1 Lineare Abbildungen

Definition 1 *Sei K ein Körper und V und W K-Vektoräume. Eine Abbildung $f : V \to W$ heisst linear (oder Homomoprhismus), wenn gilt:*

- $f(v_1 + v_2) = f(v_1) + f(v_2) \quad \forall v_1, v_2 \in V$
- $f(\lambda v) = \lambda f(v) \quad \forall \lambda \in K, v \in V$

Mit $\mathrm{Hom}(V, W)$ *bezeichnen wir die Menge aller linearen Abbildungen $V \to W$.*

Beispiele 1 (a) Sei $A = (a_{i,j}) \in M(m \times n, K)$. Setze

$$\ell_A : K^n \to K^m, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \cdot x.$$

ℓ_A ist eine lineare Abbildung. Wir werden später zeigen, dass jede lineare Abbildung $K^n \to K^m$ von dieser Form ist, d.h. für jedes $f \in \mathrm{Hom}(K^n, K^m)$ gibt es genau ein $A \in M(m \times n, K)$ mit $f = \ell_A$.

(b) Für $a, b \in \mathbb{R}$ mit $a < b$ sei $C^0([a, b])$ die Menge aller stetigen Funktionen $[a, b] \to \mathbb{R}$ und $C^1([a, b])$ die Menge aller stetig diffenzierbaren Funktionen $[a, b] \to \mathbb{R}$. Die Teilmengen $C^0([a, b]), C^1([a, b])$ sind Untervektorräume von $\mathrm{Abb}\,([a, b], \mathbb{R})$. Der Differenzialoperator

$$\frac{d}{dx} : C^1([a, b]) \to C^0([a, b]), \quad \varphi \longmapsto \frac{d\varphi}{dx} = \varphi'$$

ist linear, denn es gilt ja $(\varphi_1 + \varphi_2)' = \varphi_1' + \varphi_2'$ und $(\lambda\varphi)' = \lambda\varphi')$ für alle $\varphi, \varphi_1, \varphi_2 \in C^1([a, b])$ und $\lambda \in \mathbb{R}$.

(c) Das Integral $\int_a^b : C^0([a, b]) \to \mathbb{R}, \varphi \longmapsto \int_a^b \varphi(x)dx$ ist linear.

(d) Keine lineare Abbildungen sind z.B.:

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 2x + y + 1 \\ 3x + 4y + 5 \end{pmatrix} g : \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \sin(x) \\ e^y \end{pmatrix}$$

Denn nach Beispiel 1 ist jede lineare Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$ von der Form

$$\mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

für geeignete $a, b, c, d \in \mathbb{R}$.

Die Darstellungsmatrix einer linearen Abbildung Seien V und W endlichdimensionale K-Vektorräume und sei $f : V \to W$ linear. Wir fixieren Basen $\underline{v} = (v_1 \dots v_n)$ und $\underline{w} = (w_1 \dots w_n)$ von V und W.

Für jedes $j \in \{1, \dots, n\}$ gibt es eindeutig bestimmte $a_{1j}, a_{2j}, \dots, a_{mj} \in K$ so dass

$$f(v_j) = a_{1j}w_1 + \dots + a_{mj}w_n$$

Die Matrix $M_{\underline{w}}^{\underline{v}}(f) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ heisst Darstellungsmatrix von f bezüglich $\underline{v}$ und $\underline{w}$. Für $j = 1, \dots, n$ gilt also: Die j-te Spalte von $M_{\underline{w}}^{\underline{v}}(f)$ besteht aus den Koordinaten von $f(v_j)$ bzgl. der Basis $(w_1 \dots w_n)$.

Beispiel 1 Sei $\mathbb{R}[x]_{\leq n}$ die Menge der Polynome mit Koeffizienten in $\mathbb{R}$ vom Grad $\leq n$, d.h.

$$\mathbb{R}[x]_{\leq n} = \{f(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0 \mid a_0, \dots, a_n \in \mathbb{R}\}.$$

$\mathbb{R}[x]_{\leq n}$ ist ein $n+1$–dimensionaler Unterraum von $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$. Z.B. ist das $n+1$–Tupel der Monome $(x^0 = 1, x^1, x^2, \dots, x^n)$ eine Basis von $\mathbb{R}[x]_{\leq n}$. Die Abbildung $D_n = \dfrac{d}{dx} : \mathbb{R}[x]_{\leq n} \to \mathbb{R}[x]_{\leq (n-1)}, P(x) \longmapsto P'(x)$ ist linear.

Wir wollen die Darstellungsmatrix von D_3 bzgl. der Basen $\underline{v}: = (1, x^1, x^2, x^3)$ und $\underline{w}: = (1, x^1, x^2)$ bestimmen. Da

$$D_3(1) = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2, \quad D_3(x) = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2,$$

$$D_3(x^2) = 2x^1 = 0x^0 + 2x^1 + 0x^2, \quad D_3(x^3) = 3x^2 = 0x^0 + 0x^1 + 3x^2$$

folgt

$$M_{\underline{w}}^{\underline{v}}(D_3) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Satz 2 *Seien V und W endlichdimensionale K-Vektorräume mit fest gewählten Basen $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_m)$. Die Abbildung*

$$\mathrm{Hom}(V, W) \to M(m \times n, K),\, f \longmapsto M_{\underline{w}}^{\underline{v}}(f) \tag{1}$$

ist bijektiv.

Beweis. 1. Injektivität: Seien $f, g : V \longmapsto W$ linear mit

$$M_{\underline{w}}^{\underline{v}}(f) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} = M_{\underline{w}}^{\underline{v}}(g).$$

Folglich gilt

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i = g(v_j) \quad \forall j \in \{1, \ldots, n\}$$

Sei $v \in V$. Wir wollen zeigen $f(v) = g(v)$. Da $(v_1, \ldots, v_n)$ eine Basis von V ist, gibt es eindeutig bestimmte $\lambda_1, \ldots, \lambda_n \in K$ mit $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$.

$$\begin{aligned} f(v) = f(\lambda_1 v_1 + \ldots + \lambda_n v_n) &= \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) \\ &= \lambda_1 g(v_1) + \ldots + \lambda_n g(v_n) \\ &= g(\lambda_1 v_1 + \ldots + \lambda_n v_n) = g(v) \end{aligned}$$

$\Rightarrow f = g \Rightarrow$ Injektivität.

2. Surjektivität: Sei $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \in M(m \times n, K)$

Gesucht: $f \in \mathrm{Hom}(V, W)$ mit $M_{\underline{w}}^{\underline{v}}(f) = A$.

Definiere zunächst

$$f(v_j) := a_{1j} w_1 + \ldots + a_{mj} w_m \quad \forall j \in \{1, \ldots, n\}$$

Um $f(v)$ für einen beliebigen Vektor $v \in V$ zu definieren, stellen wir v als Linearkombionation der $v_1, \ldots, v_n$ dar:

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n.$$

Dabei sind $\lambda_1, \ldots, \lambda_n \in K$ eindeutig bestimmt. Wir setzen:

$$f(v) \colon = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n)$$

Man überprüft leicht, dass die so definierte Abbildung $f : V \to W$ linear ist und $M^{\underline{v}}_{\underline{w}}(f) = A$ gilt. $\square$

Beispiel 2 Sei $V = K^n, W = K^m$ und seien $\underline{e} = (e_1, \ldots, e_n), \underline{e'} = (e_1, \ldots, e_m)$ die Standardbasen von K^n und K^m, d.h.

$$\underline{e} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right)$$

Man rechnet leicht nach, dass die beiden Abbildungen

$$\mathrm{Hom}(K^n, K^m) \to M(m \times n, K), f \longmapsto M^{\underline{e'}}_{\underline{e}}(f)$$

$$M(m \times n, K) \to \mathrm{Hom}(K^n, K^m), A \longmapsto \ell_A$$

zueinander invers sind. Insbesondere ist die zweite Abbildung $A \mapsto \ell_A$ bijektiv.

Bemerkung 3 (a) Da die Spalten der Darstellungsmatrix die Koordinaten der Bilder $f(v_1), \ldots, f(v_n)$ bzgl. der gegebenen Basis von W sind, lässt sich die Injektivität der Abbildung (1) wie folgt interpretieren: Eine lineare Abbildung $f : V \to W$ ist eindeutig durch die Bilder der Basisvektoren $f(v_1), \ldots, f(v_n)$ festgelegt. Die Surjektivität von (1) bedeutet: Zu jeder Wahl von n Vektoren $w_1, \ldots, w_n \in W$ gibt es eine lineare Abbildung $f : V \to W$ mit $f(v_1) = w_1, \ldots, f(v_n) = w_n$. Zusammenfassend erhalten wir: *Sei V ein endlich-dimensionaler und W ein beliebiger K-Vektorraum. Sei $(v_1, \ldots, v_n)$ eine Basis von V und seien $w_1, \ldots, w_n$ beliebige Vektoren in W. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, n$.*

(b) Sei K ein Körper und V und W K-Vektoräume. Für $f, g \in \mathrm{Hom}(V, W)$ und $\lambda \in K$ verifiziert man leicht, dass die Abbildungen

$$f + g : V \to W, v \mapsto (f + g)(v) := f(v) + g(v), \tag{2}$$

$$\lambda f : V \to W, v \mapsto (\lambda f)(v) := \lambda f(v) \tag{3}$$

wieder linear sind. Mit dieser Addition und skalaren Multiplikation wird die Menge $\mathrm{Hom}_K(V, W)$ selbst zu einem K-Vektorraum. Sind V und W endlich-dimensional mit Basen $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_m)$ so kann man leicht nachrechnen, dass die Abbildung (1) linear ist.

Der nächste Satz besagt vereinfacht ausgedrückt, dass die Hintereinanderschaltung von linearen Abbildungen dem Produkt der Darstellungsmatrizen entspricht.

Satz 4 *Seien U, V und W endlich-dimensionale K-Vektorräume mit Basen $\underline{u} = (u_1, \ldots, u_n), \underline{v} = (v_1, \ldots, v_m)$ und $\underline{w} = (w_1, \ldots, w_l)$ und seien $f : U \to V$ und $g : V \to W$ lineare Abbildungen. Dann ist $g \circ f : U \to W$ linear und es gilt:*

$$M^{\underline{u}}_{\underline{w}}(g \circ f) = M^{\underline{v}}_{\underline{w}}(g) M^{\underline{u}}_{\underline{v}}(f).$$

Beweis. Sei $M^{\underline{u}}_{\underline{v}}(f) = (b_{jk})$, $M^{\underline{v}}_{\underline{w}}(g) = (a_{ij})$ und $M^{\underline{u}}_{\underline{w}}(g \circ f) = (c_{ik})$, d.h. es gilt:

$$f(u_k) = \sum_{j=1}^{m} b_{jk} v_j, \quad g(v_j) = \sum_{i=1}^{l} a_{ij} w_i, \quad \text{und}$$

$$(g \circ f)(u_k) = g(f(u_k)) = \sum_{i=1}^{l} c_{ik} w_i.$$

Es folgt:

$$\sum_{i=1}^{l} \left(\sum_{j=1}^{m} a_{ij} b_{jk} \right) w_i = \sum_{j=1}^{m} b_{jk} \left(\sum_{i=1}^{l} a_{ij} w_i \right) =$$

$$\sum_{j=1}^{m} b_{jk} g(v_j) = g\left(\sum_{j=1}^{m} b_{jk} v_j \right) = g(f(u_k)) = \sum_{i=1}^{l} c_{ik} w_i$$

und damit $\sum_{j=1}^{m} a_{ij} b_{jk} = c_{ik}$ für alle $i = 1, \ldots, l$ und $k = 1, \ldots, n$. $\square$

Transformationsmatrizen Sei V ein n-dimensionaler K-Vektorraum und $\underline{v} = (v_1, \ldots, v_n)$, $\underline{w} = (w_1, \ldots, w_n)$ zwei Basen von V. Die Darstellungsmatrix der Identität $\mathrm{id}_V : V \to V, v \mapsto v$ bzgl. $\underline{v}$ und $\underline{w}$ bezeichnen wir mit

$$T^{\underline{v}}_{\underline{w}} : = M^{\underline{v}}_{\underline{w}}(\mathrm{id}_V).$$

Sie heisst *Transformationsmatrix* des Basiswechsels. Sei v ein beliebiger Vektor in V. Mit Hilfe von $T^{\underline{v}}_{\underline{w}}$ lassen sich die Koordinaten $(x_1, \ldots, x_n)$ von v bzgl. $(v_1, \ldots, v_n)$ in die Koordinaten $(y_1, \ldots, y_n)$ bzgl. $(w_1, \ldots, w_n)$ umrechnen:

$$T^{\underline{v}}_{\underline{w}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Ist $T^{\underline{v}}_{\underline{w}} = (a_{ij})$, also

$$v_j = \mathrm{id}_V(v_j) = \sum_{i=1}^n a_{ij} w_i$$

so folgt aus $v = x_1 v_1 + \ldots x_n v_n = y_1 w_1 + \ldots y_n w_n$ nämlich

$$\sum_{i=1}^n y_i w_i = v = \sum_{j=1}^n x_j v_j = \sum_{j=1}^n x_j \sum_{i=1}^n a_{ij} w_i = \sum_{i=1}^n (\sum_{j=1}^n a_{ij} x_j) w_i$$

und damit $y_i = \sum_{j=1}^n a_{ij} x_j$.

Lemma 5 *Sei V ein endlich-dimensionaler K-Vektorraum und seien $\underline{u} = (u_1, \ldots, u_n), \underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_n)$ Basen von V. Dann gilt:*

(a) $T^{\underline{u}}_{\underline{w}} = T^{\underline{v}}_{\underline{w}} \cdot T^{\underline{u}}_{\underline{v}}$

(b) Die Transformationsmatrix $T^{\underline{v}}_{\underline{w}}$ ist invertierbar mit Inversem $(T^{\underline{v}}_{\underline{w}})^{-1} = T^{\underline{w}}_{\underline{v}}$.

Beweis. (a) folgt sofort aus Satz 4 für $U = W = V$ und $f = g = \mathrm{id}_V$.

(b) Offenbar gilt $T^{\underline{v}}_{\underline{v}} = E_n = T^{\underline{w}}_{\underline{w}}$. Mit (a) ergibt sich

$$T^{\underline{v}}_{\underline{w}} \cdot T^{\underline{w}}_{\underline{v}} = T^{\underline{w}}_{\underline{w}}, \quad T^{\underline{w}}_{\underline{v}} \cdot T^{\underline{v}}_{\underline{w}} = T^{\underline{v}}_{\underline{v}}.$$

□

1.0.1 Beispiel: Im $\mathbb{R}^2$ seien die Basen $\underline{v} = (\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}), \underline{w} = (\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}3\\1\end{pmatrix})$ und der Vektor

$$v = \begin{pmatrix}1\\-1\end{pmatrix} = 5 \cdot \begin{pmatrix}2\\1\end{pmatrix} - 3 \cdot \begin{pmatrix}3\\2\end{pmatrix}$$

gegeben. Wir wollen die Koordinaten von v bzgl. $\underline{w}$ berechnen. Sei $\underline{e} = (\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix})$ die Standardbasis. Für die Transformationsmatrix $T^{\underline{v}}_{\underline{w}}$ gilt nach Lemma 5:

$$T^{\underline{v}}_{\underline{w}} = T^{\underline{e}}_{\underline{w}} T^{\underline{v}}_{\underline{e}} = (T^{\underline{w}}_{\underline{e}})^{-1} T^{\underline{v}}_{\underline{e}}$$

Ferner gilt $T^{\underline{v}}_{\underline{e}} = \begin{pmatrix}2 & 3\\1 & 2\end{pmatrix}$, denn die erste Spalte von $T^{\underline{v}}_{\underline{e}}$ besteht aus den Koordinaten von $\begin{pmatrix}2\\1\end{pmatrix}$ und die zweite Spalte aus den Koordinaten von $\begin{pmatrix}3\\2\end{pmatrix}$ bzgl. der Standardbasis $\underline{e}$. Ebenso gilt $T^{\underline{w}}_{\underline{e}} = \begin{pmatrix}1 & 3\\0 & 1\end{pmatrix}$.

$$\Rightarrow \quad T^{\underline{v}}_{\underline{w}} = \begin{pmatrix}1 & 3\\0 & 1\end{pmatrix}^{-1} \cdot \begin{pmatrix}2 & 3\\1 & 2\end{pmatrix} = \begin{pmatrix}1 & -3\\0 & 1\end{pmatrix} \cdot \begin{pmatrix}2 & 3\\1 & 2\end{pmatrix} = \begin{pmatrix}-1 & -3\\1 & 2\end{pmatrix}.$$

Wegen $\begin{pmatrix}-1 & -3\\1 & 2\end{pmatrix}\begin{pmatrix}5\\-3\end{pmatrix} = \begin{pmatrix}4\\-1\end{pmatrix}$ gilt also

$$v = 4 \cdot \begin{pmatrix}1\\0\end{pmatrix} - 1 \begin{pmatrix}3\\1\end{pmatrix}$$

Wir wollen jetzt die Frage untersuchen, wie sich die Darstellungsmatrix $M^{\underline{v}}_{\underline{w}}(f)$ einer linearen Abbildung $f : V \to W$ ändert, wenn man von den Basen $\underline{v}$ von V und $\underline{w}$ von W zu zwei neuen Basen $\underline{v'}$ und $\underline{w'}$ übergeht.

Satz 6 *Es gilt:*

$$M^{\underline{v'}}_{\underline{w'}}(f) = T^{\underline{w}}_{\underline{w'}} \cdot M^{\underline{v}}_{\underline{w}}(f) \cdot T^{\underline{v'}}_{\underline{v}}$$

Die Darstellungsmatrizen von f bezüglich zweier Basenpaare unterscheidet sich also durch Links- und Rechtsmultiplikation mit gewissen Transformationsmatrizen.

Beweis. Das folgt wegen $f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$ aus Satz 4. □

Determinanten von Endomorphismen Sei V ein K-Vektorraum. Man nennt eine lineare Abbildung $f : V \to V$ auch einen *Endomorphismus.*

Lemma und Definition 7 *Sei V ein endlich-dimensionalen K-Vektorraum, $f : V \to V$ ein Endomorphismus und $\underline{v}$ eine Basis von V. Der Skalar*

$$\det(f) \colon = \det M_{\underline{v}}^{\underline{v}}(f) \in K$$

hängt nicht von der Wahl von $\underline{v}$ ab. Man nennt $\det(f)$ die Determinante von f.

Beweis. Sei $\underline{w}$ eine weitere Basis von V. Dann ist zu zeigen:

$$\det M_{\underline{v}}^{\underline{v}}(f) = \det M_{\underline{w}}^{\underline{w}}(f).$$

Nach Satz 6 und Lemma 5 gilt:

$$M_{\underline{w}}^{\underline{w}}(f) = T_{\underline{w}}^{\underline{v}} \cdot M_{\underline{v}}^{\underline{v}}(f) \cdot T_{\underline{v}}^{\underline{w}} = (T_{\underline{v}}^{\underline{w}})^{-1} \cdot M_{\underline{v}}^{\underline{v}}(f) T_{\underline{v}}^{\underline{w}}.$$

Es folgt:

$$\begin{aligned} \det M_{\underline{w}}^{\underline{w}}(f) &= \det((T_{\underline{v}}^{\underline{w}})^{-1}) \det M_{\underline{v}}^{\underline{v}}(f) \det T_{\underline{v}}^{\underline{w}} \\ &= (\det(T_{\underline{v}}^{\underline{w}}))^{-1} \det T_{\underline{v}}^{\underline{w}} \det M_{\underline{v}}^{\underline{v}}(f) \\ &= \det M_{\underline{v}}^{\underline{v}}(f). \end{aligned}$$

$\square$

Isomorphismen

Definition 8 *Seien V und W K-Vektorräume.*

(a) Eine bijektive, lineare Abbildung $f : V \to W$ heisst Isomorphismus (von Vektorräumen).

(b) V heisst isomorph zu W (in Zeichen: $V \cong W$), wenn es einen Isomorphsmus $f : V \to W$ gibt.

Bemerkungen 9 (a) Sei $f : V \to W$ ein Isomorphismus von K-Vektorräumen. Man kann leicht sehen, dass die Umkehrabbildung $f^{-1} : W \to V$ wieder linear ist. Also gilt: $V \cong W \Leftrightarrow W \cong V$.

(b) Man kann leicht nachrechnen, dass die Abbildung (1) auch linear ist, d.h. (1) ist ein Isomorphismus.

Lemma 10 *Seien V und W endlich-dimensionale K-Vektorräume. Seien $\underline{v}$ und $\underline{w}$ Basen von V und W. Eine lineare Abbildung $f : V \to W$ ist genau dann ein Isomorphismus, wenn gilt $\dim(V) = \dim(W)$ und wenn die Darstellungsmatrix $M^{\underline{v}}_{\underline{w}}(f)$ invertierbar ist.*

Beweis. Sei $n = \dim(V)$ und $m = \dim(W)$. Ist $f : V \to W$ ein Isomorphismus, so gilt für die Darstellungsmatrizen $A\colon = M^{\underline{v}}_{\underline{w}}(f)$, $B =\colon = M^{\underline{w}}_{\underline{v}}(f^{-1})$:

$$M^{\underline{v}}_{\underline{w}}(f) \cdot M^{\underline{w}}_{\underline{v}}(f^{-1}) = M^{\underline{v}}_{\underline{v}}(f \circ f^{-1}) = M^{\underline{v}}_{\underline{v}}(\mathrm{id}_V) = E_m$$

und ebenso $M^{\underline{w}}_{\underline{v}}(f^{-1}) \cdot M^{\underline{v}}_{\underline{w}}(f) = E_n$. Also ist $m = n$ und $M^{\underline{v}}_{\underline{w}}(f)$ ist invertierbar mit Inverser $M^{\underline{w}}_{\underline{v}}(f^{-1})$.

Ist umgekehrt $m = n$ und $A = M^{\underline{v}}_{\underline{w}}(f)$ invertierbar mit Inversem B, so gibt es aufgrund der Surjektivität von (1) eine lineare Abbildung $g : W \to V$ mit Darstellungsmatrix $M^{\underline{w}}_{\underline{v}}(g) = B$. Wegen

$$M^{\underline{v}}_{\underline{v}}(g \circ f) = M^{\underline{w}}_{\underline{v}}(g) \cdot M^{\underline{v}}_{\underline{w}}(f) = B \cdot A = E_n = M^{\underline{v}}_{\underline{v}}(\mathrm{id}_V)$$

folgt $g \circ f = \mathrm{id}_V$ (da die Abbildung (1) injektiv ist). Analog zeigt man $f \circ g = \mathrm{id}_W$. Also ist f ein Isomorphismus mit Umkehrabbildung g. □

Satz 11 *Seien V und W zwei endlich-dimensionale K-Vektorräume. Dann gilt:*

$$\dim V = \dim W \quad \Longleftrightarrow \quad V \cong W$$

Beweis. "$\Rightarrow$" folgt aus Lemma 10.

"$\Leftarrow$": Sei $n = \dim V = \dim W$. Wir wählen Basen $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_n)$ von V und W. Nach Bemerkung 3 gibt es eine lineare Abbildung $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, n$, d.h. $M^{\underline{v}}_{\underline{v}}(f) = E_n$. Explizit ist f gegeben durch

$$f(\lambda_1 v_1 + \ldots + \lambda_n v_n) = \lambda_1 w_1 + \ldots + \lambda_n w_n$$

Nach Lemma 10 ist f ein Isomorphismus. □

Folgerung 12 *Seien V und W endlich-dimensionale K-Vektorräume. Dann ist $\mathrm{Hom}(V, W)$ ebenfalls endlich-dimensional und es gilt:*

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

3

Dimensionsformel für lineare Abbildungen

Definition 13 *Sei $f : V \to W$ eine lineare Abbildung.*

$$\mathrm{Kern}(f) \colon = f^{-1}(\{0\}) = \{v \in V \mid f(v) = 0\}$$

heisst der Kern von f und

$$\mathrm{Bild}(f) \colon = f(V) = \{w \in W \,\big|\, \exists\ v \in V :\ f(v) = w\}$$

heisst das Bild von f. Ist $\mathrm{Bild}(f)$ *endlich-dimensional, so heisst* $\mathrm{Rang}(f) \colon = \dim \mathrm{Bild}(f)$ *der Rang von f.*

Lemma 14 *Sei $f : V \to W$ eine lineare Abbildung. (a)* $\mathrm{Kern}(f)$ *ist ein Untervektorraum von V. Es gilt:*

$$\mathrm{Kern}(f) = \{0\} \quad \Leftrightarrow \quad f \text{ ist injektiv}$$

(b) $\mathrm{Bild}(f)$ *ist ein Untervektorraum von W.*

Beweis. (a) Für $v, w \in \mathrm{Kern}(f)$ und $\lambda \in K$ gilt

$$f(v + w) = f(v) + f(w) = 0 + 0 = 0, \quad f(\lambda v) = \lambda f(v) = 0$$

also $v + w, \lambda v \in \mathrm{Kern}\, f$.

Ist $\mathrm{Kern}(f) = \{0\}$ und sind $v_1, v_2 \in V$ so gilt

$$f(v_1) = f(v_2) \Rightarrow f(v_1 - v_2) = f(v_1) - f(v_2) = 0 \Rightarrow v_1 - v_2 = 0 \Rightarrow v_1 = v_2.$$

Also ist f injektiv. □

Satz 15 *(Dimensionsformel für lineare Abbildungen) Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Ist V endlich-dimensional, so gilt:*

$$\dim(V) = \dim \mathrm{Kern}(f) + \mathrm{Rang}(f).$$

Beweis. Sei $r \colon = \dim \mathrm{Kern}\, f \leq n \colon = \dim V$, und sei $(v_1, \ldots, v_r)$ eine Basis von $\mathrm{Kern}\, f$, die wir zu einer Basis $(v_1, \ldots, v_n)$ von V ergänzen. Sei $v = \lambda_1 v_1 + \ldots + \lambda_n v_n \in V$. Wegen

$$f(v) = \lambda_1 f(v_1) + \ldots + \lambda_n f(v_n) = \lambda_{r+1} f(v_{r+1}) + \ldots + \lambda_n f(v_n)$$

ist $(f(v_{r+1}),\ldots,f(v_n))$ ein Erzeugendensystem von $\operatorname{Bild} f$. Die Vektoren $(f(v_{r+1}),\ldots,f(v_n))$ sind überdies linear unabhängig, denn aus

$$\lambda_{r+1}f(v_{r+1})+\ldots+\lambda_n f(v_n)=0$$

folgt $f(\lambda_{r+1}v_{r+1}+\ldots+\lambda_n v_n)=0$, also $\lambda_{r+1}v_{r+1}+\ldots+\lambda_n v_n\in\operatorname{Kern} f$ und damit

$$\lambda_1 v_1+\ldots+\lambda_r v_r=\lambda_{r+1}v_{r+1}+\ldots+\lambda_n v_n$$

für gewisse $\lambda_1,\ldots,\lambda_r\in K$. Da $(v_1,\ldots,v_n)$ eine Basis ist, folgt

$$\lambda_1=\ldots=\lambda_r=\lambda_{r+1}=\ldots=\lambda_n=0.$$

Also ist $(f(v_{r+1}),\ldots,f(v_n))$ eine Basis von $\operatorname{Bild} f$ und es folgt

$$\operatorname{Rang} f=n-r=\dim(V)-\dim\operatorname{Kern} f$$

□

Folgerung 16 *Seien V und W endlich-dimensionale K-Vektorräume mit $\dim(V)=\dim(W)$ und sei $f:V\to W$ eine lineare Abbildung. Dann sind die folgenden Bedingungen äquivalent:*

(i) f ist injektiv.

(ii) f ist surjektiv.

(iii) f ist ein Isomorphismus.

Beweis. f ist injektiv $\iff$ $\operatorname{Kern} f=0$ $\iff$ $\dim\operatorname{Kern} f=0$ $\overset{15}{\iff}$ $\operatorname{Rang}(f)=\dim(V)=\dim(W)$ $\iff$ $\operatorname{Bild}(f)=W$ $\iff$
$\iff$ f ist surjektiv. □

Folgerung 17 *Seien $f:U\to V$ und $g:V\to W$ lineare Abbildungen zwischen K-Vektorräumen. Ist V endlich-dimensional, so gilt*

$$\operatorname{Rang}(g\circ f)\leq\min(\operatorname{Rang}(f),\operatorname{Rang}(g)).$$

Beweis. Wir müssen zeigen, dass gilt:

$$\operatorname{Rang}(g\circ f)\leq\operatorname{Rang}(f),\qquad\operatorname{Rang}(g\circ f)\leq\operatorname{Rang}(g).$$

Da $\operatorname{Bild}(g\circ f)\subseteq\operatorname{Bild}(g)$ gilt zunächst $\operatorname{Rang}(g\circ f)\leq\operatorname{Rang}(g)$. Wenden wir die Dimensionsformel auf die Einschränkung $g|_{\operatorname{Bild}(f)}$ von g auf den Unterraum

Bild(f) von V an (also auf die lineare Abbildung Bild(f) $\to W, v \mapsto g(v)$), so erhalten wir wegen Bild($g\,|_{\mathrm{Bild}(f)} = g(f(V)) = \mathrm{Bild}(g \circ f)$ andererseits

$$\begin{aligned}\mathrm{Rang}(g \circ f) = \mathrm{Rang}(g\,|_{\mathrm{Bild}(f)}) &= \dim(\mathrm{Bild}(f)) - \dim(\mathrm{Kern}(g\,|_{\mathrm{Bild}(f)}) \\ &\leq \dim(\mathrm{Bild}(f)) = \mathrm{Rang}(f).\end{aligned}$$

□

Der Rang einer Matrix Sei $A = (a_{ij})$ eine $m \times n$-Matrix mit Einträgen in einem Körper K. Der *Rang von A* ist definiert als der Rang der zugehörigen linearen Abbildung

$$\ell_A : K^n \to K^m, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

d.h. $\mathrm{Rang}(A) := \mathrm{Rang}(\ell_A)$.

Es seien $z_1, \ldots, z_m \in K^n$ die Zeilen und $a_1, \ldots, a_n \in M(m \times 1, K)$ die Spalten von A. Die lineare Hülle $L(z_1, \ldots, z_m)$ von $z_1, \ldots, z_m$ heisst der *Zeilenraum* und $L(a_1, \ldots, a_n)$ heisst der *Spaltenraum* von A. Es gilt:

$$\begin{aligned}\text{Spaltenraum } A &= \{x_1 a_1 + \ldots + x_n a_n \mid x_1, \ldots, x_n \in K\} \\ &= \{A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \bigm| \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n\} \\ &= \mathrm{Bild}(\ell_A)\end{aligned}$$

Identifizieren wir den Vektorraum der Zeilenvektoren $M(1 \times m, K)$ mit $K^m = M(m \times 1, K)$ so gilt entsprechend

$$\text{Zeilenraum } A = \{yA \bigm| y \in K^m\} = \mathrm{Bild}(\ell_{A^t}).$$

Der Zeilen- (bzw. Spalten-)rang von A ist die Dimension des Zeilen- (bzw. Spalten-)raums von A. Es gilt also $\mathrm{Rang}(A) = \mathrm{Spaltenrang}(A)$. Wir werden gleich zeigen (siehe Folgerung 21), dass der Zeilenrang immer gleich dem Spaltenrang (= Rang) ist.

Die folgende Bemerkung ist nützlich für die Bestimmung des Rangs einer Matrix.

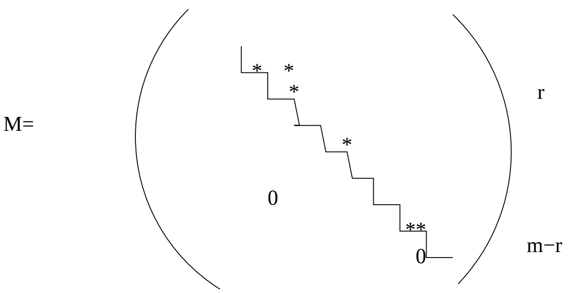
Lemma 18 *Elementare Zeilen- (bzw. Spalten-)umformungen ändern den Zeilen- (bzw. Spalten-)rang nicht.*

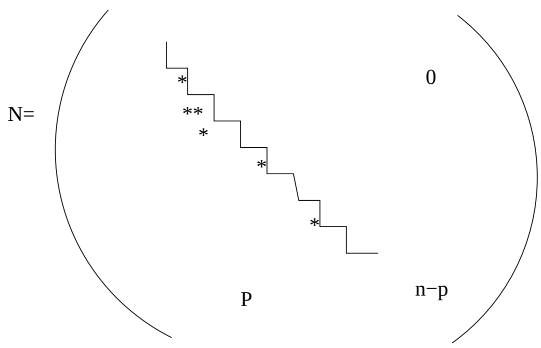
Beweis. Für Zeilenumformung vom Typ 3: Sind $z_1, \ldots, z_m \in K^n$ die Zeilen von $A = (a_{ij}) \in M(m \times n, K)$ und entsteht $\tilde{A}$ aus A durch Addition des λ-fachen der i-ten Zeile zur j-ten Zeile so gilt:

$$L(z_1, \ldots, z_i, \ldots, z_j + \lambda z_i, \ldots, z_m) = L(z_1, \ldots, z_i, \ldots, z_j, \ldots, z_m),$$

d.h. der Zeilenraum – und folglich auch der Zeilenrang – ändert sich nicht. □

Wird nun die $m \times n$-Matrix A durch elementare Zeilenumformungen in eine Matrix M in Zeilenstufenform umgewandelt (und in analoger Weise mit Spaltenumformungen in eine Matrix N in *Spaltenstufenform*),





dann zeigt die folgende Bemerkung, dass der Zeilenrang von $A = r$ ist (und entsprechend der Spaltenrang von $A = p$ ist).

Lemma 19 *Ist $A = (a_{ij})$ eine $m \times n$-Matrix in Zeilenstufenform, so sind die von Null verschiedenen Zeilen von A linear unabhängig.*

Eine entsprechende Aussage gilt für Matrizen in Spaltenstufenform.

Beweis. Sind $z_1 = (0, \ldots, 0, a_{1j_1}, \ldots, a_{1n}), z_2 = (0, \ldots, 0, a_{2j_2}, \ldots, a_{2n}), \ldots, z_r = (0, \ldots, 0, a_{rj_r}, \ldots, a_{rn})$ die von Null verschiedenen Zeilen mit $a_{1j_1}, \ldots, a_{rj_r} \neq 0$ und $j_1 < \ldots < j_r$ so folgt aus

$$\lambda_1 z_1 + \ldots + \lambda_r z_r = 0$$

durch Betrachtung der j_1-ten Komponente zunächst $\lambda_1 a_{1j_1} = 0$, also $\lambda_1 = 0$; dann $\lambda_2 a_{2j_2} = 0$, also $\lambda_2 = 0$ etc. $\square$

Für eine Matrix $A \in M(m \times n, K)$ sei $\mathrm{Kern}(A) :\,= \mathrm{Kern}(\ell_A)$. Kern$(A)$ ist also die Lösungsmenge des homogenen linearen Gleichungssystems

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4}$$

Lemma 20 Zeilenrang$(A) + \dim(\mathrm{Kern}\, A) = n$.

Beweis. Da sich die linke Seite der Gleichung unter elementaren Zeilenumformungen nicht ändert, können wir o.B.d.A. annehmen, dass A in Zeilenstufenform ist.

Ist r die Anzahl der von 0 verschiedenen Zeilen von A, so gilt: Zeilenrang$(A) = r$ und $\dim(\mathrm{Kern}\, A) = n - r$ (letzteres, da das lineare Gleichungssystem (4) $n - r$ freie Variable besitzt). $\square$

Folgerung 21 *"Zeilenrang = Spaltenrang".*

Beweis. Nach Satz 15 und Folgerung 21 gilt:

$$\begin{aligned} \mathrm{Zeilenrang}(A) &= n - \dim(\mathrm{Kern}\, A) = \dim(K^n) - \dim(\mathrm{Kern}(\ell_A)) \\ &= \mathrm{Rang}(\ell_A) = \mathrm{Spaltenrang}(A) \end{aligned}$$

$\square$

Die folgende Satz folgt sofort aus der entsprechenden Aussage (Folgerung 16) für lineare Abbildungen:

Satz 22 *Für eine $n \times n$-Matrix A sind äquivalent:*

(i) A ist invertierbar.

(ii) Rang$(A) = n$.

(iii) Kern$(A) = \{0\}$.

Bemerkung 23 Wir können ein lineares Gleichunssystem

$$(*)\qquad \begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

auch als Matrizengleichung $Ax = b$ mit $x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ schreiben. Die Matrix

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M(m \times n, K)$$

heisst Koeffizientenmatrix des Gleichungssystems. Ist $m = n$ und A invertierbar, so hat $(*)$ die eindeutig bestimmte Lösung $x = A^{-1}b$. Im allgemeinen gilt das folgende Lösbarkeitskriterium:

$(*)$ ist genau dann lösbar, wenn gilt: $\operatorname{Rang}(A, b) = \operatorname{Rang} A$

Dabei entsteht (A, b) aus der Matrix A, indem man noch die Spalte b anhängt:

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

Die Matrix (A, b) heisst *erweiterte Koeffizientenmatrix.*

2 Euklidische und unitäre Vektorräume

In diesem Abschnitt betrachten wir reelle und komplexe Vektorräume mit Skalarprodukt. Dieses erlaubt uns die Länge eines Vektors zu definieren und (im Fall eines reellen Vektorraums) den Winkel zwischen zwei Vektoren zu erklären.

Beispiel 3 Wir beginnen mit dem aus der Schulmathematik bekannten Beispiel des (Standard-)Skalarprodukts in der Anschauungsebene $V = \mathbb{R}^2$. Sei $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Nach dem Satz des Pythagoras ist die Länge $\|v\|$ von v gegeben durch

$$\|v\| = \sqrt{x_1^2 + x_2^2}$$

Sei $w \in \mathbb{R}^2$ ein weiterer Vektor. Die Länge von $v + w$ lässt sich durch die Längen von v und w und den Winkel $\phi \in [0, \pi]$ zwischen v und w ausdrücken:

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\cos(\phi)\|v\|\|w\|. \tag{5}$$

Der Term $< v, w >: = \cos(\phi)\|v\|\|w\|$ wird als das *Skalarprodukt* von v und w bezeichnet. Ist $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ so gilt

$$\|v + w\|^2 = (x_1 + y_1)^2 + (x_2 + y_2)^2 = (x_1^2 + x_2^2) + (y_1^2 + y_2^2) + 2(x_1 y_1 + x_2 y_2)$$

und damit

$$< \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} >= x_1 y_1 + x_2 y_2. \tag{6}$$

Es gilt also

$$\|v\| = \sqrt{< v, v >}, \qquad \text{und} \qquad \cos(\phi) = \frac{< v, w >}{\|v\|\|w\|}$$

d.h. sowohl die Länge eines Vektors v als auch der Winkel zwischen zwei Vektoren $v, w \neq 0$ im $\mathbb{R}^2$ lassen sich in Termen des Skalarprodukts beschreiben.

Das (Standard-)Skalarprodukt im $\mathbb{R}^n$ Analog zu (6) ist das (Standard-)Skalarprodukt zweier Vektoren im $\mathbb{R}^n$ (für $n \geq 3$) erklärt.

Definition 24 *Seien* $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. *Das (Standard-)Skalarprodukt ist definiert durch*

$$< v, w >: = v^t \cdot w = x_1 y_1 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}$$

Man rechnet leicht nach, dass dieses Produkt die folgenden Eigenschaften besitzt.

Lemma 25 *Für alle* $v, v_1, v_2, w, w_1, w_2 \in \mathbb{R}^n$ *und* $\lambda \in \mathbb{R}$ *gilt:*

(a) $< v_1 + v_2, w >=< v_1, w > + < v_1, w >, \quad < \lambda v, w >= \lambda < v, w >,$
$< v, w_1 + w_2 >=< v, w_1 > + < v, w_2 >, \quad < v, \lambda w >= \lambda < v, w >.$
(Bilinearität)

(b) $< v, w >=< w, v >$ *(Symmetrie).*

(c) $< v, v >\geq 0$ *und* $< v, v >= 0$ *genau dann wenn* $v = 0$ *ist (man nennt diese Eigenschaft positive Definitheit).*

Wegen Eigenschaft (a) können wir die nicht-negative Wurzel aus $< v, v >$ ziehen. Die Zahl

$$\|v\| :\ = \sqrt{< v, v >}$$

heißt die *Länge* oder *(euklidische) Norm* von $v \in \mathbb{R}^n$.

Seien v, w zwei, vom Nullvektor verschiedene Vektoren im $\mathbb{R}^n$. Wir wollen jetzt den *Winkel* $\phi \in [0, \pi$ zwischen v und w erklären. Motiviert durch das Beispiel des $\mathbb{R}^2$ definieren wir ϕ als die eindeutig bestimmte Zahl in $[0, \pi]$ mit

$$\cos(\phi) = \frac{< v, w >}{\|v\|\|w\|}. \tag{7}$$

Beachten Sie, dass der Cosinus auf dem Intervall $[0, \pi]$ streng monoton fallend (und stetig) ist und folglich $\cos|_{[0,\pi]}: [0, \pi] \to [-1, 1]$ nach dem Zwischenwertsatz bijektiv ist. Das die rechte Seite von (7) zwischen -1 und 1 liegt ergibt sich aus dem folgenden Lemma:

Lemma 26 *(Cauchy-Schwarzsche Ungleichung) Für* $v, w \in \mathbb{R}^n$ *gilt:*

$$|< v, w >| \ \leq \ \|v\|\|w\|. \tag{8}$$

Beweis. Für $w = 0$ sind beide Seiten $= 0$.

Sei also $w \neq 0$. Nach Lemma 25 gilt für alle $\lambda \in \mathbb{R}$:

$$\begin{array}{rlll} 0 & \leq & < v - \lambda w, v - \lambda w > & \text{positive Definitheit} \\ & = & < v, v - \lambda w > -\lambda < w, v - \lambda w > & \text{Bilinearität} \\ & = & < v, v > -\lambda < v, w > -\lambda < w, v > -\lambda^2 < w, w > & \text{Bilinearität} \\ & = & < v, v > -2\lambda < v, w > -\lambda^2 < w, w > & \text{Symmetrie.} \end{array}$$

Wähle speziell $\lambda = \frac{<v,w>}{<w,w>}$ (man beachte, dass der Nenner $< w, w >> 0$, da $w \neq 0$) Wir erhalten:

$$0 \ \leq < v, v > -2(\frac{< v, w >}{< w, w >}) < v, w > +(\frac{< v, w >}{< w, w >})^2 < w, w >=$$

$$=< v, v > -\frac{< v, w >^2}{< w, w >}.$$

Multipliziert man diese Ungleichung mit $< w, w >$ so erhält man:

$$< v, w >^2 \leq < v, v >< w, w > .$$

und damit (8) durch ziehen der Wurzeln. $\square$

Skalarprodukt auf beliebigen reellen Vektoräumen Wie im letzten Abschnitt erläutert wurde, lässt sich mit Hilfe des Standardskalarprodukt Längen- und Winkelmessung im $\mathbb{R}^n$ durchführen. Um in beliebigen reellen Vektorräumen auch Begriffe wie "die Länge eines Vektors" oder den "Winkel zwischen zwei Vektoren" zu erklären, führen wir jetzt Skalarprodukte axiomatisch ein (ohne Bezugnahme auf die Wahl eines Koordinatensystems).

Definition 27 *Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung*

$$< \ , \ >: V \times V \longrightarrow \mathbb{R}, (v, w) \mapsto < v, w >$$

heisst Skalarprodukt auf V, wenn $< \ , \ >$ bilinear, symmetrisch und positiv definit ist, d.h. wenn gilt:

(a) $< v_1 + v_2, w >=< v_1, w > + < v_1, w >, \quad < \lambda v, w >= \lambda < v, w >,$
$< v, w_1 + w_2 >=< v, w_1 > + < v, w_2 >, \quad < v, \lambda w >= \lambda < v, w > .$

(b) $< v, w >=< w, v >$

(c) $< v, v > \geq 0 \quad , \quad < v, v >= 0 \Leftrightarrow v = 0$

für alle $v, v_1, v_2, w, w_1, w_2 \in V$ und $\lambda \in \mathbb{R}$.

Ein $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt auf V heisst euklidischer Vektorraum.

Beispiele 2 (a) Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und $(v_1, \ldots, v_n)$ eine Basis von V. Für $v = x_1v_1 + \ldots + x_nv_n, w = y_1v_1 + \ldots + y_nv_n \in V$ setzten wir

$$< v, w > : = \sum_{i=1}^{n} x_i y_i.$$

Dann ist $< \ , \ >$ ein Skalarprodukt auf V und es gilt

$$< v_i, v_j >= \begin{cases} 1 & \text{wenn } i = j \text{ ist;} \\ 0 & \text{wenn } i \neq j \text{ ist.} \end{cases}$$

(b) Für $a, b \in \mathbb{R}$ mit $a < b$ sei $C([a, b])$ der reelle Vektorraum aller stetigen Funktionen $f : [a, b] \to \mathbb{R}$. Durch

$$< f, g > : = \int_a^b f(t)g(t)\, dt.$$

wird ein Skalarprodukt auf $C([a, b])$ definiert.

Definition 28 *Sei V ein euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$. Die Abbildung*

$$\| \, . \, \| : V \longrightarrow \mathbb{R}, v \mapsto \|v\| : = \sqrt{< v, v >}$$

heisst die (durch $< \ , \ >$ induzierte) Norm.

Lemma 29 *Sei V ein euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$ und zugehöriger Norm $\| \, . \, \|$. Dann gilt für alle $v, w \in V$ und $\lambda \in \mathbb{R}$:*

(a) (Cauchy-Schwarzsche Ungleichung) $\quad | < v, w > | \leq \|v\| \|w\|.$

(b) $\|v\| \geq 0, \qquad \|v\| = 0 \quad \Leftrightarrow \quad v = 0.$

(b) $\|\lambda v\| = |\lambda| \, \|v\|.$

(c) (Dreiecksungleichung) $\|v + w\| \leq \|v\| + \|w\|.$

Beweis. (a) Der Beweis von Lemma 26 lässt sich wortwörtlich auf den Fall eines beliebigen Skalarprodukts übertragen.

(b), (c) sind trivial.

(d) $\qquad \|v + w\|^2 =< v + w, v + w >=$

$$=< v, v > + < v, w > + < w, v > + < w, w > \leq \|v\|^2 + 2 < v, w > + \|w\|^2 \leq$$

$$\leq \|v\|^2 + 2| < v, w > | + \|w\|^2 \overset{(a)}{\leq} \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

□

Teil (a) des folgenden einfachen Lemma zeigt, dass man das Skalarprodukt aus der Norm rekonstruieren kann.

Lemma 30 *Sei V ein euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$ und induzierter Norm $\| \, . \, \|$. Für $v, w \in V$ gilt:*

(a) $< v, w >= \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2).$

(b) (Parallelogrammgleichung) $\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2).$

Bemerkungen 31 (a) Sei V ein reeller Vektorraum. Eine Abbildung $\| \, . \, \| : V \to \mathbb{R}, v \mapsto \|v\|$ mit den Eigenschaften (b), (c) und (d) aus Lemma 29 heisst eine *Norm auf V*. Ein Skalarprodukt $< \ , \ >$ definiert also eine Norm.

Es gibt aber Normen, die nicht von Skalarprodukten herrühren (Beispiel: $\|v\|_1 \colon = |x_1| + \ldots + |x_n|$, $v = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$; vgl. Übung ..., Blatt ...).

(b) Sei V ein $\mathbb{R}$-Vektorraum und $\| \,.\, \|$ eine Norm auf V. Für $v, w \in V$ setzen wir $d(v, w) \colon = \|v - w\|$. Man nennt die Abbildung $(v, w) \mapsto d(v, w)$ den (aus $\| \,.\, \|$ erhaltenen) *Abstand* (oder *Metrik*). Für $u, v, w \in V$ gilt:

$$\begin{array}{lll} (M1) & d(v,w) \geq 0, & d(v,w) = 0 \Leftrightarrow v = w \\ (M2) & d(v,w) = d(w,v) & \\ (M3) & d(v,w) \leq d(v,u) + d(u,w) & \text{(Dreiecksungleichung)} \end{array}$$

Eine Menge M zusammen mit einer Abbildung $d : M \times M \to \mathbb{R}$ mit den Eigenschaften (1)–(3) nennt man einen *metrischen Raum* (die Abbildung d heisst eine *Metrik auf M*). Man beachtet, dass nicht jede Metrik auf einem $\mathbb{R}$-Vektorraum aus einer Norm entsteht (ein ganz einfaches Beispiel ist die Metrik $d(v, w) \colon = \left\{ \begin{array}{ll} 0 & \text{wenn } v = w \text{ ist;} \\ 1 & \text{wenn } v \neq w \text{ ist.} \end{array} \right)$.

Definition 32 *Sei V ein euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$.*
(a) Seien $v, w \in V - \{0\}$. Nach Lemma 29 (a) gilt $-1 \leq \frac{<v,w>}{\|v\|\|w\|} \leq 1$. Es gibt daher ein eindeutig bestimmtes $\phi \in [0, \pi]$ mit $\cos(\phi) = \frac{<v,w>}{\|v\|\|w\|}$. Man nennt ϕ den Winkel zwischen v und w.

(b) Zwei Vektoren v, w heissen orthogonal oder senkrecht aufeinander stehend (Notation: $v \perp w$) falls gilt $< v, w >= 0$ (wenn $v \neq 0 \neq w$ ist also $v \perp w \Leftrightarrow \phi = \frac{\pi}{2}$).

(c) Sei U ein Untervektorraum von V. Die Menge $U^\perp \colon = \{v \in V \mid < v, u >= 0 \, \forall \, u \in U\}$ heisst das orthogonale Komplement von U.

Bemerkungen 33 (a) Für $v, w \in V$ gilt nach Lemma 30 (a):

$$v \perp w \quad \Longleftrightarrow \quad \|v + w\|^2 = \|v\|^2 + \|w\|^2 \qquad \text{(Satz des Pythagoras)}$$

(b) Sei $U \subseteq V$ ein Untervektorraum. $U^\perp$ ist ebenfalls ein Untervektorraum. Ist V endlich-dimensional so gilt $\dim(U) + \dim(U^\perp) = \dim(V)$ und $(U^\perp)^\perp = U$ (den Beweis werden wir später nachholen).

Definition 34 *Sei V, $< \ , \ >$ wie oben. Ein n-Tupel $(v_1, \ldots, v_n)$ von Vektoren in V heisst Orthonormalsystem, wenn für alle $i, j \in \{1, \ldots, n\}$ gilt:*

$$< v_i, v_j >= \left\{ \begin{array}{ll} 1 & \textit{wenn } i = j \textit{ ist;} \\ 0 & \textit{wenn } i \neq j \textit{ ist.} \end{array} \right.$$

d.h. die Vektoren $v_1, \dots, v_n$ stehen paarweise senkrecht aufeinander und haben alle die Länge 1 *(wir nennen einen solchen Vektor auch normiert).*

$(v_1, \dots, v_n)$ heisst Orthonormalbasis, wenn $(v_1, \dots, v_n)$ ein Orthonormalsystem und eine Basis von V ist.

Beispiele 3 (a) Die Standardbasis $(e_1, \dots, e_n)$ ist eine Orthonormalbasis im $\mathbb{R}^n$ (bzgl. des Standardskalarprodukts).

(b) Sei $V = C([0, 2\pi])$ mit dem Skalarprodukt

$$< f, g > : \; = \frac{1}{2\pi} \int_0^{2\pi} f(t)g(t)\,dt.$$

(vgl. Beispiel 2 (b)). Für $k, l \in \mathbb{N}$ gelten:

$$(9) \qquad \int_0^{2\pi} \cos(kt)\sin(lt)\,dt = 0$$

$$(10) \qquad \int_0^{2\pi} \sin(kt)\sin(lt)\,dt = \int_0^{2\pi} \cos(kt)\cos(lt)\,dt = 0 \quad \text{falls } k \neq l.$$

$$(11) \qquad \int_0^{2\pi} \sin(kt)^2\,dt = \int_0^{2\pi} \cos(kt)^2\,dt = \pi$$

Also bilden je n Funktionen aus der Menge

$$\{1\} \cup \{\sin(kt) \mid k \in \mathbb{N}\} \cup \{\sin(lt) \mid l \in \mathbb{N}\}$$

ein Orthonormalsystem (1 bezeichnet hier die konstante Funktion $[0, 2\pi] \to \mathbb{R}, t \mapsto 1$).

Satz 35 *Sei V ein euklidischer Vektorraum mit Skalarprodukt $< \; , \; >$ und sei $(v_1, \dots, v_n)$ ein n-Tupel von Vektoren in V.*

(a) Ist $(v_1, \dots, v_n)$ ein Orthonormalsystem, so ist $(v_1, \dots, v_n)$ linear unabhängig.

(b) Sei $(v_1, \dots, v_n)$ ein Orthonormalsystem und $U : \; = L(v_1, \dots, v_n)$. Dann gilt für jedes $v \in V$:

$$v - \sum_{i=1}^{n} < v, v_i > v_i \in U^{\perp}$$

(c) Ist $(v_1, \dots, v_n)$ eine Orthonormalbasis, so gilt $v = \sum_{i=1}^{n} < v, v_i > v_i$ für alle $v \in V$ (d.h. $< v, v_1 >, \dots, < v, v_n > \in \mathbb{R}$ sind die Koordinaten von v bzgl. $(v_1, \dots, v_n)$).

Beweis. (a) Aus $\lambda_1 v_1 + \ldots + \lambda_1 v_n = 0$ folgt:

$$0 =< v_i, \lambda_1 v_1 + \ldots + \lambda_n v_n, v_i >= \sum_{j=1}^{n} \lambda_j < v_i, v_j >= \lambda_i$$

für alle $i \in \{1, \ldots, n\}$.

(b) Für $j \in \{1, \ldots, n\}$ gilt:

$$< v - \sum_{i=1}^{n} < v, v_i > v_i, v_j >=< v, v_j > - \sum_{i=1}^{n} < v, v_i >< v_i, v_j >$$
$$=< v, v_j > - < v, v_j >= 0$$

d.h. $v - \sum_{i=1}^{n} < v, v_i > v_i$ steht senkrecht auf jedem der Vektoren $v_1, \ldots, v_n$. Folglich steht $v - \sum_{i=1}^{n} < v, v_i > v_i$ senkrecht auf jeder Linearkombination von $(v_1, \ldots, v_n)$.

(c) Sei jetzt $(v_1, \ldots, v_n)$ eine Orthonormalbasis und $v \in V$. Nach Teil (b) liegt $v - \sum_{i=1}^{n} < v, v_i > v_i$ im orthogonalen Komplement von $L(v_1, \ldots, v_n)$, d.h. der Vektor $v - \sum_{i=1}^{n} < v, v_i > v_i$ steht senkrecht auf jedem $w \in V$. Also auch auf sich selbst:

$$< v - \sum_{i=1}^{n} < v, v_i > v_i, v - \sum_{i=1}^{n} < v, v_i > v_i >= 0$$

und daher (aufgrund der positiven Definitheit des Skalarprodukts)
$v = \sum_{i=1}^{n} < v, v_i > v_i$. □

Im nächsten Satz lernen wir ein Verfahren kennen, dass uns erlaubt aus einer gegebenen Basis von V eine Orthonormalbasis zu gewinnen. Insbesondere folgt, dass jeder endlich-dimensionale euklidische Vektorraum eine Orthonormalbasis besitzt.

Satz 36 *(Gram-Schmidtsches Orthonormalisierungsverfahren) Sei V ein euklidischer Vektorraum mit Skalarprodukt $< \; , \; >$ und sei $(w_1, \ldots, w_n)$ ein linear unabhängiges n-Tupel von Vektoren in V. Dann gibt es ein Orthonormalsystem $(v_1, \ldots, v_n)$ mit*

$$L(v_1, \ldots, v_i) \quad = \quad L(w_1, \ldots, w_i)$$

für $i = 1, \ldots, n$. Insbesondere gilt: Ist $(w_1, \ldots, w_n)$ eine Basis, so ist $(v_1, \ldots, v_n)$ eine Orthonormalbasis.

Beweis. Wir beweisen die Behauptung durch Induktion über n.

Im Fall $n = 1$ setzen wir $v_1 = \frac{1}{\|w_1\|} w_1$. Wegen $\|v_1\| = \frac{1}{\|w_1\|}\|w_1\| = 1$ ist v_1 normiert und es gilt $L(v_1) = L(w_1)$.

Induktionsschluss $n - 1 \to n$ (für $n \geq 2$): Nach der Induktionsvoraussetzung gibt es ein Orthonormalsystem $(v_1, \ldots, v_{n-1})$ mit $L(v_1, \ldots, v_i) = L(w_1, \ldots, w_i)$ für $i = 1, \ldots, n - 1$. Setze

$$\widetilde{v}_n \colon = w_n - \sum_{i=1}^{n-1} < w_n, v_i > v_i.$$

Nach Satz 35 (b) und der Induktionsvoraussetzung gilt:

$$\widetilde{v}_n \in L(v_1, \ldots, v_{n-1})^{\perp} = L(w_1, \ldots, w_{n-1})^{\perp}$$

Es ist $\widetilde{v}_n \neq 0$, denn sonst würde $w_n = \sum_{i=1}^{n-1} < w_n, v_i > v_i \in L(v_1, \ldots, v_{n-1}) = L(w_1, \ldots, w_{n-1})$ gelten. Das steht im Widerspruch zur linearen Unabhängigkeit von $(w_1, \ldots, w_n)$.

Es folgt $\|\widetilde{v}_n\| \neq 0$. Wir setzen $v_n = \frac{1}{\|\widetilde{v}_n\|}\widetilde{v}_n$, so dass $\|v_n\| = 1$ ist. Wegen $\widetilde{v}_n \perp v_i$ gilt auch $v_n \perp v_i$ für $i = 1, \ldots, n - 1$, d.h. $(v_1, \ldots, v_n)$ ist ein Orthonormalsystem. Da sich $\widetilde{v}_n$ und w_n nur um eine Linearkombination von $(v_1, \ldots, v_{n-1})$ unterscheiden erhalten wir schliesslich

$$\begin{aligned} L(v_1, \ldots, v_{n-1}, v_n) &= L(v_1, \ldots, v_{n-1}, \widetilde{v}_n) = L(v_1, \ldots, v_{n-1}, w_n) \\ &= L(w_1, \ldots, w_{n-1}, w_n). \end{aligned}$$

□

Bemerkung 37 Um eine Orthonormalbasis aus einer vorgegebenen Basis $(w_1, \ldots, w_n)$ zu konstruieren kann man also den folgenden Algorithmus benutzen. Man definiert rekursiv die beiden endlichen Folgen $\widetilde{v}_1, \ldots, \widetilde{v}_n$ und $v_1, \ldots, v_n$ durch

$$\widetilde{v}_1 \colon = w_1, \quad \widetilde{v}_i \colon = w_i - \sum_{i=1}^{n-1} < w_n, v_i > v_i \qquad \text{für } i = 1, \ldots, n-1 \text{ und}$$

$$v_i \colon = \frac{1}{\|\widetilde{v}_i\|}\widetilde{v}_i \qquad \text{für } i = 1, \ldots, n.$$

Der obige Beweis zeigt, dass $\|\widetilde{v}_i\| \neq 0$ für alle $i \in \{1, \ldots, n\}$. $(v_1, \ldots, v_n)$ ist dann die gesuchte Orthonormalbasis.

Skalarprodukt auf komplexen Vektoräumen

Definition 38 *Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung*

$$< \ , \ >: V \times V \longrightarrow \mathbb{C}, (v, w) \mapsto < v, w >$$

heisst hermitesches Skalarprodukt auf V, wenn gilt:

(a) Für alle $v, v_1, v_2, w, w_1, w_2 \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned} &< v_1 + v_2, w > = < v_1, w > + < v_1, w >, \quad &< \lambda v, w > = \lambda < v, w >, \\ &< v, w_1 + w_2 > = < v, w_1 > + < v, w_2 >, \quad &< v, \lambda w > = \overline{\lambda} < v, w > . \end{aligned}$$

(hierbei bezeichnet $\bar{z}$ das komplex Konjugierte von $z \in \mathbb{C}$).

(b) $< \ , \ >$ ist hermitesch, d.h. $< v, w > = \overline{< w, v >}$ für alle $v, w \in V$.

(c) $< \ , \ >$ ist positiv definit, d.h. für $v \in V$ gilt:

$$< v, v > \geq 0, \qquad < v, v > = 0 \Leftrightarrow v = 0$$

Ein $\mathbb{C}$-Vektorraum V zusammen mit einem hermiteschen Skalarprodukt auf V heisst unitärer Vektorraum.

Bemerkung 39 Die Eigenschaft (a) wird als *Sesquilinearität* bezeichnet. $< \ , \ >$ ist also linear im ersten und *semilinear* im zweiten Argument.

Beispiele 4 *Seien $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$. Durch*

$$< z, w >: \ = z_1 \bar{w}_1 + \ldots + z_n \bar{w}_n = \sum_{i=1}^{n} z_i \bar{w}_i \in \mathbb{C}$$

ist auf $V = C^n$ ein hermitesches Skalarprodukt definiert.

Sei V ein unitärer Vektorraum mit hermiteschem Skalarprodukt $< \ , \ >$. Die Norm $\|v\|$ eines Vektors $v \in V$ ist wieder durch $\|v\| = \sqrt{< v, v >}$ definiert (man beachte, dass $< v, v >$ eine nicht-negative reelle Zahl ist). Es gelten analog die Aussagen von Lemma 29. Wir beweisen hier nur die Cauchy-Schwarzschen-Ungleichung für ein hermitesches Skalarprodukt.

Zunächst bemerken wir, dass man V auch als reellen Vektorraum auffassen kann indem man die Skalarmultiplikation $(\lambda, v) \mapsto \lambda v$ einfach auf reelle Skalare λ einschränkt. Wir benutzen die Bezeichnung $V_{\mathbb{R}}$, wenn wir V nur als reellen Vektorraum betrachten. Man rechnet leicht nach, dass der Realteil $< v, w >_{\mathbb{R}} := \mathrm{Re}(< v, w >)$ ein Skalarprodukt auf $V_{\mathbb{R}}$ ist. Die Cauchy-Schwarzsche-Ungleichung gilt also für $< \ , \ >_{\mathbb{R}}$, d.h. für $v, w \in V$ gilt:

$$| < v, w >_{\mathbb{R}} | \leq \sqrt{< v, v >_{\mathbb{R}}}\sqrt{< w, w >_{\mathbb{R}}}. \tag{12}$$

Wegen $< v, v > \in \mathbb{R}$, also $< v, v >_{\mathbb{R}} = < v, v >$ (und ebenso $< w, w >_{\mathbb{R}} = < w, w >$) ist die rechte Seite $= \|v\|\|w\|$. Ist $< v, w > \in \mathbb{R}$, so ist die linke Seite der Ungleichung (12) $= | < v, w > |$. Falls $< v, w > \notin \mathbb{R}$ setzen wir $\lambda := \frac{|<v,w>|}{<v,w>} \in \mathbb{C}$ und $v^* := \lambda v$. Dann gilt:

$$< v^*, w > = \lambda < v, w > = | < v, w > | \in \mathbb{R}, \ < v^*, v^* > = \lambda\overline{\lambda} < v, v > = < v, v >$$

da $\lambda\overline{\lambda} = |\lambda|^2 = 1$. Wir erhalten:

$$| < v, w > | = < v^*, w > = < v^*, w >_{\mathbb{R}} \overset{(12)}{\leq} \|v^*\|\|w\| = \|v\|\|w\|.$$

Hierbei haben wir (12) für v^* und w (anstelle von v und w) benutzt. $\square$

Man definiert die Begriffe *orthogonal, othogonales Komplement, Orthonormalsystem* und *Orthonormalbasis* für unitäre Vektorräume ganz analog wie im Fall von euklidischen Vektorräumen. Die Sätze 35 und 36 gelten wortwörtlich auch für unitäre Vektorräume.

3 Eigenwerte und Eigenvektoren

Sei K ein Körper, V ein endlich-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Zu einer Basis $\underline{v} = (v_1, \ldots, v_n)$ von V betrachten wir die Darstellungsmatrix $M^{\underline{v}}_{\underline{v}}(f)$ von f.

Wir wollen jetzt untersuchen, ob es möglich ist $\underline{v}$ so zu wählen, dass $M^{\underline{v}}_{\underline{v}}(f)$ eine besonders einfache Gestalt besitzt. Gibt es z.B. eine Basis $\underline{v}$, so dass $M^{\underline{v}}_{\underline{v}}(f)$ eine Diagonalmatrix ist:

$$M^{\underline{v}}_{\underline{v}}(f) = \begin{pmatrix} * & & & \\ & & & 0 \\ & & \ddots & \\ & 0 & & \\ & & & * \end{pmatrix}?$$

Diese Fragestellung lässt sich auch im Rahmen der Matrizenrechnung formulieren. Gegeben sei dazu eine quadratische Matrix $A \in M(n \times n, K)$. Gibt es eine invertierbare Matrix $B \in GL_n(K)$, so dass BAB^{-1} Diagonalgestalt besitzt:

$$BAB^{-1} = \begin{pmatrix} * & & \\ & \ddots & \\ & & * \end{pmatrix} ?$$

Wir wollen in diesem Abschnitt untersuchen unter welchen Bedingungen an f (bzw. A) dies möglich ist.

Definition 40 *(a) Sei V ein K-Vektorraum und $f : V \to V$ ein Endomorphismus. Ein Skalar $\lambda \in K$ heisst Eigenwert von f, wenn es einen Vektor $v \in V - \{0\}$ gibt mit $f(v) = \lambda v$. In diesem Fall nennt man v einen Eigenvektor von f zum Eigenwert λ.*

(b) Sei $A = (a_{ij}) \in M(n \times n, K)$. Ein $\lambda \in K$ heisst Eigenwert von A, wenn λ ein Eigenwert der zu A gehörigen linearen Abbildung $\ell_A : K^n \to K^n$ ist, d.h. wenn es einen Vektor $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n - \{0\}$ gibt mit

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Lemma 41 *Sei V ein endlich-dimensionaler K-Vektorraum, $f : V \to V$ ein Endomorphismus und $\lambda \in K$. Sei $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V und $A := M^{\underline{v}}_{\underline{v}}(f)$. Die folgenden Bedingungen sind äquivalent:*

(i) λ ist Eigenwert von f.

(ii) $\mathrm{Kern}(\lambda \cdot \mathrm{id}_V - f) \neq \{0\}$

(iii) $\det(\lambda \cdot \mathrm{id}_V - f) = 0$

(iv) $\det(\lambda E_n - A) = 0.$

Beweis. (i) $\iff$ (ii) Es gilt:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } f &\iff \exists v \in V, v \neq 0 : \quad f(v) = \lambda v \\ &\iff \exists\, v \in V - \{0\} : (f - \lambda id_V)(v) = 0 \\ &\iff \mathrm{Kern}(\lambda\, \mathrm{id}_V - f) \neq \{0\}. \end{aligned}$$

nicht (ii) $\iff$ nicht (iii)

$$\begin{aligned}\operatorname{Kern}(\lambda \operatorname{id}_V - f) = 0 &\iff \lambda \operatorname{id}_V - f \text{ ist injektiv}\\ &\iff \lambda \operatorname{id}_V - f \text{ ist ein Isomorphismus}\\ &\iff \det(\lambda \operatorname{id}_V - f) \neq 0.\end{aligned}$$

(iii) $\iff$ (iv) Nach Definition gilt:

$$\begin{aligned}\det(\lambda \operatorname{id}_V - f) &= \det(M_{\underline{v}}^{\underline{v}}(\lambda \operatorname{id}_V - f))\\ &= \det(\lambda M_{\underline{v}}^{\underline{v}}(\operatorname{id}_V) - M_{\underline{v}}^{\underline{v}}(f))\\ &= \det(\lambda E_n - A).\end{aligned}$$

□

Beispiele 5 (a) Sei $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$. Der Vektor $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist Eigenvektor von A zum Eigenwert 3.

(b) Sei $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in M(2\times 2, \mathbb{R})$. Wegen $\det(\lambda E_2 - A) = \det \begin{pmatrix} \lambda - 1 & -1 \\ -1 & \lambda \end{pmatrix} = \lambda^2 - \lambda - 1$ gilt

$$\det(\lambda E_2 - B) = 0 \quad \iff \quad \lambda = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1 \pm \sqrt{5}}{2}$$

d.h. B hat die beiden Eigenwerte $\lambda_1 = \frac{1+\sqrt{5}}{2}$ und $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

Definition 42 *(a) Sei V ein endlich-dimensionaler K-Vektorraum, $f : V \to V$ ein Endomorphismus und $\lambda \in K$. Der Untervektorraum*

$$\operatorname{Eig}(f, \lambda) :\, = \operatorname{Kern}(\lambda\lambda \operatorname{id}_V - f)$$

heisst der Eigenraum von f zu $\lambda \in K$. Ein Skalar λ ist also genau dann ein Eigenwert von f, wenn gilt $\operatorname{Eig}(f, \lambda) \neq \{0\}$. *In diesem Fall sind die Vektoren $v \in \operatorname{Eig}(f, \lambda) - \{0\}$ die Eigenvektoren von f.*

Sei λ ein Eigenwert von f. Die Dimension des Eigenraums $\operatorname{Eig}(f, \lambda)$ *wird als die geometrische Vielfachheit von λ bezeichnet.*

(b) Für $A \in M(n \times n, K)$ ist der Eigenraum von A zu $\lambda \in K$ definiert als der Eigenraum der zugehörigen linearen Abbildung:

$$\operatorname{Eig}(A, \lambda) :\, = \operatorname{Eig}(\ell_A, \lambda).$$

Ist λ ein Eigenwert von A so nennt man $\dim(\operatorname{Eig}(A, \lambda))$ *wieder die geometrische Vielfachheit von λ.*

Bemerkung 43 (a) Sei $f : V \to V$ wie in Lemma 41. Für $\lambda = 0$ gilt insbesondere:

0 ist Eigenwert von $f \iff \text{Kern}(f) \neq 0 \iff f$ ist kein Isomorphismus.

(b) Sei $A \in M(n \times n, K)$, $\lambda \in K$ und sei $\underline{e} = (e_1, \ldots, e_n)$ die Standardbasis von K^n. Wegen $M_{\underline{e}}^{\underline{e}}(\ell_A) = A$ erhalten wir:

λ ist Eigenwert von $A \iff \det(\lambda E_n - A) = 0$.

Das Charakteristische Polynom einer quadratischen Matrix. Sei $A = (a_{ij}) \in M(n \times n, K)$ und $t \in K$ setzen wir $\chi_A(t) \colon = \det(tE_n - A)$. Für $n = 2$ gilt

$$\begin{aligned}\chi_A(t) &= \det \begin{pmatrix} t - a_{11} & -a_{12} \\ -a_{21} & t - a_{22} \end{pmatrix} \\ &= (t - a_{11})(t - a_{22}) + a_{21}a_{12} = t^2 - (a_{11} + a_{22})t + \det(A)\end{aligned}$$

und für $n = 3$

$$\chi_A(t) = \det \begin{pmatrix} t - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & t - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & t - a_{33} \end{pmatrix} = t^3 + at^2 + bt + c$$

mit

$$\begin{aligned} a &= a_{11} + a_{22} + a_{33} \\ b &= a_{11}a_{22} + a_{11}a_{33} + a_{22}a_{33} - a_{12}a_{21} - a_{13}a_{31} - a_{23}a_{32} \\ c &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} \\ &= \det(A), \end{aligned}$$

d.h. für $n = 2$ und $n = 3$ ist $\chi_A(t)$ ein normiertes Polynom vom Grad n. Mit Hilfe der Leibniz Formel (Satz 1.3.26) kann man leicht zeigen, dass allgemein gilt:

Satz 44 *Für $A \in M(n \times n, K)$ ist $\chi_A(t) = \det(tE_n - A)$ ein normiertes Polynom vom Grad n. Man nennt χ_A das charakteristische Polynom von A.*

Nach Lemma 41 ist $\lambda \in K$ genau dann ein Eigenwert von A, wenn gilt $\det(\lambda E_n - A) = 0$.

Folgerung 45 *Die Eigenwerte von $A \in M(n \times n, K)$ sind die Nullstellen des charakteristischen Polynoms. Insbesondere besitzt A höchstens n verschiedene Eigenwerte.*

Ähnliche Matrizen

Definition 46 *Zwei Matrizen $A, B \in M(n \times n, K)$ heissen ähnlich (Notation: $A \sim B$), wenn es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ gibt mit*

$$S^{-1}AS = B$$

Bemerkung 47 A und B sind genau dann ähnlich, wenn das folgende Diagramm von linearen Abbildungen kommutiert:

$$\begin{tikzcd} K^n \arrow[r, "\ell_A"] & K^n \\ K^n \arrow[u, "\cong"', "\ell_S"] \arrow[r, "\ell_B"] & K^n \arrow[u, "\cong"', "\ell_S"] \end{tikzcd}$$

Lemma 48 *Für $A, B, C \in M(n \times n, K)$ gilt:*

(a) $A \sim A$

(b) $A \sim B \Rightarrow B \sim A$

(c) $A \sim B$ *und* $B \sim C \Rightarrow A \sim C$

Bemerkung 49 Man nennt eine Beziehung (oder Relation) zwischen Elementen einer Menge mit den Eigenschaften (a) – (c) aus Lemma 48 eine *Aquivalenzrelation*.

Beweis von (c). Falls gilt $A \sim B$ und $B \sim C$, so gibt es $S, T \in \mathrm{GL}_n(K)$ mit $B = S^{-1}AS$ und $C = T^{-1}BT$. Es folgt:

$$C = T^{-1}S^{-1}AST = (ST)^{-1}A(ST)$$

also $A \sim C$. □

Lemma 50 *Seien $A, B \in M(n \times n, K)$ ähnlich. Dann gilt:*

$$\chi_A(T) = \chi_B(T)$$

Beweis. Sei $S \in GL_n(K)$ mit $B = S^{-1}AS$. Dann gilt:

$$tE_n - B = tE_n - S^{-1}AS = tS^{-1}E_nS - S^{-1}AS = S^{-1}(tE_n - A)S$$

und folglich

$$\begin{aligned} \chi_B(t) &= \det(tE_n - B) = \det(S^{-1}(tE_n - A)S) \\ &= \det(S)^{-1}\det(tE_n - A)\det(S) = \det(tE_n - A) = \chi_A(t). \end{aligned}$$

□

Das charkteristische Polynom eines Endomorphismus

Satz und Definition 51 *Sei V ein n-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Für $t \in K$ setzen wir*

$$\chi_f(t) \colon = \det(t\,\mathrm{id}_V - f).$$

(a) $\chi_f(t)$ ist ein normiertes Polynom vom Grad n.

(b) Ist $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V und $A := M_{\underline{v}}^{\underline{v}}(f)$ so gilt

$$\chi_f(t) = \chi_A(t).$$

Man nennt $\chi_f(t)$ das charakteristische Polynom von f.

(c) $\lambda \in K$ ist genau dann ein Eigenwert von f, wenn $\chi_f(\lambda) = 0$ ist. Insbesondere besitzt f höchstens $n = \dim(V)$ verschiedene Eigenwerte.

Beweis. Wir zeigen zunächst (b). Wegen $M_{\underline{v}}^{\underline{v}}(t\,\mathrm{id}_V - f) = tM_{\underline{v}}^{\underline{v}}(\mathrm{id}_V) - M_{\underline{v}}^{\underline{v}}(f) = tE_n - A$ gilt nach Definition 7:

$$\chi_f(t) = \det(t\,\mathrm{id}_V - f) = \det(M_{\underline{v}}^{\underline{v}}(t\,\mathrm{id}_V - f)) = \chi_A(t).$$

(a) folgt aus (b) und Satz 44. Da ein Polynom vom Grad n höchstens n Nullstellen besitzt folgt aus (c) aus (a) und Lemma 41. □

Wir zeigen jetzt dass Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind. Das liefert einen weiteren Beweis der zweietn Aussage (c) im obigen Satz.

Lemma 52 *Sei V ein K-Vektorraum und $f : V \to V$ ein Endomorphismus. Seien $v_1, \ldots, v_r$ Eigenvektoren von f zu verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda \in K$ (d.h. es gilt $\lambda \neq \lambda_j$ für $i \neq j$ und $f(v_i) = \lambda_i v_i$ für $i = 1, \ldots, r$). Dann ist $(v_1, \ldots, v_r)$ linear unabhängig.*

Beweis. Seien $x_1 \ldots x_r \in K$ mit

$$x_1 v_1 + \ldots + x_r v_r = 0 \tag{13}$$

Wenden wir f auf diese Gleichung an, so folgt

$$0 = f(x_1 v_1 + \ldots + x_r v_r) = x_1 f(v_1) + \ldots + x_r f(v_r) = (x_1\lambda_1)v_1 + \ldots + (x_r\lambda_r)v_r$$

Durch wiederholtes Anwenden von f erhalten wir

$$\begin{aligned} 0 &= f^m(x_1v_1 + \ldots + x_rv_r) = f^{m-1}((x_1\lambda_1)v_1 + \ldots + (x_r\lambda_r)v_r) \\ &= \ldots = (x_1\lambda_1^m)v_1 + \ldots + (x_r\lambda_r^m)v_r \end{aligned}$$

für alle natürlichen Zahlen m.

Für ein beliebiges Polynom $P(t) = a_nt^n + \ldots + a_1t + \ldots + a_0$ mit Koeffizienten in K ergibt sich

$$\begin{aligned} (x_1P(\lambda_1))v_1 + \ldots + (x_rP(\lambda_r))v_r &= \sum_{i=1}^{r}(x_iP(\lambda_i))v_i = \sum_{i=1}^{r}((\sum_{j=0}^{n} x_ia_j\lambda_i^j))v_i \\ &= \sum_{j=0}^{n} a_j(\sum_{i=1}^{r}(x_i\lambda_i^j)v_i) = 0 \end{aligned}$$

Wir wählen jetzt speziell für $P(t)$ das Polynom

$$P(t) = \frac{\prod_{j=2}^{r}(t-\lambda_j)}{\prod_{j=2}^{r}(\lambda_i-\lambda_j)}$$

Man beachte, dass $P(\lambda_1) = 1$ und $P(\lambda_i) = 0$ für $i = 2, \ldots, r$. Folglich erhalten wir

$$0 = (x_1P(\lambda_1))v_1 + \ldots + (x_rP(\lambda_r))v_r = x_1v_1$$

Da v_1 ein Eigenvektor von f, also insbesonder nicht der Nullvektor ist, folgt $x_1 = 0$.

Analog zeigt man: $x_2 = \ldots = x_r = 0$. Also ist $(v_1, \ldots, v_r)$ linear unabhängig. $\square$

Folgerung 53 *Sei V ein n-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Seien $\lambda_1, \ldots, \lambda_r \in K$ die verschiedenen Eigenwerte von V. Dann gilt*

$$\dim(\mathrm{Eig}(f,\lambda_1)) + \ldots + \dim(\mathrm{Eig}(f,\lambda_r)) \leq n$$

d.h. die Summe der geometrischen Vielfachheiten von f ist höchstens $= n$.

Beweis. Für $i = 1, \ldots, r$ sei $m_i \colon = \dim(\mathrm{Eig}(f,\lambda_i))$. Sei

$$\begin{array}{ll} (v_1^{(1)}, \ldots, v_{m_1}^{(1)}) & \text{eine Basis von } \mathrm{Eig}(f,\lambda_1) \\ (v_1^{(2)}, \ldots, v_{m_1}^{(2)}) & \text{eine Basis von } \mathrm{Eig}(f,\lambda_2) \\ \vdots & \\ (v_1^{(r)}, \ldots, v_{m_r}^{(r)}) & \text{eine Basis von } \mathrm{Eig}(f,\lambda_r). \end{array}$$

Wir zeigen, dass

$$(v_1^{(1)}, \ldots, v_{m_1}^{(1)}, v_1^{(2)}, \ldots, v_{m_2}^{(2)}, \ldots, v_1^{(r)}, \ldots, v_{m_r}^{(r)}) \tag{14}$$

linear unabhängig ist. Seien dazu $x_1^{(1)}, \ldots, x_{m_1}^{(1)}, \ldots, x_1^{(r)}, \ldots, x_{m_r}^{(r)} \in K$ mit

$$\sum_{i=1}^{r} \sum_{j=1}^{m_i} x_j^{(i)} v_j^{(i)} = 0$$

Wir zeigen zunächst:

$$\sum_{j=1}^{m_i} x_j^{(i)} v_j^{(i)} = 0 \quad \text{für } i = 1, \ldots, r \tag{15}$$

Setze $w_i := \sum_{j=1}^{m_i} x_j^{(i)} v_j^{(i)} \in \mathrm{Eig}(f, \lambda_i)$ für $i = 1, \ldots, r$. Angenommen (15) gilt nicht, d.h. einige der w_i's sind $\neq 0$. Es seien etwa $w_1, w_2, \ldots, w_s$ alle $\neq 0$ und $w_{s+1} = \ldots = w_r = 0$. Dann sind $w_1, w_2, \ldots, w_s$ Eigenvektoren von f (zu verschiedenen Eigenwerten) und es gilt

$$w_1 + \ldots + w_s = w_1 + \ldots + w_r = 0$$

Das widerspricht Lemma 52. Also gilt $w_1 = \ldots = w_r = 0$.

Da $(v_1^{(i)}, \ldots, v_{m_1}^{(i)})$ linear unabhängig ist folgt aus (15):

$$x_1^{(i)} = \ldots = x_{m_i}^{(i)} = 0 \quad \text{für alle } i = 1, \ldots, r.$$

Das zeigt, dass das $m_1 + m_2 + \ldots + m_r$-Tupel (14) linear unabhängig ist. Es folgt $m_1 + m_2 + \ldots + m_r \leq \dim(V) = n$. □

Satz und Definition 54 *(a) Sei V ein n-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Die folgenden Bedingungen sind äquivalent:*

(i) V besitzt eine Basis aus Eigenvektoren von f.

(ii) Es gibt eine Basis $\underline{v} = (v_1, \ldots, v_n)$ von V so dass für die Darstellungsmatrix gilt:

$$M_{\underline{v}}^{\underline{v}}(f) = \begin{pmatrix} \lambda_1 & & & \\ & & & 0 \\ & & \ddots & \\ & 0 & & \\ & & & \lambda_n \end{pmatrix}$$

(iii) Die Summe der geometrische Vielfachheiten ist $= n$*, d.h. sind* $\lambda_1, \ldots, \lambda_r$ *die verschiedenen Eigenwerte von* f *so gilt*

$$\dim(\mathrm{Eig}(f, \lambda_1)) + \ldots + \dim(\mathrm{Eig}(f, \lambda_r)) = n$$

f *heißt diagonalisierbar, wenn es die obigen äquivalenten Bedingungen erfüllt.*

(b) Eine quadratische Matrix $A \in M(n \times n, K)$ *heisst diagonalisierbar, wenn eine der folgenden äquivalenten Bedingungen gilt:*

(i) Die lineare Abbildung $\ell_A : K^n \to K^n, x \mapsto Ax$ *ist diagonalisierbar.*

(ii) A *ist ähnlich zu einer Diagonalmatrix ist.*

Beweis. (a) (i) $\Leftrightarrow$ (ii) Offenbar besteht eine Basis $\underline{v} = (v_1, \ldots, v_n)$ genau dann aus Eigenvektoren von f wenn $M_{\underline{v}}^{\underline{v}}(f)$ eine Diagonalmatrix ist.

(i) $\Leftrightarrow$ (iii) klar.

(iii) $\Leftrightarrow$ (i) Wir benutzen die Bezeichnung aus dem Beweis von Folgerung 53. Wenn gilt $m_1 + m_2 + \ldots + m_r = n$, so ist (14) eine Basis von V aus Eigenvektoren von f.

(b) (i) $\Leftrightarrow$ (ii) Sei $(v_1, \ldots, v_n)$ eine Basis von K^n aus Eigenvektoren von ℓ_A, d.h. es gelte $Av_i = \lambda_i v_i, \lambda_i \in K$ für $i = 1, \ldots, n$. Sei $S \in M(n \times n, K)$ die Matrix mit den Spalten $v_1, \ldots, v_n$. Dann ist S invertierbar (da die Spalten eine Basis bilden) und die Spalten des Produkts AS sind $\lambda_1 v_1, \ldots, \lambda_n v_n$. Es folgt

$$AS = S \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad \Rightarrow \quad S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

(ii) $\Leftrightarrow$ (i) Ist umgekehrt $S^{-1}AS$ eine Diagonalmatrix für $S \in \mathrm{GL}_n(K)$ so bilden die Spalten von S eine Basis aus Eigenvektoren. $\square$

Folgerung 55 *Sei* V *ein* n*-dimensionaler* K*-Vektorraum und* $f : V \to V$ *ein Endomorphismus.*

(a) Sei f *diagonalisierbar so zerfällt* $\chi_f(T)$ *in ein Produkt von Linearfaktoren. Genauer gilt: Sind* $\lambda_1, \ldots, \lambda_r \in K$ *die verschiedenen Eigenwerte von* f *mit den geometrischen Vielfachheiten* $m_1, \ldots, m_r$ *so gilt:*

$$\chi_f(t) = (t - \lambda_1)^{m_1} \cdot \ldots \cdot (t - \lambda_r)^{m_r}.$$

(b) Wenn f genau n verschiedene Eigenwerte besitzt (d.h. $\chi_f(T)$ zerfällt in ein Produkt von Linearfaktoren $\chi_f(T) = \prod_{i=1}^n (t-\lambda_i)$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_n$), so ist f diagonalisierbar.

Beweis. (a) Mit den Bezeichnungen von Folgerung 53 gilt für die Darstellungsmatrix $A\colon = M_{\underline{v}}^{\underline{v}}(f)$ von f bzgl. der Basis $\underline{v} = (v_1^{(1)}, \ldots, v_{m_1}^{(1)}, \ldots, v_1^{(r)}, \ldots, v_{m_r}^{(r)})$ von V:

$$A = \begin{pmatrix} \lambda_1 E_{m_1} & & & \\ & \lambda_2 E_{m_2} & & \\ & & \ddots & \\ & & & \lambda_r E_{m_r} \end{pmatrix}$$

$$\Rightarrow \quad \chi_f(t) = \chi_A(t) = \det \begin{pmatrix} (t-\lambda_1) E_{m_1} & & & \\ & (t-\lambda_2) E_{m_2} & & \\ & & \ddots & \\ & & & (t-\lambda_r) E_{m_r} \end{pmatrix}$$

$$= \prod_{i=1}^{r} (t-\lambda_i)^{m_i}$$

(b) Wir nehmen an, dass f genau n verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$ besizt. Für $i = 1, \ldots, n$ sei $m_i = \dim(\mathrm{Eig}(f, \lambda_i))$ die geometrische Vielfachheit von λ_i. Wegen $\mathrm{Eig}(f, \lambda_i) \neq \{0\}$ gilt $m_i \geq 1$ und damit nach Folgerung 53

$$n \geq m_1 + \ldots + m_n \geq n$$

also $m_1 + \ldots + m_n = n$. Nach Satz 54 (a) (iii) ist f diagonalisierbar. □

Beispiele 6 (a) Die Matrix $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$ ist diagonalisierbar, da $\chi_A(t) = (t-3)(t-2)$ in ein Produkt von verschiedene Linearfaktoren zerfällt. Der Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist Eigenvektor zum Eigenwert 3 und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist Eigenvektor zum Eigenwert 2. Für $T\colon = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ gilt:

$$AT = T \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \quad \Rightarrow T^{-1}AT = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

(b) Sei $B := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Da $\chi_B(t) = t^2 + 1$ keine reellen Nullstellen besitzt ist B (aufgefasst als Matrix mit reellen Einträgen) nicht diagonalisierbar.

Als Matrix in $M(2 \times 2, \mathbb{C}$ ist B dagegen diagonalisierbar, denn $\chi_B(t) \in \mathbb{C}[t]$ zerfällt in verschiedene Linearfaktoren: $\chi_B(t) = (t+i)(t-i)$. In der Tat gilt für $S \colon = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$:

$$S^{-1}BS = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

(b) Die Matrix $C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist nicht diagonalisierbar. Wegen $\chi_C(t) = (t-1)^2$ ist $\lambda = 1$ der einzige Eigenwert von B. Da

$$\mathrm{Eig}(C,1) = \mathrm{Kern}(E_2 - C) = \mathrm{Kern}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = L(\begin{pmatrix} 1 \\ 0 \end{pmatrix})$$

eindimensional ist ist die Summe der geometrische Vielfachheiten 1.

Definition 56 *Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionaler K-Vektorraums und sei $\lambda \in K$ ein Eigenwert von f. Dann gibt es eine Zerlegung*

$$\chi_f(t) = (t-\lambda)^\mu \cdot g(T) \text{ mit } g(\lambda) \neq 0$$

mit eindeutig bestimmten $\mu \in \mathbb{N}, \mu \geq 1$ und $g(t) \in K[t], g(\lambda) \neq 0$. Die Zahl $\mu = \mu(\chi_f, \lambda)$ heisst die algebraische Vielfachheit von λ.

Ohne Beweis bemerken wir:

Satz 57 *Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionaler K-Vektorraums und sei $\lambda \in K$ ein Eigenwert von f. Sei m (bzw. μ) die geometrische (bzw. algebraische) Vielfachheit von f. Dann gilt:*

$$1 \leq m \leq \mu$$

Beispiel 4

$$A = \begin{pmatrix} 1 & 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

mit $A_1 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ und $A_2 = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 2 & 7 \\ 0 & 0 & 2 \end{pmatrix}$.

$$\Rightarrow \quad \chi_A = \chi_{A_1}\chi_{A_2} = (T-1)^2 \cdot (T-2)^3$$

A besitzt also zwei Eigenwerte mit den algebraischen Vielfachheiten

$$\mu(\chi_A, 1) = 2, \quad \mu(\chi_A, 2) = 3.$$

Es gilt ferner

$$\mathrm{Rang}(E_5 - A) = \mathrm{Rang}\begin{pmatrix} 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = 4$$

$$\Rightarrow \quad \dim(\mathrm{Eig}(A, 1)) = 5 - \mathrm{Rang}(E_5 - A) = 1$$

Ebenso zeigt man:

$$\mathrm{Rang}(2E_5 - A) = 4 \quad \Rightarrow \quad \mathrm{Eig}(A, 2) = 1.$$

Die geometrischen Vielfachheiten sind also beide $= 1$.

4 Endomorphismen von Euklidischen Vektorräumen

Sei V ein endlich-dimensionaler euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$ und zugehörige Norm $\|v\| = \sqrt{< v, v >}$. In diesem Abschnitt untersuchen wir speziell Endomorphismen von V die in einer Beziehung zum Skalarprodukt stehen. Wir betrachten die folgenden Typen von Endomorphismen:

- Orthogonale Projektionen.
- Orthogonale Abbildungen (das sind Abbildungen die Längen und Winkel erhalten).
- Selbstadjungierte Endomorphismen.

Orthogonale Projektionen Sei $U \subseteq V$ ein Untervektorraum. Die orthogonale Projektion $p_U : V \to V$ ist eine lineare Abbildung mit der Eigenschaft $p_U(u) = u$ für alle $u \in U$ und $p_U(w) = 0$ für alle $w \in U^\perp$ (p_U ist durch diese Eigenschaften eindeutig charakterisiert). Bevor wir die Definition von p_U angeben benötigen wir den Begriff der *direkten Summe* zweier Untervektorräume.

Lemma und Definition 58 *Seien $U_1, U_2 \subseteq V$ Untervektorräume von V. Die folgenden Bedingungen sind äquivalent:*

(i) $V = U_1 + U_2$, $U_1 \cap U_2 = \{0\}$

(ii) Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung der Gestalt $v = u_1 + u_2$ mit $u_1 \in U_1$, $u_2 \in U_2$.

In diesem Fall sagt man, dass V die direkte Summe von U_1 und U_2 ist (Notation: $V = U_1 \oplus U_2$).

Beweis. (i) $\Rightarrow$ (ii): $V = U_1 + U_2$ bedeutet, dass sich jeder Vektor $v \in V$ in der Form $v = u_1 + u_2$ mit $u_1 \in U_1$, $u_2 \in U_2$ darstellen lässt. Zu zeigen ist, dass diese Darstellung eindeutig ist. Sei also $v = u_1' + u_2'$ eine weitere Darstellung mit $u_1' \in U_1$, $u_2' \in U_2$. Dann gilt

$$u_1 + u_2 = u_1' + u_2' \quad \Rightarrow u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2$$

Die Voraussetzung $U_1 \cap U_2 = \{0\}$ liefert also $u_1 - u_1' = u_2' - u_2 = 0$, d.h. $u_1 = u_1'$ und $u_2' = u_2$.

(ii) $\Rightarrow$ (i): Übung. $\square$

Beispiel 5 Sei

$$V = \mathbb{R}^3, \quad U_1 = L\left(\begin{pmatrix}1\\1\\0\end{pmatrix}\right), \quad U_2 = \left\{\begin{pmatrix}0\\y\\z\end{pmatrix} \mid y, z, \in \mathbb{R}\right\} \quad (y\text{-}z\text{-Ebene}).$$

Dann gilt $U_1 \cap U_2 = \{0\}$, $\dim(U_1) = 1$, $\dim(U_2) = 2$. Mit Hilfe der Dimensionsformel für Untervektorräume folgt $U_1 + U_2 = V$. Also ist V direkte Summe von U_1 und U_2.

Lemma 59 *Sei $U \subseteq V$ ein Untervektorraum. Dann ist V direkte Summe von U und $U^\perp$.*

Beweis. Für $u \in U \cap U^\perp$ gilt $u \perp u$, d.h. $< u, u >= 0$ und damit $u = 0$. Folglich: $U \cap U^\perp = \{0\}$.

Wir zeigen jetzt, dass $V = U + U^\perp$ gilt. Sei $(v_1, \ldots, v_r)$ eine Orthonormalbasis von U. Für ein beliebiges $v \in V$ setzen wir:

$$u\colon = \sum_{i=1}^{r} < v, v_i > v_i, \quad w\colon = v - \sum_{i=1}^{r} < v, v_i > v_i$$

Dann gilt $v = u + w$, $u \in L(v_1, \ldots, v_r) = U$ und (nach Satz 35 (c)) $w \in U^\perp$. $\square$

Definition 60 *Sei U ein Untervektorraum von V. Nach Lemma 58 und 59 lässt sich jedes $v \in V$ darstellen in der Form $v = u + w$ mit eindeutig bestimmten $u \in U$, $w \in U^\perp$. Man nennt die Abbildung*

$$p_U : V \to V, v \mapsto u$$

die orthogonale (oder senkrechte) Projektion von V auf U.

Satz 61 *(Eigenschaften der orthogonalen Projektion) Sei $v \in V$ und $(v_1, \ldots, v_r)$ eine Orthonormalbasis von U. Es gilt:*

(a) p_U ist linear mit $p_U \circ p_U = p_U$;

(b) $p_U(v) \in U$ und $v - p_U(v) \in U^\perp$;

(c) $p_U(v) = v \iff v \in U, \quad p_U(v) = 0 \iff v \in U^\perp$;

(d) $p_U(v) = \sum_{i=1}^{r} < v, v_i > v_i$.

Beweis. (a) und (b) folgen sofort aus der Definition von p_U. $\square$

Folgerung 62 *Sei V ein endlich-dimensional euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$ und sei $U \subseteq V$ ein Untervektorraum. Dann gilt:*

(a) $\dim(U) + \dim(U^\perp) = \dim(V)$.

(b) $(U^\perp)^\perp = U$.

Beweis. (a) folgt wegen $\mathrm{Kern}(p_U) = U^\perp$, $\mathrm{Bild}(p_U) = U$ aus der Dimensionsformel für lineare Abbildungen.

(b) Sei $u \in U$. Wegen $< u, w >= 0$ für alle $w \in U^\perp$ gilt $u \in (U^\perp)^\perp$, d.h. $U \subseteq (U^\perp)^\perp$. Wenden wir (a) für $U^\perp$ anstelle von U an so folgt ferner

$$\begin{aligned} \dim(U^\perp)^\perp &= \dim(V) - \dim(U^\perp) = (\dim(U) + \dim(U^\perp)) - \dim(U^\perp) \\ &= \dim(U) \end{aligned}$$

und damit $(U^\perp)^\perp = U$. $\square$

Orthogonale Abbildungen

Definition 63 *Eine lineare Abbildung $f : V \to V$ heisst orthogonale Abbildung wenn gilt:*

$$< f(v), f(w) > = < v, w > \qquad \textit{für alle } v, w \in V.$$

Lemma 64 Sei $f : V \to V$ eine orthogonale Abbildung. Dann gilt:

(a) f erhält Längen, d.h. $\|f(v)\| = \|v\| \quad \forall\, v \in V$.

(b) f erhält Längen und Winkel, d.h. sind $v, w \in V - \{0\}$ und $\varphi \in [0, \pi]$ der Winkel zwischen v und w, so ist φ auch der Winkel zwischen $f(v)$ und $f(w)$.

(c) Insbesondere: $v \perp w \Rightarrow f(v) \perp f(w)$.

(d) f ist ein Isomorphismus und f^{-1} ist wieder eine orthogonale Abbildung.

(e) f hat höchstens 1 und -1 als Eigenwert.

Beweis. (a) $\quad \|f(v)\| = \sqrt{< f(v), f(v) >} = \sqrt{< v, v >} = \|v\|$.

(b) Folgt aus

$$\frac{< v, w >}{\|v\|\,\|w\|} = \frac{< f(v), f(v) >}{\|v\|\,\|w\|} \overset{(a)}{=} \frac{< f(v), f(v) >}{\|f(v)\|\,\|f(w)\|}.$$

(c) klar.

(d) Wegen

$$v \in \mathrm{Kern}(f) \quad \Leftrightarrow \quad f(v) = 0 \quad \Leftrightarrow \quad 0 = \|f(v)\| \overset{(a)}{=} \|v\| \Leftrightarrow \qquad \Leftrightarrow \quad v = 0$$

ist f injektiv. Nach Folgerung 16 ist f ein Isomorphismus. Für f^{-1} gilt dann

$$< f^{-1}(v), f^{-1}(w) >=< f(f^{-1}(v)), f(f^{-1}(w)) >=< v, w >$$

d.h. f^{-1} ist ebenfalls orthogonal.

(e) Sei $v \in V - \{0\}$ ein Eigenvektor von f zum Eigenwert $\lambda \in \mathbb{R}$. Dann gilt

$$\begin{aligned} \lambda^2 < v, v > &=< \lambda v, \lambda v >=< f(v), f(v) >=< v, v > \\ &\Rightarrow \lambda^2 = 1 \quad \Rightarrow \quad \lambda = 1 \quad \text{oder} \quad \lambda = -1. \end{aligned}$$

□

Beispiele 7 (a) Sei $V = \mathbb{R}^2$ mit dem Standard-Skalarprodukt:

$$< \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} >: = x_1y_1 + x_2y_2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Sei $\varphi \in [0, 2\pi): = \{t \in \mathbb{R} \mid 0 \leq t < 2\pi\}$ und $f : V \to V$ die Drehung um den Winkel φ. Die Darstellungsmatrix $M^{\underline{e}}_{\underline{e}}(f)$ von f bzgl. der Standardbasis $\underline{e} = (e_1, e_2) = (\begin{pmatrix} 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \end{pmatrix})$ ist dann die Matrix

$$D(\varphi): = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

d.h. $f = \ell_{D(\varphi)}$ (offensichtlich erhält f Längen und Winkel).

Behauptung: f ist orthogonal.

Beweis. Es gilt

$$D(\varphi)^t \cdot D(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = E_2$$

(da $\sin^2\varphi \cos^2\varphi = 1$). Für $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ gilt:

$$\begin{aligned} < f(v), f(w) > &=< D(\varphi) \cdot v, D(\varphi) \cdot w > \\ &= (D(\varphi) \cdot v)^t \cdot D(\varphi) \cdot w = v^t \cdot D(\varphi)^t \cdot D(\varphi) \cdot w \\ &= v^t \cdot w =< v, w > . \end{aligned}$$

Also ist f orthogonal.

(b) Sei V ein beliebiger endlich-dimensionaler euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$. Sei $w \in V$ mit $\|w\| = 1$. Die Abbildung $s_w : V \to V, v \mapsto v - 2 < v, w > w$ ist orthogonal, denn für $v_1, v_2 \in V$ gilt wegen $< w, w >= 1$:

$$< s_w(v_1), s_w(v_2) >=< v_1 - 2 < v_1, w > w, v_2 - 2 < v_2, w > w >$$

$$=< v_1, v_2 > - < 2 < v_1, w > w, v_2 > - < v_1, 2 < v_2, w > w > + < 2 < v_1, w > w, 2 < v_2, w > w >$$

$$=< v_1, v_2 > -2 < v_1, w >< w, v_2 > -2 < v_2, w >< v_1, w > +4 < v_1, w >< v_2, w >< w, w >$$

$$=< v_1, v_2 > .$$

Die Abbildung s_w heisst *orthogonale Spiegelung* an der zu v orthogonalen Hyperebene $L(w)^\perp$ (falls gilt $\dim(V) = 2$, so ist $L(w)^\perp$ eindimensional und heisst dann auch *Spiegelungsachse*).

Satz 65 *Sei V ein endlich-dimensionaler euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$ und sei $\underline{v} = (v_1, \ldots, v_n)$ eine Orthonormalbasis von V. Sei $f : V \to V$ ein Endomorphismus und $A := M_{\underline{v}}^{\underline{v}}(f)$. Folgende Bedingungen sind äquivalent:*

(i) f ist orthognal;

(ii) $AA^t = A^t A = E_n$;

(iii) $(f(v_1), \ldots, f(v_n))$ ist eine Orthonormalbasis;

(iv) $\|f(v)\| = \|v\|$ für alle $v \in V$.

Beweis. Nach Satz 35 gilt:

$$f(v_j) = \sum_{i=1}^{n} < f(v_j), v_i > v_i$$

d.h. für die Einträge a_{ij} von A gilt $a_{ij} =< v_i, f(v_j) >$. Für $i, j \in \{1, \ldots, n\}$ erhalten wir

$$\begin{aligned}\sum_{k=1}^{n} a_{ki} a_{kj} &= \sum_{k=1}^{n} < v_k, f(v_i) >< v_k, f(v_j) > \\ &=< \sum_{k=1}^{n} < v_k, f(v_i) > v_k, f(v_j) >=< f(v_i), f(v_j) >\end{aligned}$$

$$A^t \cdot A = (< f(v_i), f(v_j) >) \tag{16}$$

(i) $\Rightarrow$ (ii) Ist f orthogonal so gilt

$$< f(v_i), f(v_j) >=< v_i, v_j >= \delta_{ij} : = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j, \end{cases}$$

also $A^t A = E_n$. Somit ist A invertierbar mit Inverser $A^{-1} = A^t$ und es folgt $AA^t = AA^{-1} = E_n$.

(ii) $\Rightarrow$ (iii) Wegen (16) folgt aus $A^t A = E_n$:

$$< f(v_i), f(v_j) >= \sum_{k=1}^{n} a_{ki} a_{kj} = \delta_{ij}.$$

(iii) $\Rightarrow$ (iv) Für $v = \sum\limits_{i=1}^{n} < v, v_i > v_i$ gilt:

$$
\begin{aligned}
< f(v), f(v) > &= \sum_{i,j=1}^{n} < v, v_i >< v, v_j >< f(v_i), f(v_j) > \\
&= \sum_{i,j=1}^{n} < v, v_i >< v, v_j > \delta_{ij} \\
&= \sum_{i=1}^{n} < v, v_i >^2 = < v, v >
\end{aligned}
$$

und damit $\|f(v)\| = \|v\|$ für alle $v \in V$.

(iv) $\Rightarrow$ (i) Nach Lemma 30 gilt:

$$
\begin{aligned}
< f(v), f(w) > &= \frac{1}{2}\left(\|f(v) + f(w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2\right) \\
&= \frac{1}{2}\left(\|f(v + w)\|^2 - \|f(v)\|^2 - \|f(2)\|^2\right) \\
&= \frac{1}{2}\left(\|v + w\|^2 - \|v\|^2 - \|w\|^2\right) = < v, w > .
\end{aligned}
$$

$\square$

Definition 66 *Eine Matrix $A \in M(n \times n, \mathbb{R})$ heisst orthogonal, wenn gilt: $AA^t = A^tA = E_n$. Die Menge aller orthogonalen Matrizen wird mit $O(n)$ bezeichnet.*

Bemerkungen 67 (a) $A \in M(n{\times}n, \mathbb{R})$ ist genau dann orthogonal wenn $\ell_A : \mathbb{R}^n \to \mathbb{R}^n$ eine orthogonale Abbildung ist (bzgl. des Standardskalarprodukts auf $\mathbb{R}^n$). Also gilt: A ist genau dann orthogonal, wenn die Spalten (oder Zeilen) von A eine Orthonormalbasis bilden.

(b) $O(n)$ ist eine Untergruppe von $GL_n(\mathbb{R})$.

(c) $A \in O(n) \Rightarrow \det(A) = \pm 1$ (denn $AA^t = E_n \Rightarrow \det(A)^2 = \det(A)\det(A^t)$ $= \det(E_n) = 1$).

Lemma 68 *Sei $f : V \to V$ eine orthogonale Abbildung und $U \subseteq V$ ein Untervektorraum. Dann gilt:*

$$f(U) \subseteq U \quad \Rightarrow \quad f(U^\perp) \subseteq U^\perp$$

Beweis. Da f ein Isomorphismus ist folgt aus $f(U) \subseteq U$ sogar $f(U) = U$. Für $w \in U^\perp$ und $u \in U$ also $u = f(u')$ mit $u' = f^{-1}(u) \in U$ ergibt sich:

$$\begin{aligned} < f(w), u > =&< f(w), f(f^{-1}(u)) > =< w, f^{-1}(u) >= 0 \qquad \forall\, u \in U \\ &\Rightarrow f(w) \in U^\perp \end{aligned}$$

□

Wir untersuchen jetzt orthogonale Abbildungen von 2- und 3-dimensionalen euklidischen Räumen.

Satz 69 *Sei V ein 2-dimensionaler euklidischer Vektorraum, sei (v_1, v_2) eine Orthonormalbasis von V und sei $f : V \to V$ eine orthogonale Abbildung. Dann gilt für die Darstellungsmatrix $A = M^{\underline{v}}_{\underline{v}}(f)$:*

(a) Ist $\det A = 1$ so gibt es genau ein $\varphi \in [0, 2\pi)$ mit

$$A = D(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

d.h. f ist eine Drehung um den Winkel φ um das Zentrum 0.

(b) Ist $\det(A) = -1$, so gibt es ein $\varphi \in [0, 2\pi)$ mit

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

In diesem Fall gibt es eine Orthonormalbasis $\underline{w} = (w_1, w_2)$, so dass

$$M^{\underline{w}}_{\underline{w}}(f) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

d.h. $f = s_{w_1}$ ist eine orthogonale Spiegelung mit Spiegelungsachse $L(w_1)^\perp = L(w_2)$.

Beweis. Sei $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Da die Spalten von A eine Orthonormalbasis im $\mathbb{R}^2$ bilden gilt:

$$\| \begin{pmatrix} a \\ b \end{pmatrix} \|^2 = a^2 + b^2 = 1 = c^2 + d^2 = \| \begin{pmatrix} c \\ d \end{pmatrix} \|^2, \quad \begin{pmatrix} a \\ b \end{pmatrix} \perp \begin{pmatrix} c \\ d \end{pmatrix}.$$

Schreiben wir die komplexe Zahl $a+bi$ in Polarkoordinatendarstellung $a+bi = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$ mit $\varphi \in [0, 2\pi)$ so gilt $r = \sqrt{a^2+b^2} = 1$ und damit $a = \cos\varphi, b = \sin\varphi)$. Da $\begin{pmatrix} c \\ d \end{pmatrix}$ ein Vektor der Länge 1 ist, der auf $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$ senkrecht steht. Davon gibt es nur zwei nämlich $\begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$ und $\begin{pmatrix} \sin\varphi \\ -\cos\varphi \end{pmatrix}$. Also gilt:

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} \sin\varphi \\ -\cos\varphi \end{pmatrix}$$

Der erste Fall tritt genau dann auf, wenn $\det A = 1$ ist und der zweite wenn $\det A = -1$ ist.

Wenn $\det A = -1$ ist, also $A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$ so gilt für das charakteristische Polynom von f:

$$\chi_f(t) = \chi_A(t) = t^2 - 1 = (t-1)(t+1)$$

Nach Folgerung 55 ist f in diesem Fall diagonalisierbar (mit den Eigenwerten 1 und -1). Wir wählen in den beiden Eigenräumen jeweils einen Eigenvektor $w_1 \in \mathrm{Eig}(f, -1)$ und $w_2 \in \mathrm{Eig}(f, 1)$ der Länge 1. Es gilt:

$$\begin{aligned} < w_1, w_2 > &= < f(w_1), f(w_2) > = < -w_1, w_2 > = - < w_1, w_2 > \\ \Rightarrow \quad & < w_1, w_2 > = 0. \end{aligned}$$

Das zeigt, dass (w_1, w_2) eine Orthonormalbasis ist mit $M_{\underline{w}}^{\underline{w}}(f) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

□

Satz 70 *Sei V ein 3-dimensionaler euklidischer Vektorraum und $f : V \to V$ eine orthogonale Abbildung.*

(a) Ist $\det(f) = 1$, so gibt es eine Orthonormalbasis $\underline{v} = (v_1, v_2, v_3)$, so dass die Darstellungsmatrix $A := M_{\underline{v}}^{\underline{v}}(f)$ die folgende Gestalt besitzt:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

für ein $\varphi \in [0, 2\pi)$ (d.h. f ist eine Drehung um den Winkel φ mit Drehachse $L(v_1)$).

(b) Ist $\det(f) = -1$*, so gibt es eine Orthonormalbasis* $\underline{v} = (v_1, v_2, v_3)$ *mit*

$$A = M_{\underline{v}}^{\underline{v}}(f) = \begin{pmatrix} \cos\varphi & -sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

für ein $\varphi \in [0, 2\pi)$ *(in diesem Fall ist* f *eine Drehspiegelung, d.h. eine Drehung um die Achse* $L(v_3)$ *um den Winkel* φ *und eine Spiegelung an der von* v_1 *und* v_2 *aufgespannten Ebene* $L(v_1, v_2)$*).*

Beweis. Wegen $\deg(\chi_f) = 3$ besitzt χ_f eine reelle Nullstelle λ (das folgt aus dem Zwischenwertsatz). Nach Lemma 64 (c) gilt $\lambda = \pm 1$.

1. Fall: $\det(f) = 1$ und $\lambda = 1$.

Sei $v_1 \in \mathrm{Eig}(f, 1)$ mit $\|v_1\| = 1$ und sei $U{:} = L(v_1)$. Wegen $f(v_1) = v_1$ gilt $f(U) = U$ und damit nach Lemma 68 auch $f(U^\perp) \subseteq U^\perp$. Sei $g : U^\perp \to U^\perp$ die Einschränkung von f auf $U^\perp$, d.h. $g(w){:} = f(w)$ für $w \in U^\perp$. Nach Folgerung 62 gilt $\dim(U^\perp) = \dim(V) - \dim(U) = 2$. Wir können also Satz 69 auf die orthogonale Abbildung g anwenden: Ist (v_2, v_3) eine Orthonormalbasis von $U^\perp$ so gibt es ein $\varphi \in [0, 2\pi)$, so dass für $B{:} = M_{(v_2,v_3)}^{(v_2,v_3)}(g)$ gilt:

$$B = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad \text{oder} \quad B = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

je nachdem ob $\det(B) = 1$ oder $\det(B) = -1$ gilt.

Sei $\underline{v}{:} = (v_1, v_2, v_3)$. Da $v_1 \perp v_i$ für $i = 2, 3$ ist $\underline{v}$ eine Orthonormalbasis von V und es gilt für $A = M_{\underline{v}}^{\underline{v}}(f)$:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & \sin\varphi & -\cos\varphi \end{pmatrix}$$

Da nach Voraussetzung $\det(f) = \det(A) = 1$ tritt der zweite Fall nicht auf.

1. Fall: $\det(f) = 1$ und $\lambda = -1$.

Wie oben wählen wir $v_1 \in \mathrm{Eig}(f, 1)$ mit $\|v_1\| = 1$. Wieder gilt $f(L(v_1)) = L(v_1)$ und folglich $f(L(v_1)^\perp) = L(v_1)^\perp$. Ist (w_2, w_3) eine Orthonormalbasis von $L(v_1)^\perp$ und $\underline{w}{:} = (v_1, w_2, w_3)$ so zeigt dieselbe Argumentation wie im ersten Fall, dass

$$M_{\underline{w}}^{\underline{w}}(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \quad \text{oder} \quad = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & \sin\varphi & -\cos\varphi \end{pmatrix}$$

Die Bedingung $\det(f) = 1$ impliziert, dass diesmal nur der zweite Fall auftritt. Nach Satz 69 (b) können wir aber eine zweite Orthonormalbasis (v_2, v_3) von $L(v_1)^\perp$ so wählen, dass für $\underline{v}\colon = (v_1, v_2, v_3)$ gilt

$$M^{(v_1,v_2,v_3)}_{(v_1,v_2,v_3)}(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Schlieslich gilt für $\underline{v}\colon = (v_3, v_1, v_2)$ und $\varphi\colon = \pi$:

$$M^{\underline{v}}_{\underline{v}}(f) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

In den übrigen Fällen ($\det(f) = -1$, $\lambda = 1$) und ($\det(f) = -1$, $\lambda = -1$) kann man ähnlich argumentieren. $\square$

Bemerkungen 71 (a) Für eine orthogonale Abbildung $f : V \to V$ auf einem euklidischen Vektorraum V beliebiger Dimension n gilt die folgende Verallgemeinerung der Sätze 69 und 70: Es gibt eine Orthonormalbasis $\underline{v} = (v_1, \dots, v_n)$, so dass die Darstellungsmatrix von f bzgl. $\underline{v}$ die folgende Gestalt besitzt:

$$M^{\underline{v}}_{\underline{v}}(f) = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& D(\varphi_1) && \\ &&&&&&& \ddots & \\ &&&&&&&& D(\varphi_r) \end{pmatrix}$$

mit

$$D(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

(b) Die entsprechende Aussage für eine orthogonale Matrix $A \in O(n)$ lautet:

Es gibt ein $T \in O(n)$ mit

$$T^t A T = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & D(\varphi_1) & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & D(\varphi_r) \end{pmatrix}$$

Symmetrische Matrizen und selbstadjungierter Endomorphismus
Sei V ein endlich-dimensionaler euklidischer Vektorraum mit Skalarprodukt $< \ , \ >$.

Definition 72 *Ein Endomorphismus $f : V \to V$ heisst selbstadjungiert, wenn gilt:*

$$< f(v), w >=< v, f(w) > \qquad \forall\ v, w \in V.$$

Satz 73 *Sei $f : V \to V$ ein Endomorphismus, $\underline{v} = (v_1, \ldots, v_n)$ eine Orthonormalbasis von V und $A := M_{\underline{v}}^{\underline{v}}(f)$. Folgende Bedingungen sind äquivalent:*

(i) f ist selbstadjungiert;

(ii) A ist symmetrisch, d.h. $A = A^t$.

Beweis. Wie im Beweis von Satz 65 gezeigt wurde, gilt für die Einträge a_{ij} von A: $a_{ij} =< v_i, f(v_j) >$. Ist f selbstadjungiert so folgt

$$a_{ij} =< v_i, f(v_j) >=< f(v_i), v_j >=< v_j, f(v_i) >= a_{ji}.$$

also $A^t = A$.

Sei umgekehrt $A = A^t$ und seien $v, w \in V$. Wir schreiben v und w als Linearkombination in der Basis $\underline{v}$:

$$v = \sum_{i=1}^{n} x_i v_i, \quad w = \sum_{j=1}^{n} y_j v_j$$

$$
\begin{aligned}
< f(v), w > &= < \sum_{i=1}^{n} x_i f(v_i), \sum_{j=1}^{n} y_j v_j > = \\
&= \sum_{i,j} x_i y_j < f(v_i), v_j > = \sum_{i,j} x_i y_j a_{ij} \\
&= \sum_{i,j} x_i y_j a_{ji} = \sum_{i,j} y_i x_j a_{ij} = < v, f(w) > .
\end{aligned}
$$

□

Wir bezeichnen mit $\mathrm{Sym}(n, \mathbb{R})$ die Teilmenge der symmetrischen $n \times n$ Matrizen in $M(n \times n, \mathbb{R})$.

Lemma 74 *Sei $A \in \mathrm{Sym}(n, \mathbb{R})$. Dann besitzt A einen reellen Eigenwert.*

Beweis. Sei $\partial\mathbb{B} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ (hier bezeichnet $\| \ldots \|_2$ die Norm bzgl. des Standardskalarprodukts $< \ , \ >$).

Wir benutzen ohne Beweis folgendes Resultat aus der Analysis:

Es gibt ein $v \in \partial\mathbb{B}$ mit

$$
< x, Ax > \leq < v, Av > =: \lambda \qquad \forall\ x \in \partial\mathbb{B} \tag{17}
$$

Aus (17) folgt:

$$
< x, Ax > \leq < v, Av >< x, x > = \lambda < x, x > \qquad \forall\ x \in \mathbb{R}^n
$$

Für $x \in \mathbb{R}^n$ und $t \in \mathbb{R}$ setze

$$
f(t) := < v + tx, A(v + tx) > - \lambda < v + tx, v + tx >
$$

f ist stetig und hat ein Maximum bei $t = 0$. Folglich gilt $f'(0) = 0 = f(0)$. Wegen

$$
\begin{aligned}
f(t) = &< v, Av > + 2t < Av, x > + t^2 < x, Ax > \\
&- \lambda(< v, v > + 2t < v, x > + t^2 < x, x >)
\end{aligned}
$$

gilt für die Ableitung $f'(0)$:

$$
0 = f'(0) = 2 < Av, x > - 2\lambda < v, x > = 2 < Av - \lambda v, x >
$$

d.h. $Av - \lambda v$ steht senkrecht auf allen Vektoren $x \in \mathbb{R}^n$.

$$\Rightarrow Av - \lambda v = 0.$$

$\square$

Folgerung 75 *Sei $f : V \to V$ selbstadjungiert. Dann besitzt f einen reellen Eigenwert.*

Satz 76 *(Hauptachsentransformation, Spektralsatz) Sei $f : V \to V$ selbstadjungiert. Dann gibt es eine Orthonormalbasis $\underline{v} = (v_1, \dots, v_n)$ von V so dass gilt:*

$$M_{\underline{v}}^{\underline{v}}(f) = \begin{pmatrix} \lambda_1 & & & \\ & & 0 & \\ & & \ddots & \\ & 0 & & \\ & & & \lambda_n \end{pmatrix}$$

Insbesondere ist f diagonalisierbar.

Zum Beweis benötigen wir noch folgendes einfache Lemma:

Lemma 77 *Sei $f : V \to V$ selbstadjungiert und $U \subseteq V$ ein Untervektorraum mit $f(U) \subseteq U$. Dann gilt: $f(U^\perp) \subseteq U^\perp$.*

Beweis von Satz 76. Durch Induktion nach $\dim(V) = n$.

Nach Folgerung 75 besitzt f einen Eigenwert $\lambda_1 \in \mathbb{R}$. Wir wähle $v_1 \in Eig(f, \lambda_1)$ normiert und setzten $U :\, = L(v_1)^\perp$. Wegen $f(L(v_1)) \subseteq L(v_1)$ gilt nach obigem Lemma $f(U^\perp) \subseteq U^\perp$. Ausserdem gilt: $\dim(U^\perp) = n - 1$. Nach der Induktionvoraussetzung (angewendet auf $f|_{U^\perp}$) besitzt $U^\perp$ eine Orthonormalbasis $(v_2, \dots, v_n)$ aus Eigenvektoren von f. Das n-Tupel $(v_1, v_2, \dots, v_n)$ ist daher eine Orthonormalbasis von V aus Eigenvektoren von f. $\square$

Folgerung 78 *Sei $A \in \mathrm{Sym}(n, \mathbb{R})$. Es gibt ein $T \in O(n)$ mit*

$$T^t A T = \begin{pmatrix} \lambda_1 & & & \\ & & 0 & \\ & & \ddots & \\ & 0 & & \\ & & & \lambda_n \end{pmatrix}.$$