K}) \implies A^T \in GL(n, \mathbb{K})$, *und* $(A^T)^{-1} = (A^{-1})^T$.

(c) $A, B \in GL(n, \mathbb{K}) \implies A\,B \in GL(n, \mathbb{K})$, *und* $(A\,B)^{-1} = B^{-1}\,A^{-1}$.

Beweis: Zu (a): Folgt aus Satz 2.8.2 (b). Zu (b): Folgt aus $(A^{-1})^T\,A^T = (A\,A^{-1})^T = I^T = I$, wobei Satz 2.3.3 (h) benutzt wurde. Zu (c): Folgt aus $(B^{-1}\,A^{-1})\,(A\,B) = B^{-1}\,(A^{-1}\,A)\,B = B^{-1}\,B = I$. □

Aufgabe 2.8.6 *Zeige: Wenn A invertierbar ist, dann ist auch A^{-1} invertierbar, und die inverse Matrix von A^{-1} ist gleich A.*

Kapitel 3

Lineare Gleichungssysteme

3.1 Das Gaußsche Eliminationsverfahren

Definition 3.1.1 *Im Folgenden seien stets eine Matrix*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathbb{K}^{m \times n}$$

und ein Vektor $b = (b_1, b_2, \dots, b_m)^T \in \mathbb{K}^m$ *fest gewählt. Für einen weiteren Vektor* $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{K}^n$ *ist die Gleichung* $A\,x = b$ *äquivalent mit den* m linearen Gleichungen in n Unbekannten $x_1, \dots, x_n$

$$\begin{aligned} a_{11}\,x_1 &+ a_{12}\,x_2 + \dots + a_{1n}\,x_n &= b_1\,, \\ a_{21}\,x_1 &+ a_{22}\,x_2 + \dots + a_{2n}\,x_n &= b_2\,, \\ \vdots & \qquad\qquad\qquad\qquad \vdots & \vdots \\ a_{m1}\,x_1 &+ a_{m2}\,x_2 + \dots + a_{mn}\,x_n &= b_m\,. \end{aligned}$$

Deshalb nennen wir $A\,x = b$ *auch ein* (inhomogenes) lineares Gleichungssystem*. Die Matrix* A *heißt auch* Koeffizientenmatrix *des Gleichungssystems, und* b *heißt* Inhomogenitätenvektor*. Die Gleichung* $A\,x = 0$ *heißt die* zugehörige homogene Gleichung *oder* homogenes Gleichungssystem*. Wir schreiben auch* $L(A, b)$ *für die Menge aller Vektoren* x*, welche die Gleichung* $A\,x = b$ *erfüllen, und* $L(A)$ *für die Lösungsmenge der homogenen Gleichung* $A\,x = 0$*. Jedes* $x \in L(A) \setminus \{0\}$ *heißt eine* nicht-triviale *Lösung des homogenen Gleichungssystems. Wir nennen das inhomogene System* universell lösbar*, wenn es für jedes* $b \in \mathbb{K}^m$ *ein* $x \in L(A, b)$ *gibt. Wir sagen, dass das inhomogene System für ein* $b \in \mathbb{K}^m$ eindeutig lösbar *ist, wenn es genau ein* $x \in L(A, b)$ *gibt. Die Matrix*

$$(A, b) = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]$$

heißt die erweiterte Matrix *des Gleichungssystems* $A\,x = b$*.*

Bemerkung 3.1.2 (Gaußsches Eliminationsverfahren) *Es ist leicht einzusehen, dass die Lösungsmenge* $L(A, b)$ *bei elementaren Zeilenoperationen für die erweiterte Matrix unverändert bleibt. Da z. B. das Vertauschen zweier Spalten von* A *einer Vertauschung der entsprechenden Unbekannten entspricht,* *wollen wir uns beim Berechnen der Lösungen von $Ax = b$ auf Zeilenoperationen beschränken; beachte aber, dass diese immer auf die erweiterte Matrix anzuwenden sind. Wir gehen dabei folgendermaßen vor:*

1. Falls $A = 0$ ist, breche ab. Sonst weiter mit 2.

2. Suche die Spalte von A mit der niedrigsten Nummer, welche nicht nur Nullen enthält; ihre Nummer sei gleich k, und j sei so, dass $a_{jk} \neq 0$ ist. Vertausche in der erweiterten Matrix die Zeile Nr. j mit der ersten. Subtrahiere danach Vielfache der ersten von allen folgenden Zeilen, um in der k-ten Spalte alle Elemente mit Nummern (j, k), $2 \leq j \leq m$, zu 0 zu machen. Ist dies geschehen, hat die neue erweiterte Matrix die Form

$$\left[\begin{array}{ccc|c|c|c} 0 & \dots & 0 & \tilde{a}_{1k} & \tilde{a}_{1,k+1} \dots \tilde{a}_{1n} & \tilde{b}_1 \\ \hline 0 & \dots & 0 & 0 & & \\ \vdots & & \vdots & \vdots & \tilde{A} & \tilde{b} \\ 0 & \dots & 0 & 0 & & \end{array}\right],$$

   wobei $\tilde{a}_{1k} \neq 0$ ist. Danach wenden wir dieselben Schritte 1. und 2. auf die kleinere Matrix $(\tilde{A}, \tilde{b})$ wieder an.

Das oben beschriebene Verfahren führt in endlich vielen Schritten zu einer Endmatrix $(\hat{A}, \hat{b})$ in Zeilenstufenform, und wir können folgendes ablesen:

(a) Da wir nur Zeilenoperationen benutzt haben, ist $\operatorname{rang} A = \operatorname{rang} \hat{A}$ und $\operatorname{rang}(A, b) = \operatorname{rang}(\hat{A}, \hat{b})$.

(b) Eventuell enthalten einige der untersten Zeilen von $(\hat{A}, \hat{b})$ nur noch Nullen; diese sind für die Bestimmung der Lösungsmenge ohne Bedeutung und können deshalb weggelassen werden. Die Anzahl der verbliebenen Gleichungen ist dann gleich $\operatorname{rang}(A, b)$.

(c) Wenn danach in der untersten Zeile von $(\hat{A}, \hat{b})$ nur rechts im Inhomogenitätenteil eine von 0 verschiedene Zahl steht, ergibt sich ein Widerspruch und die Lösungsmenge ist deshalb leer. Dies geschieht offenbar genau dann, wenn der Rang von A kleiner als der Rang der erweiterten Matrix (A, b) ist.

(d) Wenn $\operatorname{rang} A = \operatorname{rang}(A, b))$ ist, ist in der letzten Zeile von $\hat{A}$ mindestens ein Element von 0 verschieden. Wir können daher in der untersten Gleichung alle der Unbekannten bis auf eine beliebig wählen und dann die Gleichung nach der verbleibenden Unbekannten auflösen. Setzt man die erhaltenen Unbekannten in die darüberstehenden Gleichungen ein, kann man danach mit der vorausgehenden Gleichung genauso verfahren.

(e) Da die Anzahl der so zu lösenden Gleichungen gleich $\operatorname{rang} A = \operatorname{rang}(A, b))$ ist, können wir insgesamt für $n - \operatorname{rang} A$ Unbekannte beliebige *freie Parameter* einsetzen, während die übrigen bestimmt sind, aber natürlich von der Wahl dieser Parameter abhängen.

Insgesamt kann man auf diese Weise die allgemeine Lösung des Gleichungssystems ermitteln. Sie kann leer sein oder nur aus einem Vektor bestehen, kann aber auch einen oder mehrere freie Parameter enthalten. *Wir wollen dies noch in einigen Beispielen deutlicher sehen.*

Aufgabe 3.1.3 *Löse*

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & - & x_3 & = & 0\,, \\ & & x_2 & + & x_3 & = & 1\,, \\ x_1 & - & x_2 & - & x_3 & = & -2\,. \end{array}$$

Aufgabe 3.1.4 *Löse*

$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & - & x_3 & = & 0\,, \\ & & x_2 & + & x_3 & = & 1\,, \\ x_1 & - & x_2 & - & 2x_3 & = & -2\,. \end{array}$$

Aufgabe 3.1.5 *Löse*

$$\begin{array}{rcrcrcr} 2\,x_1 & + & x_2 & - & x_3 & = & 0\,, \\ & & x_2 & + & x_3 & = & 1\,, \\ x_1 & - & x_2 & - & 2\,x_3 & = & -3/2\,. \end{array}$$

Aufgabe 3.1.6 (Lösen von linearen GLS mit MAPLE) *Führe folgende* MAPLE-*Kommandos aus:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3,5, [[1,0,0,1,-1], [1,1,-1,2,0], [2,-1,-1,0,0]]);
> b := Vector(3, [0,1,0]);
> LinearSolve(A,b);
```

Interpretiere das Resultat, und variiere die Sequenz so, dass sie die linearen Gleichungssysteme aus den vorangegangen Aufgaben löst.

3.2 Struktur der Lösungsmenge

Wir wissen bereits, wie man ein lineares Gleichungssystem löst, wollen aber nun allgemeine Aussagen über die Lösungsmenge machen.

Satz 3.2.1 (Struktur der Lösungsmenge) *Für A und b wie in obiger Definition gilt:*

(a) $x_p \in L(A,b)\,,\quad x_h \in L(A) \quad\Longrightarrow\quad x_p \;+\; x_h \in L(A,b).$

(b) $x_1, x_2 \in L(A,b) \quad\Longrightarrow\quad x_1 - x_2 \in L(A).$

(c) $x_1, x_2 \in L(A),\ \alpha, \beta \in \mathbb{K} \quad\Longrightarrow\quad \alpha\, x_1 + \beta\, x_2 \in L(A).$

In anderen Worten bedeutet das: Die Menge $L(A)$ ist ein Unterraum von $\mathbb{K}^n$, und falls $L(A,b) \neq \emptyset$ ist, gilt für jedes $x_0 \in L(A,b)$, dass $L(A,b) = x_0 + L(A)$ ist. Dies wiederum heißt, dass $L(A,b)$ eine lineare Mannigfaltigkeit ist und die gleiche Dimension wie der Unterraum $L(A)$ hat.

Beweis: Zu (a): $A\,(x_p + x_h) = A\,x_p + A\,x_h = b + 0 = b$. Zu (b): $A\,(x_1 - x_2) = A\,x_1 - A\,x_2 = b - b = 0$. Zu (c): $A\,(\alpha\,x_1 \;+\; \beta\,x_2) \;=\; \alpha\,A\,x_1 \;+\; \beta\,A\,x_2 \;=\; 0.$ □

Satz 3.2.2 *Es ist stets* $\dim L(A) = n - \operatorname{rang} A$.

Beweis: In Kapitel 6 werden wir sehen, dass die Matrix A eine sogenannte lineare Abbildung von von $\mathbb{K}^n$ nach $\mathbb{K}^m$ definiert, und ihr Rang ist gleich der Dimension des Bildes dieser Abbildung; deshalb folgt die Behauptung aus der allgemeinen Gleichung in Satz 6.2.2 (b). Man kann aber die Gültigkeit der Behauptung auch aus dem Gaußschen Eliminationsverfahren zur Berechnung der Lösungen ablesen: Die allgemeine Lösung des homogenen Systems enthält $s = n - \operatorname{rang} A$ freie Parameter. Wählt man einen davon gleich 1 und die übrigen gleich 0, so erhält man s linear unabhängige Vektoren $x_1, \ldots, x_2 \in L(A)$. Die allgemeine Lösung von $A\,x = 0$ ist eine Linearkombination dieser Lösungen, also ist das System $(x_1, \ldots, x_s)$ sogar eine Basis von $L(A)$. □

Satz 3.2.3 *Ein inhomogenes lineares Gleichungssystem ist genau dann lösbar, wenn der Rang der Koeffizientenmatrix gleich dem der erweiterten Matrix ist.*

Beweis: Dies folgt bereits aus dem Gaußschen Eliminationsverfahren, kann aber auch folgendermaßen bewiesen werden: Klar ist, dass $\operatorname{rang} A \leq \operatorname{rang}(A, b)$ ist. Falls ein $x \in \mathbb{K}^n$ existiert, für welches $A\,x = b$ ist, dann ist b eine Linearkombination der Spalten von A, und deshalb kann der Rang von (A, b) nicht größer als der von A sein. Umgekehrt, falls $\operatorname{rang} A = \operatorname{rang}(A, b)$ ist, dann muss b eine Linearkombination der Spalten von A sein, und deshalb gibt es ein $x \in \mathbb{K}^n$ mit $A\,x = b$. □

Satz 3.2.4 *Es gelten folgende Aussagen:*

(a) *Genau dann hat das homogene Gleichungssystem eine nicht-triviale Lösung, wenn* $\operatorname{rang} A < n$ *ist.*

(b) *Genau dann ist das inhomogene Gleichungssystem universell, also für jedes* $b \in \mathbb{K}^m$ *lösbar, wenn* $\operatorname{rang} A = m$ *ist.*

(c) *Genau dann ist das inhomogene Gleichungssystem für jedes* $b \in \mathbb{K}^m$ *eindeutig lösbar, wenn* $\operatorname{rang} A = n = m$ *ist, d. h., wenn* A *quadratisch und invertierbar ist. Ist dies der Fall, so ist die eindeutige Lösung von* $A\,x = b$ *gegeben als* $x = A^{-1}\,b$.

Beweis: Folgt aus den vorangegangenen Resultaten. □

3.3 Berechnen der inversen Matrix

Die Berechnung der inversen Matrix einer Matrix $A \in \mathbb{K}^{n \times n}$ besteht in dem Auffinden einer Matrix X derselben Größe, für welche $A\,X = I$ ist. Dies gilt nach Aufgabe 2.2.2 genau dann, wenn für $k = 1, \ldots, n$ die k-te Spalte von X das lineare Gleichungssystem $A\,x = e_k$ löst, wobei e_k den k-ten Vektor der kanonischen Basis von $\mathbb{K}^n$ bezeichnet. Also können die Spalten von X mit dem Gaußschen Eliminaionsverfahren berechnet werden. Dies geschieht am effektivsten so, dass man die Matrix $(A, I) \in \mathbb{K}^{n \times 2n}$ durch elementare Zeilentransformationen solange umformt, bis die Form (I, X) erreicht ist; dies lässt sich genau dann erreichen, wenn A tatsächlich invertierbar ist. Ist diese Form dann erzielt, so ist die Matrix X gerade die Inverse zu A. Um einzusehen, dass dies richtig ist, beachte dass die Lösungsmenge einer Matrixgleichung der Form $A\,X = B$ mit Matrizen A, X, B passender Größe bei elementaren Zeilenumformungen von (A, B) nicht geändert wird.

Aufgabe 3.3.1 *Zeige: Eine zweireihige quadratische Matrix*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

mit $\Delta = a\,d - b\,c \neq 0$ *ist invertierbar, und es gilt*

$$A^{-1} = \frac{1}{\Delta} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Finde eine einfache Weise, sich diese Formel einzuprägen. Wir werden im nächsten Kapitel sehen, dass die Zahl $a\,d - b\,c$ *gerade die Determinante von* A *ist. Beachte aber, dass es für größere Matrizen keine so elementare Formel für die inverse Matrix gibt.*

Aufgabe 3.3.2 *Berechne die inverse Matrix zu*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Aufgabe 3.3.3 (Inverse Matrix mit MAPLE) *Die folgende Sequenz von* MAPLE-*Kommandos berechnet inverse Matrizen:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(5, 5, [ [1, 0, 0, 1, -1], [1, 1, -1, 2, 0], [2, -1, -1, 0, 0], [-1, 1, 0, 2, 1], [-1, 2, 3, 1, 0] ]);
> MatrixInverse(A);
> B := Matrix(2, 2, [ [1, 1], [-1, -1] ]);
> MatrixInverse(B);
```

Interpretiere das Resultat, und variiere die Sequenz so, dass sie die Inverse der Matrix aus der vorangegangenen Aufgabe berechnet.

Aufgabe 3.3.4 *Zeige: Eine obere bzw. untere Dreiecksmatrix A ist genau dann invertierbar, wenn ihre Diagonalelemente alle von* 0 *verschieden sind, und dann ist die inverse Matrix ebenfalls eine obere bzw. untere Dreiecksmatrix. Die Diagonalelemente von A^{-1} sind dabei gerade die Kehrwerte der Diagonalelemente von A, und die Außerdiagonalelemente können rekursiv, d. h., nacheinander, berechnet werden.*

Aufgabe 3.3.5 (Dreiecksmatrixinversion mit MAPLE) *Die folgende Sequenz*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(5, 5, [ [a_11, a_12, a_13, a_14, a_15], [0, a_22, a_23, a_24, a_25], [0, 0, a_33,
  a_34, a_34], [0, 0, 0, a_44, a_45], [0, 0, 0, 0, a_55] ]);
> MatrixInverse(A);
```

veranlasst MAPLE, *die Inverse einer Dreiecksmatrix mit fünf Zeilen und Spalten formelmäßig anzugeben. Wo liegt hier eigentlich eine Schwäche von* MAPLE *versteckt?*

Kapitel 4

Determinanten

4.1 Gruppen und Permutationen

Definition 4.1.1 *Eine Menge G, zusammen mit einer Abbildung, auch Verknüpfung oder Operation genannt,*

$$\circ \ : \ G \times G \longrightarrow G, \qquad (g_1, g_2) \longmapsto g_1 \circ g_2,$$

heißt eine Gruppe*, wenn folgende Axiome gelten:*

(G1) $\forall\, g_1, g_2, g_3 \in G : \quad g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$ (Assoziativgesetz)

(G2) $\exists\, e \in G \quad \forall\, g \in G : \quad e \circ g = g$ (Existenz eines Einselements)

(G3) $\forall\, g \in G \quad \exists\, \tilde{g} \in G : \quad \tilde{g} \circ g = e$ (Existenz eines Inversen)

Da in den Axiomen das Einselement bzw. das Inverse auf der linken Seite steht, spricht man auch genauer von einer Linkseins und einem Linksinversen. Falls zusätzlich noch gilt

(G4) $\forall\, g_1, g_2 \in G : \quad g_1 \circ g_2 = g_2 \circ g_1$ (Kommutativgesetz)

dann sprechen wir von einer kommutativen *oder* abelschen *Gruppe. Die Anzahl der Elemente einer Gruppe G heißt auch die* Ordnung *der Gruppe.*

Eine bijektive Abbildung $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ heißt Permutation *der Zahlen $1, \dots, n$. Die Menge aller Permutationen von $1, \dots, n$ heißt die* symmetrische Gruppe S_n*. Für $\sigma_1, \sigma_2 \in S_n$ heißt die Hintereinanderausführung $\sigma_1 \circ \sigma_2$ auch* das Produkt *von σ_1, σ_2; beachte genau die Reihenfolge, da die Hintereinanderausführung von Abbildungen im Allgemeinen nicht kommutativ ist.*

Bemerkung 4.1.2 *Man kann zeigen, dass in jeder Gruppe eine Linkseins immer eine Rechtseins ist, d. h., aus* (G2) *folgt $g \circ e = g$ für alle $g \in G$. Genauso ist ein Linksinverses stets auch Rechtsinverses. Außerdem gibt es nur ein Einselement und zu jedem $g \in G$ nur ein Inverses. Der Beweis soll hier ausgelassen werden, wird aber in der Vorlesung* Algebra *behandelt. Für das Inverse zu g schreibt man auch g^{-1}. Man fasst also die Verknüpfung in einer Gruppe üblicherweise als eine Art von Multiplikation auf und spricht deshalb von einer multiplikativen, oder einer multiplikativ geschriebenen Gruppe. Manchmal ist es aber sinnvoller, die Verknüpfung „$\circ$“ als Addition „$+$“ zu schreiben, was insbesondere dann häufig, aber leider nicht immer, geschieht, wenn die Gruppe abelsch ist.*

Aufgabe 4.1.3 *Zeige: In jeder Gruppe G gilt die Regel* $(g_1 \circ g_2)^{-1} = g_2^{-1} \circ g_1^{-1}$ *für alle* $g_1, g_2 \in G$.

Aufgabe 4.1.4 *Sei G eine Gruppe. Zeige:*

(a) *Für jedes* $g \in G$ *gilt* $(g^{-1})^{-1} = g$.

(b) *Für* $g_1, g_2 \in G$ *gilt* $g_1^{-1} = g_2^{-1} \iff g_1 = g_2$.

Aufgabe 4.1.5 *Sei I eine beliebige nicht-leere Menge. Zeige, dass die Menge aller Bijektionen von I auf sich bezgl. der Hintereinanderausführung eine Gruppe bilden. Zeige für den Fall, dass I mindestens drei Elemente hat, dass diese Gruppe nicht bijektiv ist. Bestimme für den Fall, dass I eine endliche Mange ist, die Anzahl der verschiedenen Bijektionen von I in sich, also die Ordnung der Gruppe aller Bijektionen.*

Beispiel 4.1.6 *Jeder Vektorraum ist bezgl. der Addition eine abelsche Gruppe. Die Menge* $GL(n, \mathbb{K})$ *der* n*-reihigen invertierbaren Matrizen ist eine Gruppe bezüglich der Matrixmultiplikation, und für* $n \geq 2$ *ist diese Gruppe nicht kommutativ. Die Menge* S_n *ist ebenfalls eine Gruppe bezüglich der Hintereinanderausführung, und für* $n \geq 3$ *ist sie nicht kommutativ; vergleiche hierzu* Aufgabe 4.1.5.

Beispiel 4.1.7 *Eine Permutation* σ *ist durch die Festlegung aller ihrer Bilder* $\sigma(1), \ldots, \sigma(n)$ *eindeutig festgelegt. Wir schreiben deshalb etwa*

$$\sigma \;=\; \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix} \;=\; \big(\sigma(1), \sigma(2), \ldots, \sigma(n)\big)\,.$$

Ist z. B. $n = 4$, *und wählt man*

$$\sigma_1 \;=\; \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad \sigma_2 \;=\; \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix},$$

so ist das Produkt von σ_1 *und* σ_2 *gleich*

$$\sigma_1 \circ \sigma_2 \;=\; \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

Satz 4.1.8 *In jeder Gruppe G gilt:*

(a) *Die Zuordnung* $g \longmapsto g^{-1}$ *ist eine bijektive Abbildung von G in sich selber.*

(b) *Ist ein* $g_0 \in G$ *fest gewählt, so ist auch die Abbildung* $g \longmapsto g_0 \circ g$ *bijetiv von G in G.*

Beweis: Zu (a): Unter Benutzung von Aufgabe 4.1.4 schließen wir wie folgt: Es gilt $(g^{-1})^{-1} = g$, woraus die Surjektivität folgt, und wegen $g_1^{-1} = g_2^{-1} \iff g_1 = g_2$ folgt die Injektivität. Zu (b): Es ist $\tilde{g} = g_0 \circ g$ genau dann, wenn $g = g_0^{-1} \circ \tilde{g}$ ist, und daher gilt die Behauptung. □

Aufgabe 4.1.9 *Zeige durch Induktion über n: Wenn für eine Permutation* $\sigma \in S_n$ *immer* $j \leq \sigma(j)$ *für alle* $j = 1, \ldots, n$ *gilt, dann folgt* $\sigma(j) = j$.

Aufgabe 4.1.10 (Permutationen und MAPLE) MAPLE *kann mit Permutationen arbeiten. Allerdings wird hier die* Darstellung durch elementfremde Zykeln *benutzt: Das Symbol* $[a_1, a_2, \ldots, a_s]$ *bedeutet hier eine Abbildung, welche* a_1 *auf* a_2, a_2 *auf* a_3 *u. s. w., und* a_s *wieder auf* a_1 *abildet. Wenn alle* a_k *verschiedene Zahlen aus* $\{1, \ldots, n\}$ *sind, dann nennt man* $[a_1, a_2, \ldots, a_s]$ *einen* Zykel. *Wenn man vereinbart, dass alle* $a \in \{1, \ldots, n\}$, *die in der Liste* $[a_1, a_2, \ldots, a_s]$ *nicht auftreten, auf sich selber abgebildet werden, dann legt jeder solche Zykel eine Permutation aus* S_n *fest. Man überlegt sich jetzt leicht, dass* *jede Permutation durch endlich viele* paarweise disjunkte *Zykeln definiert werden kann. Z. B. bedeutet in* MAPLE *das Symbol* $[[1,2,3],\ [4,5]]$ *die Permutation, welche die* 1 *auf die* 2, *die* 2 *auf die* 3, *die* 3 *auf die* 1, *die* 4 *auf die* 5, *und die* 5 *auf die* 4 *abbildet. Dies ist also ein Element von* S_5, *kann aber auch als Element von* S_n *mit* $n \geq 6$ *aufgefasst werden, da ja die Regel gilt, dass alle Zahlen* k, *die in den Zykeln nicht vorkommen, also hier alle* $k \geq 6$, *auf sich selber abgebildet werden. Die Befehle*

> restart;

> with(group);

> mulperms([[1,2,3], [4,5]], [[2,3], [1,5]]);

berechnen das Produkt der beiden angegebenen Permutationen.

4.2 Vorzeichen von Permutationen

Definition 4.2.1 *Sei* $\sigma \in S_n$. *Ein Zahlenpaar* (j,k), $1 \leq j < k \leq n$, *heißt eine* Inversion *oder ein* Fehlstand *von* σ, *wenn* $\sigma(j) > \sigma(k)$ *ist. Die Anzahl solcher Fehlstände sei* $m(\sigma)$. *Wir setzen das* Vorzeichen *oder* Signum $\operatorname{sgn}(\sigma)$ *gleich* 1, *bzw. gleich* -1, *wenn die Anzahl aller Inversionen von* σ *gerade, bzw. ungerade, ist. Anders ausgedrückt ist also* $\operatorname{sgn}(\sigma) = (-1)^{m(\sigma)}$ *Eine Permutation* $\tau \in S_n$ *heißt eine* Transposition, *falls ein Zahlenpaar* (j,k), $1 \leq j < k \leq n$, *existiert, welches durch* τ *vertauscht wird, während alle anderen Zahlen von* τ *festgelassen werden; d. h. genauer*

$$\tau(j) = k\,, \quad \tau(k) = j\,, \quad \tau(\ell) = \ell \quad \forall\ \ell \neq j,k\,,\ 1 \leq \ell \leq n\,.$$

Die identische Abbildung $id:\ \{1,\ldots,n\} \longrightarrow \{1,\ldots,n\}$ *heißt auch* identische Permutation.

Beispiel 4.2.2 *Die identische Permutation id besitzt keinen Fehlstand, also ist* $\operatorname{sgn}(id) = 1$. *Für* $\sigma = (1,2,4,3)$ *gibt es genau einen Fehlstand, nämlich das Paar* $(3,4)$. *Deshalb ist* $\operatorname{sgn}(\sigma) = -1$.

Aufgabe 4.2.3 (Vorzeichen und MAPLE) *Das Vorzeichen einer Permutation kann mit Hilfe von* MAPLE *indirekt berechnet werden, da für elementfremde Zykeln die sogenannte* Parität *gleich dem Vorzeichen ist und mit dem Kommando* > parity *berechnet werden kann. Überprüfe dies mit der Sequenz*

> restart;

> with(group);

> parity([[2,3], [1,5]]);

> parity([[1,2,3], [4,5]]);

Satz 4.2.4 (Regeln für das Vorzeichen) *Für alle* $n \geq 2$ *und alle* $\sigma, \tau \in S_n$ *gilt:*

(a) $\quad \operatorname{sgn}(\sigma) \;=\; \prod_{1 \leq j < k \leq n} \dfrac{\sigma(j) - \sigma(k)}{j-k}\;.$

(b) $\quad \forall\ \sigma_1, \sigma_2 \in S_n: \quad \operatorname{sgn}(\sigma_1 \circ \sigma_2) \;=\; \operatorname{sgn}(\sigma_1)\ \operatorname{sgn}(\sigma_2).$

(c) $\quad \forall\ \sigma \in S_n: \quad \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma).$

Beweis: Zu (a): Nach Definition von $m(\sigma)$ gilt

$$\begin{aligned}
\prod_{1\le j<k\le n} (\sigma(j)-\sigma(k)) &= \prod_{\substack{1\le j<k\le n\\ \sigma(j)>\sigma(k)}} (\sigma(j)-\sigma(k)) \prod_{\substack{1\le j<k\le n\\ \sigma(j)<\sigma(k)}} (\sigma(j)-\sigma(k))\\
&= (-1)^{m(\sigma)} \prod_{\substack{j>k\\ \sigma(j)<\sigma(k)}} (\sigma(j)-\sigma(k)) \prod_{\substack{j<k\\ \sigma(j)<\sigma(k)}} (\sigma(j)-\sigma(k))\\
&= (-1)^{m(\sigma)} \prod_{1\le\sigma(j)<\sigma(k)\le n} (\sigma(j)-\sigma(k))\\
&= (-1)^{m(\sigma)} \prod_{1\le j<k\le n} (j-k)\,,
\end{aligned}$$

woraus die Behauptung folgt. Zu (b): Unter Verwendung von (a) zeigt man

$$\begin{aligned}
\operatorname{sgn}(\sigma\circ\tau) &= \prod_{j<k} \frac{\sigma(\tau(j))-\sigma(\tau(k))}{\tau(j)-\tau(k)}\,\frac{\tau(j)-\tau(k)}{j-k}\\
&= \operatorname{sgn}(\tau) \prod_{j<k} \frac{\sigma(\tau(j))-\sigma(\tau(k))}{\tau(j)-\tau(k)}\\
&= \operatorname{sgn}(\tau) \prod_{\substack{j<k\\ \tau(j)<\tau(k)}} \frac{\sigma(\tau(j))-\sigma(\tau(k))}{\tau(j)-\tau(k)} \prod_{\substack{j<k\\ \tau(j)>\tau(k)}} \frac{\sigma(\tau(j))-\sigma(\tau(k))}{\tau(j)-\tau(k)}\\
&= \operatorname{sgn}(\tau) \prod_{\substack{j<k\\ \tau(j)<\tau(k)}} \frac{\sigma(\tau(j))-\sigma(\tau(k))}{\tau(j)-\tau(k)} \prod_{\substack{j>k\\ \tau(j)<\tau(k)}} \frac{\sigma(\tau(j))-\sigma(\tau(k)}{\tau(j)-\tau(k)}\\
&= \operatorname{sgn}(\tau) \prod_{\tau(j)<\tau(k)} \frac{\sigma(\tau(j))-\sigma(\tau(k)}{\tau(j)-\tau(k)} \;=\; \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)\,.
\end{aligned}$$

Zu (c): Mit (b) folgt $1=\operatorname{sgn}(id)=\operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1})$, und daraus folgt die Behauptung. □

Aufgabe 4.2.5 *Zeige: Alle Transpositionen $\tau\in S_n$ haben Vorzeichen -1, und es gilt $\tau^{-1}=\tau$.*

Aufgabe 4.2.6 *Berechne das Vorzeichen der Permutationen $\sigma_1=(2,1,3,4)$ und $\sigma_2=(2,3,4,1)$.*

Lösung: σ_1 hat nur den Fehlstand $(1,2)$, also ist $\operatorname{sgn}\sigma_1=-1$. Für σ_2 sind $(1,4)$, $(2,4)$ und $(3,4)$ Fehlstände, und deshalb ist $\operatorname{sgn}\sigma_2=-1$. □

Satz 4.2.7 *Zu jeder Permutation $\sigma\in S_n$ existieren endlich viele Transpositionen $\tau_1,\ldots,\tau_m\in S_n$ mit*

$$\sigma \;=\; \tau_1\circ\ldots\circ\tau_m\,.$$

Beweis: Wir benutzen Induktion über n: Für $n=1$ ist nichts zu zeigen. Sei jetzt $n\ge 2$, und sei $k=\sigma(n)$. Falls $k<n$ ist, gibt es eine Transposition τ, welche n und k vertauscht. Falls $k=n$ ist, sei $\tau=id$ gesetzt. Die Permutation $\tilde\sigma=\tau\circ\sigma$ erfüllt dann $\tilde\sigma(n)=n$ und kann deshalb auch als Permutation in S_{n-1} angesehen werden. Nach Induktonshypothese gibt es dann $\tau_1,\ldots,\tau_j\in S_{n-1}$ mit $\tilde\sigma\;=\;\tau_1\circ\ldots\circ\tau_j$. Wenn wir die τ_k zu Permutationen in S_n machen, indem wir $\tau_k(n)=n$ setzen, folgt die Behauptung. □

4.3 Definition der Determinante

Definition 4.3.1 *Für $n \in \mathbb{N}$ und $A \in \mathbb{K}^{n\times n}$ heißt die Zahl*

$$\det A \;=\; \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\; a_{1\sigma(1)}\, a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$$

die Determinante *von A.*

Beispiel 4.3.2 *Für $n = 1$ ist $A = [a]$ und $\det A = a$. Für $n = 2$ gibt es zwei Permutationen in S_2, nämlich*

$$\sigma_1 = id = (1,2)\,, \quad \sigma_2 \;=\; (2,1)\,.$$

Das Vorzeichen von σ_2 ist -1, und deshalb folgt

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \;= a_{11}\, a_{22} \;-\; a_{12}\, a_{21}\,.$$

Im Fall $n = 3$ hat S_3 sechs Elemente, nämlich

$$\sigma_1 = id = (1,2,3) \quad \sigma_2 = (2,3,1)\,, \quad \sigma_3 = (3,1,2)\,,$$

$$\sigma_4 = (1,3,2) \quad \sigma_5 = (2,1,3)\,, \quad \sigma_6 = (3,2,1)\,,$$

wobei die ersten bzw. letzten drei Permutationen das Vorzeichen 1 bzw. -1 haben. Deshalb ist

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \;=\; \begin{array}{l} a_{11}\, a_{22}\, a_{33} \;+\; a_{12}\, a_{23}\, a_{31} \;+\; a_{13}\, a_{21}\, a_{32} \\ -\, a_{11}\, a_{23}\, a_{32} \;-\; a_{12}\, a_{21}\, a_{33} \;-\; a_{13}\, a_{22}\, a_{31}\,. \end{array}$$

Diese Formel heißt auch Sarrussche Regel. *Für $n = 4$ enthält die in der Definition stehende Formel für die Determinante bereits 24 Terme und ist für die Berechnung ungeeignet! Wir werden aber einen effektiven Algorithmus zur Berechnung der Determinante kennen lernen, und auch für $n = 3$ werden wir Determinanten ausschließlich mit diesem Algorithmus berechnen.*

Aufgabe 4.3.3 *Berechne die Determinante der Matrizen*

$$A \;=\; \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\,, \quad B \;=\; \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\,.$$

Aufgabe 4.3.4 *Nach Definition ist die Determinante eine endliche Summe von Termen, die bis auf ein Vorzeichen von der Form $a_{1\sigma(1)}\, a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$ sind. Zeige, dass dieses Produkt auch in der Form*

$$a_{\sigma^{-1}(1)\,1}\, a_{\sigma^{-1}(2)\,2} \cdot \ldots \cdot a_{\sigma^{-1}(n)\,n}$$

geschrieben werden kann. Leite daraus ab, dass in jedem solchen Produkt genau ein Element aus jeder Zeile, aber auch aus jeder Spalte, von A steht.

Aufgabe 4.3.5 *Beweise die Formel*

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \;=\; a_{11}\, a_{22}\, a_{33}$$

und überlege, ob eine analoge Formel auch für größere Matrizen gilt.

Lösung: Da $a_{21} = a_{31} = a_{32} = 0$ sind, folgt die Behauptung aus der Sarrusschen Regel. Für beliebiges n sei A so, dass $a_{jk} = 0$ ist sobald $j > k$ ist. Das bedeutet, dass in der Definition der Determinante nur solche Terme stehen bleiben, für die immer $j \le \sigma(j)$ ist. Dies ist aber nach Aufgabe 4.1.9 nur für die identische Permutation der Fall, und deshalb ist

$$\det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} = a_{11}\, a_{22} \cdot \ldots \cdot a_{nn} \,.$$

□

Aufgabe 4.3.6 *Überprüfe mit Hilfe der Definition folgende Regel für Determinanten:* Falls eine Zeile oder eine Spalte einer quadratischen Matrix A nur Nullen enthält, ist $\det A = 0$.

Aufgabe 4.3.7 *Zeige allein mit der Definition der Determinante, dass*

$$\det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{bmatrix} = a_{11} \, \det \begin{bmatrix} a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} \end{bmatrix} .$$

4.4 Rechenregeln für Determinanten

Für die folgenden Rechenregeln seien $a_1, \ldots, a_n \in \mathbb{K}^n$ die Spalten einer Matrix A, und wir schreiben auch $\det[a_1, \ldots, a_n]$ an Stelle von $\det A$. Dadurch wird „det“ eine Abbildung von $\mathbb{K}^n \times \ldots \times \mathbb{K}^n$ (mit n Faktoren) nach $\mathbb{K}$. Eine solche Abbildung heißt auch eine n-stellige Form auf $\mathbb{K}^n$.

Satz 4.4.1 (Rechenregeln für Determinanten) *Für $a, b, a_1, \ldots, a_n \in \mathbb{K}^n$ und $\lambda, \mu \in \mathbb{K}$ gilt:*

(a) *Für Transpositionen $\tau \in S_n$ gilt $\det[a_{\tau(1)}, \ldots, a_{\tau(n)}] = -\det[a_1, \ldots, a_n]$. In Worten ausgedrückt heißt das: Beim Vertauschen zweier Spalten wechselt die Determinante das Vorzeichen.*

(b) $\det[\lambda\, a + \mu\, b, a_2, \ldots, a_n] = \lambda \det[a, a_2, \ldots, a_n] + \mu \det[b, a_2, \ldots, a_n]$. *Diese Eigenschaft besagt in Worten, dass die Determinante linear in der ersten Spalte ist.*

Beweis: Zu (a): Nach Definition ist

$$\begin{aligned} \det[a_{\tau(1)}, \ldots, a_{\tau(n)}] &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\; a_{1\sigma(\tau(1))} \cdot \ldots \cdot a_{n\sigma(\tau(n))} \\ &= -\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma \circ \tau)\; a_{1\sigma(\tau(1))} \cdot \ldots \cdot a_{n\sigma(\tau(n))} \,. \end{aligned}$$

Da nach Satz 4.1.8 (b) mit σ auch $\sigma \circ \tau$ einmal die Gruppe S_n durchläuft, folgt die Behauptung. Zu (b): Folgt direkt aus der Definition der Determinante. □

Definition 4.4.2 *Aussage* (b) *des vorstehenden Satzes besagt, dass die Determinante linear in der ersten Spalte ist, und wegen* (a) *gilt Gleiches für jede andere Spalte. Man sagt kurz dass die Determinante eine* Multilinearform *ist. Wegen* Satz 4.4.1 (a) *nennt man die Determinante auch* alternierend*.*

Satz 4.4.3 *Für jedes $A \in \mathbb{K}^{n\times n}$ gilt:*

(a) *Falls A eine obere oder untere Dreiecksmatrix ist, dann ist ihre Determinante gleich dem Produkt ihrer Diagonalelemente.*

(b) $\det A = \det A^T, \quad \det I = 1.$

Beweis: Zu (a): Wurde für obere Dreiecksmatrizen in Aufgabe 4.3.5 bewiesen und folgt für untere aus (b). Zu (b): Nach Definition der Determinante ist

$$\begin{aligned} \det A^T &= \sum_\sigma \operatorname{sgn}(\sigma) \prod_j a_{\sigma(j)j} = \sum_\sigma \prod_j \operatorname{sgn}(\sigma) a_{j\sigma^{-1}(j)} \\ &= \sum_{\sigma^{-1}} \operatorname{sgn}(\sigma^{-1}) \prod_j a_{j\sigma^{-1}(j)}\,, \end{aligned}$$

und da nach Satz 4.1.8 (a) mit σ auch σ^{-1} die Gruppe S_n durchläuft, folgt die Behauptung. □

Wegen ihrer Wichtigkeit formulieren wir noch einmal folgende Regeln für Determinanten, die sich sofort aus Satz 4.4.1 ergeben:

Korollar zu Satz 4.4.1 *Für jedes $A \in \mathbb{K}^{n\times n}$ gilt:*

(a) *Wenn B aus A durch Vertauschen zweier Spalten entsteht, so ist* $\det B = -\det A$.

(b) *Entsteht B aus A durch Multiplikation einer Spalte mit einem Faktor $\lambda \in \mathbb{C}$, so ist* $\det B = \lambda \det A$.

(c) *Entsteht B aus A, indem man ein Vielfaches einer Spalte zu einer anderen adddiert, so ist* $\det B = \det A$.

Wegen Satz 4.4.3 (b) *gelten die Regeln* (a) – (c) *auch für die Zeilen an Stelle der Spalten von A.*

Aufgabe 4.4.4 *Zeige: Wenn eine quadratische Matrix zwei gleiche Zeilen oder zwei gleiche Spalten enthält, dann ist ihre Determinante gleich 0.*

Bemerkung 4.4.5 (Berechnung einer Determinante) *Mit Hilfe der obigen Rechenregeln kann man auch größere Determinanten relativ leicht berechnen. Dabei benutzt man folgenden Algorithmus, welcher wie beim Lösen von linearen Gleichungssystemen abläuft und wieder als* Gaußsches Eliminationsverfahren *bezeichnet wird: Falls die erste Spalte von A nur Nullen enthält, ist* $\det A = 0$. *Sonst suche in der ersten Spalte der Matrix A ein Element $a_{j1} \neq 0$ und vertausche die j-te Zeile von A mit der ersten, außer falls $j = 1$ ist. Danach ziehe Vielfache der ersten Zeile von allen anderen ab, so dass alle anderen Elemente der ersten Spalte durch Nullen ersetzt werden. Durch diese Schritte wird aus A eine Matrix B der Form*

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & b_{22} & b_{23} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{n2} & b_{n3} & \dots & b_{nn} \end{bmatrix},$$

und für diesen Fall gilt $\det A = \det B$ *falls wir keine Zeilen vertauschen mussten bzw.* $\det A = -\det B$ *sonst. Wegen* Aufgabe 4.3.7 *gilt*

$$\det B = b_{11} \det \begin{bmatrix} b_{22} & \dots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{n2} & \dots & b_{nn} \end{bmatrix}.$$

Das bedeutet, dass wir die Berechnung der Determinante von A auf die Berechnung einer Determinante einer kleineren Matrix zurückgeführt haben! Diese Determinante kann wieder nach dem gleichen Schema berechnet werden.

Aufgabe 4.4.6 *Berechne die Determinante der Matrix*

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 \\ -1 & 1 & -1 & 0 \end{bmatrix}.$$

Bemerkung 4.4.7 *In dem oben eingeführten Verfahren zur Berechnung einer Determinante kann man auch die entsprechenden Operationen für Spalten verwenden, wenn dies vorteilhaft erscheint. Außerdem folgt aus der Definition, dass die Determinante eine ganze Zahl ist, falls die Elemente der Matrix ganzzahlig sind. Dies kann zur Kontrolle der Rechnung genutzt werden.*

Aufgabe 4.4.8 (Determinanten mit MAPLE) *Führe die folgenden* MAPLE-*Kommandos aus:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(5, 5, [ [1, 0, 0, 1, -1], [1, 1, -1, 2, 0], [2, -1, -1, 0, 0], [-1 ,1, 0, 2, 1], [-1, 2, 3, 1, 0] ]);
> Determinant(A);
> B := Matrix(2,2, [[a,b], [c,d]]);
> Determinant(B);
> C := Matrix(2, 3, [[1,1,1], [1, -1, 0]]);
> Determinant(C);
```

Verwende ähnliche Sequenzen zur Berechnung der Determinante der Matrix aus der vorangegangenen Aufgabe.

4.5 Weitere Eigenschaften von Determinanten

Die wichtigste Anwendung der Determinante besteht in folgender Charakterisierung der Invertierbarkeit von Matrizen:

Proposition 4.5.1 *Eine Matrix $A \in \mathbb{K}^{n\times n}$ ist genau dann invertierbar, wenn $\det A \neq 0$ ist.*

Beweis: Elementare Zeilen- und Spaltentransformationen verändern den Rang von A nicht, während die Determinante sich höchstens um von 0 verschiedene Faktoren ändert. Mit diesen Transformationen kann man A auf obere Dreiecksgestalt bringen, und eine Dreiecksmatrix A ist nach Aufgabe 3.3.4 genau dann invertierbar, wenn alle Diagonalelemente von 0 verschieden sind. Daraus folgt die Behauptung. □

Satz 4.5.2 (Determinantenmultiplikationssatz) *Für zwei quadratische n-reihige Matrizen A und B gilt immer*

$$\det(A\,B) \;=\; \det A\ \det B\,.$$

Beweis: Falls B nicht invertierbar, also $\operatorname{rang} B < n$, ist, folgt $\operatorname{rang}(A\,B) < n$, und deshalb folgt mit der vorstehenden Proposition dass $\det(A\,B) = \det B = 0$ ist. Also gilt die Behauptung in diesem Fall. Falls B invertierbar ist, folgt aus Aufgabe 2.8.3, dass B ein Produkt von endlich vielen Elementarmatrizen ist. Daher genügt es, einen Beweis für den Fall zu geben, dass B selber eine Elementarmatrix ist; der allgemeine Fall folgt dann durch Induktion. Für eine Elementarmatrix B überprüft man aber die Behauptung mit den Rechenregeln für Determinanten. □

Aufgabe 4.5.3 *Zeige für jede invertierbare Matrix A die Regel* $\det A^{-1} = (\det A)^{-1}$.

Lemma 4.5.4 (Dreieckig geblockte Matrizen) *Seien A_{11} vom Typ $n_1 \times n_1$, A_{22} vom Typ $n_2 \times n_2$ und A_{12} vom Typ $n_1 \times n_2$. Dann gilt*

$$\det \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad = \quad \det A_{11} \ \det A_{22}$$

Beweis: Durch elementare Zeilentransformationen, angewandt auf $[A_{11}, A_{12}]$ und A_{22}, kann die linke Matrix so umgeformt werden, dass A_{11} und A_{22} obere Dreiecksgestalt haben, ohne dass die Blockdreiecksform der Gesamtmatrix verändert wird. Daraus folgt die Behauptung. □

Aufgabe 4.5.5 *Formuliere selber Verallgemeinerungen von obigem Lemma auf den Fall von oberen Blockdreiecksmatrizen mit mehr als zwei Diagonalblöcken, und beweise diese Aussagen mit Induktion. Was kann man für untere Blockdreiecksmatrizen sagen?*

Behauptung 4.5.6 (Vandermondesche Determinante)
Für alle $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gilt die Gleichung:

$$\det \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{bmatrix} \quad = \prod_{1 \le j < k \le n} (\lambda_k - \lambda_j)\,.$$

Insbesondere heißt das, dass die linksstehende Matrix, auch Vandermondesche Matrix *genannt, genau dann invertierbar ist, wenn alle λ_j verschieden sind. Vergleiche diese Art von Matrix auch mit* Aufgabe 2.2.4.

Beweis: Falls zwei der λ_j gleich sind, hat die linksstehende Matrix zwei gleiche Spalten, und deshalb ist ihre Determinante gleich 0. Also gilt die Gleichung für diesen Fall, und wir können jetzt annehmen dass alle λ_j verschieden sind, also höchstens ein $\lambda_j = 0$ ist. Außerdem erkennt man, dass die Gleichung bei Vertauschen von Spalten der Matrix, also Umnumerierung der λ_j, richtig bleibt, und deshalb können wir sogar annehmen, dass $\lambda_j \neq 0$ für $j = 2, \ldots, n$ gilt. Ersetzt man jetzt λ_1 durch eine Variable t, so erkennt man aus der Definition der Determinante, dass die Vandermondesche Determinante ein Polynom $p(t)$ vom Grade $n-1$ ist, welches nach der Vorüberlegung die Werte $\lambda_2, \ldots, \lambda_n$ als Nullstellen hat. Deshalb folgt durch Abdividieren dieser Nullstellen

$$p(t) \; = \; c\,(\lambda_2 - t) \cdot \ldots \cdot (\lambda_n - t)$$

mit $c \in \mathbb{K} \setminus \{0\}$. Für $t = 0$ ergibt sich durch Anwenden von Lemma 4.5.4

$$c\,\lambda_2 \cdot \ldots \cdot \lambda_n \; = \; \det \begin{bmatrix} \lambda_2 & \ldots & \lambda_n \\ \lambda_2^2 & \ldots & \lambda_n^2 \\ \vdots & & \vdots \\ \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{bmatrix}\,.$$

Durch Teilen der $(j-1)$-ten Spalte der rechtsstehenden Matrix durch den Faktor λ_j $(\neq 0)$ für $j = 2, \ldots, n$ folgt, dass c gleich dem Wert der Vandermondeschen Deteminante der $n-1$ Skalare $\lambda_2, \ldots, \lambda_n$ ist, und durch Induktion über n folgt deshalb die Behauptung. □

Aufgabe 4.5.7 (Vandermonde-Matrizen und MAPLE)
Führe folgende MAPLE-*Befehle aus:*

```
> restart;

> with(LinearAlgebra);

> A := VandermondeMatrix([1,2,3,4]);

> Determinant(A);

> B := VandermondeMatrix([a,b,c,d]);

> det := Determinant(B);

> factor(det);
```

Vergleiche mit der obigen Behauptung. Wo liegt hier eine kleine Schwäche von MAPLE*?*

4.6 Entwicklungssätze

Definition 4.6.1 *Sei $n \geq 2$, und sei A eine n-reihige quadratische Matrix. Die Determinante der Matrix, die aus A durch Streichen der j-ten Zeile und der k-ten Spalte entsteht, wird mit d_{jk} bezeichnet und heißt* ein Minor *von A. Die Zahl $A_{jk} = (-1)^{j+k}\, d_{jk}$ heißt dann das* algebraische Komplement *zu a_{jk}.*

Aufgabe 4.6.2 *Zeige durch geeignetes Vertauschen von Zeilen und Spalten: Das algebraische Komplement A_{jk} ist die Determinante derjenigen Matrix, welche man durch Ersetzen der k-ten Spalte von A durch den j-ten kanonischen Basisvektor enthält. Wie heißt die entsprechende Aussage für Zeilen?*

Satz 4.6.3 *Für jedes $A = [a_{jk}] \in \mathbb{K}^{n\times n}$ gilt*

(a) $$\sum_{\ell=1}^{n} A_{\ell k}\; a_{\ell j} \;=\; \delta_{jk} \det A \quad \textit{für} \quad 1 \leq j, k \leq n.$$

(b) $$\sum_{\ell=1}^{n} a_{j\ell}\; A_{k\ell} \;=\; \delta_{jk} \det A \quad \textit{für} \quad 1 \leq j, k \leq n.$$

Beweis: Benutzt man die Linearität der Determinante in jeder ihrer Spalten sowie Aufgabe 4.6.2, so sieht man dass $\sum_{\ell=1}^{n} A_{\ell k}\; a_{\ell j}$ die Determinante der Matrix ist, die man aus A durch Ersetzen der k-ten durch die j-te Spalte erhält. Ist $j \neq k$, so enthält diese Matrix zwei gleiche Spalten und ihre Determinante verschwindet; ist $j = k$, so ist diese Matrix gleich A. Also gilt (a). Aussage (b) wird genauso durch Betrachten der Zeilen, oder aber durch Transposition bewiesen. □

Die folgenden beiden Ergebnisse sind direkte Konsequenzen bzw. Umformulierungen des letzten Satzes:

Korollar zu Satz 4.6.3 *Falls* $\det A \neq 0$, *also* A *invertierbar ist, gilt*

$$A^{-1} \;=\; \frac{1}{\det A} \quad \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}.$$

Mit anderen Worten: Die Matrix aus den algebraischen Komplementen von A *ist gleich dem Transponierten der inversen Matrix multipliziert mit* $\det A$.

Korollar zu Satz 4.6.3 (Entwicklungssatz nach einer Zeile bzw. Spalte) *Sei* A *eine quadratische* n*-reihige Matrix mit* $n \geq 2$, *und sei ein* $k \in \{1, \dots, n\}$ *ausgewählt. Dann gelten die Gleichungen*

$$\det A \;=\; \sum_{j=1}^{n} (-1)^{j+k} \, a_{jk} \, d_{jk} \quad = \sum_{j=1}^{n} (-1)^{j+k} \, a_{kj} \, d_{kj} \,.$$

Bemerkung 4.6.4 *Wir haben nur Minoren der Größe* $(n-1) \times (n-1)$ *betrachtet. Wenn man mehrere Zeilen und Spalten einer Matrix streicht und dann die Determinante der Restmatrix bildet, so erhält man die übrigen Minoren niedrigerer Größe. Auch für solche gibt es einen Entwicklungssatz für die Determinante, den man z. B. bei* Gawronski [2] *nachlesen kann.*

Aufgabe 4.6.5 *Berechne mit Hilfe des Entwicklungssatzes die Determinante von*

$$A \;=\; \begin{bmatrix} -1 & 1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 4 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & -1 \end{bmatrix}.$$

Aufgabe 4.6.6 (Cramersche Regel) *Sei* A *eine quadratische* n*-reihige Matrix mit den Spalten* a_k, *und sei* $\det A \neq 0$. *Dann ist das lineare Gleichungssystem* $A\,x = b$ *für jedes* $b \in \mathbb{K}^n$ *eindeutig lösbar. Zeige: Die Koordinaten* x_k *dieses Vektors* x *sind gegeben durch*

$$x_k \;=\; \frac{1}{\det A} \; \det[a_1, \dots, a_{k-1}, b, a_{k+1}, \dots, a_n] \qquad \forall\; k = 1, \dots, n\,.$$

Kapitel 5

Euklidische und unitäre Räume

5.1 Definition des Skalarprodukts

Definition 5.1.1 *Sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung*

$$\langle\,\cdot\,,\,\cdot\,\rangle :\ V \times V \longrightarrow \mathbb{K}\,, \qquad (v_1, v_2) \longmapsto \langle v_1, v_2\rangle \in \mathbb{K}$$

heißt ein Skalarprodukt *oder* inneres Produkt *auf V, wenn folgende Regeln gelten:*

(S1) $\forall\ v \in V :\quad \langle v, v\rangle \ \geq\ 0\,;\quad \langle v, v\rangle \ =\ 0 \iff v = 0$.

(S2) $\forall\ v_1, v_2 \in V :\quad \langle v_1, v_2\rangle \ =\ \overline{\langle v_2, v_1\rangle}$.

(S3) $\forall\ v, v_1, v_2 \in V :\quad \langle v, v_1 + v_2\rangle \ =\ \langle v, v_1\rangle \ +\ \langle v, v_2\rangle$.

(S4) $\forall\ v_1, v_2 \in V,\ \lambda \in \mathbb{K} :\quad \langle v_1, \lambda\, v_2\rangle \ =\ \lambda\,\langle v_1, v_2\rangle$.

Falls ein solches Skalarprodukt auf V gegeben ist, nennt man V im Falle von $\mathbb{K} = \mathbb{R}$ auch euklidischen Raum*, im anderen Fall einen* unitären Raum*. In beiden Fällen spricht man auch von einem* Prä-Hilbertraum *oder* einem Raum mit Skalarprodukt*.*

Bemerkung 5.1.2 *Aus den Axiomen folgt sofort die weitere Regel*

$$\forall\ v_1, v_2 \in V,\ \lambda \in \mathbb{K} :\quad \langle \lambda\, v_1, v_2\rangle \ =\ \overline{\lambda}\,\langle v_1, v_2\rangle\,.$$

Im Fall $\mathbb{K} = \mathbb{R}$ ist ein Skalarprodukt immer eine reelle Zahl, und dann bedeutet (S2) *einfach $\langle v_1, v_2\rangle = \langle v_2, v_1\rangle$ für alle $v_1, v_2 \in V$.*

Beispiel 5.1.3 *Auf dem Vektorraum $V = \{0\}$ ist die Nullabbildung ein inneres Produkt. Falls V igendein Raum mit Skalarprodukt ist, so ist jeder Unterraum U von V wieder ein Raum mit Skalarprodukt, wenn man einfach die Restriktion des Skalarproduktes auf $U \times U$ betrachtet. Für Vektoren $x = (x_1, \ldots, x_n)^T, y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ ist*

$$\langle x, y\rangle \ =\ \sum_{k=1}^{n} x_k\, y_k \ =\ x^T y$$

ein inneres Produkt, das sogenannte kanonische Skalarprodukt *in $\mathbb{R}^n$. In $\mathbb{C}^n$ ist*

$$\langle x, y\rangle \ =\ \sum_{k=1}^{n} \overline{x_k}\ y_k \ =\ \overline{x}^T y$$

das entsprechende kanonische innere Produkt. Wenn nichts anderes gesagt ist, betrachten wir in $\mathbb{K}^n$ *im Folgenden immer dieses kanonische Skalarprodukt. Für zwei Funktionen* $f, g \in C[a,b]$ *ist ein kanonisches inneres Produkt durch*

$$\langle f, g \rangle \;=\; \int_a^b \overline{f(x)}\, g(x)\, dx$$

gegeben, wobei der Querstrich wegfallen kann, falls die Werte der Funktionen reell sind. Häufig gebraucht werden aber auch gewichtete innere Produkte *von Funktionen. Dabei ist eine feste* Gewichtsfunktion k *gegeben, die bis auf endlich viele Punkte positiv und stetig auf* $[a,b]$ *ist, und man setzt*

$$\langle f, g \rangle \;=\; \int_a^b \overline{f(x)}\, g(x)\, k(x)\, dx\,.$$

Auf der Menge der über $[a,b]$ *integrierbaren Funktionen erhält man durch die obigen Festsetzungen* keine *inneren Produkte, da* $\langle f, f \rangle = 0$ *sein kann ohne dass* $f(x) = 0$ *ist für alle* $x \in [a,b]$.

Aufgabe 5.1.4 *Sei* V *ein Vektorraum mit innerem Produkt. Zeige* $\langle 0, v \rangle = \langle v, 0 \rangle = 0$ *für alle* $v \in V$.

Aufgabe 5.1.5 *Für* $A \in \mathbb{K}^{n\times n}$ *sei* $\langle x, y \rangle_A = \overline{x}^T A\, y$ *gesetzt, für alle* $x, y \in \mathbb{K}^n$. *Für* $A = I$ *ist also* $\langle x, y \rangle_A$ *das kanonische Skalarprodukt in* $\mathbb{K}^n$. *Überprüfe, welche der Regeln für ein inneres Produkt für jedes* A *gelten, und gib ein Beispiel für* A, *für welches nicht alle Regeln erfüllt sind.*

Aufgabe 5.1.6 *Sei* J *eine nicht-leere Menge, und sei* $\mathbb{K}_J$ *die Menge aller Abbildungen* f *von* J *in* $\mathbb{K}$ *mit* $f(j) \neq 0$ *höchstens für endlich viele* $j \in J$. *Zeige, dass* $\mathbb{K}_J$ *ein Vektorraum über* $\mathbb{K}$ *ist, und dass durch*

$$\langle f, g \rangle = \sum_{j \in J} \overline{f(j)}\, g(j)$$

ein inneres Produkt auf $\mathbb{K}_J$ *definiert wird.*

Satz 5.1.7 *Sei* V *ein Raum über* $\mathbb{K}$ *mit einer Basis* $(v_j, j \in J)$, *und sei für* $v_k = \sum_{j\in J} \alpha_j^{(k)}\, v_j$

$$\langle v_1, v_2 \rangle_J \;=\; \sum_{j\in J} \overline{\alpha_j^{(1)}}\, \alpha_j^{(2)}\,.$$

Dann ist $\langle\,\cdot\,,\,\cdot\,\rangle_J$ *ein inneres Produkt auf* W.

Beweis: Kann einfach mit der Definition des inneren Produktes nachgeprüft werden! □

Aufgabe 5.1.8 *Zeige, dass man in jedem endlich-dimensionalen Vektorraum ein inneres Produkt definieren kann.*

5.2 Die Norm eines Vektors

Im Folgenden sei V immer ein Raum mit Skalarprodukt.

Definition 5.2.1 *Für jedes* $v \in V$ *heißt*

$$\|v\| \;=\; \sqrt{\langle v, v \rangle}$$

die Norm *oder* Länge *von* v. *Für* $v_1, v_2 \in V$ *heißt* $\|v_1 - v_2\|$ *auch* Abstand zwischen v_1 *und* v_2.

Aufgabe 5.2.2 *Zeige für beliebige $v_1, v_2 \in V$:*

(a) *Falls $\mathbb{K} = \mathbb{R}$ ist, gilt*

$$\langle v_1, v_2 \rangle \;=\; \|(v_1 + v_2)/2\|^2 \;-\; \|(v_1 - v_2)/2\|^2 \,.$$

(b) *Falls $\mathbb{K} = \mathbb{C}$ ist, gilt*

$$\begin{aligned} \langle v_1, v_2 \rangle \;&=\; \|(v_1 + v_2)/2\|^2 \;-\; \|(v_1 - v_2)/2\|^2 \\ &\quad + i\,\|(v_1 + i\,v_2)/2\|^2 \;-\; i\,\|(v_1 - i\,v_2)/2\|^2 \,. \end{aligned}$$

Satz 5.2.3 (Cauchy-Schwarzsche Ungleichung) *Für $v_1, v_2 \in V$ gilt immer*

$$\big|\langle v_1, v_2 \rangle\big| \;\le\; \|v_1\|\;\|v_2\| \;,$$

und das Gleichheitszeichen gilt genau dann, wenn v_1 und v_2 linear abhängig sind.

Beweis: Für $v_2 = 0$ ist die Behauptung korrekt, da nach Aufgabe 5.1.4 $\langle v, 0 \rangle = 0$ ist für alle $v \in V$. Sei jetzt $v_2 \neq 0$. Falls v_1, v_2 linear abhängig sind, muss $v_1 = \alpha\, v_2$ für ein $\alpha \in \mathbb{K}$ gelten, und dann folgt $|\langle v_1, v_2 \rangle| = |\overline{\alpha}\,\langle v_2, v_2 \rangle| = |\alpha|\,\|v_2\|^2$, $\|v_1\| = |\alpha|\,\|v_2\|$, und deshalb gilt die Ungleichung mit dem Gleichheitszeichen. Falls die Vektoren linear unabhängig sind, ist $v_1 - \lambda\, v_2 \neq 0$ für alle $\lambda \in \mathbb{K}$. Dann folgt mit den Regeln eines Skalarproduktes $0 < \langle v_1 - \lambda\, v_2, v_1 - \lambda\, v_2 \rangle = \langle v_1, v_1 \rangle - \lambda\,\langle v_1, v_2 \rangle - \overline{\lambda}\,\langle v_2, v_1 \rangle + \lambda\,\overline{\lambda}\,\langle v_2, v_2 \rangle$. Wählt man speziell $\lambda = \langle v_2, v_1 \rangle / \|v_2\|^2$, so folgt hieraus

$$0 \;<\; \|v_1\|^2 \;-\; 2\,\frac{|\langle v_1, v_2 \rangle|^2}{\|v_2\|^2} \;+\; \frac{|\langle v_1, v_2 \rangle|^2}{\|v_2\|^4}\,\|v_2\|^2 \;=\; \|v_1\|^2 \;-\; \frac{|\langle v_1, v_2 \rangle|^2}{\|v_2\|^2} \,,$$

woraus die behauptete Ungleichung mit dem „$<$“-Zeichen folgt. □

Korollar zu Satz 5.2.3 (Eigenschaften der Norm) *Für die Abbildung $v \longmapsto \|v\|$ gilt immer*

(N1) $\forall\, v \in V: \quad \|v\| \ge 0\,, \quad \|v\| = 0 \iff v = 0$ (Definitheit)

(N2) $\forall\, v \in V\,,\ \lambda \in \mathbb{K}: \quad \|\lambda\, v\| = |\lambda|\,\|v\|$ (Homogenität)

(N3) $\forall\, v_1, v_2 \in V: \quad \|v_1 + v_2\| \;\le\; \|v_1\| \;+\; \|v_2\|$ (Dreiecksungleichung)

Beweis: Aus obigem Satz und den Rechenregeln für ein Skalarprodukt folgt $\|v_1 + v_2\|^2 = \langle v_1 + v_2, v_1 + v_2 \rangle = \langle v_1, v_1 \rangle + \langle v_1, v_2 \rangle + \langle v_2, v_1 \rangle + \langle v_2, v_2 \rangle \le \|v_1\|^2 + 2\,\|v_1\|\,\|v_2\| + \|v_2\|^2 = (\|v_1\| + \|v_2\|)^2$, und das ist die Dreiecksungleichung. Die anderen beiden Behauptungen folgen unmittelbar aus den Eigenschaften eines Skalarproduktes. □

Aufgabe 5.2.4 *Benutze die Cauchy-Schwarzsche Ungleichung bzw. die Dreiecksungleichung, um folgendes zu beweisen: Für alle $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{C}$ gilt immer*

$$\Big|\sum_{j=1}^{n} a_j\, b_j\Big| \;\le\; \sum_{j=1}^{n} |a_j\, b_j| \;\le\; \Big(\sum_{j=1}^{n} |a_j|^2\Big)^{1/2} \Big(\sum_{j=1}^{n} |b_j|^2\Big)^{1/2} \,,$$

$$\Big(\sum_{j=1}^{n} |a_j + b_j|^2\Big)^{1/2} \;\le\; \Big(\sum_{j=1}^{n} |a_j|^2\Big)^{1/2} \;+\; \Big(\sum_{j=1}^{n} |b_j|^2\Big)^{1/2} \,.$$

Aufgabe 5.2.5 *Mit ℓ_2 bezeichnet man die Menge aller komplexen Zahlenfolgen $a = (a_1, a_2, a_3, \ldots)$ mit der Eigenschaft, dass die Reihe $\sum_{j=1}^{\infty} |a_j|^2$ konvergiert. Benutze die Ungleichungen aus der vorangegangenen Aufgabe, um zu zeigen:*

(a) ℓ_2 ist ein Vektorraum über $\mathbb{C}$.

(b) Definiert man für Folgen $a = (a_1, a_2, a_3, \ldots)$, $b = (b_1, b_2, b_3, \ldots,)\in \ell_2$

$$\langle a, b\rangle \;=\; \sum_{j=1}^{\infty} a_j\, b_j\,,$$

so ist die Reihe immer absolut konvergent, und $\langle\,\cdot\,,\,\cdot\,\rangle$ ist ein Skalarprodukt auf ℓ_2.

Der Raum ℓ_2 ist unendlich-dimensional und kann als natürliche Verallgemeinerung der Räume $\mathbb{C}^n$ angesehen werden. Die Folgen e_k, welche an der k-ten Stelle eine 1 und an den übrigen lauter Nullen stehen haben, entsprechen den Vektoren der kanonischen Basis von $\mathbb{C}^n$. Finde heraus, warum diese Folgen keine Basis von ℓ_2 bilden.

Definition 5.2.6 *Sei V ein Vektorraum über $\mathbb{K}$. Eine beliebige Abbildung $\|\cdot\|: \; V \longrightarrow \mathbb{R}$ heißt immer dann* eine Norm auf V*, wenn die drei Axiome* (N1) – (N3) *erfüllt sind. Wenn auf V eine solche Norm gegeben ist, nennt man V auch* normierten *Raum. Ein Raum mit Skalarprodukt ist also immer auch ein normierter Raum, aber nicht unbedingt umgekehrt, denn auf V kann eine Norm gegeben sein, welche nicht von einem Skalarprodukt herrührt.*

5.3 Orthogonalität und Winkel

Definition 5.3.1 *Zwei Vektoren v_1, v_2 in einem Raum über $\mathbb{K}$ mit Skalarprodukt V heißen* orthogonal*, falls $\langle v_1, v_2\rangle = 0$ ist. Zwei nicht-leere Teilmengen $A, B \subset V$ heißen zueinander orthogonal, falls jeder Vektor aus A zu jedem aus B orthogonal ist. Ist $A \subset V$ nicht-leer, so heißt die Menge $A^\perp$ aller Vektoren aus V, welche zu allen Vektoren aus A orthogonal sind, das* orthogonale Komplement *von A. In anderen Worten: $A^\perp$ ist die größte Teilmenge von V, so dass A und $A^\perp$ zueinander orthogonal sind. Falls $\mathbb{K} = \mathbb{R}$, also das Skalarprodukt von Vektoren immer reell ist, gibt es zu $v_1, v_2 \in V \setminus \{0\}$ genau einen Winkel $\varphi \in [0, \pi]$ mit*

$$\cos\varphi \;=\; \frac{\langle v_1, v_2\rangle}{\|v_1\|\,\|v_2\|}\,,$$

und wir nennen dieses φ den Winkel *zwischen v_1 und v_2. Offenbar sind die Vektoren genau dann orthogonal, wenn der Winkel zwischen ihnen gleich $\pi/2$ ist. Falls $V = \mathbb{R}^2$ oder $= \mathbb{R}^3$ ist, stimmt diese Definition eines Winkels mit der Anschauung überein, was aber hier keine Rolle spielt. Beachte jedenfalls, dass für $\mathbb{K} = \mathbb{C}$ zwar kein Winkel zwischen zwei Vektoren definiert ist, dass aber der Begriff der Orthogonalität trotzdem sinnvoll ist.*

Satz 5.3.2 *Sei V ein Raum über $\mathbb{K}$ mit Skalarprodukt, und sei $A \subset V$ nicht leer. Dann gilt:*

(a) *$A^\perp$ ist ein Unterraum von V.*

(b) *$(A^\perp)^\perp \supset A$.*

(c) *$\mathcal{L}(A) \cap A^\perp = \{0\}$.*

Beweis: Zu (a): Es ist immer $0 \in A^\perp$, also ist $A^\perp \neq \emptyset$. Weiter folgt aus den Rechenregeln für das Skalarprodukt, dass die Unterraumaxiome für $A^\perp$ erfüllt sind. Teil (b) ist klar nach Definition von $A^\perp$. Zu (c): Sei $u \in \mathcal{L}(A)$, dann gibt es $u_1, \ldots, u_m \in A$ und $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ mit $u = \sum_{j=1}^m \alpha_j\, u_j$. Also folgt für $v \in A^\perp$, dass $\langle v, u\rangle = \sum_{j=1}^m \alpha_j\, \langle v, u_j\rangle = 0$ ist, weil v zu allen u_j orthogonal ist. Wenn u auch zu $A^\perp$ gehört, können wir $v = u$ wählen und erhalten $\langle u, u\rangle = 0$, also $u = 0$. □

Aufgabe 5.3.3 (Orthogonalität und MAPLE) *Finde diejenigen* MAPLE-*Kommandos im Paket* LinearAlgebra, *welche etwas mit Winkel, Norm und/oder Orthogonalität in* $\mathbb{K}^n$ *zu tun haben.*

5.4 Orthogonalsysteme

Definition 5.4.1 *Ein System* $(v_j, j \in J)$ *von Vektoren in einem Raum* V *über* $\mathbb{K}$ *mit Skalarprodukt heißt* ein Orthogonalsystem, *falls keiner der Vektoren der Nullvektor ist, und wenn*

$$\forall\, j,k \in J : \langle v_j, v_k \rangle \;=\; 0 \quad \text{falls } j \neq k\,.$$

Falls zusätzlich gilt $\|v_j\| = 1$ *für alle* $j \in J$, *dann sprechen wir von einem* Orthonormalsystem. *Beachte, dass das leere System immer ein Orthonormalsystem ist. Falls* $(v_j, j \in J)$ *zusätzlich noch eine Basis von* V *sind, sprechen wir von einer* Orthogonalbasis *bzw. einer* Orthonormalbasis *von* V.

Beispiel 5.4.2 *In* $\mathbb{K}^n$ *ist die kanonische Basis eine Orthonormalbasis. Der Vektorraum* $V = \{0\}$ *besitzt ebenfalls eine Orthonormalbasis, nämlich das leere System. In* ℓ_2 *sind die Folgen* e_k *zwar ein Orthonormalsystem, aber keine Basis; vergleiche hierzu* Aufgabe 5.2.5. *Allgemein sind in* $\mathbb{K}_J$ *wie in* Aufgabe 5.1.6 *die Funktionen* f_k *mit* $f_k(j) = \delta_{kj}$ *für* $j,k \in J$ *eine Orthonormalbasis.*

Aufgabe 5.4.3 *Sei* V *ein Raum über* $\mathbb{K}$ *mit einer Basis* $(v_j, j \in J)$, *und sei in* V *ein inneres Produkt wie in* Satz 5.1.7 *eingeführt. Zeige, dass dann* $(v_j, j \in J)$ *eine Orthonormalbasis ist.*

Lemma 5.4.4 *Ein Orthogonalsystem* $(v_j, j \in J)$ *ist immer linear unabhängig. Falls* $(v_j, j \in J)$ *sogar ein Orthonormalsystem ist, und falls* $v = \sum_{j\in J} \alpha_j\, v_j$ *ist, wobei nur endlich viele* α_j *von* 0 *verschieden sind, so folgt*

$$\|v\|^2 \;=\; \sum_{j\in J} |\alpha_j|^2\,. \tag{5.4.1}$$

Beweis: Seien $j_1,\dots,j_n \in J$. Aus $0 = \sum_{k=1}^n \lambda_k\, v_{j_k}$ folgt $0 = \langle v_{j_\nu}, \sum_{k=1}^n \lambda_k\, v_{j_k}\rangle = \sum_{k=1}^n \lambda_k\, \langle v_{j_\nu}, v_{j_k}\rangle = \lambda_\nu\, \|v_{j_\nu}\|^2$ für jedes $\nu = 1,\dots,n$, und da $v_{j_\nu} \neq 0$ ist, folgt hieraus $\lambda_\nu = 0$. Das ist die lineare Unabhängigkeit. Mit den Regeln für das Skalarprodukt folgt weiter

$$\|v\|^2 \;=\; \sum_{j\in J}\sum_{k\in J} \overline{\alpha_j}\,\alpha_k\, \langle v_j, v_k\rangle\,,$$

und wegen $\langle v_j, v_k\rangle = \delta_{jk}$ folgt daraus (5.4.1). □

Satz 5.4.5 (Gram-Schmidtsches Orthogonalisierungsverfahren)
Sei V *ein Raum über* $\mathbb{K}$ *mit Skalarprodukt, sei* $(w_1,\dots,w_n)$ *ein linear unabhängiges System in* V, *und sei das System* $(v_1,\dots,v_n)$ *durch folgende Rekursionsgleichungen definiert:*

$$\tilde v_k \;=\; w_k \;-\; \sum_{j=1}^{k-1} \langle v_j, w_k\rangle\, v_j\,, \quad v_k \;=\; \frac{1}{\|\tilde v_k\|}\,\tilde v_k\,, \qquad 1 \le k \le n\,,$$

wobei für $k=1$ *die leere Summe wie üblich als* 0 *zu interpretieren ist. Dann ist* $(v_1,\dots,v_n)$ *ein Orthonormalsystem, und*

$$\mathcal{L}(v_1,\dots,v_k) \;=\; \mathcal{L}(w_1,\dots,w_k) \qquad \forall\, k = 1,\dots,n\,.$$

Beweis: Wir zeigen per Induktion, dass das System $(v_1, \ldots, v_k)$ immer ein Orthonormalsystem ist, und dass $\mathcal{L}(v_1, \ldots, v_k) = \mathcal{L}(w_1, \ldots, w_k)$ gilt für alle $k = 1, \ldots, n$. Dies ist sicher richtig für $k = 1$, und wenn es für irgendein $k \geq 1$ stimmt, dann folgt für alle $\nu = 1, \ldots, k$:

$$\langle v_\nu, \tilde{v}_{k+1} \rangle = \langle v_\nu, w_{k+1} \rangle - \sum_{j=1}^{k} \langle v_j, w_{k+1} \rangle \, \langle v_\nu, v_j \rangle = 0\,,$$

denn $\langle v_\nu, v_j \rangle = \delta_{\nu j}$. Also ist $(v_1, \ldots, v_{k+1})$ ein Orthonormalsystem und somit linear unabhängig. Außerdem gilt $v_{k+1} \in \mathcal{L}(w_1, \ldots, w_{k+1})$, und daher folgt $\mathcal{L}(v_1, \ldots, v_{k+1}) \subset \mathcal{L}(w_1, \ldots, w_{k+1})$. Da aber die Dimensionen der linearen Hüllen beide gleich $k + 1$ sind, folgt sogar die Gleichheit. □

Aufgabe 5.4.6 *Orthogonalisiere die Vektoren*

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Das heißt genauer: Berechne die drei Vektoren v_1, v_2, v_3 aus dem obigen Satz.

Aufgabe 5.4.7 (Orthogonalisieren mit MAPLE) *Verwende den* MAPLE-*Befehl* > GramSchmidt*, um Vektoren in $\mathbb{K}^n$ zu orthogonalisieren.*

Aufgabe 5.4.8 *Sei U ein Unterraum von $\mathbb{C}^n$, und sei $(u_1, \ldots, u_m)$ eine Basis von U, welche nur aus reellen Vektoren u_j besteht; eine solche Basis existiert nach* Aufgabe 1.7.2*, falls mit jedem $u \in U$ auch $\overline{u} \in U$ ist. Schließe mit dem Gram-Schmidtschen Orthogonalisierungsverfahren, dass es dann sogar eine Orthonormalbasis von U gibt, welche nur aus reellen Vektoren u_j besteht.*

Bemerkung 5.4.9 *Das Gram-Schmidtsche Orthogonalisierungsverfahren kann nicht nur auf endlich viele, sondern auch auf eine Folge, also abzählbar unendlich viele, Vektoren $(w_j, j \in \mathbb{N})$ angewandt werden. Man erhält dann wieder eine Folge $(v_j, j \in \mathbb{N})$ mit der Eigenschaft $\mathcal{L}(v_1, \ldots, v_k) = \mathcal{L}(w_1, \ldots, w_k)$ für alle $k \in \mathbb{N}$. Wenn man das Orthogonalisierungsverfahren anwendet, muss man auch nicht unbedingt wissen, ob die zu orthogonalisierenden Vektoren linear unabhängig sind: Genau dann, wenn im k-ten Schritt der Vektor $\tilde{v}_k$ der Nullvektor ist, war w_k von den vorhergehenden Vektoren linear abhängig, und dann kann man das System um diesen Vektor w_k verkürzen.*

Korollar zu Satz 5.4.5 *Ein endlich-dimensionaler Raum mit Skalarprodukt V besitzt immer eine Orthonormalbasis.*

Beweis: Falls $V = \{0\}$ ist, ist das leere System Orthonormalbasis. Sonst sei $(w_1, \ldots, w_n)$ eine Basis von V. Mit Satz 5.4.5 erhält man dann ein Orthonormalsystem $(v_1, \ldots, v_n)$ mit $\mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(w_1, \ldots, w_n) = V$, und das war zu zeigen. □

Aufgabe 5.4.10 (Orthogonalpolynome) *Betrachte irgendein Skalarprodukt auf dem Raum $\mathbb{C}[t]$ aller Polynome mit komplexen Koeffizienten. Zeige: Es gibt ein Orthogonalsystem $(p_j, j \in \mathbb{N}_0)$ von Polynomen, wobei p_j den Grad j hat, und diese Polynome sind bis auf einen konstanten Faktor eindeutig bestimmt. Überprüfe für das Skalarprodukt*

$$\langle p, q \rangle \;=\; \int_{-1}^{1} p(t)\, q(t)\, dt\,,$$

dass diese zugehörigen Orthogonalpolynome *die Darstellung*

$$p_j(t) \;=\; \frac{1}{2^j\, j!}\, \frac{d^j}{dt^j}\, (t^2 - 1)^j \qquad \forall\; j \in \mathbb{N}_0$$

haben. Diese Orthogonalpolynome heißen auch Legendresche Polynome. *Es gibt viele andere wichtige Beispiele von Orthogonalpolynomen für andere innere Produkte auf $\mathbb{C}[t]$, die in der Physik und/oder Technik eine wichtige Rolle spielen.*

5.5 Beste Approximation, orthogonale Projektion und Fourierkoeffizienten

Definition 5.5.1 *Sei V ein Raum über $\mathbb{K}$ mit Skalarprodukt, und sei U ein Unterraum von V. Falls es zu $v \in V$ einen Vektor $u_0 \in U$ gibt, für welchen*

$$\|v \;-\; u_0\| \;\leq\; \|v - u\| \qquad \forall\, u \in U\,,$$

dann heißt u_0 beste Approximation *für V in dem Unterraum U. Falls es zu $v \in V$ einen Vektor $u_0 \in U$ gibt, für welchen $v - u_0 \in U^{\perp}$ ist, dann nennen wir u_0* orthogonale Projektion *von v auf den Unterraum U. Wir werden beweisen, dass für endlich-dimensionales U sowohl eine orthogonale Projektion als auch eine beste Approximation existieren und sogar übereinstimmen.*

Satz 5.5.2 (Beste Approximation)
Sei V ein Raum mit Skalarprodukt, und seien $v_1, \ldots, v_n$ ein Orthogonalsystem in V. Dann gilt für jedes $v \in V$:

(a) *Für $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ ist $\|v \;-\; \sum_{k=1}^{n} \alpha_k\, v_k\|$ genau dann minimal, wenn*

$$\alpha_k \;=\; \frac{\langle v_k, v\rangle}{\langle v_k, v_k\rangle} \;=\; \frac{\langle v_k, v\rangle}{\|v_k\|^2} \qquad \forall\; k \;=\; 1, \ldots, n\,. \tag{5.5.1}$$

(b) *Es gilt die* Besselsche Ungleichung

$$\sum_{k=1}^{n} \frac{|\langle v_k, v\rangle|^2}{\|v_k\|^2} \;\leq\; \|v\|^2\,. \tag{5.5.2}$$

(c) *Genau dann gilt die* Parsevalsche Gleichung

$$\sum_{k=1}^{n} \frac{|\langle v_k, v\rangle|^2}{\|v_k\|^2} \;=\; \|v\|^2\,, \tag{5.5.3}$$

wenn $v \in \mathcal{L}(v_1, \ldots, v_n)$ ist, und dann ist

$$v \;=\; \sum_{k=1}^{n} \frac{\langle v_k, v\rangle}{\|v_k\|^2}\; v_k\,.$$

Beweis: Es folgt mit den Regeln eines Skalarproduktes:

$$\begin{aligned} 0 \;\;\leq\;\; \|v \;-\; \sum_{k=1}^{n} \alpha_k\, v_k\|^2 \;&=\; \|v\|^2 - \sum_{k=1}^{n} \alpha_k\,\langle v, v_k\rangle - \sum_{k=1}^{n} \overline{\alpha_k}\,\langle v_k, v\rangle \\ &\qquad + \sum_{j=1}^{n}\sum_{k=1}^{n} \overline{\alpha_j}\,\alpha_k\,\langle v_j, v_k\rangle \\ =\;\; \|v\|^2 \;-\; \sum_{k=1}^{n} |\langle v_k, v\rangle|^2/\|v_k\|^2 \;&+\; \sum_{k=1}^{n} \Big|\alpha_k\,\|v_k\| - \;\langle v_k, v\rangle/\|v_k\|\,\Big|^2. \end{aligned}$$

Daraus liest man (a) ab. Setzt man speziell $\alpha_k = \langle v_k, v\rangle/\|v_k\|^2$ für $k = 1, \ldots, n$, so folgt (b). Falls $v \in \mathcal{L}(v_1, \ldots, v_n)$ ist, gibt es $\alpha_1, \ldots, \alpha_n$ derart, dass $0 = \|v - \sum_{k=1}^n \alpha_k v_k\|^2$ ist, und dann folgt aus (a) und (b), dass $\alpha_k = \langle v_k, v\rangle/\|v_k\|^2$ für $k = 1, \ldots, n$ gilt. Also muss die Parsevalsche Gleichung gelten. Falls umgekehrt die Parsevalsche Gleichung gilt, folgt wegen

$$\left\| v - \sum_{k=1}^n \frac{\langle v_k, v\rangle}{\|v_k\|^2} v_k \right\|^2 = \|v\|^2 - \sum_{k=1}^n |\langle v_k, v\rangle|^2/\|v_k\|^2 = 0$$

dass $v \in \mathcal{L}(v_1, \ldots, v_n)$ ist. □

Aufgabe 5.5.3 (Orthogonale Projektion mit MAPLE) *Finde selber heraus, was die folgenden Befehle mit der Berechnung der besten Approximation zu tun haben:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3, 2, [[1,1],[1,1],[-1, 0]]);
> a := Vector([1, 2, 3]);
> x := LeastSquares(A,a);
```

Definition 5.5.4 *Sei $(v_j, j \in J)$ ein beliebiges Orthogonalsystem in einem Raum mit Skalarprodukt V. Für ein $v \in V$ nennt man die Zahlen*

$$\frac{\langle v_j, v\rangle}{\|v_j\|^2} \qquad \forall\, j \in J$$

die Fourierkoeffizienten von v bzgl. des Orthogonalsystems $(v_j, j \in J)$.

Korollar zu Satz 5.5.2 (Orthogonale Projektion) *Sei V ein Raum mit Skalarprodukt, und sei U ein endlich-dimensionaler Unterraum von V. Dann gilt:*

(a) $V = U \oplus U^\perp$.

(b) *Für jedes $v \in V$ ist die beste Approximation von v in U gleich der orthogonalen Projektion von v auf U.*

(c) *Ist $u_1, \ldots, u_n$ eine Orthogonalbasis von U, so ist*

$$u = \sum_{k=1}^n \frac{\langle u_k, v\rangle}{\|u_k\|^2} u_k \tag{5.5.4}$$

die orthogonale Projektion von v auf U.

Beweis: Sei $v \in V$ und u wie in (5.5.4). Dann folgt $\langle u_j, u\rangle = \langle u_j, v\rangle$, also $\langle v - u, u_j\rangle = 0$ für alle $j = 1, \ldots, n$. Daher ist $v - u \in U^\perp$, und da dies für jedes $v \in V$ gilt, folgt $V = U + U^\perp$. Dass die Summe direkt ist, folgt bereits aus Satz 5.3.2. Also gilt (a), aber auch (c). Behauptung (b) folgt direkt aus dem Satz über die beste Approximation. □

Aufgabe 5.5.5 *Überprüfe, dass die Vektoren $v_1 = (1, 0, 1)^T$, $v_2 = (0, 1, 0)^T$ ein Orthogonalsystem in $\mathbb{R}^3$ mit dem kanonischen Skalarprodukt bilden. Berechne die orthogonale Projektion von $v = (1, 1, 0)^T$ auf den von v_1, v_2 aufgespannten Unterraum.*

Kapitel 6

Lineare Abbildungen

Wenn nichts anderes gesagt wird, seien im Folgenden V und W zwei beliebige Vektorräume über *demselben Körper* $\mathbb{K}$.

6.1 Definition und elementare Eigenschaften

Definition 6.1.1 *Eine Abbildung* $T: V \longrightarrow W,\ v \longmapsto T(v) = T\,v$ *heißt* lineare Abbildung von V nach W, *oder einfach* linear, *falls folgendes gilt:*

(L1) $\forall\, v_1, v_2 \in V: \quad T(v_1 + v_2) \;=\; T\,v_1 + T\,v_2.$ (Additivität)

(L2) $\forall\, \lambda \in \mathbb{K},\ v \in V: \quad T(\lambda\, v) \;=\; \lambda\, T\, v.$ (Homogenität)

Beide Aussagen zusammen sind offenbar äquivalent zu der einen Aussage

(L) $\forall\, \lambda_1, \lambda_2 \in \mathbb{K},\ v_1, v_2 \in V: \quad T(\lambda_1\, v_1 + \lambda_2\, v_2) \;=\; \lambda_1\, T\, v_1 + \lambda_2\, T\, v.$

Die Menge aller linearen Abbildungen von V nach W wird mit $L(V, W)$ bezeichnet. Falls $W = V$ ist, sei $L(V, V) = L(V)$ geschrieben. Abbildungen in $L(V)$ heißen auch Endomorphismen *in V.*

Beispiel 6.1.2 *Folgende Abbildungen sind linear:*

(a) *Für $A \in \mathbb{K}^{m\times n}$ und $x \in \mathbb{K}^n$ ist $A\,x \in \mathbb{K}^m$, und aus* Satz 2.3.3 (a), (b) *folgt, dass die Abbildung $x \longmapsto A\,x$ linear ist. Also legt jede Matrix A in* kanonischer Weise *eine lineare Abbildung fest, die wir der Einfachheit halber meistens ebenfalls mit A bezeichnen.*

(b) *Für beliebige Vektorräume V, W über $\mathbb{K}$ ist die* Nullabbildung $v \longmapsto 0$ *immer linear. Für $W = V$ ist auch die* identische Abbildung *oder die* Identität $v \longmapsto v$ *linear.*

(c) *Sei V endlich-dimensional, und sei $(v_1, \ldots, v_n)$ eine Basis von V. Dann gibt es nach* Proposition 1.5.6 *zu jedem $v \in V$ eindeutig bestimmte Zahlen $\alpha_k \in \mathbb{K}$ mit $v = \sum_{k=1}^{n} \alpha_k\, v_k$. Die Abbildung $v \longmapsto (\alpha_1, \ldots, \alpha_n)$ von V nach $\mathbb{K}^n$ ist linear und injektiv. Da für $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ der Vektor $v = \sum_{j=1}^{n} \alpha_j\, v_j$ zu V gehört, ist die Abbildung sogar surjektiv. Wir nennen sie auch* die Koordinatenabbildung *zur Basis $(v_1, \ldots, v_n)$.*

(d) *Für $V = W = \mathbb{K}[t]$ ist die Abbildung $p \longmapsto p'$, die jedem Polynom seine Ableitung zuordnet, linear, also ein Endomorphismus. Für beliebige $a < b$ ist die Abbildung $p \longmapsto \int_a^b p(t)\,dt$ linear von $\mathbb{K}[t]$ nach $\mathbb{K}$.*

Satz 6.1.3 *Für jedes $T \in L(V,W)$ gilt:*

(a) *Für jede Wahl von $\lambda_j \in \mathbb{K}$ und $v_j \in V$, $j = 1, \ldots, n$, gilt*

$$T\Big(\sum_{j=1}^{n} \lambda_j\, v_j\Big) \;=\; \sum_{j=1}^{n} \lambda_j\, T(v_j)\,.$$

(b) *Stets gilt $T(0) = 0$, wobei links bzw. rechts der Nullvektor von V bzw. W steht.*

(c) *Falls ein System $(v_j, j \in J)$ linear abhängig in V ist, dann ist das System $(T(v_j), j \in J)$ linear abhängig in W.*

Beweis: Zu (a): Folgt durch Induktion über n aus der Definition der linearen Abbildung. Zu (b): $0 = 0+0$, also ist $T(0) = T(0)+T(0)$, und daraus folgt $T(0) = 0$. Zu (c): Nach Definition der linearen Abhängigkeit gibt es $j_1, \ldots, j_n \in J$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, wobei nicht alle λ_k verschwinden, so dass $0 = \sum_{k=1}^{n} \lambda_k\, v_{j_k}$ ist. Wegen (a) und (b) folgt hieraus $0 = \sum_{k=1}^{n} \lambda_k\, T\big(v_{j_k}\big)$, woraus sich die Behauptung ergibt. □

Aufgabe 6.1.4 *Funktionen $f : D \longrightarrow \mathbb{R}$, mit $D \subset \mathbb{R}$, kann man sich über ihren Graphen in $\mathbb{R}^2$ veranschaulichen. Finde heraus, wie man an diesem Graphen erkennt, ob eine solche Funktion eine lineare Abbildung in obigem Sinne ist.*

6.2 Kern und Bild

Definition 6.2.1 *Für $T \in L(V,W)$ heißen* $\operatorname{Kern} T = \{v \in V :\ T\,v = 0\}$ *der* Kern von T, *und* $\operatorname{Bild} T = \{w \in W :\ \exists v \in V \text{mit } T\,v = w\}$ *das* Bild *von T. Für $A \subset V$ sei $T(A) = \{T\,v :\ v \in A\}$ gesetzt. Insbesondere ist also $T(V) = \operatorname{Bild} T$.*

Satz 6.2.2 *Für jedes $T \in L(V,W)$ gilt:*

(a) $\operatorname{Kern} T$ *ist ein Unterraum von V,* $\operatorname{Bild} T$ *ist ein Unterraum von W.*

(b) $$\dim \operatorname{Kern} T \;+\; \dim \operatorname{Bild} T \;=\; \dim V.$$

(c) *Genau dann ist T injektiv, wenn* $\operatorname{Kern} T = \{0\}$.

(d) *Falls W endlich-dimensional ist, dann ist T genau dann surjektiv, wenn* $\dim \operatorname{Bild} T = \dim W$ *ist.*

Beweis: Zu (a): Für $v_1, v_2 \in \operatorname{Kern} T$ und $\lambda_1, \lambda_2 \in \mathbb{K}$ gilt $T(\lambda_1\, v_1 + \lambda_2\, v_2) = \lambda_1\, T\, v_1 + \lambda_2\, T\, v_2 = 0$, da $T\,v_1 = T\,v_2 = 0$ ist. Also folgt $\lambda_1\, v_1 + \lambda_2\, v_2 \in \operatorname{Kern} T$. Für $w_1, w_2 \in \operatorname{Bild} T$ gibt es $v_1, v_2 \in V$ mit $T\,v_j = w_j$, und für λ_1, λ_2 wie oben folgt $\lambda_1\, w_1 + \lambda_2\, w_2 = T(\lambda_1\, v_1 + \lambda_2\, v_2) \in \operatorname{Bild} T$. Zu (b): Falls $\dim \operatorname{Kern} T = \infty$ ist, muss auch V unendlich-dimensional sein, und es ist nichts zu zeigen. Falls $\dim \operatorname{Bild} T = \infty$ ist, gibt es ein abzählbar-unendliches System $(w_j, j \in \mathbb{N})$ in $\operatorname{Bild} T$, welches linear unabhängig ist, und zu jedem $j \in \mathbb{N}$ existiert ein $v_j \in V$ mit $T\,v_j = w_j$. Nach Satz 6.1.3 (c) muss $(v_j, j \in \mathbb{N})$ in V linear unabhängig sein, und deshalb ist auch $\dim V = \infty$, so dass nichts mehr zu zeigen ist. Seien jetzt $\operatorname{Kern} T$ und $\operatorname{Bild} T$ beide endlich-dimensional, und sei $(v_1, \ldots, v_n)$ bzw. $(w_1, \ldots, w_m)$ Basis von $\operatorname{Kern} T$ bzw. $\operatorname{Bild} T$. Zu jedem w_j existiert ein Urbild in V, welches wir mit v_{n+j} bezeichnen. Der Beweis von (b) ist erbracht, wenn wir nachweisen, dass $(v_1, \ldots, v_{n+m})$ eine Basis von V ist. Dazu seien $\lambda_j \in \mathbb{K}$ so, dass $0 = \sum_{j=1}^{n+m} \lambda_j\, v_j$ gilt. Dann folgt wegen $T\,v_j = 0$ für $j = 1, \ldots, n$, dass $0 = \sum_{j=1}^{m} \lambda_{n+j}\, w_j$ ist. Wegen der linearen Unabhängigkeit von $(w_1, \ldots, w_m)$ folgt hieraus $\lambda_k = 0$ für $k = n+1, \ldots, n+m$. Also ergibt sich $0 = \sum_{j=1}^{n} \lambda_j\, v_j$, woraus aber wegen der linearen Unabhängigkeit von $(v_1, \ldots, v_n)$ folgt, dass auch $\lambda_k = 0$ für $k = 1, \ldots, n$ gilt. Das zeigt insgesamt die lineare Unabhängigkeit von $(v_1, \ldots, v_{n+m})$. Sei jetzt $v \in V$, und sei $w = T\,v$. Dann gibt es

$\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ mit $w = \sum_{j=1}^m \alpha_j\, w_j$. Für $\tilde v = v - \sum_{j=1}^m \alpha_j\, v_{n+j}$ folgt dann $T\,\tilde v = w - \sum_{j=1}^m \alpha_j\, w_j = 0$, und somit ist $\tilde v \in \operatorname{Kern} T$. Daher gibt es $\beta_1, \ldots, \beta_n \in \mathbb{K}$ mit $\tilde v = \sum_{j=1}^n \beta_j\, v_j$, und somit ist gezeigt, dass $(v_1, \ldots, v_{n+m})$ ein Erzeugendensystem für V ist. Die Beweise der Teile (c) und (d) stellen wir als Übungsaufgabe. □

Aufgabe 6.2.3 *Beweise die Teile* (c), (d) *der vorausgegangenen Satzes.*

Aufgabe 6.2.4 *Bestimme Kern und Bild der Abbildungen aus* Beispiel 6.1.2.

Definition 6.2.5 *Die Dimension des Bildes von $T \in L(V, W)$ heißt auch* der Rang von T. *Für $w \in W$ schreiben wir $T^{-1}(w)$ für die Menge aller $v \in V$ mit $T\,v = w$. Also besteht $T^{-1}(w)$ aus allen Lösungen der Gleichung $T\,v = w$, und $T^{-1}(0)$ ist gerade der Kern von T. Beachte aber, dass diese Bezeichung nicht etwa impliziert, dass T injektiv oder sogar bijektiv ist!*

Aufgabe 6.2.6 *Eine Matrix $A \in \mathbb{K}^{m\times n}$ legt nach* Beispiel 6.1.2 *durch $x \longmapsto A\,x$ eine lineare Abbildung T_A von $\mathbb{K}^n$ in $\mathbb{K}^m$ fest. Zeige dass der Rang dieser linearen Abbildung T_A dasselbe ist wie der Rang der Matrix A.*

Satz 6.2.7 *Sei $T \in L(V, W)$, und sei $w \in W$ fest gewählt. Dann gilt:*

(a) *Genau dann ist $T^{-1}(w) \neq \emptyset$, wenn $w \in \operatorname{Bild} T$ ist.*

(b) *Für $w \in \operatorname{Bild} T$ ist $T^{-1}(w)$ eine lineare Mannigfaltigkeit; genauer: Ist $v_0 \in T^{-1}(w)$ beliebig ausgewählt, so ist*

$$T^{-1}(w) \;=\; v_0 \;+\; \operatorname{Kern} T\,.$$

Insbesondere ist die Dimension von $T^{-1}(w)$ immer gleich der Dimension von $\operatorname{Kern} T$.

Beweis: Aussage (a) ist klar nach Definition von $T^{-1}(w)$. Für (b) sei $v \in \operatorname{Kern} T$; dann ist $T\,(v_0 + v) = T\,v_0 + T\,v \;=\; w + 0 \;=\; w$, und deshalb ist $v + v_0 \in T^{-1}(w)$. Umgekehrt sei $v \in T^{-1}(w)$, dann ist $T\,(v - v_0) = T\,v - T\,v_0 = w - w = 0$, also $v - v_0 \in \operatorname{Kern} T$. □

Aufgabe 6.2.8 *Vergleiche die Ergebnisse dieses Abschnittes mit denen aus* Abschnitt 3.2 *über die Struktur der Lösungsmenge eines linearen Gleichungssystems.*

6.3 Lineare Abbildungen und Basen

Besonders in unendlich-dimensionalen Vektorräumen ist es nicht sofort klar, ob und wieviele lineare Abbildungen es gibt. Wir wollen die Frage für den Fall beantworten, dass wir eine Basis des Urbildraums V kennen und zeigen, dass eine lineare Abbildung durch Angabe ihrer Bilder der Basisvektoren eindeutig festgelegt ist, dass man aber andererseits diese Bilder *beliebig vorgeben kann.*

Satz 6.3.1 (Existenz von linearen Abbildungen) *Sei $(v_j, j \in J)$ eine Basis von V, und sei $(w_j, j \in J)$ ein System von Vektoren aus W, welches also „von der gleichen Größe“ wie die Basis, aber sonst beliebig ist. Dann gibt es genau ein $T \in L(V, W)$ mit*

$$T\,v_j \;=\; w_j \qquad \forall\; j \in J\,. \tag{6.3.1}$$

Beweis: Ist $v \in V$ gegeben, so gibt es eindeutig bestimmte $\alpha_j \in \mathbb{K}$ mit $v = \sum_{j \in J} \alpha_j v_j$, wobei nur endlich viele α_j nicht gleich 0 sind. Wenn $T \in L(V, W)$ und (6.3.1) erfüllt ist, dann muss gelten $T v = \sum_{j \in J} \alpha_j T v_j = \sum_{j \in J} \alpha_j w_j$. Umgekehrt definieren wir jetzt eine Abbildung T von V nach W durch $T v = \sum_{j \in J} \alpha_j w_j$. Für $\tilde{v} = \sum_{j \in J} \tilde{\alpha}_j v_j \in V$ und $\alpha, \beta \in \mathbb{K}$ gilt dann $\alpha v + \beta \tilde{v} = \sum_{j \in J} \big(\alpha \alpha_j + \beta \tilde{\alpha}_j\big) v_j$, und deshalb folgt $T(\alpha v + \beta \tilde{v}) = \sum_{j \in J} \big(\alpha \alpha_j + \beta \tilde{\alpha}_j\big) w_j = \alpha T v + \beta T \tilde{v}$. Also ist die Abbildung T linear. $\square$

Aufgabe 6.3.2 *Zeige: Ist* $\dim V = n < \infty$ *und* $\dim W = m < \infty$*, so folgt* $\dim L(V, W) = n\, m$.

Definition 6.3.3 *Eine bijektive lineare Abbildung* $T : V \longrightarrow W$ *heißt* ein Isomorphismus. *Falls ein solcher Isomorphismus existiert, so heißen die Vektorräume* V *und* W *auch* isomorph. *Falls* $W = V$ *ist, sagen wir statt Isomorphismus auch* Automorphismus.

Korollar zu Satz 6.3.1 *Unter den Voraussetzungen von* Satz 6.3.1 *ist die lineare Abbildung* T *genau dann ein Isomorphismus. wenn* $(w_j, j \in J)$ *eine Basis von* W *ist.*

Beweis: Sei $(w_j, j \in J)$ eine Basis. Dann gibt es zu $w \in W$ eindeutig bestimmte $\alpha_j \in \mathbb{K}$ mit $w = \sum_{j \in J} \alpha_j w_j$, wobei nur endlich viele α_j nicht gleich 0 sind. Also gibt es genau ein $v \in V$, und zwar $v = \sum_{j \in J} \alpha_j v_j$, für welches $w = T v$ gilt. Deshalb ist T ein Isomorphismus. Umgekehrt, wenn T ein Isomorphismus ist, gibt es zu jedem $w \in W$ genau ein Urbild $v \in V$ mit $T v = w$. Zu diesem v gibt es eindeutig bestimmte $\alpha_j \in \mathbb{K}$ mit $v = \sum_{j \in J} \alpha_j v_j$, wobei nur endlich viele α_j nicht gleich 0 sind, und es folgt $w = \sum_{j \in J} \alpha_j w_j$. Daher ist $(w_j, j \in J)$ ein Erzeugendensystem. Setzt man $0 = \sum_{j \in J} \alpha_j w_j$ und beachtet dass $T v = 0$ genau für $v = 0$ gilt, so folgt analog die lineare Unabhängigkeit des Systems $(w_j, j \in J)$. $\square$

Beispiel 6.3.4 *Sei* V *ein* n*-dimensionaler Vektorraum über* $\mathbb{K}$*, und sei das System* $(v_1, \ldots, v_n)$ *eine Basis von* V*. Die Koordinatenabbildung*

$$v = \sum_{j=1}^{n} \alpha_j v_j \quad \longmapsto \quad (\alpha_1, \ldots, \alpha_n)^T$$

ist nach Aufgabe 6.1.2 *ein Isomorphismus von* V *auf* $\mathbb{K}^n$*. Also ist jeder* n*-dimensionale Vektorraum zu* $\mathbb{K}^n$ *isomorph, und die Koordinatenabbildung heißt auch* kanonischer Isomorphismus. *Beachte aber, dass man erst von* dem *kanonischen Isomorphismus sprechen kann, nachdem man eine Basis für* V *gewählt hat.*

Proposition 6.3.5 *Eine Hintereinanderausführung zweier linearer Abbildungen ist ebenfalls linear. Die Umkehrabbildung eines Isomorphismus ist wieder ein Isomorphismus.*

Beweis: Seien U, V, W Vektorräume über $\mathbb{K}$, und seien $T_1 \in L(U, V)$, $T_2 \in L(V, W)$, sowie $T = T_2 \circ T_1$. Für $u_1, u_2 \in U$, $\lambda_1, \lambda_2 \in \mathbb{K}$ gilt dann

$$\begin{aligned} T(\lambda_1 v_1 + \lambda_2 v_2) &= T_2\big(T_1(\lambda_1 v_1 + \lambda_2 v_2)\big) \;=\; T_2\big(\lambda_1 T(v_1) + \lambda_2 T(v_2)\big) \\ &= \lambda_1 T_2\big(T_1(v_1)\big) + \lambda_2 T_2\big(T_1(v_2)\big) \;=\; \lambda_1 T(v_1) + \lambda_2 T(v_2)\,. \end{aligned}$$

Also ist $T \in L(U, W)$. Sei jetzt $T \in L(V, W)$ ein Isomorphismus, also bijektiv. Für $v_1, v_2 \in V$, $\lambda_1, \lambda_2 \in \mathbb{K}$ gilt dann

$$T(\lambda_1 v_1 + \lambda_2 v_2) \;=\; \lambda_1 T(v_1) + \lambda_2 T(v_2)\,.$$

Mit $w_1 = T(v_1)$, $w_2 = T(v_2)$, also $v_1 = T^{-1}(w_1)$, $v_2 = T^{-1}(w_2)$, folgt aus dieser Gleichung durch Anwendung von T^{-1} auf beide Seiten

$$T^{-1}(\lambda_1\, w_1 + \lambda_2\, w_2) \;=\; \lambda_1\, v_1 + \lambda_2\, v_2 \;=\; \lambda_1\, T^{-1}(w_1) + \lambda_2\, T^{-1}(w_2)\,.$$

Also ist T^{-1} linear. □

Aufgabe 6.3.6 *Überprüfe, dass der Begriff der Isomorphie eine Äquivalenzrelation auf der Menge aller Vektorräume ist.*

Satz 6.3.7 *Sei V ein Raum über $\mathbb{K}$ mit Skalarprodukt, und sei W ein zu V isomorpher Vektorraum über $\mathbb{K}$. Für einen Isomorphismus $T \in L(V, W)$ definieren wir*

$$\langle w_1, w_2\rangle_T \;=\; \langle T^{-1}\, w_1, T^{-1}\, w_2\rangle \qquad \forall\; w_1, w_2 \in W\,.$$

Dann ist $\langle\,\cdot\,,\,\cdot\,\rangle_T$ ein inneres Produkt auf W.

Beweis: Für $w_1, w_2 \in W$ und $\lambda \in \mathbb{K}$ gilt wegen der Linearität von T^{-1}, dass $T^{-1}(w_1 + w_2) = T^{-1}\, w_1 + T^{-1}\, w_2$ und $T^{-1}(\lambda\, w_2) = \lambda\, T^{-1}\, w_2$ ist, und daraus folgen leicht die vier Axiome des inneren Produktes für $\langle\,\cdot\,,\,\cdot\,\rangle_T$. □

6.4 Isomorphie und Dimension

Satz 6.4.1 *Für beliebige Vektorräume V und W über dem gleichen Körper $\mathbb{K}$ gilt:*

(a) *Falls es ein injektives $T \in L(V, W)$ gibt, folgt* $\dim W \geq \dim V$.

(b) *Falls V und W isomorph sind, folgt* $\dim V = \dim W$.

(c) *Falls* $\dim V = \dim W < \infty$ *ist, sind V und W isomorph.*

Beweis: Zu (a): Sei ein $T \in L(V, W)$ injektiv, und sei $(v_j, j \in J)$ ein linear unabhängiges System in V. Dann folgt für beliebige $j_1, \ldots, j_n \in J$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, dass $0 = \sum_{k=1}^n \lambda_k\, T\, v_{j_k} = T\big(\sum_{k=1}^n \lambda_k\, v_{j_k}\big)$ genau dann gilt, wenn $0 = \sum_{k=1}^n \lambda_k\, v_{j_k}$ ist, denn der Kern von T besteht nur aus dem Nullvektor. Daraus folgt, dass alle $\lambda_k = 0$ sein müssen. Also ist das System der Bilder $(T\, v_j, j \in J)$ linear unabhängig in W. Wenn V unendlich-dimensional ist, gibt es ein abzählbar-unendliches System in V, und dann auch in W, und somit ist auch W unendlich-dimensional. Im anderen Fall gibt es in V eine Basis $(v_1, \ldots, v_n)$, mit $n = \dim V$, und dann ist $(T\, v_1, \ldots, T\, v_n)$ jedenfalls linear unabhängig in W. Also ist $\dim W \geq n$. Zu (b): Sei $T \in L(V, W)$ ein Isomorphismus. Dann folgt aus (a), dass $\dim W \geq \dim V$ ist. Da $T^{-1} \in L(W, V)$ ebenfalls injektiv ist, folgt auch die umgekehrte Ungleichung. Zu (c): Wähle Basen von V und W; diese enthalten nach Voraussetzung gleich viel Vektoren. Aus dem Korollar zu Satz 6.3.1 folgt dann die Behauptung. □

Wegen der Wichtigkeit formulieren wir noch folgende direkte Konsequenz dieses Satzes:

Korollar zu Satz 6.4.1 *Zwei endlich-dimensionale Vektorräume über $\mathbb{K}$ sind genau dann isomorph, wenn sie dieselbe Dimension haben.*

Aufgabe 6.4.2 *Sei* $\dim V = \dim W < \infty$. *Zeige: Wenn ein $T \in L(V, W)$ surjektiv bzw. injektiv ist, dann ist es bereits ein Isomorphismus.*

6.5 Der Dualraum, die duale Abbildung

Definition 6.5.1 *Für jeden Vektorraum V über $\mathbb{K}$ heißt $L(V,\mathbb{K})$, also die Menge aller linearen Abbildungen von V in den Körper $\mathbb{K}$, der* zu V (algebraisch) duale Raum*, oder kurz* der Dualraum *zu V, und wir schreiben V^* für diesen Raum. Jedes $f \in V^*$ heißt auch* lineares Funktional *oder* Linearform *auf V. Für ein $T \in L(V,W)$ und $f \in W^*$ ist $f\circ T \in V^*$. Die Abbildung $f \longmapsto f\circ T$ bildet also W^* nach V^* ab und wird mit T^* bezeichnet und die* zu T duale Abbildung *genannt.*

Aufgabe 6.5.2 *Sei V endlich-dimensional. Zeige: Für jedes $v \in V\setminus\{0\}$ gibt es ein $f \in V^*$ mit $f(v) \neq 0$.*

Lemma 6.5.3 *Für jedes $T \in L(V,W)$ ist die duale Abbildung T^* linear, also $T^* \in L(W^*,V^*)$.*

Beweis: Für $f_1, f_2 \in W^*$ und $\alpha_1, \alpha_2 \in V$ gilt nach Definition von T^*

$$T^*(\alpha_1 f_1 + \alpha_2 f_2) \;=\; (\alpha_1 f_1 + \alpha_2 f_2)\circ T \;=\; \alpha_1 T^* f_1 + \alpha_2 T^* f_2\,,$$

und das war zu zeigen. □

Satz 6.5.4 *Wenn V endlich-dimensional ist, gilt* $\dim V = \dim V^*$*, und insbesondere sind V und V^* isomorph.*

Beweis: Wenn $\dim V = 0$ ist, wenn also V nur aus dem Nullvektor besteht, dann ist die einzige lineare Abbildung von V in $\mathbb{K}$ die Nullabbildung. Also gilt in diesem Fall die Behauptung. Sei jetzt $1 \leq \dim V = n < \infty$, und sei $(v_1,\ldots,v_n)$ eine Basis von V. Aus Satz 6.3.1 folgt die Existenz von $f_1,\ldots,f_n \in V^*$ mit der Eigenschaft $f_k(v_j) = \delta_{jk}$ für $1 \leq j,k \leq n$. Für ein beliebiges $f \in V^*$ seien dann $f(v_j) = \alpha_j$ für $j = 1,\ldots,n$. Für $g = f - \sum_{j=1}^n \alpha_j f_j$ folgt dann $f(v_k) = 0$ für $k = 1,\ldots,n$, und deshalb ist g die Nullabbildung. Also ist $(f_1,\ldots,f_n)$ Erzeugendensystem von V^*. Für die lineare Unabhängigkeit sei $0 = \sum_{j=1}^n \alpha_j f_j$, d. h. genauer, $0 = \sum_{j=1}^n \alpha_j f_j(v)$ für alle $v \in V$. Einsetzen von v_k zeigt dann $0 = \lambda_k$, und daher ist $(f_1,\ldots,f_n)$ Basis von V^*. □

Satz 6.5.5 *Wenn V und W beide endlich-dimensional sind, dann ist die Abbildung $T \longmapsto T^*$ ein Isomorphismus von $L(V,W)$ auf $L(W^*,V^*)$.*

Beweis: Für $T_1, T_2 \in L(V,W)$ und $\alpha_1, \alpha_2 \in \mathbb{K}$ sei $T = \alpha_1 T_1 + \alpha_2 T_2$. Dann gilt für jedes $f \in W^*$:

$$f\circ T \;=\; \alpha_1 f\circ T_1 \;+\; \alpha_2 f\circ T_2 \;=\; \alpha_1 T_1^*(f) \;+\; \alpha_2 T_2^*(f)\,,$$

und deshalb ist $T^* = \alpha_1 T_1^* + \alpha_2 T_2^*$. Also ist $T \longmapsto T^*$ eine lineare Abbildung von $L(V,W)$ in $L(W^*,V^*)$. Wenn $T^* = 0$, d. h., gleich der Nullabbildung ist, dann folgt per Definition $f(T(v)) = 0$ für alle $f \in W^*$ und alle $v \in V$. Dies impliziert aber nach Aufgabe 6.5.2 dass $T(v) = 0$ ist für alle $v \in V$, und daher ist T die Nullabbildung in $L(V,W)$. Also ist die Abbildung $T \longmapsto T^*$ injektiv. Aus Aufgabe 6.3.2 folgt aber dass $L(V,W)$ und $L(W^*,V^*)$ die gleiche Dimension haben, und deshalb ist $T \longmapsto T^*$ auch surjektiv. □

6.6 Die adjungierte Abbildung

Im Folgenden betrachten wir endlich-dimensionale Vektorräume mit inneren Produkten und führen die sogenannte adjungierte Abbildung ein. Diese hängt eng mit der dualen Abbildung zusammen.

Definition 6.6.1 *Sei V und W Vektorräume über $\mathbb{K}$. Eine Abbildung $T : V \longrightarrow W$ heißt* semilinear, *wenn*

- $\forall\, v_1, v_2 \in V : \quad T(v_1 + v_2) \;=\; T\,v_1 + T\,v_2.$
- $\forall\, \lambda \in \mathbb{K}\,,\; v \in V : \quad T(\lambda\, v) \;=\; \overline{\lambda}\, T\, v.$

Im Fall $\mathbb{K} = \mathbb{R}$ ist also Semilinearität das gleiche wie Linearität, aber nicht wenn $\mathbb{K} = \mathbb{C}$ ist.

Lemma 6.6.2 *Sei V ein endlich-dimensionaler Raum über $\mathbb{K}$ mit Skalarprodukt. Dann gibt es zu jedem $f \in V^*$ genau ein $w_f \in V$ so, dass*

$$f(v) \;=\; \langle w_f, v\rangle \qquad \forall\, v \in V\,.$$

Die Abbildung $f \longmapsto w_f$ ist bijektiv und semilinear.

Beweis: Sei $(v_1, \ldots, v_n)$ eine Orthonormalbasis von V. Dann hat jedes $v \in V$ die Darstellung $v = \sum_{k=1}^{n} \langle v_k, v\rangle\, v_k$. Für $f \in V^*$ setzen wir $\alpha_k = f(v_k)$ für $k = 1, \ldots, n$. Dann folgt $f(v) = \sum_{k=1}^{n} \alpha_k\, \langle v_k, v\rangle = \langle w_f, v\rangle$ mit $w_f = \sum_{k=1}^{n} \overline{\alpha_k}\, v_k$. Also gibt es ein w_f wie behauptet. Dieses w_f ist durch f eindeutig festgelegt, denn aus $\langle w_1 - w_2, v\rangle = 0$ für alle $v \in V$ folgt $w_1 = w_2$. Die Semilinearität von $f \longmapsto w_f = \sum_{k=1}^{n} \overline{\alpha_k}\, v_k$ kann leicht nachgeprüft werden. □

Bemerkung 6.6.3 *Wir betrachten jetzt zwei endlich-dimensionale Räume V und W über $\mathbb{K}$ mit Skalarprodukten, sowie ein $T \in L(V, W)$. Die duale Abbildung T^* bildet dann W^* nach V^* ab. Zu $f \in W^*$ gehört dann nach* Lemma 6.6.2 *genau ein $w_f \in W$, und zu $g = T^* f \in V^*$ gibt es aus demselben Grund ein $w_g \in V$. Wir können also eine Abbildung $T^{ad} : W \longrightarrow V$ definieren durch*

$$T^{ad}(w_f) \;=\; w_g\,, \quad g = T^* f\,, \qquad \forall\, f \in W^*\,.$$

Die Abbildungen $f \longmapsto w_f$ und $g \longmapsto w_g$ sind beide semilinear, aber die Abbildung $w_f \longmapsto T^{ad}(w_f) = w_g$ ist linear, wie man leicht überprüft. Mit Hilfe der Definition von T^* folgt für die Abbildung T^{ad} die Beziehung

$$\langle T^{ad}\, w, v\rangle \;=\; \langle w, T\, v\rangle \qquad \forall\, v \in V\,,\; w \in W\,. \tag{6.6.1}$$

Wir bemerken noch, dass die Zuordnung $T \longmapsto T^{ad}$ eine bijektive semilineare Abbildung von $L(V, W)$ auf $L(W, V)$ ist.

Definition 6.6.4 *Seien V und W Räume über $\mathbb{K}$ mit Skalarprodukten, und sei $T \in L(V, W)$. Eine Abbildung $T^{ad} \in L(W, V)$ heißt* zu T adjungiert, *falls* (6.6.1) *gilt. Es folgt leicht, dass es zu T höchstens eine adjungierte Abbildung geben kann, und aus der obigen Bemerkung folgt die Existenz von T^{ad} falls V und W beide endlich-dimensional sind. Ein Endomorphismus $T \in L(V)$ heißt* selbstadjungiert, *falls $T^{ad} = T$ ist, falls also immer gilt*

$$\langle T\, v_1, v_2\rangle \;=\; \langle v_1, T\, v_2\rangle \qquad \forall\, v_1, v_2 \in V\,.$$

Wir nennen einen Endomorphismus $T \in L(V)$ auch normal, *falls die adjungierte Abbildung T^{ad} existiert und mit T kommutiert, d. h., falls gilt*

$$T^{ad} \circ T \;=\; T \circ T^{ad}\,.$$

Offenbar ist jeder selbstadjungierte Endomorphismus auch normal.

Aufgabe 6.6.5 (Rechenregeln für adjungierte Abbildungen)
Seien V und W Räume über $\mathbb{K}$ mit Skalarprodukten. Zeige folgende Regeln:

(a) *Wenn zu $T \in L(V,W)$ eine adjungierte Abbildung $T^{ad} \in L(W,V)$ existiert, dann ist die zu T^{ad} adjungierte Abbildung wieder gleich T.*

(b) *Seien $\lambda_1, \lambda_2 \in \mathbb{K}$ und $T_1, T_2 \in L(V,W)$. Wenn zu T_1 bzw. T_2 adjungierte Abbildungen $T_1^{ad}, T_2^{ad} \in L(W,V)$ existieren, dann existiert auch die adjungierte Abbildung für $\lambda_1 T_1 + \lambda_2 T_2$, und*

$$\big(\lambda_1 T_1 + \lambda_2 T_2\big)^{ad} \;=\; \overline{\lambda_1}\, T_1^{ad} + \overline{\lambda_2}\, T_2^{ad}\,.$$

(c) *Die identische Abbildung id sowie die Nullabbildung auf V sind selbstadjungiert.*

Für die Hintereinanderausführung von Abbildungen vergleiche die folgende Proposition.

Proposition 6.6.6 *Seien U, V, W Räume über $\mathbb{K}$ mit Skalarprodukten, und seien $T_1 \in L(U,V)$ und $T_2 \in L(V,W)$ mit den adjungierten Abbildungen $T_1^{ad} \in L(V,U)$ und $T_2^{ad} \in L(W,V)$. Dann ist $T_1^{ad} \circ T_2^{ad}$ die adjungierte Abbildung zu $T_2 \circ T_1$.*

Beweis: Für $u \in U$, $v \in V$, $w \in W$ gilt nach Definition der adjungierten Abbildungen

$$\langle T_1^{ad}\, v, u\rangle \;=\; \langle v, T_1\, u\rangle\,, \qquad \langle T_2^{ad}\, w, v\rangle \;=\; \langle w, T_2\, v\rangle\,.$$

Setzt man links $v = T_2^{ad}\, w$, so folgt $\langle T_1^{ad}\, T_2^{ad}\, w, u\rangle = \langle T_2^{ad}\, w, T_1\, u\rangle$, und wenn man dann in der zweiten Gleichung $v = T_1\, u$ einsetzt, so folgt $\langle T_2^{ad}\, w, T_1\, u\rangle = \langle w, T_2\, T_1\, u\rangle$, und das war zu zeigen. □

Aufgabe 6.6.7 *Sei V ein Raum über $\mathbb{K}$ mit Skalarprodukt, und sei $T \in L(V)$ selbstadjungiert. Zeige: $\langle v, T\, v\rangle$ ist immer eine reelle Zahl, auch wenn $\mathbb{K} = \mathbb{C}$ ist.*

Aufgabe 6.6.8 *Zeige: Für $V = \mathbb{K}^n$ mit dem kanonischen Skalarprodukt, und $A \in \mathbb{K}^{n \times n}$ ist die Abbildung $x \longmapsto A\, x$ genau dann selbstadjungiert, wenn $\overline{A}^T = A$ ist.*

6.7 Längentreue Abbildungen

Definition 6.7.1 *Seien V, W Räume über $\mathbb{K}$ mit Skalarprodukten. Ein $T \in L(V,W)$ heißt* längentreu, *falls*

$$\forall\ v \in V: \quad \|T\, v\| \;=\; \|v\|\,,$$

und längen- und winkeltreu, *falls*

$$\forall\ v_1, v_2 \in V: \quad \langle T\, v_1, T\, v_2\rangle \;=\; \langle v_1, v_2\rangle\,.$$

Beachte, dass in beiden Bedingungen rechts die Norm bzw. das Skalarprodukt in V, links die Norm bzw. das Skalarprodukt in W steht. Offenbar ist jede längen- und winkeltreue Abbildung auch längentreu. Dass auch die Umkehrung dieser Aussage gilt, folgt aus dem nächsten Satz. Für $\mathbb{K} = \mathbb{R}$ bzw. $\mathbb{K} = \mathbb{C}$ nennt man eine längen- und winkeltreue Abbildung auch orthogonal *bzw.* unitär.

Satz 6.7.2 *Seien V, W Räume über $\mathbb{K}$ mit Skalarprodukten. Dann ist jedes längentreue $T \in L(V,W)$ auch injektiv und längen- und winkeltreu.*

Beweis: Die Injektivität folgt leicht, weil aus $0 = T\, v$ folgt dass $0 = \|T\, v\|$ ist, und bei Längentreue gilt dies genau dann, wenn $\|v\| = 0$ ist. Die Winkeltreue folgt aus Aufgabe 5.2.2. □

Satz 6.7.3 *Seien V, W Räume über $\mathbb{K}$ mit Skalarprodukten, und sei $(v_j, j \in J)$ eine Orthonormalbasis von V. Genau dann ist ein $T \in L(V, W)$ längentreu, wenn das System $(T\,v_j, j \in J)$ wieder ein Orthonormalsystem ist.*

Beweis: Sei T längentreu. Dann folgt für $j, k \in J$, dass $\langle T\,v_j, T\,v_k\rangle = \langle v_j, v_k\rangle = \delta_{jk}$ ist, und deshalb ist $(T\,v_j, j \in J)$ ein Orthonormalsystem. Umgekehrt, sei $(T\,v_j, j \in J)$ ein Orthonormalsystem. Für $v \in V$ gibt es $j_1, \ldots, j_n \in J$ und $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ mit $v = \sum_{k=1}^n \alpha_k\, v_{j_k}$, und aus (5.4.1) folgt $\|v\|^2 = \sum_{k=1}^n |\alpha_k|^2$. Für $w = T\,v = \sum_{k=1}^n \alpha_k\, T\,v_{j_k}$ folgt genauso $\|T\,v\|^2 = \sum_{k=1}^n |\alpha_k|^2$, und deshalb ist T längentreu. □

Satz 6.7.4 *Seien V, W Räume über $\mathbb{K}$ mit Skalarprodukten. Dann besitzt jedes längentreue und surjektive $T \in L(V, W)$ eine adjungierte Abbildung, und $T^{ad} = T^{-1}$. Im Fall $W = V$ ist T auch normal.*

Beweis: Eine längentreue und surjektive Abbildung ist bijektiv, und somit existiert $T^{-1} \in L(W, V)$. Mit $T\,v_1 = w_1$, also $v_1 = T^{-1}\,w_1$, folgt aus der Definition der Winkeltreue (was ja dasselbe wie die Längentreue ist), dass

$$\langle w_1, T\,v_2\rangle \;=\; \langle T^{-1}\,w_1, v_2\rangle \qquad \forall\; w_1 \in W\,,\;\; v_2 \in V\,.$$

Bis auf die Bezeichnungen ist dies gleich (6.6.1), wenn man T^{ad} durch T^{-1} ersetzt, und das war zu zeigen, denn die Normalität für $W = V$ folgt, da $T \circ T^{-1} = T^{-1} \circ T$ ist. □

Aufgabe 6.7.5 *Zeige: Für eine Matrix $A \in \mathbb{K}^{m\times n}$ ist die Abbildung $x \longmapsto A\,x$ genau dann längentreu von $\mathbb{K}^n$ in $\mathbb{K}^m$, wenn die Spalten von A ein Orthonormalsystem in $\mathbb{K}^m$ bezüglich des kanonischen Skalarprodukts bilden. Warum folgt daraus $m \geq n$?*

Aufgabe 6.7.6 *Seien jetzt $V = W = \mathbb{R}^2$ mit dem kanonischen Skalarprodukt, und sei $A \in \mathbb{R}^{2\times 2}$. Zeige: Die lineare Abbildung $x \longmapsto A\,x$ von $\mathbb{R}^2$ in sich ist genau dann längentreu, wenn es ein $\phi \in \mathbb{R}$ gibt, für welches*

$$A \;=\; \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \qquad \text{oder} \qquad A \;=\; \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix}\,.$$

Wodurch unterscheiden sich diese beiden Fälle? Interpretiere beide Fälle geometrisch.

6.8 Definite Endomorphismen

Definition 6.8.1 *Sei V ein Raum über $\mathbb{K}$ mit Skalarprodukt. Für einen selbstadjungierten Endomorphismus $T \in L(V)$ ist $\langle v, T\,v\rangle$ für alle $v \in V$ nach* Aufgabe 6.6.7 *eine reelle Zahl. Wir nennen T*

(a) *positiv definit*, falls $\langle v, T\,v\rangle > 0$ gilt für alle $v \in V \setminus \{0\}$,

(b) *negativ definit*, falls $\langle v, T\,v\rangle < 0$ gilt für alle $v \in V \setminus \{0\}$,

(c) *positiv semidefinit*, falls $\langle v, T\,v\rangle \geq 0$ gilt für alle $v \in V$,

(d) *negativ semidefinit*, falls $\langle v, T\,v\rangle \leq 0$ gilt für alle $v \in V$,

(e) *indefinit*, falls $\langle v_1, T\,v_1\rangle > 0$ gilt für mindestens ein $v_1 \in V$, während $\langle v_2, T\,v_2\rangle < 0$ für ein anderes $v_2 \in V$ gilt.

Beachte, dass für jeden sebstadjungierten Endomorphismus immer einer dieser Fälle eintritt. Die Menge $L(V)$ ist also die Vereinigung von fünf Teilmengen, die jeweils alle Endomorphismen mit derselben Definitheitseigenschaft *umfassen. Dabei sind die positiv bzw. negativ definiten Endomorphismen in der Menge der positiv bzw. negativ semidefiniten enthalten.*

Aufgabe 6.8.2 *Zeige dass die identische Abbildung auf V positiv definit ist.*

Aufgabe 6.8.3 *Für $a, c \in \mathbb{R}$, $b \in \mathbb{C}$ sei*

$$A = \begin{bmatrix} a & b \\ \overline{b} & c \end{bmatrix}.$$

Zeige, dass die Abbildung $x \longmapsto A\,x$ von $\mathbb{C}^2$ in sich selbstadjungiert ist, und stelle fest, für welche a, b, c diese Abbildung positiv oder negativ definit oder semidefinit bzw. indefinit ist.

Satz 6.8.4 *Sei V ein Raum über $\mathbb{K}$ mit Skalarprodukt, und seien $T \in L(V)$ selbstadjungiert und $U \in L(V)$ längentreu und surjektiv. Dann ist auch $U^{-1}\,T\,U$ selbstadjungiert, und T und $U^{-1}\,T\,U$ haben dieselbe Definitheitseigenschaft.*

Beweis: Nach Satz 6.7.4 ist $U^{ad} = U^{-1}$, und daher folgt aus Proposition 6.6.6 die Selbstadjungiertheit von $U^{-1}\,T\,U$. Weiter folgt für $v_1, v_2 \in V$ und $\tilde{v}_j = U\,v_j$, dass

$$\langle v_1, U^{-1}\,T\,U\,v_2\rangle = \overline{\langle U^{ad}\,T\,U\,v_2, v_1\rangle} = \overline{\langle T\,U\,v_2, U\,v_1\rangle} = \langle \tilde{v}_1, T\,\tilde{v}_2\rangle\,.$$

Für $v_2 = v_1 = v$, also $\tilde{v}_2 = \tilde{v}_1 = \tilde{v}$, folgt hieraus $\langle v, U^{-1}\,T\,U\,v\rangle = \langle \tilde{v}, T\,\tilde{v}\rangle$. Da U bijektiv ist, folgt die Behauptung. □

6.9 Eigenwerte und Eigenvektoren

Definition 6.9.1 *Sei V Vektorraum über $\mathbb{K}$. Ein $\lambda \in \mathbb{K}$ heißt* ein Eigenwert *eines Endomorphismus $T \in L(V)$, falls ein $v \in V \setminus \{0\}$ existiert, für welches*

$$T\,v = \lambda\,v \tag{6.9.1}$$

ist. Jeder solche Vektor v heißt dann ein Eigenvektor *von T zum Eigenwert λ. Die Menge aller $v \in V$, einschließlich des Nullvektors, welche die Gleichung* (6.9.1) *erfüllen, heißt* der Eigenraum von T *zum Eigenwert λ. Beachte, dass die Definition des Eigenraums formal auch sinnvoll ist, wenn λ kein Eigenwert von T ist – dann allerdings besteht der Eigenraum nur aus dem Nullvektor. Ein Endomorphismus T heißt* diagonalisierbar, *falls es eine Basis von V gibt, welche aus Eigenvektoren von T besteht.*

Bemerkung 6.9.2 *Genau dann ist λ Eigenwert eines Endomorphismus T, wenn die Abbildung $T - \lambda\,id$ nicht injektiv ist; dabei soll id die identische Abbildung auf V sein, also $id\,v = v$ für alle $v \in V$ erfüllen. Der Eigenraum von T zum Eigenwert λ ist also genau der Kern der Abildung $T - \lambda\,id$. In aufbauenden Vorlesungen, z. B. über* Funktionalanalysis, *nennt man λ einen* Spektralwert *von T, wenn $T - \lambda\,id$ nicht bijektiv ist. Also ist ein Eigenwert immer auch Spektralwert, aber im Allgemeinen gilt die Umkehrung nicht. Wegen* Aufgabe 6.4.2 *gilt aber:*

> *Wenn V endlich-dimensional ist, dann ist jeder Spektralwert auch ein Eigenwert.*

Aufgabe 6.9.3 *Sei V Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$ bijektiv, also ein Isomorphismus. Zeige: Genau dann ist $\lambda \in \mathbb{K}$ ein Eigenwert von T, wenn λ^{-1} ein Eigenwert von T^{-1} ist. Insbesondere ist $\lambda = 0$ niemals Eigenwert einer injektiven Abbildung.*

Proposition 6.9.4 *Sei T ein Endomorphismus eines Vektorraumes V über $\mathbb{K}$, und sei $(v_j, j \in J)$ ein System von Eigenvektoren von T zu paarweise verschiedenen Eigenwerten λ_j; genauer gelte*

$$\forall\, j \in J: \quad T\,v_j = \lambda_j\,v_j\,, \qquad \forall\, j, k \in J: \quad \big(j \neq k \Longrightarrow \lambda_j \neq \lambda_k\big)\,.$$

Dann ist das System $(v_j, j \in J)$ linear unabhängig.

Beweis: Für den Beweis genügt es wegen Aufgabe 1.4.5, ein endliches System $(v_1, \dots, v_n)$ zu betrachten, und wir geben den Beweis durch Induktion über n. Für $n = 1$ ist die Behauptung klar, da ein Eigenvektor nie der Nullvektor ist. Sei jetzt also $n \geq 2$, und seien $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ so, dass $\sum_{k=1}^n \alpha_k \, v_k = 0$ ist. Dann folgt $0 = T(0) = \sum_{k=1}^n \alpha_k \, T(v_k) = \sum_{k=1}^n \alpha_k \, \lambda_k \, v_k$. Durch Kombination dieser Gleichungen folgt

$$0 \;=\; \Big(\sum_{k=1}^{n} \alpha_k \, \lambda_k \, v_k \Big) \;-\; \lambda_n \, \Big(\sum_{k=1}^{n} \alpha_k \, v_k \Big) \;=\; \sum_{k=1}^{n-1} \alpha_k \, \big(\lambda_k - \lambda_n \big) \, v_k \, .$$

Wegen der Induktionshypothese ist das System $(v_1, \dots, v_{n-1})$ linear unabhängig, und daher folgt $0 = \alpha_k \, \big(\lambda_k - \lambda_n \big)$ für $1 \leq k \leq n-1$. Nach Voraussetzung ist $\lambda_k \neq \lambda_n$ für $k \leq n-1$, und somit folgt $0 = \alpha_k$ für $k \leq n-1$. Dann ist aber $\alpha_n \, v_n = 0$, woraus wegen $v_n \neq 0$ folgt, dass auch $\alpha_n = 0$ ist. □

Satz 6.9.5 *Sei T ein Endomorphismus eines Raumes V mit Skalarprodukt.*

(a) *Wenn T längentreu ist, so ist $|\lambda| = 1$ für alle Eigenwerte λ von T.*

(b) *Wenn T selbstadjungiert ist, dann sind alle Eigenwerte von T reell, und Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.*

(c) *Wenn T positiv bzw. negativ definit ist, dann sind alle Eigenwerte von T positiv bzw. negativ, wenn T positiv bzw. negativ semidefinit ist, dann sind alle Eigenwerte von T nichtnegativ bzw. nichtpositiv.*

(d) *Wenn T einen positiven und einen negativen Eigenwert hat, dann ist T indefinit.*

(e) *Wenn T normal und v ein Eigenvektor von T zum Eigenwert λ ist, dann ist v auch Eigenvektor von T^{ad} zum Eigenwert $\overline{\lambda}$. Außerdem stehen Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.*

Beweis: Aus $T\,v = \lambda\, v$ folgt $\|T\,v\| = |\lambda| \, \|v\|$, und daraus folgt (a). Zu (b): Aus (6.6.1) folgt für ein selbstadjungiertes T die Gleichung $\langle T\,v_1, v_2 \rangle = \langle v_1, T\,v_2 \rangle$ für alle $v_1, v_2 \in V$. Wenn $T\,v = \lambda\, v$ ist, folgt $\lambda \, \langle v, v \rangle = \langle v, \lambda v \rangle = \langle v, T\,v \rangle = \langle T\,v, v \rangle = \overline{\lambda} \, \langle v, v \rangle$, und da $\langle v, v \rangle = \|v\|^2 \neq 0$ ist, folgt hieraus $\overline{\lambda} = \lambda$, also $\lambda \in \mathbb{R}$. Sei jetzt $T\,v_j = \lambda_j \, v_j$ für $j = 1$ und $j = 2$. Dann folgt (da ja $\lambda_j \in \mathbb{R}$ sein muss)

$$(\lambda_1 - \lambda_2) \, \langle v_1, v_2 \rangle \;=\; \langle \lambda_1 \, v_1, v_2 \rangle \;-\; \langle v_1, \lambda_2 \, v_2 \rangle \;=\; \langle T\,v_1, v_2 \rangle \;-\; \langle v_1, T\,v_2 \rangle \, .$$

Da T selbstadjungiert ist, verschwindet die rechte Seite dieser Gleichung, und falls $\lambda_1 \neq \lambda_2$ ist, folgt $\langle v_1, v_2 \rangle = 0$. Zu (c),(d): Folgt aus der jeweiligen Definition und der Gleichung $\lambda \, \|v\|^2 = \lambda \, \langle v, v \rangle = \langle v, T\,v \rangle$ für jeden Eigenvektor v zum Eigenwert λ. Zu (e): Wenn T normal ist, dann zeigt man mit Aufgabe 6.6.5, dass auch $T - \lambda\, id$ normal ist, denn $(T - \lambda\, id)^{ad} = T^{ad} - \overline{\lambda}\, id$, und die identische Abbildung kommutiert mit jedem Endomorphismus. Deshalb reicht es, wenn wir den Fall $\lambda = 0$, also $T\,v = 0$, betrachten. In diesem Fall gilt aber

$$\langle T^{ad}\, v, T^{ad}\, v \rangle \;=\; \langle v, T \circ T^{ad}\, v \rangle \;=\; \langle v, T^{ad} \circ T\, v \rangle \;=\; \langle T\,v, T\,v \rangle \, .$$

Daraus ergibt sich aber, dass $T\,v = 0$ ist genau wenn $T^{ad}\, v = 0$ ausfällt. Sei jetzt wieder $T\,v_j = \lambda_j \, v_j$, also auch $T^{ad}\, v_j = \overline{\lambda}_j \, v_j$ für $j = 1$ und $j = 2$. Dann folgt (da jetzt die Eigenwerte nicht immer reell sind)

$$(\lambda_1 - \lambda_2) \, \langle v_1, v_2 \rangle \;=\; \langle \overline{\lambda}_1 \, v_1, v_2 \rangle \;-\; \langle v_1, \lambda_2 \, v_2 \rangle \;=\; \langle T^{ad}\, v_1, v_2 \rangle \;-\; \langle v_1, T\,v_2 \rangle \, .$$

Aus der Definition der adjungierten Abbildung folgt aber dann dass die rechte Seite verschwindet. □

Kapitel 7

Lineare Abbildungen und Matrizen

In diesem Kapitel seien V und W immer endlich-dimensionale Vektorräume über $\mathbb{K}$, und keiner von beiden habe die Dimension 0, also beide enthalten mehr als nur den Nullvektor.

7.1 Die Darstellungsmatrix

Definition 7.1.1 *Seien $(v_1, \ldots, v_n)$ bzw. $(w_1, \ldots, w_m)$ fest gewählte Basen von V bzw. W. Für jede lineare Abbildung $T \in L(V, W)$ gibt es eindeutig bestimmte Zahlen $a_{jk} \in \mathbb{K}$ mit*

$$T\, v_k \;=\; \sum_{j=1}^{m} a_{jk}\, w_j \qquad \forall\; k = 1, \ldots, n\,. \tag{7.1.1}$$

Die hierdurch definierte Matrix $A = [a_{jk}] \in \mathbb{K}^{m \times n}$ heißt die Darstellungsmatrix für T *bezüglich der gewählten Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$.*

Beispiel 7.1.2 *Im Fall $V = W$, also für einen Endomorphismus $T \in L(V)$, ist es sinnvoll, wengleich auch nicht zwangsläufig, wenn wir nur eine Basis $(v_1, \ldots, v_n)$ wählen, wenn also $w_j = v_j$ ist für alle $j = 1, \ldots, n$. In dieser Situation ist die Darstellungsmatrix für die identische Abbildung gerade die Einheitsmatrix. Für $V = \mathbb{K}^n$ und $W = \mathbb{K}^m$ definiert jede Matrix $A \in \mathbb{K}^{m \times n}$ eine lineare Abbildung $x \longmapsto A\,x$ zwischen diesen Räumen. Wählt man in beiden Räumen die kanonischen Basen, so ist die Darstellungsmatrix dieser Abbildung gleich A.*

Aufgabe 7.1.3 *Zeige: Ein Endomorphismus $T \in L(V)$ ist genau dann diagonalisierbar, wenn es eine Basis von V gibt, bezüglich der die Darstellungsmatrix von T diagonal ist.*

Aufgabe 7.1.4 *Sei $V = W = \mathbb{K}_n[t]$ der Raum der Polynome vom Grad $\leq n$ mit Koeffizienten in $\mathbb{K}$, und sei $T\,p = p'$, die Ableitung von p, für $p \in \mathbb{K}_n[t]$. Finde die Darstellungsmatrix von T bezüglich der kanonischen Basis der Monome.*

Satz 7.1.5 *Seien $(v_1, \ldots, v_n)$ bzw. $(w_1, \ldots, w_m)$ fest gewählte Basen von V bzw. W. Für jede lineare Abbildung $T \in L(V, W)$ sei $A = A_T = [a_{jk}]$ die Darstellungsmatrix bezüglich dieser Basen. Dann ist die Zuordnung $T \longmapsto A_T$ ein Isomorphismus von $L(V, W)$ auf $\mathbb{K}^{m \times n}$, und für $v = \sum_{k=1}^n \alpha_k\, v_k$ sowie $w = \sum_{j=1}^m \beta_j\, w_j$ gilt $T\,v = w$ genau dann, wenn*

$$\beta_j \;=\; \sum_{k=1}^{n} a_{jk}\, \alpha_k \qquad \forall\; j = 1, \ldots, m\,.$$

Beweis: Nach Satz 6.3.1 gibt es zu jedem $A \in \mathbb{K}^{m\times n}$ ein $T \in L(V,W)$ mit (7.1.1), und deshalb ist die Zuordnung $T \longmapsto A_T$ surjektiv. Sind $T_1, T_2 \in L(V,W)$ mit Darstellungsmatrizen A_1 und A_2, und sind $\alpha_1, \alpha_2 \in \mathbb{K}$, so gilt $(\alpha_1 T_1 + \alpha_2 T_2)\, v_k = \alpha_1 T_1\, v_1 + \alpha_2 T_2\, v_k$ für alle $k = 1, \ldots, n$, und daher ist $\alpha_1 A_1 + \alpha_2 A_2$ die Darstellungsmatrix zu $\alpha_1 T_1 + \alpha_2 T_2$. Also ist die Zuordnung $T \longmapsto A_T$ linear. Wegen $\dim L(V,W) = \dim \mathbb{K}^{m\times n} = n\,m$ folgt mit Aufgabe 6.4.2, dass die Abbildung $T \longmapsto A_T$ auch injektiv sein muss. Der Rest des Beweises ist Inhalt der nächsten Übungsaufgabe. □

Aufgabe 7.1.6 *Zeige unter den Voraussetzungen und mit den Bezeichnungen des obigen Satzes: Ist $v = \sum_{k=1}^n \alpha_k\, v_k$ und $T\, v = \sum_{j=1}^m \beta_j\, w_j$, so folgt $\beta_j = \sum_{k=1}^n a_{jk}\, \alpha_k$, für $j = 1, \ldots, m$.*

Bemerkung 7.1.7 (Kommutatives Diagramm) *Den Inhalt des obigen Satzes kann man sich am besten folgendermaßen einprägen: Nach Wahl der beiden Basen von V und W betrachten wir die beiden kanonischen Isomorphismen $\phi_n : V \longrightarrow \mathbb{K}^n$ und $\phi_m : W \longrightarrow \mathbb{K}^m$, die jedem Vektor $v = \sum_{k=1}^n \alpha_k\, v_k$ und $w = \sum_{j=1}^m \beta_j\, w_j$ ihre Koordinatenvektoren $\phi_n(v) = (\alpha_1, \ldots, \alpha_n)^T$ und $\phi_m(w) = (\beta_1, \ldots, \beta_m)^T$ zuweisen. Die Darstellungsmatrix A bildet $\mathbb{K}^n$ nach $\mathbb{K}^m$ ab, und es gilt*

$$A\, \phi_n(v) \;=\; \phi_m(T\, v) \qquad \forall\, v \in V\,.$$

Dies wird auch durch folgendes kommutatives Diagramm *veranschaulicht:*

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \phi_n \Big\downarrow & & \Big\downarrow \phi_m \\ \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \end{array}$$

Beachte aber unbedingt, dass die Matrix A und auch die kanonischen Isomorphismen erst nach Wahl der beiden Basen in V und W festliegen; wie diese Größen sich bei Basiswechsel verändern, wird noch zu untersuchen sein.

Satz 7.1.8 *Seien U, V, W drei endlich-dimensionale Vektorräume über $\mathbb{K}$ mit Basen $(u_1, \ldots, u_r)$ bzw. $(v_1, \ldots, v_n)$ bzw. $(w_1, \ldots, w_m)$. Seien weiter $T_1 \in L(U,V)$, $T_2 \in L(V,W)$ und $T = T_2 \circ T_1$. Wenn dann $A_1 = [a^{(1)}_{jk}]$ und $A_2 = [a^{(2)}_{jk}]$ die Darstellungsmatrizen zu T_1 bzw. T_2 sind, so ist die Darstellungmatrix zu T gleich dem Produkt $A_2\, A_1$, jeweils bezüglich der entsprechenden Basen.*

Beweis: Nach Definition der darstellenden Matrizen gilt

$$T_1\, u_k \;=\; \sum_{j=1}^n a^{(1)}_{jk}\, v_j \quad (k = 1, \ldots, r)\,, \quad T_2\, v_k \;=\; \sum_{j=1}^m a^{(2)}_{jk}\, w_j \quad (k = 1, \ldots, n)\,.$$

Hieraus folgt sofort

$$T\, u_k \;=\; \sum_{\nu=1}^n a^{(1)}_{\nu k}\, T_2\, v_\nu \;=\; \sum_{j=1}^m \Big(\sum_{\nu=1}^n a^{(1)}_{j\nu}\, a^{(2)}_{\nu k}\Big)\, w_j \qquad (1 \le k \le r)\,.$$

Dies ist die Behauptung. □

Korollar zu Satz 7.1.8 *Seien V und W Vektorräume über $\mathbb{K}$ mit derselben, endlichen, Dimension $n \in \mathbb{N}$, und seien $(v_1, \ldots, v_n)$ bzw. $(w_1, \ldots, w_n)$ fest gewählte Basen von V bzw. W. Eine lineare Abbildung $T \in L(V,W)$ ist genau dann ein Isomorphismus, wenn ihre Darstellungsmatrix A invertierbar ist, und die Darstellungsmatrix zu $T^{-1} \in L(W,V)$ bzgl. der Basen $(w_1, \ldots, w_n)$ und $(v_1, \ldots, v_n)$ ist gleich A^{-1}.*

Beweis: Sei T ein Isomorphismus, und sei die Darstellungsmatrix von T^{-1} mit B bezeichnet. Wegen $v_k = T^{-1}(T\, v_k)$ für alle $k = 1, \ldots, n$ folgt dann, dass die Darstellungmatrix zu $T^{-1} \circ T$ die Einheitsmatrix sein muss, und deshalb folgt mit Satz 7.1.8 die Gleichung $B\,A = I$. Also ist A invertierbar und $B = A^{-1}$. Umgekehrt, wenn A invertierbar ist, und wenn $\tilde{T} \in L(W,V)$ die Darstellungsmatrix A^{-1} hat, so folgt $\tilde{T}(T\, v_k) = v_k$ für alle $k = 1, \ldots, n$. Daraus folgt aber, dass T bijektiv und $\tilde{T} = T^{-1}$ sein muss. □

7.2 Basiswechsel

Lemma 7.2.1 *Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei das System $(v_1, \ldots, v_n)$ eine Basis von V. Für ein $B = [b_{jk}] \in \mathbb{K}^{n \times n}$ seien $\tilde{v}_k = \sum_{j=1}^n b_{jk}\, v_j$ für $k = 1, \ldots, n$ gesetzt. Genau dann ist $(\tilde{v}_1, \ldots, \tilde{v}_n)$ ebenfalls Basis von V, wenn B invertierbar ist.*

Beweis: Es ist $0 = \sum_{k=1}^n \lambda_k\, \tilde{v}_k = \sum_{j=1}^n \bigl(\sum_{k=1}^n b_{jk}\, \lambda_k\bigr)\, v_j$ genau dann wenn $\sum_{k=1}^n b_{jk}\, \lambda_k = 0$ ist für alle $j = 1, \ldots, n$. Dies bedeutet, dass die Linearkombination der Spalten von B mit den Koeffizienten λ_k den Nullvektor ergibt. Die Invertierbarkeit von B ist gleichbedeutend mit der linearen Unabhängigkeit der Spalten von B, und dies ist somit dasselbe wie die lineare Unabhängigkeit von $(\tilde{v}_1, \ldots, \tilde{v}_n)$. □

Satz 7.2.2 (Basiswechsel) *Seien V und W zwei endlich-dimensionale Vektorräume über $\mathbb{K}$, und seien $(v_1, \ldots, v_n)$ und $(\tilde{v}_1, \ldots, \tilde{v}_n)$ zwei Basen von V, sowie $(w_1, \ldots, w_m)$ und $(\tilde{w}_1, \ldots, \tilde{w}_m)$ zwei Basen von W. Seien $B = [b_{jk}] \in \mathbb{K}^{n \times n}$ bzw. $C = [c_{jk}] \in \mathbb{K}^{m \times m}$ so, dass*

$$\tilde{v}_k \;=\; \sum_{j=1}^n b_{jk}\, v_j \quad (1 \leq k \leq n)\,, \qquad \tilde{w}_k \;=\; \sum_{j=1}^m c_{jk}\, w_j \quad (1 \leq k \leq m)\,. \tag{7.2.1}$$

Schließlich sei $T \in L(V, W)$, und A bzw. $\tilde{A}$ seien die Darstellungsmatrizen zu T, einmal bezüglich $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$, sowie zum anderen bezüglich $(\tilde{v}_1, \ldots, \tilde{v}_n)$ und $(\tilde{w}_1, \ldots, \tilde{w}_m)$. Dann gilt

$$\tilde{A} \;=\; C^{-1}\, A\, B\,.$$

Beweis: Nach Definition der Darstellungsmatrizen gilt $T\, v_k = \sum_{j=1}^m a_{jk}\, w_j$, $T\, \tilde{v}_k = \sum_{j=1}^m \tilde{a}_{jk}\, \tilde{w}_j$, jeweils für $k = 1, \ldots, n$. Einsetzen von (7.2.1) ergibt die Gleichung

$$\sum_{j=1}^m \Bigl(\sum_{\nu=1}^n a_{j\nu}\, b_{\nu k}\Bigr)\, w_j \;=\; \sum_{j=1}^m \Bigl(\sum_{\mu=1}^m c_{j\mu}\, \tilde{a}_{\mu k}\Bigr)\, w_j \qquad \forall\; k = 1, \ldots, n\,.$$

Bringt man beide Summen auf eine Seite, so folgt unter Benutzung der linearen Unabhängigkeit von $(w_1, \ldots, w_m)$, dass $0 = \sum_{\nu=1}^n a_{j\nu}\, b_{\nu k} - \sum_{\mu=1}^m c_{j\mu}\, \tilde{a}_{\mu k}$ ist für alle $j = 1, \ldots, m$ und $k = 1, \ldots, n$. Dies ist die Behauptung. □

Definition 7.2.3 *Nach dem Satz vom Basiswechsel sind die beiden Darstellungsmatrizen derselben linearen Abbildung zu verschiedenen Paaren von Basen von V und W zueinander äquivalent. Umgekehrt folgt mit* Lemma 7.2.1*, dass mit einer Darstellungsmatrix von T auch jede zu dieser äquivalente Matrix eine Darstellungsmatrix zu anderen Basen ist. Wir sagen deshalb:* Die Darstellungsmatrix einer linearen Abbildungen zwischen endlich-dimensionalen Vektorräumen ist bis auf Äquivalenz eindeutig bestimmt.

7.3 Ähnliche Matrizen

Wie schon früher gesagt, kann man bei Endomorphismen $T \in L(V)$ sinnvollerweise nur eine Basis, nämlich die in V, wählen, um die Darstellungsmatrix von T zu vereinfachen. Dies führt zu folgendem Begriff der Ähnlichkeit von Matrizen:

Definition 7.3.1 *Matrizen $A, \tilde{A} \in \mathbb{K}^{n \times n}$ heißen* ähnlich zueinander*, wenn es eine invertierbare Matrix $B \in \mathbb{K}^{n \times n}$ gibt, für die*

$$\tilde{A} \;=\; B^{-1}\, A\, B\,.$$

Aufgabe 7.3.2 (Ähnliche Matrizen mit MAPLE) MAPLE *kann prüfen, ob Matrizen ähnlich sind. Überprüfe dies an Hand der Kommandos*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3, 3, [ [1, 0, 1], [0, 1, 1], [2, -1, 1] ]);
> B := Matrix(3, 3, [ [1, 1, 2], [0, 1, 1], [0, 1, 1] ]);
> IsSimilar(A,B);
```

Was hier genau geschieht, und wie man auch „zu Fuß“ die Ähnlichkeit von Matrizen zeigen oder widerlegen kann, wird erst später klarer werden.

Aufgabe 7.3.3 *Zeige, dass der Begriff der Ähnlichkeit eine Äquivalenzrelation auf* $\mathbb{K}^{n\times n}$ *darstellt. Zeige weiter, dass ähnliche Matrizen immer dieselbe Determinante haben.*

Als einfache Folgerung aus dem Satz vom Basiswechsel im letzten Abschnitt erhalten wir das analoge Resultat für Endomorphismen:

Satz 7.3.4 (Basiswechsel bei Endomorphismen)
Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$, und seien $(v_1, \ldots, v_n)$ und $(\tilde{v}_1, \ldots, \tilde{v}_n)$ zwei Basen von V. Sei $B = [b_{jk}] \in \mathbb{K}^{n\times n}$ so, dass

$$\tilde{v}_k \;=\; \sum_{j=1}^{n} b_{jk}\, v_j \quad (1 \le k \le n)\,. \tag{7.3.1}$$

Schließlich sei $T \in L(V)$, und A bzw. $\tilde{A}$ seien die Darstellungsmatrizen zu T, einmal bezüglich $(v_1, \ldots, v_n)$, sowie zum anderen bezüglich $(\tilde{v}_1, \ldots, \tilde{v}_n)$. Dann gilt

$$\tilde{A} \;=\; B^{-1}\, A\, B\,.$$

Definition 7.3.5 (Determinante eines Endomorphismus) *Ein Endomorphismus $T \in L(V)$, für einen endlich-dimensionalen Raum V, hat zwar unterschiedliche Darstellungsmatrizen, aber alle Darstellungsmatrizen von T sind zueinander ähnlich und haben insbesondere dieselbe Determinante. Deshalb können wir* die Determinante von T *als die Determinante einer der Darstellungsmatrizen von T definieren.*

Aufgabe 7.3.6 *Finde die Determinante des Endomorphismus $p \longmapsto p'$ von $\mathbb{K}_n[t]$, $n \in \mathbb{N}_0$, in sich.*

7.4 Hermitesche, normale und unitäre Matrizen

Satz 7.4.1 *Gegeben seien endlich-dimensionale Räume V und W über $\mathbb{K}$ mit Skalarprodukten. Sei $A \in \mathbb{K}^{m\times n}$ die Darstellungsmatrix für ein $T \in L(V, W)$ bzgl. zweier Orthonormalbasen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ von V bzw. W. Dann hat die adjungierte Abbildung T^{ad} die Darstellungsmatrix $\overline{A}^T$ bzgl. der Basen $(w_1, \ldots, w_m)$ und $(v_1, \ldots, v_n)$.*

Beweis: Sei die Darstellungsmatrix von T^{ad} mit $B = [b_{jk}]$ bezeichnet. Dann gilt nach Definition der Darstellungsmatrizen für alle $k = 1, \ldots, n$ und $\nu = 1, \ldots, m$

$$T\, v_k \;=\; \sum_{j=1}^{m} a_{jk}\, w_j\,, \qquad T^{ad}\, w_\nu \;=\; \sum_{\mu=1}^{n} b_{\mu\nu}\, v_\mu\,.$$

Mit den Regeln für Skalarprodukte folgt dann $\langle T^{ad}\, w_\nu, v_k\rangle = \sum_{\mu=1}^{n} \overline{b_{\mu\nu}}\, \langle v_\mu, v_k\rangle = \overline{b_{k\nu}}$, $\langle w_\nu, T\, v_k\rangle = \sum_{j=1}^{m} a_{jk}\, \langle w_\nu, w_j\rangle = a_{\nu k}$, und aus (6.6.1) folgt die Behauptung. □

Definition 7.4.2 *Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt* hermitesch, *falls $\overline{A}^T = A$ ist. Im Fall $\mathbb{K} = \mathbb{R}$ nennt man eine solche Matrix auch* symmetrisch. *Wir nennen A* normale Matrix, *wenn $\overline{A}^T A = A\, \overline{A}^T$ ist.*

Aufgabe 7.4.3 *Zeige, dass eine Diagonalmatrix immer normal ist. Zeige weiter, dass die Diagonalelemente einer hermiteschen Matrix immer reelle Zahlen sind.*

Korollar zu Satz 7.4.1 *Sei V ein endlich-dimensionaler Raum mit Skalarprodukt. Genau dann ist ein Endomorphismus $T \in L(V)$ selbstadjungiert, wenn seine Darstellungsmatrix bezüglich irgendeiner Orthonormalbasis von V hermitesch ist. Weiter ist T genau dann normal, wenn seine Darstellungsmatrix bezüglich irgendeiner Orthonormalbasis von V normal ist.*

Definition 7.4.4 *Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt* unitär, *falls $\overline{A}^T A = I$ ist. Insbesondere ist eine unitäre Matrix invertierbar, und $A^{-1} = \overline{A}^T$. Im Fall $\mathbb{K} = \mathbb{R}$ nennt man eine solche Matrix auch* orthogonal.

Aufgabe 7.4.5 (MAPLE und unitäre Matrizen) *Untersuche, ob* MAPLE *eine Matrix auf Orthogonalität und Unitarität prüfen kann.*

Aufgabe 7.4.6 *Sei A eine quadratische n-reihige Matrix, und sei in $\mathbb{K}^n$ das kanonische innere Produkt betrachtet. Zeige: Genau dann ist A unitär, wenn die Spalten von A ein Orthonormalsystem in $\mathbb{K}^n$ bilden. In welchem Sinn gilt dasselbe auch für die Zeilen? Zeige weiter: Ist A unitär, und ist $y = A\, x$ für ein $x \in \mathbb{K}^n$, so folgt $\|y\|^2 = \|x\|^2$, d. h., die Abbildung $x \longmapsto A\, x$ ist längentreu.*

Satz 7.4.7 *Gegeben seien endlich-dimensionale Räume V und W über $\mathbb{K}$ mit Skalarprodukten, und sei $n = \dim V = \dim W \geq 1$. Sei $A \in \mathbb{K}^{n\times n}$ die Darstellungsmatrix für ein $T \in L(V, W)$ bzgl. zweier Orthonormalbasen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$ von V bzw. W. Genau dann ist T längentreu, wenn A unitär ist.*

Beweis: Wenn T längentreu ist, folgt $\langle T\, v_j, T\, v_k\rangle = \langle v_j, v_k\rangle = \delta_{jk}$ für $1 \leq j, k \leq n$. Wegen $T\, v_j = \sum_{\nu=1}^{n} a_{\nu j}\, w_\nu$ für $j = 1, \ldots, n$ folgt daraus

$$\delta_{jk} \;=\; \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} \overline{a}_{\nu j}\, a_{\mu k}\,,$$

was die eine Richtung der Behauptung ist. Umgekehrt, sei A invertierbar und $\overline{A}^T = A^{-1}$. Dann folgt für $v = \sum_{k=1}^{n} \alpha_k v_j$, also $T\, v = \sum_{j=1}^{n} \beta_j\, w_j$ mit $\beta_j = \sum_{k=1}^{n} a_{jk}\, \alpha_k$: Wegen Aufgabe 5.4.1 ist $\|T\, v\|^2 = \sum_{j=1}^{n} |\beta_j|^2$ und $\|v\|^2 = \sum_{k=1}^{n} |\alpha_k|^2$, und aus Aufgabe 7.4.6 folgt $\|T\, v\| = \|v\|$, also die Längentreue. □

Aufgabe 7.4.8 *Beweise folgendes Analogon zu* Lemma 7.2.1:

Lemma 7.4.9 *Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ mit innerem Produkt, und sei das System $(v_1, \ldots, v_n)$ eine Orthonormalbasis von V. Für ein $U = [u_{jk}] \in \mathbb{K}^{n\times n}$ seien $\tilde v_k = \sum_{j=1}^n u_{jk}\, v_j$ für $k = 1, \ldots, n$ gesetzt. Genau dann ist $(\tilde v_1, \ldots, \tilde v_n)$ ebenfalls Orthonormalbasis von V, wenn U unitär ist.*

Aufgabe 7.4.10 *Zeige: Wenn V ein endlich-dimensionaler Raum über $\mathbb{K}$ mit innerem Produkt ist, dann gilt für jeden längentreuen Automorphismus von V immer $|\det T| = 1$ gilt. Schließe hieraus für den Fall $\mathbb{K} = \mathbb{R}$, dass es zwei Arten von längentreuen Automorphismen gibt, und vergleiche dies mit* Abschnitt 9.2.

7.5 Charakteristisches Polynom

Definition 7.5.1 *Für $A \in \mathbb{K}^{n\times n}$ heißt*

$$p_A(t) \;=\; \det\big(A - t\, I\big) \tag{7.5.1}$$

das charakteristische Polynom *der Matrix A. Aus der Definition der Determinante folgt sofort, dass p_A ein Polynom n-ten Grades in der Veränderlichen t ist, und genauer gilt*

$$p_A(t) \;=\; (-1)^n\, t^n \;+\; (-1)^{n-1}\, t^{n-1}\,\mathrm{tr}\, A \;\pm\; \ldots \;+\; \det A\,,$$

wobei $\mathrm{tr}\, A = a_{11} + \ldots + a_{nn}$ *die sogenannte* Spur *von A bezeichnet.*

Aufgabe 7.5.2 *Zeige, dass ähnliche Matrizen immer das gleiche charakteristische Polynom haben.*

Aufgabe 7.5.3 (MAPLE und charakteristische Polynome) *Das charakteristische Polynom kann bei kleineren Matrizen leicht berechnet werden. Überprüfe dies an den Beispielen*

$$A \;=\; \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}, \qquad B \;=\; \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Benutze hierzu auch die folgenden MAPLE-*Kommandos:*

```
> restart;
> with(LinearAlgebra);
> A := Matrix(3, 3, [ [1, 0, 1], [0, 1, 1], [2, -1, 1] ]);
> B := Matrix(3, 3, [ [1, 1, 2], [0, 1, 1], [0, 1, 1] ]);
> PA := CharacteristicPolynomial(A,t);
> factor(PA);
> PB := CharacteristicPolynomial(B,t);
> factor(PB);
```

Beachte aber, dass das von MAPLE *berechnete charakteristische Polynom gleich* $\det(t\, I - A)$ *ist. Ob man dieses oder das von uns benutzte Polynom vorzieht, ist Ansichtssache. Was bedeutet wohl das letzte Kommando?*

Aufgabe 7.5.4 *Seien $n \in \mathbb{N}$ und $a_0, \ldots, a_{n-1} \in \mathbb{K}$. Finde das charakteristische Polynom der Matrix*

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ a_0 & a_1 & a_2 & \ldots & a_{n-2} & a_{n-1} \end{bmatrix} \tag{7.5.2}$$

durch Entwicklung nach der letzten Zeile. Schließe dann, dass es zu jedem $p \in \mathbb{K}[t]$ vom Grad n und mit höchstem Koeffizienten gleich $(-1)^n$ eine Matrix $A \in \mathbb{K}^{n \times n}$ gibt, deren charakteristisches Polynom gerade gleich p ist.

Bemerkung 7.5.5 (Fundamentalsatz der Algebra) *Ein Resultat, welches wir hier benutzen aber nicht beweisen wollen, ist das folgende:*

- Jedes nicht konstante Polynom $p \in \mathbb{K}[t]$ hat in dem Körper $\mathbb{C}$ der komplexen Zahlen mindestens eine Nullstelle.

Hat man eine Nullstelle λ eines Polynoms p gefunden, so kann man p in der Form

$$p(t) \;=\; (t - \lambda)\, q(t)$$

schreiben, wobei q wieder ein Polynom ist. Dieses ist entweder konstant oder hat selbst wieder eine Nullstelle. Iteriert man diese Schlussweise, so folgt die Darstellung

$$p(t) \;=\; a\,(t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)\,,$$

wobei n der Grad des Polynoms p und $a \neq 0$ sein höchster Koeffizient ist. Die Werte $\lambda_1, \ldots, \lambda_n$ sind, nicht notwendigerweise verschiedene, komplexe Zahlen. Wenn unter den Zahlen $\lambda_1, \ldots, \lambda_n$ eine Zahl λ genau s-mal auftritt, dann heißt s die Vielfachheit *der Nullstelle λ. Es gilt dann offenbar*

$$p(t) \;=\; (t - \lambda)^s\, q(t)$$

mit einem Polynom q, für welches $q(\lambda) \neq 0$ ist. Die Vielfachheit s einer Nullstelle λ erkennt man auch daran, dass gilt:

$$0 \;=\; p(\lambda) \;=\; p'(\lambda) \;= \ldots = p^{(s-1)}(\lambda)\,, \quad p^{(s)}(\lambda) \neq 0\,.$$

Es ist wichtig zu beachten, dass die Nullstellen eines Polynoms alle oder teilweise nicht-reell sein können, selbst wenn $\mathbb{K} = \mathbb{R}$ ist, d. h., wenn die Koeffizienten des Polynoms reell sind. Es gilt aber folgende Aussage:

- Wenn $p \in \mathbb{R}[t]$ ist, also reelle Koeffizienten hat, dann ist für jede Nullstelle λ auch die konjugiert komplexe Zahl $\overline{\lambda}$ eine Nullstelle von p, und beide haben dieselbe Vielfachheit.

Somit treten die nicht-reellen Nullstellen eines $p \in \mathbb{R}[t]$ immer in Paaren von zueinander konjugiert-komplexen Zahlen auf.

Aufgabe 7.5.6 *Zeige: Ein Polynom $p \in \mathbb{R}[t]$ mit ungeradem Grad hat mindestens eine reelle Nullstelle.*

Definition 7.5.7 *Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$. Wenn ein Endomorphismus $T \in L(V)$ gegeben ist, hängt zwar seine Darstellungsmatrix, nicht aber deren charakteristisches Polynom, von der Wahl einer Basis von V ab, und wir sprechen deshalb auch vom* charakteristischen Polynom des Endomorphismus *T und schreiben auch p_T für dieses Polynom. Zur Berechnung dieses charakteristischen Polynoms müssen wir aber eine konkrete Basis von V finden, dann die entsprechende Darstellungsmatrix bestimmen, und für diese p_A berechnen.*

Aufgabe 7.5.8 *Berechne das charakteristische Polynom des Endomorphismus* $p \longmapsto p'$ *von* $\mathbb{K}_n[t]$, $n \in \mathbb{N}_0$, *in sich.*

Satz 7.5.9 *Sei* $n \in \mathbb{N}$, *sei* V *ein* n*-dimensionaler Vektorraum über* $\mathbb{K}$, *und sei* $T \in L(V)$ *mit charakteristischem Polynom* p_T.

(a) *Genau dann ist ein* $\lambda \in \mathbb{K}$ *Eigenwert von* T, *wenn* $p_T(\lambda) = 0$ *ist.*

(b) *Ist* $(v_1, \ldots, v_n)$ *eine Basis von* V, *und ist* A *die Darstellungsmatrix von* T *zu dieser Basis, so ist ein* $v = \sum_{k=1}^n \alpha_k v_k \in V$ *genau dann ein Eigenvektor von* T *zu einem Eigenwert* $\lambda \in \mathbb{K}$, *wenn der Vektor* $(\alpha_1, \ldots, \alpha_n)^T$ *eine nicht-triviale Lösung des homogenen linearen Gleichungssystems*

$$\big(A - \lambda I\big)\, x = 0 \tag{7.5.3}$$

ist. Anders ausgedrückt liegt ein solches v *genau dann im Eigenraum von* T *zum Eigenwert* λ, *wenn* $(\alpha_1, \ldots, \alpha_n)^T$ *eine beliebige Lösung von* (7.5.3) *ist.*

(c) *Wenn das charakteristische Polynom* n *verschiedene Nullstellen in* $\mathbb{K}$ *hat, ist* T *diagonalisierbar.*

(d) *Wenn* V *ein Raum mit Skalarprodukt und* T *selbstadjungiert ist, dann sind alle Nullstellen des charakteristischen Polynoms von* T *reell.*

Beweis: Genau wenn $\det(A - \lambda I) = 0$ ist, besitzt (7.5.3) eine nicht-triviale Lösung, und daher folgt (a) aus (b). Zu (b): Nach Satz 7.1.5 gilt $T v = w = \sum_{k=1}^n \beta_k v_k$ genau dann, wenn $a = (\alpha_1, \ldots, \alpha_n)^T$ und $b = (\beta_1, \ldots, \beta_n)^T$ die Gleichung $b = A a$ erfüllen. Also gilt $T v = v$ genau dann, wenn $A a = a$ gilt, d. h., wenn a die Gleichung (7.5.3) erfüllt. Aussage (c) folgt, da Eigenvektoren zu verschiedenen Eigenwerten immer linear unabhängig sind. Zu (d): Sei A die Darstellungsmatrix von T bezüglich einer Orthogonalbasis von V. Dann ist $x \longmapsto A x$ ein selbstadjungierter Endomorphismus in $\mathbb{C}^n$, und jede Nullstelle des charakteristischen Polynoms ist ein Eigenwert dieses Endomorphismus. Mit Satz 6.9.5 folgt die Behauptung. □

Bemerkung 7.5.10 *Beachte dass ein Eigenwert eines* $T \in L(V)$ *immer zum Skalarenkörper* $\mathbb{K}$ *gehören muss. Das bedeutet, dass im Fall* $\mathbb{K} = \mathbb{R}$ *ein Endomorphismus keinen Eigenwert zu haben braucht. Andererseits folgt aus dem Fundamentalsatz der Algebra, dass für* $\mathbb{K} = \mathbb{C}$ jeder Endomorphismus mindestens einen Eigenwert besitzt. *In den Übungen und Klausuren zur Vorlesung wird meistens eine Matrix gegeben sein, deren Elemente reelle Zahlen sind. Wenn dann die Aufgabe lautet, deren Eigenwerte und Eigenvektoren zu berechnen, sollen immer alle, d. h., auch die komplexen Nullstellen des charakteristischen Polynoms bestimmt und eine maximale Zahl von linear unabhängigen Lösungen von* (7.5.3) *gefunden werden. Siehe dazu auch die folgenden Aufgaben.*

Aufgabe 7.5.11 *Finde alle Eigenwerte und Eigenvektoren der Matrizen*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Entscheide, ob die Matrizen als Endomorphismen in $\mathbb{K}^n$ *diagonalisierbar sind.*

Aufgabe 7.5.12 *Sei* A *eine* n*-reihige quadratische Matrix, und sei* $p_A(x)$ *ihr charakteristisches Polynom. Seien* $\lambda_1, , \ldots, \lambda_n$ *die Eigenwerte von* A, *also die Nullstellen von* $p_A(x)$, *wobei mehrfache Nullstellen auch mehrfach aufgeschrieben seien, so dass die Zahlen* $\lambda_1, , \ldots, \lambda_n$ *nicht unbedingt alle verschieden sind. Zeige:*

$$\prod_{k=1}^n \lambda_k = \det A \qquad \sum_{k=1}^n \lambda_k = \operatorname{tr} A\,.$$

Definition 7.5.13 *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$ mit charakteristischem Polynom p_T. Nach* Satz 7.5.9 *ist jede in $\mathbb{K}$ gelegene Nullstelle von p_T ein Eigenwert von T. Die Vielfachheit dieser Nullstelle heißt* die algebraische Vielfachheit des Eigenwertes λ, *und die Dimension des Eigenraumes zum Eigenwert λ, also die maximale Anzahl von linear unabhängigen Eigenvektoren, heißt* die geometrische Vielfachheit des Eigenwertes. *Nach Definition eines Eigenwertes ist die geometrische Vielfachheit immer ≥ 1. Zu jedem $s \in \mathbb{N}$ und jedem $\sigma \in \{1, \ldots, s\}$ findet man leicht Endomorphismen, die einen Eigenwert λ mit algebraischer Vielfachheit s und geometrischer Vielfachheit σ haben. Siehe dazu die nächste Aufgabe.*

Aufgabe 7.5.14 *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei $(v_1, \ldots, v_n)$ Basis von V. Sei weiter $0 \leq \nu \leq n-1$. Dann gibt es genau einen Endomorphismus $T \in L(V)$ mit*

$$T\,v_k \;=\; \begin{cases} v_{k+1} & (1 \leq k \leq \nu)\,, \\ 0 & (\nu < k \leq n)\,. \end{cases}$$

Berechne dessen Eigenwerte und ihre algebraischen und geometrischen Vielfachheiten.

Aufgabe 7.5.15 *Berechne das charakteristische Polynom sowie alle Eigenwerte mit ihren jeweiligen geometrischen und algebraischen Vielfachhheiten für die Matrix*

$$N \;=\; \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix} \tag{7.5.4}$$

Vergleiche mit der vorangegangenen Aufgabe.

Satz 7.5.16 *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$. Für jeden Eigenwert λ von T ist die geometrische Vielfachheit höchstens gleich der algebraischen Vielfachheit.*

Beweis: Sei $(v_1, \ldots, v_s)$ Basis des Eigenraums von T zum Eigenwert λ. Nach dem Basisergänzungssatz können wir dieses System zu einer Basis $(v_1, \ldots, v_n)$ von V erweitern. Bezüglich dieser Basis hat die Darstellungsmatrix A von T die Form

$$A \;=\; [\lambda\, e_1, \ldots, \lambda\, e_s, a_{s+1}, \ldots, a_n]$$

wobei e_k der k-te Einheitsvektor der kanonischen Basis von $\mathbb{K}^n$ und die $a_j \in \mathbb{K}^n$ sind. Daraus folgt für das charakteristische Polynom p_T:

$$p_T(t) \;=\; \det(A - t\,I) \;=\; (\lambda - t)^s\; q(t)\,,$$

mit

$$q(t) \;=\; \det \begin{bmatrix} a_{s+1,s+1} - t & a_{s+1,s+2} & \ldots & a_{s+1,n-1} & a_{s+1,n} \\ a_{s+2,s+1} & a_{s+2,s+2} - t & \ldots & a_{s+2,n-1} & a_{s+2,n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,s+1} & a_{n-1,s+2} & \ldots & a_{n-1,n-1} - t & a_{n-1,n} \\ a_{n,s+1} & a_{n,s+2} & \ldots & a_{n,n-1} & a_{nn} - t \end{bmatrix}.$$

Daraus folgt, dass die algebraische Vielfachhheit des Eigenwertes mindestens s ist, und dies ist nach Definition die geometrische Vielfachhheit. $\square$

Aufgabe 7.5.17 *Zeige: Ein Endomorphismus $T \in L(V)$ eines endlich-dimensionalen Vektorraumes V über $\mathbb{K}$ ist genau dann diagonalisierbar, wenn alle Nullstellen des charakteristischen Polynoms von T in $\mathbb{K}$ liegen und daher auch Eigenwerte von T sind, und wenn zusätzlich die geometrische Vielfachheit jedes Eigenwertes gleich seiner algebraischen Vielfachheit ist.*

Aufgabe 7.5.18 (Eigenwerte und Eigenvektoren mit MAPLE)
Finde selber heraus, inwieweit MAPLE *die Eigenwerte und Eigenvektoren einer Matrix berechnen kann.*

Kapitel 8

Normalformen und Definitheit von Matrizen

In diesem Kapitel sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$, und V enthalte nicht nur den Nullvektor. Wir wollen untersuchen, wie die „einfachste Form“ der Darstellungsmatrix eines gegebenen Endomorphismus von V aussieht. Da eine Darstellungsmatrix immer bis auf Ähnlichkeit bestimmt ist, ist dies äquivalent zur Frage nach einer „einfachsten Form“ einer quadratischen Matrix A unter Ähnlichkeit, oder genauer, nach einem möglichst einfachen Repräsentanten innerhalb jeder Ähnlichkeitsklasse von Matrizen. Beachte, dass in vielen der folgenden Resultate vorausgesetzt wird, dass die Nullstellen des charakteristischen Polynoms eines Endomorphismus oder einer Matrix zu dem Skalarenkörper $\mathbb{K}$ gehören; diese Voraussetzung ist natürlich für den Fall $\mathbb{K} = \mathbb{C}$ immer erfüllt.

8.1 Trigonalisierung

Satz 8.1.1 *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$ so, dass alle Nullstellen des charakteristischen Polynoms von T in $\mathbb{K}$ liegen. Dann gibt es eine Basis $(v_1, \ldots, v_n)$ von V, bezüglich der die Darstellungsmatrix von T obere Dreiecksgestalt hat. Auf der Diagonalen dieser Darstellungsmatrix stehen dann die Eigenwerte von T entsprechend ihrer algebraischen Vielfachheit, und man kann sogar die Reihenfolge dieser Eigenwerte beliebig vorschreiben. Falls V ein Raum mit Skalarprodukt ist, kann zusätzlich $(v_1, \ldots, v_n)$ sogar als Orthonormalbasis gewählt werden.*

Beweis: Für $n = \dim V = 1$ ist nichts zu zeigen. Sei jetzt $n \geq 2$, und sei der Satz für $n-1$ schon bewiesen. Wir wählen einen beliebigen Eigenwert λ und dazu einen Eigenvektor $v \in V$ von T. Nach dem Basisergänzungssatz können wir $(v_1 = v)$ zu einer Basis $(v_1, \ldots, v_n)$ von V ergänzen, und im Fall dass V ein Skalarprodukt besitzt, kann dies sogar eine Orthonormalbasis sein. Die Darstellungsmatrix von T bezüglich dieser Basis hat dann die Form

$$A = \begin{bmatrix} \lambda & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \ldots & a_{nn} \end{bmatrix} = A_1 + A_2 \,,$$

mit

$$A_1 \;=\; \begin{bmatrix} \lambda & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad A_2 \;=\; \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Wir definieren jetzt Endomorphismen T_1 bzw. T_2 gerade so, dass ihre Darstellungsmatrizen bezüglich der gewählten Basis gleich A_1 bzw. A_2 sind. Dann gilt $T = T_1 + T_2$, und T_2 bildet den Raum $W = \mathcal{L}(v_2, \dots, v_n)$ in sich ab. Das charakteristische Polynom von T hat die Form $p(t) = (\lambda - t)\, q(t)$ mit $q(t) = \det(B - t\, I)$, wobei

$$B \;=\; \begin{bmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Man erkennt aber, dass B gerade die Darstellungsmatrix von T_2, als Abbildung von W in sich, bezüglich der Basis $(v_2, \dots, v_n)$ ist. Daher gehören alle Nullstellen des charakteristischen Polynoms von T_2 zu $\mathbb{K}$, und wir können nach Induktionshypothese schließen, dass es eine Basis, ja sogar eine Orthonormalbasis im Fall eines Raums mit Skalarprodukt, $(w_2, \dots, w_n)$ von W gibt, bezüglich der die Darstellungsmatrix von T_2 obere Dreiecksform hat, mit dem entsprechenden Zusatz für die Reihenfolge der Diagonalelemente. Dann ist $(v_1, w_2, \dots, w_n)$ eine Basis von V und sogar Orthonormalbasis bei Vorliegen eines Skalarproduktes, da ja v_1 zu jedem Vektor aus W orthogonal ist. Die Form der Darstellungsmatrizen von T_1 und T_2 (wobei T_2 jetzt wieder als Abbildung von V in sich aufgefasst wird) hat sich bei diesem Basiswechsel nicht verändert, aber die für T_2 hat weitere Nullen in allen Positionen unterhalb der Diagonalen. Zusmmengenommen zeigt das, dass die Darstellungsmatrix von T bezüglich der neuen Basis obere Dreiecksform hat. □

Definition 8.1.2 *Wir nennen zwei Matrizen $A, B \in \mathbb{K}^{n\times n}$* unitär ähnlich*, wenn es eine unitäre bzw. orthogonale Matrix $U \in \mathbb{K}^{n\times n}$ gibt, für welche $B = \overline{U}^T A U$ ist. Beachte, dass wegen $U^{-1} = \overline{U}^T$ die Matrizen B und A dann auch ähnlich zueinander sind. Wegen* Lemma 7.4.9 *sind die Darstellungsmatrizen eines Endomorphismus bezüglicher zweier Orthonormalbasen immer unitär ähnlich.*

Korollar zu Satz 8.1.1 (Schursches Lemma) *Jede Matrix $A \in \mathbb{K}^{n\times n}$, für welche alle Nullstellen des charakteristischen Polynoms in $\mathbb{K}$ liegen, ist unitär ähnlich zu einer oberen Dreiecksmatrix.*

Beweis: Die Matrix A ist die Darstellungsmatrix des Endomorphismus $x \longmapsto A\,x$ bezüglich der kanonischen Basis in $\mathbb{K}^n$. Diese ist eine Orthonormalbasis bezüglich des kanonischen Skalarproduktes in $\mathbb{K}^n$, und nach Satz 8.1.1 gibt es eine andere Orthonormalbasis in $\mathbb{K}^n$, so dass die neue Darstellungsmatrix B ober Dreiecksgestalt hat. Nach Lemma 7.4.9 wird der Übergang von der kanonischen zur neuen Orthonormalbasis gerade von einer unitären Matrix bewirkt, und daraus folgt die Behauptung. □

8.2 Der Satz von Cayley-Hamilton

Definition 8.2.1 *Sei V ein beliebiger Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$. Für eine natürliche Zahl $k \in \mathbb{N}$ definieren wir T^k als die k-malige Hintereinanderausführung der Abbildung T. Setzt man noch T^0 gleich der identischen Abbildung, so können wir dann für jedes Polynom $p(t) = \sum_{k=0}^m p_k\, t^k$ mit Koeffizienten $p_k \in \mathbb{K}$ die Abbildung*

$$p(T) \;=\; \sum_{k=0}^{m} p_k\, T^k$$

bilden und stellen fest, dass $p(T) \in L(V)$ ist. Wir erhalten somit für jedes feste $T \in L(V)$ eine Abbildung $p \longmapsto p(T)$ von $\mathbb{K}[t]$ in $L(V)$, und es ist nicht schwer zu sehen, dass diese Abbildung linear ist. Sie ist zusätzlich auch multiplikativ *im Sinne der nächsten Aufgabe.*

Aufgabe 8.2.2 *Überprüfe folgende Rechenregel: Für Polynome $p_1, p_1 \in \mathbb{K}[t]$ und $T \in L(V)$ gilt immer*

- $$p = p_1\, p_2 \quad \Longrightarrow \quad p(T) = p_1(T) \circ p_2(T).$$

Bemerkung 8.2.3 *Für eine Matrix $A \in \mathbb{K}^{n \times n}$ definiert man sinnvollerweise A^k als das k-fache Produkt von A mit sich selber und setzt $A^0 = I$. Damit ist dann auch für jedes $p \in \mathbb{K}[t]$ eine Matrix $p(A) \in \mathbb{K}^{n \times n}$ definiert. Wählt man in der Situation der obigen Definition eine Basis von V, und ist dann A die Darstellungmatrix von T, so folgt aus* Satz 7.1.8, *dass $p(A)$ die Darstellungsmatrix von $p(T)$ ist.*

Satz 8.2.4 (Cayley-Hamilton) *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$. Wenn p das charakteristische Polynom von T ist, dann folgt $p(T) = 0$, wobei 0 die Nullabbildung in $L(V)$ bezeichnet.*

Beweis: Wir unterscheiden zwei Fälle:

(a) Sei $\mathbb{K} = \mathbb{C}$. Dann können wir nach Satz 8.1.1 eine Basis $(v_1, \ldots, v_n)$ von V wählen, bezüglich der die Darstellungsmatrix $A = [a_{jk}]$ von T obere Dreiecksform hat. In diesem Fall ist (7.1.1) äquivalent zu der Beziehung

$$(T - a_{kk}\, id)\, v_k \;=\; \sum_{j=1}^{k-1} a_{jk}\, v_j \qquad \forall\, k = 1, \ldots, n\,, \tag{8.2.1}$$

wobei id wieder die identische Abbildung in V bedeutet. Für das charakteristische Polynom ergibt sich in diesem Fall die Darstellung

$$p(t) \;=\; \prod_{k=1}^{n} (a_{kk} - t)\,.$$

Wir wollen nun durch Induktion beweisen, dass für jedes $k = 1, \ldots, n$ gilt

$$\big(a_{kk}\, id - T\big) \circ \ldots \circ \big(a_{11}\, id - T\big)\, v_j \;=\; 0 \qquad \forall\, 1 \le j \le k\,. \tag{8.2.2}$$

Für $k = 1$ folgt dies aus (8.2.1). Sei jetzt der Beweis für ein $k \in \{1, \ldots, n-1\}$ schon geführt. Dann folgt jedenfalls durch Anwenden von $\big(a_{k+1,k+1}\, id - T\big)$ auf beide Seiten von(8.2.2)

$$\big(a_{k+1,k+1}\, id - T\big) \circ \big(a_{kk}\, id - T\big) \circ \ldots \circ \big(a_{11}\, id - T\big)\, v_j \;=\; 0 \quad \forall\, 1 \le j \le k\,.$$

Also ist noch der Fall $j = k+1$ zu behandeln. In diesem Fall können wir benutzen, dass die Abbildungen $\big(a_{\nu\nu}\, id - T\big)$ miteinander kommutieren, und deshalb folgt mit Hilfe von (8.2.1) und der Induktionshypothese

$$\begin{aligned}
\big(a_{k+1,k+1}\, id - T\big) \circ \big(a_{kk}\, id - T\big) \circ \ldots \circ \big(a_{11}\, id - T\big)\, v_{k+1} \quad &= \\
\big(a_{kk}\, id - T\big) \circ \ldots \circ \big(a_{11}\, id - T\big) \circ \big(a_{k+1,k+1}\, id - T\big)\, v_{k+1} \quad &= \\
-\big(a_{kk}\, id - T\big) \circ \ldots \circ \big(a_{11}\, id - T\big) \sum_{j=1}^{k} a_{j,k+1}\, v_j \quad &= \quad 0\,.
\end{aligned}$$

Also ist (8.2.2) auch für $k+1$, und deshalb für alle $k = 1, \ldots, n$ richtig. Wendet man dies an für den Fall $k = n$, so folgt $p(T)\, v_j = 0$ für alle $j = 1, \ldots, n$. Da jedes $v \in V$ eine Linearkombination der Basisvektoren v_j ist, folgt daraus wiederum $p(T)\, v = 0$ für alle $v \in V$, und das war zu zeigen.

(b) Sei jetzt $\mathbb{K} = \mathbb{R}$. Sei A die Darstellungsmatrix von T zu irgendeiner Basis von V. Dann ist $x \longmapsto A\, x$ ein Endomorphismus von $\mathbb{C}^n$ mit dem gleichen charakteristischen Polynom p wie T. Aus (a) folgt deshalb $p(A) = 0$, und da $p(A)$ die Darstellungsmatrix von $p(T)$ ist, folgt hieraus wiederum $p(T) = 0$. □

Definition 8.2.5 (Minimalpolynom) *Nach dem Satz von Cayley-Hamilton gibt es zu jedem Endomorphismus T eines n-dimensionalen Vektorraums V, mit $n \in \mathbb{N}$, ein nicht-triviales Polynom p, für welches $p(T) = 0$ ist. In der Menge aller dieser Polynome gibt es immer eines mit minimalem Grad und höchstem Koeffizienten 1; wie wir gleich zeigen, ist dieses Polynom durch T eindeutig festgelegt, und es heißt* Minimalpolynom zu *T.*

Lemma 8.2.6 *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und sei $T \in L(V)$. Dann gibt es ein Polynom $p \in \mathbb{K}[t] \setminus \{0\}$, welches folgende beiden Eigenschaften hat:*

(a) $p(T) = 0$.

(b) *Ist $p_1 \in \mathbb{K}[t]$ so, dass $p_1(T) = 0$ ist, so gibt es ein $q \in \mathbb{K}[t]$ mit $p_1 = q\,p$.*

Das Polynom p ist durch diese beiden Eigenschaften bis auf einen konstanten Faktor eindeutig festgelegt und hat unter allen $p_1 \in \mathbb{K}[t] \setminus \{0\}$ mit der Eigenschaft $p_1(T) = 0$ den kleinsten Grad.

Beweis: Aus dem Satz von Cayley-Hamilton folgt die Existenz von Polynomen $p \in \mathbb{K}[t] \setminus \{0\}$ mit $p(T) = 0$, und unter allen diesen *wählen* wir ein p mit minimalem Grad. Ist p_1 irgendein Polynom in $\mathbb{K}[t]$, so gibt es immer $q, r \in \mathbb{K}[t]$, für welche $p_1 = q\,p + r$ gilt, wobei r entweder das Nullpolynom ist oder einen echt kleineren Grad als p besitzt. Wenn dann auch $p_1(T) = 0$ ist, folgt $r(T) = 0$, woraus folgt dass r tatsächlich das Nullpolynom ist, denn sonst ergäbe sich ein Widerspruch zur Wahl von p. Also sind (a) und (b) erfüllt. Die Eindeutigkeit von p folgt aber ebenfalls, denn wenn p_1 und p denselben Grad haben, muss in (b) q konstant sein. □

Aufgabe 8.2.7 (Minimalpolynom und MAPLE) *Finde heraus, wie man mit* MAPLE *Minimalpolynome berechnen kann.*

Aufgabe 8.2.8 (Nullstellen des Minimalpolynoms) *Zeige, dass das Minimalpolynom eines Endomorphismus genau die gleichen Nullstellen wie sein charakteristisches Polynom hat, allerdings im Allgemeinen mit geringerer Vielfachheit. Gib Beispiele von $n \times n$- Matrizen an, für die das Minimalpolynom ersten Grades ist.*

Aufgabe 8.2.9 (Nilpotente Abbildungen) *Man nennt einen Endomorphismus $T \in L(V)$* nilpotent, *falls ein $n \in \mathbb{N}$ existiert, für welches $T^n = 0$ ist. Das kleinste $n \in \mathbb{N}_0$ mit $T^n = 0$ heißt auch* der Nilpotenzgrad *von T. Bestimme für endlich-dimensionales V das charakteristische Polynom sowie die Form des Minimalpolynoms eines nilpotenten Endomorphismus.*

8.3 Hauptachsentransformation

Satz 8.3.1 (Satz von der Hauptachsentransformation) *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ mit Skalarprodukt, und sei $T \in L(V)$ selbstadjungiert. Dann gibt es eine Orthonormalbasis in V, bezüglich der die Darstellungsmatrix von T diagonal ist. Anders ausgedrückt heißt das: Es gibt eine Orthonormalbasis von V, die aus Eigenvektoren von T besteht.*

Beweis: Nach Satz 7.5.9 (d) sind alle Nullstellen des charakteristischen Polynoms von T reell, also sicher in $\mathbb{K}$. Daher kann Satz 8.1.1 angewandt werden und sichert, dass eine Orthonormalbasis von V existiert, bezüglich der die Darstellungsmatrix von T obere Dreiecksform hat. Nach dem Korollar zu Satz 7.4.1 ist aber diese Darstellungsmatrix hermitesch. Die einzigen hermiteschen Dreiecksmatrizen sind aber die Diagonalmatrizen. Das ist die Behauptung. □

Definition 8.3.2 *Wir nennen eine Matrix $A \in \mathbb{K}^{n\times n}$* unitär diagonalisierbar*, wenn es eine unitäre bzw. orthogonale Matrix $U \in \mathbb{K}^{n\times n}$ gibt, für welche $B = \overline{U}^T A\, U$ diagonal ist.*

Korollar zu Satz 8.3.1 *Jede hermitesche Matrix A ist unitär diagonalisierbar. Genauer: Ist $n \in \mathbb{N}$, und ist $A \in \mathbb{K}^{n\times n}$ hermitesch, so gibt es eine unitäre Matrix $U \in \mathbb{K}^{n\times n}$ und eine Diagonalmatrix Λ, für welche*

$$A\,U \;=\; U\,\Lambda$$

ist. Die Diagonalelemente von Λ sind gerade die Eigenwerte von A, und die Spalten von U sind die Eigenvektoren von A, in der zu den Diagonalelementen von Λ passenden Reihenfolge.

Beweis: Die Abbildung $x \longmapsto A\,x$ ist ein selbstadjungierter Endomorphismus in $\mathbb{K}^n$, und damit folgt die Behauptung aus dem Satz von der Hauptachsentransformation. $\square$

Aufgabe 8.3.3 *Berechne ein System von 3 orthonormalen Eigenvektoren zur Matrix*

$$A \;=\; \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

8.4 Diagonalisierung normaler Endomorphismen

Satz 8.4.1 *Sei $n \in \mathbb{N}$, sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ mit Skalarprodukt, und sei $T \in L(V)$. Dann sind folgende Aussagen äquivalent:*

(a) *T ist normal, und alle Nullstellen des charakteristischen Polynoms liegen in $\mathbb{K}$.*

(b) *Es gibt eine Orthonormalbasis von V, die aus Eigenvektoren von T besteht.*

Beweis: Wenn es eine Orthonormalbasis aus Eigenvektoren von T gibt, dann ist die zugehörige Darstellungsmatrix von T diagonal, und die Diagonalelemente sind gerade die Nullstellen des charakteristischen Polynoms, liegen also in $\mathbb{K}$. Da jede Diagonalmatrix eine normale Matrix ist, folgt aus dem Korollar zu Satz 7.4.1 die Normalität von T. Also folgt (a) aus (b). Für die Umkehrung gehen wir induktiv vor: Für $n = 1$ ist nichts zu zeigen. Sei jetzt $n \geq 2$, und sei der Satz für $n-1$ bewiesen. Sei v ein beliebiger Eigenvektor von T zu einem Eigenwert λ, und sei U das orthogonale Komplement zu v, also die Menge aller $u \in V$ mit $\langle v, u\rangle = 0$. Dann gilt $\langle v, T\,u\rangle = \langle T^{ad}\, v, u\rangle$, und nach Satz 6.9.5 (e) folgt $T^{ad}\, v = \overline{\lambda}\, v$. Also folgt $\langle v, T\,u\rangle = 0$ für alle $u \in U$, und somit bildet T den $(n-1)$-dimensionalen Raum U in sich ab. Nach Induktionshypothese gibt es also in U eine Orthonormalbasis aus Eigenvektoren von T, und diese kann durch Hinzunahme von v (jedenfalls wenn man vorher noch v normiert, d. h., durch Multiplikation mit einem Faktor dafür sorgt, dass $\|v\| = 1$ ist) zu einer Orthonormalbasis von V erweitert werden. $\square$

Aufgabe 8.4.2 *Zeige: Eine normale Matrix ist genau dann hermitesch, wenn alle Nullstellen ihres charakteristischen Polynoms reelle Zahlen sind. Schließe hieraus, dass der Satz über die Hauptachsentransformation aus* Satz 8.4.1 *folgt.*

Korollar zu Satz 8.4.1 *Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann normal, wenn sie unitär ähnlich zu einer Diagonalmatrix ist.*

Beweis: Wenn A unitär änlich zu B ist, ist A genau dann normal, wenn B normal ist. Da Diagonalmatrizen immer normal sind, folgt die eine Richtung des Satzes. Wenn A eine normale Matrix ist, dann ist auch der Endomorphismus $x \longmapsto A\,x$ normal in $\mathbb{C}^n$, und nach Satz 8.4.1 gibt es eine Orthonormalbasis aus Eigenvektoren dieses Endomorphismus. Daraus ergibt sich die unitäre Ähnlichkeit von A zu einer Diagonalmatrix. $\square$

8.5 Definite Matrizen

Definition 8.5.1 *Sei $n \in \mathbb{N}$, und sei $A \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix. Die Funktion*

$$q_A(x) \;=\; \overline{x}^T A\,x \;=\; \langle x, A\,x\rangle \qquad \forall\, x \in \mathbb{K}^n\,,$$

wobei rechts das kanonische Skalarprodukt in $\mathbb{K}^n$ steht, heißt die zu A gehörige quadratische Form. *Beachte, dass die quadratische Form wegen* Aufgabe 6.6.7 *immer reell ist. Analog zu* Abschnitt 6.8 *nennen wir die Matrix A* positiv semidefinit, *falls $q_A(x) \geq 0$ für alle $x \in \mathbb{R}^n$, und* positiv definit, *falls sogar $q_A(x) > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ ist. Entsprechend heißt A* negativ semidefinit *bzw.* negativ definit, *falls $-A$ positiv semidefinit bzw. positv definit ist. Schließlich heißt A* indefinit, *falls es weder positiv noch negativ semidefinit ist, d. h., falls es $x_1, x_2 \in \mathbb{R}^n$ gibt mit $q_A(x_1) < 0$ und $q_A(x_2) > 0$.*

Aufgabe 8.5.2 *Zeige: Die Einheitsmatrix ist positiv definit; eine Diagonalmatrix ist genau dann positiv definit bzw. semidefinit, wenn alle Diagonalelemente reell und positiv bzw. nicht-negativ sind. Finde selbst die entsprechende Charakterisierung von negativ (semi)definiten und indefiniten Diagonalmatrizen.*

Aufgabe 8.5.3 (Definitheit und MAPLE) *Finde heraus, wie man mit dem Programm* MAPLE *die Definitheitseigenschaft einer Matrix prüfen kann.*

Satz 8.5.4 *Sei $n \in \mathbb{N}$, und sei $A \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix.*

(a) *Genau dann ist A positiv semidefinit, wenn alle Eigenwerte von A nicht-negativ sind.*

(b) *Genau dann ist A positiv definit, wenn alle Eigenwerte von A positiv sind.*

(c) *Genau dann ist A negativ semidefinit, wenn alle Eigenwerte von A nicht-positiv sind.*

(d) *Genau dann ist A negativ definit, wenn alle Eigenwerte von A negativ sind.*

(e) *Genau dann ist A indefinit, wenn ein Eigenwert von A positiv und ein anderer negativ ist.*

Beweis: Aus Satz 6.8.4 folgt, dass unitär ähnliche Matrizen die gleiche Definitheitseigenschaft haben. Nach dem Satz über die Hauptachsentransformation ist A unitär ähnlich zu einer Diagonalmatrix, deren Diagonalelemente gerade die Eigenwerte von A sind. Daraus folgt mit Aufgabe 8.5.2 die Behauptung. $\square$

Definition 8.5.5 *Für eine Matrix $A = [a_{jk}] \in \mathbb{K}^{n\times n}$ heißen die Zahlen*

$$\det \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{bmatrix} \qquad \forall\; k \;=\; 1, \dots, n$$

die Hauptunterdeterminanten *von A. Die Zahl k heißt dabei auch* die Ordnung *der Hauptunterdeterminante. Wenn wir zwei Zeilen von A und anschließend die zwei Spalten von A mit den gleichen Nummern vertauschen, sprechen wir auch von einer* simultanen Vertauschung von zwei Zeilen und Spalten *von A. Die Hauptunterdeterminanten jeder Matrix B, die aus A durch endlich viele simultane Vertauschungen von Zeilen und Spalten hervorgeht, heißen auch* allgemeine *Hauptunterdeterminanten von A. Anders ausgedrückt sind die allgemeinen Hauptunterdeterminanten von A gleich den Determinanten der aus A durch Streichen von Zeilen und Spalten gleicher Nummern entstehenden Matrizen.*

Satz 8.5.6 *Sei $n \in \mathbb{N}$, und sei $A \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix.*

(a) *Wenn A positiv definit ist, dann sind alle allgemeinen Hauptunterdeterminanten von A positiv; insbesondere sind also alle Diagonalelemente positiv.*

(b) *Wenn alle Hauptunterdeterminanten von A positiv sind, dann existiert genau eine obere Dreiecksmatrix B mit reellen und positiven Diagonalelementen, so dass*

$$A \;=\; \overline{B}^{T}\, B \tag{8.5.1}$$

ist, und hieraus folgt $q_{A}(x) = \|y\|^2$, mit $y = B\,x$, sodass A positiv definit ist.

Beweis: Sei A positiv definit. Für $1 \le k \le n$ sei U_k der Unterraum aller Vektoren $x = (x_1, \ldots, x_n)^{T} \in \mathbb{K}^n$ mit $x_{k+1} = \ldots = x_n = 0$. Wenn wir die quadratische Form q_{A} auf U_k einschränken, ergibt dies die quadratische Form auf $\mathbb{K}^k$ mit der Matrix

$$A_k \;=\; \begin{bmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \ldots & a_{kk} \end{bmatrix},$$

und diese Matrix muss daher auch positiv definit sein. Nach Satz 8.5.4 müssen also alle Eigenwerte von A_k positiv sein, und nach Aufgabe 7.5.12 folgt hieraus $\det A_k > 0$. Eine simultane Vertauschung von Zeilen und Spalten von A ist gleichwertig zu einer Umnumerierung der Variablen x_k und ändert nichts an der positiven Definitheit von A. Damit ist (a) bewiesen. Den Beweis von (b) führen wir induktiv: Für $n = 1$ ist nichts zu zeigen, deshalb sei jetzt $n \ge 2$ und der Beweis für Matrizen der Größe $(n-1)\times(n-1)$ schon erbracht. Wir schreiben

$$A \;=\; \left[\begin{array}{c|c} A_{n-1} & a \\ \hline \overline{a}^{T} & c \end{array}\right],$$

wobei wir die Tatsache ausnützen, dass A hermitesch ist. Nach Induktionshypothese gibt es genau eine obere Dreiecksmatrix B_{n-1} mit reellen und positiven Diagonalelementen, so dass $A_{n-1} = \overline{B}_{n-1}^{T}\, B_{n-1}$. Es folgt

$$q_{A}(x) \;=\; \overline{y}^{T}\, C\, y\,, \qquad C \;=\; \left[\begin{array}{c|c} I & b \\ \hline \overline{b}^{T} & c \end{array}\right], \quad y \;=\; \left[\begin{array}{c|c} B_{n-1} & 0 \\ \hline 0 & 1 \end{array}\right]\, x\,,$$

wobei $b = (\overline{B}_{n-1}^{T})^{-1}\, a$ ist. Mit dem Entwicklungssatz sieht man $\det C = c - \overline{b}^{T}\, b$, und aus dem Determinantenmultiplikationssatz folgt dass A und C dieselbe Determinante haben. Also folgt aus der Voraussetzung dass $\det A = c - \overline{b}^{T}\, b > 0$ ist, und wir setzen

$$d = \sqrt{c - \overline{b}^{T}\, b}\,.$$

Daraus folgt aber jetzt

$$A \;=\; \overline{B}^{T}\, B\,, \qquad B \;=\; \left[\begin{array}{c|c} B_{n-1} & b \\ \hline 0 & d \end{array}\right].$$

Offenbar ist B obere Dreiecksmatrix mit positiv-reellen Diagonalelementen, und $q_{A}(x) = \overline{y}\, y$ mit $y = B\,x$. Man erkennt auch leicht, dass dieses B durch A eindeutig bestimmt ist. Also gilt (b) auch für Matrizen der Größe $n \times n$, was zu zeigen war. □

Definition 8.5.7 *Nach dem vorausgegangenen Satz gibt es zu jeder positiv definiten Matrix A genau eine obere Dreiecksmatrix B mit reell-positiven Diagonalelementen, für die* (8.5.1) *gilt. Man nennt dies auch die* Cholesky-Zerlegung der Matrix A. *Tatsächlich besteht ein gängiges numerisches Verfahren zur Prüfung einer Matrix auf positive Definitheit in der Berechnung dieser Cholesky-Zerlegung.*

Bemerkung 8.5.8 *Wenn man von A zu $-A$ übergeht, dann bleiben alle allgemeinen Hauptunterdeterminanten mit gerader Ordnung erhalten, während die mit ungerader Ordnung ihr Vorzeichen wechseln. Daher kann man aus obigem Satz leicht Kriterien für negative Definitheit ableiten.*

Aufgabe 8.5.9 *Untersuche die Matrizen aus* Aufgabe 7.5.11 *auf Definitheit.*

Aufgabe 8.5.10 *Zeige für die Matrix*

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \qquad a, b, c \in \mathbb{R},$$

folgende Aussagen:

- *Genau dann ist A indefinit, wenn $a\,c < b^2$ ist.*
- *Genau dann ist A positiv definit, wenn $a\,c > b^2$ und $a > 0$ ist.*
- *Genau dann ist A negativ definit, wenn $a\,c > b^2$ und $a < 0$ ist.*

Daraus folgt, dass A in allen anderen Fällen semidefinit ist.

Kapitel 9

Ergänzungen

9.1 Kegelschnitte

Als eine Anwendung der quadratischen Formen wollen wir kurz folgende Aufgabe studieren:

- Gegeben sei eine reelle symmetrische Matrix $A = [a_{jk}] \in \mathbb{R}^{2\times 2}$, ein reeller Vektor $b \in \mathbb{R}^2$ sowie eine Zahl $c \in \mathbb{R}$. Untersuche die Lösungsmenge $L \subset \mathbb{R}^2$ der Gleichung

  $$x^T A\, x \;+\; b^T x \;+\; c \;=\; 0\,. \tag{9.1.1}$$

  Ausgeschrieben lautet diese Gleichung, unter Beachtung von $a_{21} = a_{12}$ wegen der Symmetrie von A,

  $$a_{11}\, x_1^2 \;+\; 2\, a_{12}\, x_1\, x_2 \;+\; a_{22}\, x_2^2 \;+\; b_1\, x_1 \;+\; b_2\, x_2 \;+\; c \;=\; 0\,.$$

  Man sieht daran, dass es sich um die allgemeinste quadratische Gleichung in zwei Variablen x_1, x_2 handelt.

Die Lösungsmenge einer solchen Gleichung ist, bis auf entartete Fälle, *ein Kegelschnitt*, d. h. die Schnittmenge (in $\mathbb{R}^3$) einer Ebene mit einem Doppelkegel; darauf soll hier aber nicht eingegangen werden.

Zur Lösung der Aufgabe beachten wir, dass nach dem Satz über die Hauptachsentransformation eine orthogonale Matrix $U \in \mathbb{R}^{2\times 2}$ existiert, für welche $\Lambda = U^T A\, U$ eine Diagonalmatrix ist. Setzt man $x = U\, y$ in (9.1.1) ein, und bezeichnet man den Vektor $b^T U$ der Einfachheit halber wieder mit b^T, so erhält man die neue Gleichung

$$y^T \Lambda\, y \;+\; b^T y \;+\; c \;=\; 0.$$

Wir können o. B. d. A. noch annehmen, dass $\det U = 1$ ist, denn sonst kann man zum Beispiel die erste Spalte von U mit einem entsprechenden Faktor multiplizieren, was die Orthogonalität bzw. Invertierbarkeit von U nicht verändert (denn im Falle einer orthogonalen Matrix ist ja $\det U = 1$ oder $\det U = -1$).

Nach Aufgabe 6.7.6 ist die Abbildung $x = U\, y$ im Fall einer orthogonalen Matrix U mit $\det U = 1$ eine Drehung von $\mathbb{R}^2$, oder anders ausgedrückt, die Einführung eines neuen rechtwinkeligen Koordinatensystems, welches gegenüber dem vorigen verdreht ist. Aus diesem Grund genügt es für die Interpretation der Lösungsmenge, wenn wir im Folgenden annehmen, dass A von vorne herein eine Diagonalmatrix ist, sodass die Gleichung (9.1.1) die Form

$$a_1\, x_1^2 \;+\; a_2\, x_2^2 \;+\; b_1\, x_1 \;+\; b_2\, x_2 \;+\; c \;=\; 0\,.$$

Wir unterscheiden jetzt die folgenden grundsätzlich verschiedenen Fälle:

1. Sei $a_1\, a_2 \neq 0$: Durch quadratische Ergänzung schreiben wir die Gleichung um in die Form
$$a_1 \left(x_1 + \frac{b_1}{2\,a_1}\right)^2 \;+\; a_2 \left(x_2 + \frac{b_2}{2\,a_2}\right)^2 \;=\; \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2} \;-\; c\,. \tag{9.1.2}$$
Mit den Abkürzungen
$$x_j^{(0)} \;=\; -\frac{b_j}{2\,a_j}\,, \qquad \lambda_j = \sqrt{\left|a_j^{-1}\right| \left|\frac{b_1^2}{4\,a_1} + \frac{b_2^2}{4\,a_2} - c\right|}$$
für $j = 1$ und $j = 2$ ergeben sich folgende Unterfälle:

   (a) Seien a_1, a_2 beide positiv: Dann ist die linke Seite von (9.1.2) immer positiv, und wir schließen:

      i. Falls $c > \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2}$ ist, ist die Lösungsmenge L der Gleichung leer.

      ii. Falls $c = \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2}$ ist, ist $L = \{x_0 = (x_1^{(0)},\, x_2^{(0)})^T\,\}$.

      iii. Falls $c < \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2}$ ist, bringen wir die Gleichung (9.1.2) auf die Form
$$\left(\frac{x_1 - x_1^{(0)}}{\lambda_1}\right)^2 \;+\; \left(\frac{x_2 - x_2^{(0)}}{\lambda_2}\right)^2 \;=\; 1\,.$$
Daraus erkennt man, dass L geometrisch eine Ellipse mit Mittelpunkt $x_1^{(0)}, x_2^{(0)}$ und achsenparallelen Halbachsen λ_1, λ_2 ist. Falls $\lambda_1 = \lambda_2 = 1$ ist (was man mit Hilfe des Sylvesterschen Trägheitssatzes erreichen kann), handelt es sich offenbar um einen Kreis mit dem angegeben Mittelpunkt und dem Radius $r = 1$.

   (b) Seien $a_1, a_2 < 0$: Diesen Unterfall kann man durch Multiplikation der Gleichung (9.1.2) mit -1 auf den ersten zurückführen.

   (c) Seien $a_1 > 0$ und $a_2 < 0$: Dann kann die linke Seite von (9.1.2) sowohl positiv als auch negativ werden. In jedem Fall gibt es folgende Möglichkeiten:

      i. Falls $c = \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2}$ ist, ist (9.1.2) äquivalent zu
$$x_2 \;=\; x_2^{(0)} \;\pm\; \frac{\lambda_2}{\lambda_1}\,(x_1 - x_1^{(0)})\,,$$
so dass L ein Paar von sich schneidenden Geraden ist.

      ii. Falls $c > \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2}$ ist, ist (9.1.2) äquivalent zu
$$\left(\frac{x_1 - x_1^{(0)}}{\lambda_1}\right)^2 \;-\; \left(\frac{x_2 - x_2^{(0)}}{\lambda_2}\right)^2 \;=\; -1\,.$$
Damit ist L geometrisch ein Hyperbelpaar, welches nach unten bzw. oben geöffnet ist.

      iii. Falls $c < \frac{b_1^2}{4\,a_1} \;+\; \frac{b_2^2}{4\,a_2}$ ist, können wir durch Vertauschen von x_1 und x_2, also geometrisch durch Spiegelung an der Geraden $x_2 = x_1$, aus dem vorangegangenen Fall ablesen, dass L wieder ein Paar von Hyperbeln ist, die aber jetzt nach rechts bzw. links geöffnet sind.

   (d) Seien $a_1 < 0$ und $a_2 > 0$: Multiplikation mit -1 führt wieder auf den vorherigen Fall.

2. Sei $a_2 = 0$, $a_1 \neq 0$: Division durch a_1 erlaubt dann anzunehmen, dass $a_1 = 1$ ist. Die Gleichung kann dann in der Form
$$\left(x_1 + b_1/2\right)^2 + b_2\, x_2 \;=\; b_1^2/4 \;-\; c$$
geschrieben werden. Es ergeben sich dann folgende Unterfälle:

   (a) Sei $b_2 = 0$: Dann ist L leer, falls $c > b_1^2/4$ ist, und sonst ergeben sich für $c = b_1^2/4$ eine und für $c < b_1^2/4$ zwei parallele Gerade.

   (b) Sei $b_2 \neq 0$: Dann stellt L eine Parabel dar, die nach unten bzw. oben geöffnet ist, wenn $b_2 > 0$ bzw. $b_2 < 0$ ist.

3. Sei $a_2 = 0$, $a_1 \neq 0$: Vertauschen von x_1 und x_2 führt diesen auf den vorangegangenen Fall zurück.
4. Seien $a_1 = a_2 = 0$: Dann liegt eine lineare Gleichung vor, dessen Lösungsmenge L entweder leer oder eine Gerade ist.

Aufgabe 9.1.1 *Drücke die oben aufgetretenen Fälle direkt durch die Definitheitseigenschaften der Matrix A in* (9.1.1) *aus.*

9.2 Drehungen und Spiegelungen

Definition 9.2.1 *Eine orthogonale Matrix $U \in \mathbb{R}^{n\times n}$ erfüllt $U^T U = I$, woraus mit dem Determinantenmultiplikationssatz folgt* $\det U = \pm 1$. *Wenn* $\det U = 1$ *ist, nennt man U eine* Drehmatrix *oder auch einfach* Drehung *in $\mathbb{R}^n$. Wenn dagegen* $\det U = -1$ *ist, spricht man von einer* Drehspiegelung.

Satz 9.2.2 *Die Menge der unitären, der orthogonalen, sowie die Menge der Drehmatrizen sind Gruppen bezüglich der Matrixmultiplikation.*

Beweis: Da für die Matrixmultiplikation immer ein Assoziativgesetz gilt, ist nur zu zeigen, dass die Einheitsmatrix zu jeder der beiden Mengen gehört, und dass mit einer Matrix auch die inverse in der Menge ist. Dies ist aber klar wegen der Definition der entsprechenden Matrizenmenge sowie dem Determinantenmultiplikationssatz. □

Eine unitäre, und speziell eine orthogonale Matix ist immer normal, und deshalb nach Satz 8.4.1 unitär ähnlich zu einer Diagonalmatrix. Dabei sind auch für orthogonale, also reelle Matrizen die Eigenwerte im Allgemeinen komplexe Zahlen vom Betrag 1. Interessanter, und relativ leicht zu beantworten, ist aber die Frage nach einer Normalform einer orthogonalen Matrix bezüglich reell-unitärer, also orthogonaler Ähnlichkeit. Da orthogonale Matrizen die Darstellungsmatrizen von längentreuen Endomorphismen eines endlich-dimensionalen Vektorraumes über $\mathbb{R}$ sind, ist die Antwort auf diese Frage gleichwertig mit dem folgenden Satz:

Satz 9.2.3 *Sei $n \geq 2$, und sei V ein n-dimensionaler euklidischer Raum, also ein Raum über $\mathbb{R}$ mit Skalarprodukt. Sei weiter $T \in L(V)$ längentreu. Dann gibt es eine Orthonormalbasis $(v_1, \dots, v_n)$ von V, für welche die Darstellungsmatrix A von T die folgende Form hat:*

$$A = \left[\begin{array}{c|c|c|c|c|c} I_p & 0 & 0 & 0 & \dots & 0 \\ \hline 0 & -I_q & 0 & 0 & \dots & 0 \\ \hline 0 & 0 & A_1 & 0 & \dots & 0 \\ \hline 0 & 0 & 0 & A_2 & \dots & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & 0 & 0 & \dots & A_m \end{array}\right],$$

wobei I_p bzw. I_q Einheitsmatrizen vom Typ $p \times p$ bzw. $q \times q$ sind, während

$$A_k = \begin{bmatrix} \cos\phi_k & \sin\phi_k \\ -\sin\phi_k & \cos\phi_k \end{bmatrix} \qquad \forall\ k = 1, \dots, m,$$

mit $\phi_k \in \mathbb{R}$. Dabei ist $p + q + 2m = n$, wobei auch $p = 0$ oder $q = 0$ oder $m = 0$ zugelassen sein soll, was bedeutet dass die entsprechenden Matrizen nicht auftreten.

Beweis: Sei A zunächst die Darstellungsmatrix von T bezüglich irgendeiner Orthonormalbasis von V. Dann ist A eine orthogonale Matrix und somit auch normal, und deshalb gibt es nach Satz 8.4.1 eine unitäre Matrix U derart, dass $\overline{U}^T A\, U = \Lambda$ eine Diagonalmatrix ist. Dabei sind die Diagonalelemente von Λ die Nullstellen des charakteristischen Polynoms von T, also eines Polynoms mit reellen Koeffizienten, und eine Nullstelle λ der Vielfachheit s tritt genau s-mal als ein Diagonalelement in Λ auf. Diese Nullstellen können im Allgemeinen beliebige komplexe Zahlen vom Betrag 1 sein – allerdings ist mit jeder nicht-reellen Nullstelle auch die konjugiert-komplexe Zahl eine Nullstelle derselben Vielfachheit. Die Spalten u_k der Matrix U bilden eine Orthonormalbasis von $\mathbb{C}^n$ und sind außerdem Eigenvektoren von A, und durch Umnumerierung der Basisvektoren, also Vertauschen der Spalten von U, können wir eine beliebige Anordnung der Diagonalelemente von Λ erreichen. Wir gehen deshalb davon aus, dass entlang der Diagonalen von Λ zuerst die Zahl 1 und dann die Zahl -1, beide entsprechend ihrer Vielfachheit, auftreten, und danach die nicht-reellen Nullstellen folgen, wobei diese in Paaren von λ und der konjugiert-komplexen Zahl $\overline{\lambda}$ auftreten. Beachte, dass es natürlich möglich ist, dass alle Diagonalelemente von Λ reell sind, oder dass die Zahl 1 oder -1 garnicht als Nullstelle auftritt.

Wenn p die Vielfachheit der Nullstelle $\lambda = 1$ des charakteristischen Polynoms von A bezeichnet, dann sind die ersten p Spalten von U gerade eine Basis des Eigenraums von A zum Eigenwert $\lambda = 1$. Dieser Eigenraum hat die Eigenschaft, dass mit u auch $\overline{u}$ zum Raum gehört, und deshalb gibt es nach Aufgabe 5.4.8 eine Orthonormalbasis aus lauter reellen Eigenvektoren. Also können wir o. B. d. A. annehmen, dass die ersten p Spalten von U reelle Vektoren sind. Analog schließt man für die nächsten q Spalten von U, welche ja eine Basis des Eigenraumes zum Eigenwert $\lambda = -1$ sind.

Sei jetzt λ_k das erste nicht-reelle Diagonalelement von A. Dann ist $|\lambda_k| = 1$, also $\lambda_k = \cos\phi_k + i\,\sin\phi_k$, mit einem $\phi_k \in \mathbb{R}$, und $\sin\phi_k \neq 0$. Wegen der gewählten Anordnung der Diagonalelemente von A ist dann das nächste Diagonalelement von A gleich $\lambda_{k+1} = \cos\phi_k - i\,\sin\phi_k$. Wenn $(v_1, \ldots, v_s)$ eine Orthonormalbasis des Eigenraumes zu λ_k bilden, dann zeigt man leicht dass $(\overline{v}_1, \ldots, \overline{v}_s)$ eine Orthonormalbasis des Eigenraumes zu $\overline{\lambda_k} = \lambda_{k+1}$ ist. Daher können wir o. B. d. A. annehmen, dass $u_{k+1} = \overline{u}_k$ ist. Wenn man u_k durch $\tilde{u}_k = (i\sqrt{2})^{-1}(u_k - u_{k-1})$ und u_{k+1} durch $\tilde{u}_{k+1} = (\sqrt{2})^{-1}(u_k + u_{k-1})$ ersetzt, so kann man überprüfen, dass die neuen, genau wie die alten, Spaltenvektoren eine Orthonormalbasis von $\mathbb{C}^n$ bilden; allerdings sind die neuen Spaltenvektoren $\tilde{u}_k$ und $\tilde{u}_{k+1}$ jetzt reell geworden. Aus $A\,u_k = \lambda_k\,u_k$ und $A\,u_{k+1} = \overline{\lambda}_k\,u_{k+1}$ folgt dann

$$A\,\tilde{u}_k \;=\; \cos\phi_k\,\tilde{u}_k + \sin\phi_k\,\tilde{u}_{k+1}\,, \qquad A\,\tilde{u}_{k+1} \;=\; -\sin\phi_k\,\tilde{u}_k + \cos\phi_k\,\tilde{u}_{k+1}\,.$$

Also tritt in der Darstellungsmatrix von T bezüglich $(u_1, \ldots, u_{k-1}, \tilde{u}_k, \tilde{u}_{k+1}, u_{k+2}, \ldots, u_n)$ nach den 1-en und (-1)-en entlang der Diagonalen gerade ein Kästchen der Form A_k auf. Genauso kann man auch für die übrigen Paare von konjugiert-komplexen Diagonalelementen schließen. Dies ist, bis auf die Numerierung der Blöcke A_k, die Behauptung. □

Aufgabe 9.2.4 *Benütze den vorausgegangenen Satz, um für den Fall $n = 3$ zu zeigen, dass zu jedem längentreuen Isomorphismus eines dreidimensionalen euklidischen Raumes V eine Orthonormalbasis von V existiert, bezüglich der die Darstellungsmatrix A von T eine der folgenden beiden Formen hat:*

$$A \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \qquad \text{oder} \qquad A \;=\; \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix}.$$

Im ersten Fall ist T also eine Drehung, im zweiten eine Drehspiegelung.

Literaturverzeichnis

[1] **G. Fischer**, *Lineare Algebra. Eine Einführung für Studienanfänger*, Vieweg Studium: Grundkurs Mathematik. Wiesbaden: Vieweg. x, 384 S., 2000.

[2] **W. Gawronski**, *Grundlagen der Linearen Algebra. Studienbuch für Studierende der Mathematik, Informatik, Physik und anderer Naturwissenschaften ab dem 1. Semester*, Studien-Text. Mathematik. Wiesbaden: Aula-Verlag. 413 S. DM 44.00 , 1996. 45

[3] **H. Grauert und H.-C. Grunau**, *Lineare Algebra und analytische Geometrie*, R. Oldenbourg Verlag, Munich, 1999.

[4] **W. Greub**, *Linear algebra*, Graduate Texts in Mathematics, Bd. 23. New York - Heidelberg -Berlin: Springer-Verlag. XVII, 451 p., 5 Figs. DM 92.00; $ 42.90 , 4. Aufl., 1981.

[5] **J. Heinhold und B. Riedmueller**, *Lineare Algebra und Analytische Geometrie. Teil 2*, München: Carl Hanser Verlag. VIII, 319 S. mit 47 Fig., 125 Beisp. und 137 Aufg. DM 44.00 , 1973.

[6] **J. Heinhold und B. Riedmüller**, *Lineare Algebra und analytische Geometrie. Teil 1*, München: Carl Hanser Verlag. VII, 215 S. mit 32 Figuren , 1971.

[7] **K. Jänich**, *Lineare Algebra. 9. Aufl.*, Springer-Lehrbuch. Berlin: Springer. xii, 271 S. EUR 19.95 (D); sFr. 31.00 , 2002.

[8] **H.-J. Kowalsky und G. O. Michler**, *Lineare Algebra. 10., völlig neu bearb. Aufl.*, De Gruyter Lehrbuch. Berlin: de Gruyter. xiv, 399 p. DM 44.00 /pbk; DM 98.00 /hbk , 1995.

[9] **R. Lingenberg**, *Lineare Algebra. Erster Teil einer Vorlesung. (B.I.-Hochschulskripten. 828/828a)*, Mannheim-Zürich: Bibliographisches Institut, XII, 161 S. , 1969.

[10] **F. Lorenz**, *Lineare Algebra.* I, Bibliographisches Institut, Mannheim, 2. Aufl., 1988.

[11] ——, *Lineare Algebra. II*, Bibliographisches Institut, Mannheim, 2. Aufl., 1989.

[12] **U. Stammbach**, *Lineare Algebra. 4. durchges. Aufl.*, Teubner Studienskripten: Mathematik. Stuttgart: B. G. Teubner. 259 S., 1994.

Index