22. Endomorphismen euklidischer und unitärer Räume

Im letzten Kapitel haben wir ausführlich Vektorräume studiert, auf denen die Zusatzstruktur eines Skalarprodukts gegeben ist. Wir wollen nun sehen, welche Vorteile und Vereinfachungen uns diese Zusatzstruktur bei der Untersuchung von Endomorphismen gibt, wenn diese in gewissem Sinne mit dem gegebenen Skalarprodukt verträglich sind. Das zentrale Resultat dieses Kapitels wird schließlich der sogenannte Spektralsatz in Abschnitt 22.C sein, der die Diagonalisierbarkeit solcher Endomorphismen garantiert und — wie wir sehen werden — so universell ist, dass er bei geschickter Anwendung auch Aussagen über Bilinearformen bzw. Sesquilinearformen und in Abschnitt 22.D sogar über lineare Abbildungen mit unterschiedlichem Start- und Zielraum machen kann.

22.A Orthogonale und unitäre Abbildungen

Die natürlichste Verträglichkeitsbedingung zwischen Morphismen und Skalarprodukten ist vermutlich die folgende.

Definition 22.1 (Orthogonale und unitäre Abbildungen und Matrizen)**.**

(a) Ein Endomorphismus $f\colon V\to V$ eines Vektorraums V mit Skalarprodukt heißt **orthogonal** (im Fall $\mathbb{K}=\mathbb{R}$) bzw. **unitär** (im Fall $\mathbb{K}=\mathbb{C}$), wenn

$$\langle f(x), f(y)\rangle = \langle x,y\rangle$$

für alle $x,y\in V$ gilt.

(b) Eine Matrix $A\in \mathrm{Mat}(n\times n,\mathbb{R})$ heißt **orthogonal**, wenn $A^{\mathsf{T}}A=E$ gilt, also wenn A invertierbar ist mit $A^{-1}=A^{\mathsf{T}}$. Wir bezeichnen die Menge aller orthogonalen $n\times n$-Matrizen mit $\mathrm{O}(n)\subset \mathrm{Mat}(n\times n,\mathbb{R})$.

Eine Matrix $A\in \mathrm{Mat}(n\times n,\mathbb{C})$ heißt **unitär**, wenn $\overline{A}^{\mathsf{T}}A=E$ gilt, also wenn A invertierbar ist mit $A^{-1}=\overline{A}^{\mathsf{T}}$. Die Menge aller unitären $n\times n$-Matrizen wird mit $\mathrm{U}(n)\subset \mathrm{Mat}(n\times n,\mathbb{C})$ bezeichnet.

Wie üblich schreiben wir diese Bedingung im Folgenden oft in beiden Fällen als $\overline{A}^{\mathsf{T}}A=E$.

Bemerkung 22.2.

(a) Ist $B=(x_1,\dots,x_n)$ eine Basis von V, so genügt es, die Bedingung eines orthogonalen bzw. unitären Morphismus $f\colon V\to V$ für alle Paare von Basisvektoren zu überprüfen: Ist nämlich $\big\langle f(x_i), f(x_j)\big\rangle = \big\langle x_i,x_j\big\rangle$ für alle $i,j=1,\dots,n$, so gilt wegen der Linearität von f und der Bilinearität bzw. Sesquilinearität des Skalarprodukts auch für alle $x=\lambda_1x_1+\dots+\lambda_nx_n$ und $y=\mu_1x_1+\dots+\mu_nx_n$

$$\langle f(x), f(y)\rangle = \sum_{i,j=1}^{n}\overline{\lambda_i}\mu_j\,\big\langle f(x_i), f(x_j)\big\rangle = \sum_{i,j=1}^{n}\overline{\lambda_i}\mu_j\,\big\langle x_i,x_j\big\rangle = \langle x,y\rangle\,.$$

(b) Eine Matrix $A=(a_{i,j})_{i,j}\in \mathrm{Mat}(n\times n,\mathbb{K})$ ist genau dann orthogonal bzw. unitär, wenn die Spalten von A eine Orthonormalbasis bezüglich des Standardskalarprodukts bilden: Nach Definition des Matrixprodukts ist nämlich

$$\overline{A}^{\mathsf{T}}A = \left(\sum_{j=1}^{n}\overline{a_{j,i}}\,a_{j,k}\right)_{i,k}.$$

Der Ausdruck $\sum_{j=1}^{n}\overline{a_{j,i}}\,a_{j,k}$ ist aber genau das Standardskalarprodukt der i-ten mit der k-ten Spalte von A. Daher bilden diese Spalten genau dann eine Orthonormalbasis bezüglich des

Standardskalarprodukts, wenn dieser Ausdruck gleich 1 für $i = k$ und 0 für $i \neq k$ ist, also wenn $\overline{A}^{\mathsf{T}} A = E$ ist.

Es wäre also vermutlich konsequenter, eine reelle Matrix A mit $A^{\mathsf{T}} A = E$ ortho*normal* statt ortho*gonal* zu nennen. Die Bezeichnung „orthogonale Matrix“ ist in der Literatur aber so üblich, dass wir hier nicht davon abweichen wollen.

Bemerkung 22.3 (Geometrische Deutung orthogonaler Abbildungen)**.** Nach Definition 22.1 erhalten orthogonale und unitäre Abbildungen Skalarprodukte, und damit auch Längen, Orthogonalität und (im reellen Fall) Winkel zwischen zwei Vektoren. Über $\mathbb{R}$ kann man sie sich daher als Drehungen, Spiegelungen und Kombinationen davon vorstellen (siehe auch Beispiel 22.7 und Aufgabe 22.12).

Wir wollen nun als Erstes zeigen, dass die oben eingeführten Bedingungen für orthogonale und unitäre Matrizen wie erwartet denen der zugehörigen Endomorphismen entsprechen, sofern es sich um Abbildungsmatrizen bezüglich einer Orthonormalbasis handelt. Die Interpretation als Drehungen bzw. Spiegelungen gilt damit also auch für orthogonale Matrizen. Hierfür benötigen wir zunächst ein kleines Hilfsresultat zur Berechnung von Abbildungs- und Basiswechselmatrizen in Vektorräumen mit Skalarprodukt.

Lemma 22.4. *Es seien V ein endlich erzeugter Vektorraum mit Skalarprodukt und $B = (x_1, \dots, x_n)$ eine Orthonormalbasis von V. Dann gilt:*

(a) *Ist $x \in V$ mit Koordinatendarstellung $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$ bezüglich B, so gilt $\lambda_i = \langle x_i, x \rangle$ für alle $i = 1, \dots, n$. Es ist also*

$$x = \langle x_1, x \rangle x_1 + \cdots + \langle x_n, x \rangle x_n$$

für alle $x \in V$.

(b) *Ist $f \colon V \to V$ ein Endomorphismus, so ist die Abbildungsmatrix von f bezüglich B gleich $A_f^B = \big(\langle x_i, f(x_j) \rangle \big)_{i,j}$.*

(c) *Ist $B' = (y_1, \dots, y_n)$ eine weitere Basis von V, so ist die zugehörige Basiswechselmatrix gleich $A^{B',B} = \big(\langle x_i, y_j \rangle \big)_{i,j}$.*

Beweis.

(a) Nehmen wir das Skalarprodukt von x_i mit der Gleichung $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$, so erhalten wir wegen $\langle x_i, x_j \rangle = 0$ und $\langle x_i, x_i \rangle = 1$ für $i \neq j$ sofort

$$\langle x_i, x \rangle = \lambda_1 \langle x_i, x_1 \rangle + \cdots + \lambda_n \langle x_i, x_n \rangle = \lambda_i.$$

(b) Nach Satz 16.23 ist der Eintrag in Zeile i und Spalte j von A_f^B genau die i-te Koordinate von $f(x_j)$ bezüglich B, nach (a) also $\langle x_i, f(x_j) \rangle$.

(c) Nach Definition 16.28 ist der Eintrag in Zeile i und Spalte j von $A^{B',B}$ genau die i-te Koordinate von y_j bezüglich B, nach (a) also $\langle x_i, y_j \rangle$. □

Folgerung 22.5. *Es sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt und B eine Orthonormalbasis von V. Dann gilt:*

(a) *Ein Endomorphismus $f \colon V \to V$ ist genau dann orthogonal bzw. unitär, wenn A_f^B orthogonal bzw. unitär ist.*

(b) *Eine weitere Basis B' von V ist genau dann auch eine Orthonormalbasis, wenn $A^{B',B}$ orthogonal bzw. unitär ist.*

Beweis. Es seien $B = (x_1, \dots, x_n)$ und $A = \big(\langle x_i, y_j \rangle\big)_{i,j}$ für Vektoren $y_1, \dots, y_n \in \mathbb{K}^n$. Dann gilt

$$\begin{aligned}
\overline{A}^{\mathsf{T}} A &= \left(\sum_{j=1}^{n} \overline{\langle x_j, y_i \rangle} \, \langle x_j, y_k \rangle \right)_{i,k} && \text{(Definition 16.5)} \\
&= \left(\left\langle \sum_{j=1}^{n} \langle x_j, y_i \rangle \, x_j, y_k \right\rangle \right)_{i,k} && \text{(Sesquilinearität des Skalarprodukts)} \\
&= (\langle y_i, y_k \rangle)_{i,k} \, . && \text{(Lemma 22.4 (a))}
\end{aligned}$$

(a) Setzen wir $y_i = f(x_i)$ für alle i und eine lineare Abbildung $f \colon V \to V$, so ist $A = A_f^B$ nach Lemma 22.4 (b). Damit ist diese Abbildungsmatrix genau dann orthogonal bzw. unitär, wenn $\big(\langle f(x_i), f(x_k) \rangle\big)_{i,k} = E = \big(\langle x_i, x_k \rangle\big)_{i,k}$ ist, also nach Bemerkung 22.2 (a) wenn f orthogonal bzw. unitär ist.

(b) Setzen wir $B' = (y_1, \dots, y_n)$, so ist $A = A^{B',B}$ nach Lemma 22.4 (c). Damit ist diese Basiswechselmatrix genau dann orthogonal bzw. unitär, wenn $\big(\langle y_i, y_k \rangle\big)_{i,k} = E$ ist, also wenn B' eine Orthonormalbasis ist. □

Bemerkung 22.6. Gemäß Definition 22.1 (b) sind orthogonale bzw. unitäre Matrizen invertierbar. Mit Folgerung 22.5 (a) bedeutet dies also, dass orthogonale bzw. unitäre Endomorphismen eines endlich erzeugten Vektorraums mit Skalarprodukt stets Isomorphismen sind.

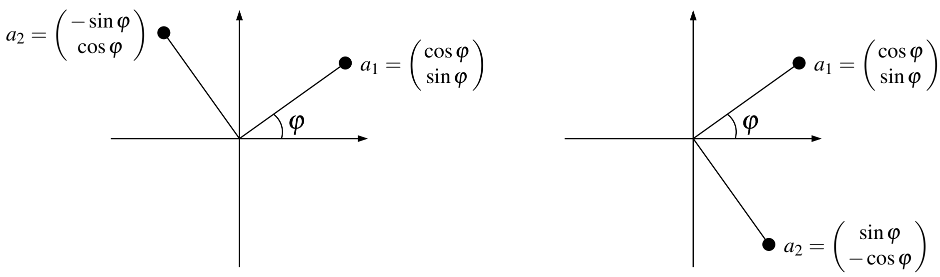
Beispiel 22.7 (O(2))**.** Nach Bemerkung 22.2 (b) ist eine Matrix $A = (a_1 \,|\, a_2) \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ mit $a_1, a_2 \in \mathbb{R}^2$ genau dann orthogonal, wenn bezüglich des Standardskalarprodukts

$$||a_1|| = 1, \quad ||a_2|| = 1 \quad \text{und} \quad a_1 \perp a_2$$

gilt. Also liegt a_1 auf dem Einheitskreis und lässt sich nach der Polarkoordinatendarstellung aus Satz 9.26 damit als

$$a_1 = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$

für ein $\varphi \in \mathbb{R}$ schreiben, und a_2 entsteht aus a_1 durch eine positive oder negative Drehung um $\frac{\pi}{2}$. Wir erhalten also die folgenden beiden Möglichkeiten:



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

In beiden Fällen werden die Einheitsvektoren e_1 und e_2 durch A auf a_1 und a_2 abgebildet. In (a) haben wir geometrisch also eine Drehung um den Winkel φ, in (b) eine Spiegelung an der Winkelhalbierenden zwischen e_1 und a_1. Damit lassen sich alle Elemente von O(2) in der Tat wie erwartet durch Drehungen und Spiegelungen beschreiben.

Lemma 22.8 (O(n) und U(n) als Gruppen)**.** *Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ orthogonal bzw. unitär, so auch AB und A^{-1}. Insbesondere sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ also Gruppen (siehe Definition 3.1) mit der Matrixmultiplikation als Verknüpfung; man nennt sie die* ***orthogonalen*** *bzw.* ***unitären Gruppen****.*

Beweis. Es gelte $A^{-1} = \overline{A}^{\mathsf{T}}$ und $B^{-1} = \overline{B}^{\mathsf{T}}$. Dann ist

$$\begin{aligned} (AB)^{-1} &= B^{-1}A^{-1} = \overline{B}^{\mathsf{T}}\overline{A}^{\mathsf{T}} = \overline{AB}^{\mathsf{T}} \\ \text{und} \quad (A^{-1})^{-1} &= (\overline{A}^{\mathsf{T}})^{-1} = \overline{A^{-1}}^{\mathsf{T}} \end{aligned}$$

nach Lemma 16.7 (d) und Lemma 16.20, d. h. AB und A^{-1} sind ebenfalls orthogonal bzw. unitär. □

Lemma 22.9 (Determinante und Eigenwerte orthogonaler und unitärer Matrizen)**.** *Für jede orthogonale oder unitäre Matrix A gilt:*

(a) $|\det A| = 1$.

(b) *Ist λ ein Eigenwert von A, so ist $|\lambda| = 1$.*

Beweis.

(a) Aus $\overline{A}^{\mathsf{T}}A = E$ ergibt sich sofort $1 = \det E = \det \overline{A}^{\mathsf{T}} \cdot \det A = \overline{\det A} \cdot \det A = |\det A|^2$.

(b) Gilt $Ax = \lambda x$ für ein $x \in \mathbb{K}^n \backslash \{0\}$, so ist auch $\overline{A}\overline{x} = \overline{\lambda}\overline{x}$, nach Transponieren also $\overline{x}^{\mathsf{T}}\overline{A}^{\mathsf{T}} = \overline{\lambda}\overline{x}^{\mathsf{T}}$. Zusammen erhalten wir wegen $\overline{A}^{\mathsf{T}}A = E$ so

$$\overline{x}^{\mathsf{T}}x = \overline{x}^{\mathsf{T}}\overline{A}^{\mathsf{T}}Ax = \overline{\lambda}\overline{x}^{\mathsf{T}} \cdot \lambda x = |\lambda|^2 \cdot \overline{x}^{\mathsf{T}}x,$$

mit $\overline{x}^{\mathsf{T}}x \neq 0$ also $|\lambda|^2 = 1$. □

Bemerkung 22.10. Nach Folgerung 22.5 (a) gelten die Aussagen aus Lemma 22.9 genauso für Endomorphismen endlich erzeugter Vektorräume mit Skalarprodukt.

Bemerkung 22.11 (Spezielle orthogonale und unitäre Matrizen)**.** Bei einer orthogonalen Matrix können die Determinante und die Eigenwerte nach Lemma 22.9 nur 1 und -1 sein. Für eine komplexe unitäre Matrix ist dagegen jede Zahl auf dem komplexen Einheitskreis möglich. In beiden Fällen spielen aber diejenigen Matrizen, deren Determinante gleich 1 ist, eine große Rolle. Man definiert daher:

(a) Eine orthogonale Matrix $A \in \mathrm{O}(n)$ heißt *spezielle orthogonale Matrix*, wenn $\det A = 1$ gilt. Die Menge der speziellen orthogonalen $n \times n$-Matrizen wird mit $\mathrm{SO}(n)$ bezeichnet.

(b) Eine unitäre Matrix $A \in \mathrm{U}(n)$ heißt *spezielle unitäre Matrix*, wenn $\det A = 1$ gilt. Man bezeichnet die Menge der speziellen unitären $n \times n$-Matrizen mit $\mathrm{SU}(n)$.

Man sieht leicht, dass auch $\mathrm{SO}(n)$ und $\mathrm{SU}(n)$ Gruppen bezüglich der Matrixmultiplikation sind, denn wenn $\det A = \det B = 1$ gilt, so ist nach Satz 18.6 ja auch $\det(AB) = \det A^{-1} = 1$ — sind also A und B speziell orthogonal bzw. unitär, so auch AB und A^{-1}.

Wir hatten in Bemerkung 22.3 ja schon gesehen, dass man sich die Elemente von $\mathrm{O}(n)$ als Drehungen bzw. Spiegelungen vorstellen kann. Dabei gilt:

- Drehungen entsprechen den Elementen von $\mathrm{SO}(n)$, da diese kontinuierlich aus der identischen Abbildung erzeugt werden können und ihre Determinante daher nicht von 1 auf -1 wechseln kann.
- Spiegelungen sind bezüglich einer geeigneten Orthonormalbasis $(x_1, \dots, x_n)$ von der Form $x_1 \mapsto -x_1$ und $x_i \mapsto x_i$ für $i \geq 2$, so dass die zugehörige Abbildungsmatrix $\mathrm{diag}(-1, 1, \dots, 1)$ ist und damit Determinante -1 hat. Damit entsprechen Spiegelungen Elementen von $\mathrm{O}(n) \backslash \mathrm{SO}(n)$.

Aufgabe 22.12 ($\mathrm{O}(3)$ und $\mathrm{O}(4)$)**.** Man zeige:

(a) Ist $A \in \mathrm{O}(3)$, so gibt es ein $\varphi \in \mathbb{R}$ und $T \in \mathrm{O}(3)$ mit

$$T^{-1}AT = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

(b) Ist $A \in \mathrm{O}(4)$, so gibt es im Allgemeinen *kein* $\varphi \in \mathbb{R}$ und $T \in \mathrm{O}(4)$ mit

$$T^{-1}AT = \begin{pmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & \pm 1 & 0 & 0 \\ 0 & 0 & \cos\varphi & -\sin\varphi \\ 0 & 0 & \sin\varphi & \cos\varphi \end{pmatrix}.$$

(Hinweis: Untersuche, ob A einen Eigenwert besitzen muss.)

Was bedeutet das Ergebnis geometrisch?

Aufgabe 22.13. Es sei f ein Endomorphismus eines euklidischen Vektorraums V. Man zeige: Gilt $||f(x)|| = ||x||$ für alle $x \in V$, erhält f also Längen von Vektoren, dann ist f bereits orthogonal.

Aufgabe 22.14. Es seien V ein euklidischer Vektorraum und $v \in V$ mit $||v|| = 1$. Wir betrachten den Endomorphismus $f\colon V \to V$, $x \mapsto x - 2\,\langle v, x\rangle\, v$.

(a) Zeige, dass f eine orthogonale Abbildung ist.

(b) Berechne das charakteristische Polynom χ_f.

(c) Wie kann man f geometrisch beschreiben?

Aufgabe 22.15 (QR-Zerlegung)**.** Zeige, dass sich jede reelle invertierbare Matrix $A \in \mathrm{GL}(n,\mathbb{R})$ eindeutig in der Form $A = QR$ schreiben lässt, wobei $Q \in \mathrm{O}(n)$ eine orthogonale Matrix und R eine obere Dreiecksmatrix mit positiven Diagonaleinträgen ist.

(Hinweis: Bezeichne die Spaltenvektoren von A bzw. Q mit $a_1, \dots, a_n \in \mathbb{R}^n$ bzw. $x_1, \dots, x_n \in \mathbb{R}^n$ und schreibe die Matrixgleichung $A = QR$ in Gleichungen für diese Spaltenvektoren um.)

Diese Zerlegung wird in der Vorlesung „Einführung in die Numerik" im zweiten Studienjahr noch eine wichtige Rolle spielen. Möchte man ein Gleichungssystem $Ax = b$ lösen, so ist dies mit der Zerlegung $A = QR$ wie oben natürlich äquivalent zum System $Rx = Q^{\mathsf{T}}b$, das sich wegen der oberen Diagonalform von R einfach durch Rückwärtseinsetzen lösen lässt. In der Regel ist dieses Verfahren für große Gleichungssysteme schneller und unanfälliger für eine Verstärkung von Rundungsfehlern als das euch bekannte Gauß-Verfahren. 52

22.B Selbstadjungierte Abbildungen

In diesem Abschnitt wollen wir eine weitere Art der Verträglichkeit eines Endomorphismus mit einem Skalarprodukt behandeln. Obwohl sie im Gegensatz zur Bedingung einer orthogonalen bzw. unitären Abbildung keine einfache geometrische Interpretation besitzt, ist sie dennoch in der Praxis sehr wichtig — vor allem, da sie einfach an der Abbildungsmatrix abzulesen ist. Um sie einzuführen, brauchen wir das folgende Konzept der adjungierten Abbildungen.

Satz und Definition 22.16 (Adjungierte Abbildung)**.** *Es sei $f\colon V \to V$ ein Endomorphismus eines endlich-dimensionalen Vektorraums V mit Skalarprodukt. Dann gilt:*

(a) *Es gibt genau einen Endomorphismus $f^*\colon V \to V$ mit*

$$\langle f(x), y\rangle = \langle x, f^*(y)\rangle$$

*für alle $x, y \in V$. Man nennt f^* die zu f **adjungierte Abbildung**.*

(b) *Ist $B = (x_1, \dots, x_n)$ eine Orthonormalbasis von V, so gilt für die Abbildungsmatrix von f^* bezüglich B*

$$A^B_{f^*} = \overline{A^B_f}^{\mathsf{T}}.$$

Beweis. Es seien $B = (x_1, \dots, x_n)$ eine Orthonormalbasis von V und $g\colon V \to V$ ein Endomorphismus. Beachte zunächst, dass mit demselben Argument wie in Bemerkung 22.2 (a) genau dann

$$\langle f(x), y\rangle = \langle x, g(y)\rangle \qquad (*)$$

für alle $x, y \in V$ ist, wenn $\langle f(x_i), x_j \rangle = \langle x_i, g(x_j) \rangle$ für alle $i, j = 1, \dots, n$ gilt. Wegen der Hermitizität des Skalarprodukts ist diese Bedingung nun aber äquivalent zu

$$(\overline{\langle x_j, f(x_i) \rangle})_{i,j} = (\langle x_i, g(x_j) \rangle)_{i,j},$$

nach Lemma 22.4 (b) also zu $\overline{A_f^B}^{\mathsf{T}} = A_g^B$. Also gibt es genau einen solchen Endomorphismus g, nämlich denjenigen, der bezüglich der Basis B zur Abbildungsmatrix $\overline{A_f^B}^{\mathsf{T}}$ gehört. Dies zeigt bereits beide Teile des Satzes. □

Bemerkung 22.17. Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{K})$.

(a) Aufgrund von Satz 22.16 (b) nennt man $\overline{A}^{\mathsf{T}}$ manchmal auch die zu A *adjungierte Matrix* und schreibt sie als A^*. Bezüglich einer Orthonormalbasis entsprechen sich dann also die Konzepte der adjungierten Abbildung und der adjungierten Matrix.

(b) Da offensichtlich $\overline{\overline{A}^{\mathsf{T}}}^{\mathsf{T}} = A$ gilt, ist nach Satz 22.16 (b) auch $(f^*)^* = f$ für jeden Endomorphismus f eines endlich erzeugten $\mathbb{K}$-Vektorraums mit Skalarprodukt. Dies bedeutet, dass nicht nur

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle, \quad \text{sondern auch} \quad \langle f^*(x), y \rangle = \langle x, (f^*)^*(y) \rangle = \langle x, f(y) \rangle$$

für alle $x, y \in V$ gilt: Man kann in einem Skalarprodukt stets die Abbildung f auf der einen Seite durch f^* auf der anderen Seite ersetzen (und umgekehrt).

Beispiel 22.18 (Adjungierte Abbildung einer orthogonalen bzw. unitären Abbildung)**.** Ist eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ orthogonal bzw. unitär, also $\overline{A}^{\mathsf{T}} A = E$, so ist $\overline{A}^{\mathsf{T}} = A^{-1}$ die adjungierte Matrix zu A. Für eine orthogonale oder unitäre Abbildung f eines endlich-dimensionalen Vektorraums mit Skalarprodukt ist dementsprechend also $f^* = f^{-1}$.

Die für diesen Abschnitt angekündigte Verträglichkeitsbedingung eines Endomorphismus mit einem Skalarprodukt besteht nun einfach darin, dass die adjungierte Abbildung mit der ursprünglichen übereinstimmt.

Definition 22.19 (Selbstadjungierte Abbildungen)**.** Ein Endomorphismus $f \colon V \to V$ eines endlich-dimensionalen Vektorraums V mit Skalarprodukt heißt **selbstadjungiert**, wenn $f^* = f$, also

$$\langle f(x), y \rangle = \langle x, f(y) \rangle$$

für alle $x, y \in V$ gilt.

Bemerkung 22.20 (Selbstadjungierte Matrizen)**.** Gemäß Satz 22.16 (b) ist ein Endomorphismus $f \colon V \to V$ genau dann selbstadjungiert, wenn die Abbildungsmatrix A_f^B bezüglich einer Orthonormalbasis B von V symmetrisch bzw. hermitesch ist, also $\overline{A_f^B}^{\mathsf{T}} = A_f^B$ gilt. Man bezeichnet symmetrische bzw. hermitesche Matrizen daher manchmal auch als *selbstadjungiert*.

Beispiel 22.21 (Selbstadjungiertheit orthogonaler Projektionen)**.** Es seien U ein Unterraum eines endlich erzeugten Vektorraums V mit Skalarprodukt und $f \colon V \to V$ die orthogonale Projektion auf U wie in Definition 21.29. Ist $(x_1, \dots, x_k)$ eine Orthonormalbasis von U, die wir zu einer Orthonormalbasis $B = (x_1, \dots, x_n)$ von V ergänzen (siehe Satz 21.31), so ist also

$$f(x_i) = x_i \text{ für } i = 1 \dots, k \quad \text{und} \quad f(x_i) = 0 \text{ für } i = k+1, \dots, n,$$

und damit $A_f^B = \mathrm{diag}(1, \dots, 1, 0, \dots, 0)$ (mit k Einträgen 1 und $n-k$ Einträgen 0) nach Satz 16.23. Da diese Matrix hermitesch ist, ist f nach Bemerkung 22.20 selbstadjungiert.

Beachte jedoch, dass orthogonale Projektionen (auch wenn man es vom Namen her anders vermuten könnte) nach Bemerkung 22.6 keine orthogonalen Abbildungen sind, da sie ja in der Regel nicht invertierbar sind, wie man an der obigen Form der Abbildungsmatrix ablesen kann.

Lemma 22.22 (Determinante und Eigenwerte hermitescher Matrizen)**.** *Für jede hermitesche Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ gilt:*

(a) $\det A \in \mathbb{R}$.

(b) *Ist λ ein Eigenwert von A, so ist $\lambda \in \mathbb{R}$.*

Beweis.

(a) Dies hatten wir bereits in Folgerung 21.34 (a) gesehen.

(b) Gilt $Ax = \lambda x$ für ein $x \in \mathbb{K}^n \backslash \{0\}$, so auch $\overline{A}\overline{x} = \overline{\lambda}\overline{x}$, und damit $\overline{x}^\mathsf{T}\overline{A}^\mathsf{T} = \overline{\lambda}\overline{x}^\mathsf{T}$. Also folgt

$$\lambda \overline{x}^\mathsf{T} x = \overline{x}^\mathsf{T} A x = \overline{x}^\mathsf{T}\overline{A}^\mathsf{T} x = \overline{\lambda}\overline{x}^\mathsf{T} x,$$

nach Division durch $\overline{x}^\mathsf{T} x \neq 0$ also $\lambda = \overline{\lambda}$ und damit $\lambda \in \mathbb{R}$. □

Um die weiteren Eigenschaften von orthogonalen, unitären und selbstadjungierten Abbildungen zu untersuchen, wenden wir nun einen kleinen Trick an: Es stellt sich heraus, dass alle diese Abbildungen Spezialfälle einer noch etwas allgemeineren Klasse von Abbildungen, den sogenannten *normalen Abbildungen*, sind, und dass man viele gemeinsame Eigenschaften orthogonaler, unitärer und selbstadjungierter Abbildungen auch für diese normalen Abbildungen zeigen kann — mit einem einzigen Beweis, der dann alle bisher betrachteten Fälle abdeckt.

Definition 22.23 (Normale Endomorphismen und Matrizen)**.**

(a) Ein Endomorphismus $f\colon V \to V$ eines endlich erzeugten Vektorraums V mit Skalarprodukt heißt **normal**, wenn $f^* \circ f = f \circ f^*$ gilt.

(b) Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ heißt **normal**, wenn $\overline{A}^\mathsf{T} A = A\overline{A}^\mathsf{T}$ gilt.

Bemerkung 22.24.

(a) Da die Verkettung von Morphismen der Matrixmultiplikation entspricht, ist ein Endomorphismus nach Satz 22.16 (b) genau dann normal, wenn seine Abbildungsmatrix bezüglich einer beliebigen Orthonormalbasis normal ist.

(b) Der Begriff eines normalen Endomorphismus ist ebenso wie der eines selbstadjungierten Endomorphismus ein theoretisches Konzept, für das es keine einfache geometrische Interpretation gibt. Auch sollte der Name „normal“ nicht darüber hinwegtäuschen, dass es für einen Endomorphismus durchaus nicht normal ist, normal zu sein: Die „meisten“ Endomorphismen sind dies nicht. Die speziellen Morphismen, die wir in diesem Kapitel betrachten wollen, haben diese Eigenschaft jedoch:

Beispiel 22.25.

(a) Eine Diagonalmatrix $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ ist normal: Wegen $\overline{A}^\mathsf{T} = \mathrm{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$ ist dann nämlich $\overline{A}^\mathsf{T} A = \mathrm{diag}(|\lambda_1|^2, \dots, |\lambda_n|^2) = A\overline{A}^\mathsf{T}$.

Beachte jedoch, dass diagonalisierbare Matrizen nicht notwendig normal sind: Für die diagonalisierbare Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{ist} \quad A^\mathsf{T} A = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{aber} \quad AA^\mathsf{T} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

(b) Ist $f\colon V \to V$ ein orthogonaler bzw. unitärer Endomorphismus eines endlich-dimensionalen Vektorraums V mit Skalarprodukt, so gilt $f^* = f^{-1}$ nach Beispiel 22.18, und daher auch

$$f^* \circ f = f^{-1} \circ f = \mathrm{id}_V = f \circ f^{-1} = f \circ f^*.$$

Ist f selbstadjungiert, also $f^* = f$, so gilt natürlich trivialerweise ebenfalls $f^* \circ f = f \circ f^*$.

Orthogonale, unitäre und selbstadjungierte Endomorphismen (und damit auch Matrizen) sind also normal.

Wir werden daher jetzt solche normalen Endomorphismen und Matrizen untersuchen — und können dann jedes Ergebnis dazu auf einen beliebigen der vorherigen Fälle von orthogonalen, unitären oder selbstadjungierten Endomorphismen anwenden. Wie immer beginnen wir mit den grundlegenden Eigenschaften solcher Morphismen.

Satz 22.26 (Eigenschaften normaler Endomorphismen)**.** *Es seien V ein endlich erzeugter Vektorraum mit Skalarprodukt und $f\colon V \to V$ ein normaler Endomorphismus. Dann gilt:*

(a) *Für alle $x \in V$ ist $||f^*(x)|| = ||f(x)||$. Insbesondere ist also $f^*(x) = 0$ genau dann wenn $f(x) = 0$, d. h. es gilt $\operatorname{Ker} f^* = \operatorname{Ker} f$.*

(b) *Ist x ein Eigenvektor von f zum Eigenwert λ, so ist x auch ein Eigenvektor von f^* zum Eigenwert $\overline{\lambda}$.*

(c) *Eigenvektoren von f zu verschiedenen Eigenwerten stehen senkrecht aufeinander.*

Beweis.

(a) Für $x \in V$ ist

$$||f^*(x)||^2 = \langle f^*(x), f^*(x)\rangle = \langle x, f(f^*(x))\rangle \overset{(*)}{=} \langle x, f^*(f(x))\rangle = \langle f(x), f(x)\rangle = ||f(x)||^2,$$

wobei wir in $(*)$ die Normalität von f ausgenutzt haben.

(b) Es sei $f(x) = \lambda x$. Dann ist

$$\begin{aligned}||f^*(x) - \overline{\lambda}x||^2 &= \left\langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \right\rangle \\ &= \langle f^*(x), f^*(x)\rangle - \lambda\,\langle x, f^*(x)\rangle - \overline{\lambda}\,\langle f^*(x), x\rangle + \lambda\overline{\lambda}\,\langle x,x\rangle \\ &\overset{\text{(a)}}{=} \langle f(x), f(x)\rangle - \lambda\,\langle f(x), x\rangle - \overline{\lambda}\,\langle x, f(x)\rangle + \lambda\overline{\lambda}\,\langle x,x\rangle \\ &= \langle f(x) - \lambda x, f(x) - \lambda x\rangle \\ &= 0\end{aligned}$$

und damit $f^*(x) - \overline{\lambda}x = 0$.

(c) Sind $x, y \in V$ und $f(x) = \lambda x$ sowie $f(y) = \mu y$ mit $\lambda \neq \mu$, so folgt

$$\lambda\,\langle x,y\rangle = \left\langle \overline{\lambda}x, y\right\rangle \overset{\text{(b)}}{=} \langle f^*(x), y\rangle = \langle x, f(y)\rangle = \langle x, \mu y\rangle = \mu\,\langle x,y\rangle\,,$$

also $(\lambda - \mu)\,\langle x,y\rangle = 0$. Wegen $\lambda \neq \mu$ ergibt sich daraus sofort $\langle x,y\rangle = 0$. □

22.C Der Spektralsatz

Wir kommen nun zum Hauptresultat dieses Kapitels. Es besagt, dass normale Endomorphismen mit zerfallendem charakteristischen Polynom stets diagonalisierbar sind, und zwar sogar mit einer Orthonormalbasis. Dieser Satz wird der Spektralsatz genannt, da die Menge aller Eigenwerte eines Endomorphismus oft als das Spektrum dieses Endomorphismus bezeichnet wird und dieser Satz durch die Diagonalisierbarkeit eine Aussage über diese Eigenwerte und ihre Vielfachheiten macht.

Satz 22.27 (**Spektralsatz**)**.** *Für einen Endomorphismus f eines endlich-dimensionalen Vektorraums V mit Skalarprodukt sind äquivalent:*

(a) *f ist normal und χ_f zerfällt in Linearfaktoren;*

(b) *V besitzt eine Orthonormalbasis aus Eigenvektoren von f.*

Insbesondere ist f dann nach Folgerung 19.29 also diagonalisierbar; man sagt für (b) auch, dass f ***orthogonal*** *bzw.* ***unitär diagonalisierbar*** *ist.*

Beweis.

(a) $\Rightarrow$ (b): Wir zeigen diese Richtung mit Induktion über $n := \dim V$; der Fall $n = 0$ ist trivial. Es sei nun also $n > 0$.

Da χ_f nach Voraussetzung in Linearfaktoren zerfällt, gibt es einen Eigenwert λ von f mit zugehörigem Eigenvektor x_1. Durch Normieren können wir diesen so wählen, dass $||x_1|| = 1$. Es sei $U = \operatorname{Lin}(x_1)$.

Wir ergänzen nun x_1 zu einer Orthonormalbasis $B = (x_1, x_2', \dots, x_n')$ von V, so dass nach Bemerkung 21.39 also $B' = (x_2', \dots, x_n')$ eine Orthonormalbasis von $U^\perp$ ist. Unser Ziel ist es, die Induktionsvoraussetzung auf die Einschränkung von f auf $U^\perp$ anzuwenden. Dazu müssen wir die folgenden Dinge überprüfen:

- $f(U^\perp) \subset U^\perp$, d. h. f lässt sich zu einem Endomorphismus $f|_{U^\perp} : U^\perp \to U^\perp$ einschränken: Für alle $x \in U^\perp$ gilt

$$\begin{aligned} \langle f(x), x_1 \rangle &= \langle x, f^*(x_1) \rangle && \text{(Satz 22.16 (a))} \\ &= \left\langle x, \overline{\lambda} x_1 \right\rangle && \text{(Satz 22.26 (b), da } f \text{ normal)} \\ &= \overline{\lambda}\, \langle x, x_1 \rangle \\ &= 0 && \text{(wegen } x \in U^\perp\text{),} \end{aligned}$$

  und damit auch $f(x) \in U^\perp$.

- Die Einschränkung $f' := f|_{U^\perp}$ erfüllt die Voraussetzung (a) des Satzes: Für die Abbildungsmatrix A_f^B gilt

$$A_f^B = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & A_{f'}^{B'} \end{array} \right)$$

  (wobei gemäß Satz 16.23 die erste Spalte aus $f(x_1) = \lambda x_1$ und die Nullen neben λ aus $f(U^\perp) \subset U^\perp$ folgen). Da f normal ist, ist nun nach Bemerkung 22.24 (a)

$$\overline{A_f^B}^{\mathsf{T}} A_f^B = A_f^B\, \overline{A_f^B}^{\mathsf{T}} \quad \Rightarrow \quad \left(\begin{array}{c|c} |\lambda|^2 & 0 \\ \hline 0 & \overline{A_{f'}^{B'}}^{\mathsf{T}} A_{f'}^{B'} \end{array} \right) = \left(\begin{array}{c|c} |\lambda|^2 & 0 \\ \hline 0 & A_{f'}^{B'}\, \overline{A_{f'}^{B'}}^{\mathsf{T}} \end{array} \right),$$

  also $A_{f'}^{B'}$ und damit f' normal; und wegen $\chi_f(t) = (t - \lambda)\, \chi_{f'}(t)$ zerfällt mit χ_f auch $\chi_{f'}$ in Linearfaktoren.

Nach der Induktionsvoraussetzung gibt es nun also eine Orthonormalbasis $(x_2, \dots, x_n)$ von $U^\perp$ aus Eigenvektoren von f (bzw. f'), so dass wie gewünscht $(x_1, \dots, x_n)$ eine Orthonormalbasis von V aus Eigenvektoren von f ist. 53

(b) $\Rightarrow$ (a): Es sei B eine Orthonormalbasis von V aus Eigenvektoren von f. Dann ist die Abbildungsmatrix $A = A_f^B$ nach Folgerung 19.29 eine Diagonalmatrix. Damit ist diese Matrix normal nach Beispiel 22.25 (a) und hat ein zerfallendes charakteristisches Polynom nach Folgerung 19.35 (a). □

Da der Spektralsatz in der Praxis sehr häufig gebraucht wird, schreiben wir ihn hier auch noch einmal in Matrixform auf:

Folgerung 22.28 (Spektralsatz in Matrixform)**.** *Für eine quadratische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ sind äquivalent:*

(a) *A ist normal und χ_A zerfällt in Linearfaktoren;*

(b) *es gibt eine orthogonale bzw. unitäre Matrix T, so dass*

$$T^{-1} A T = \overline{T}^{\mathsf{T}} A T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix (mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von A auf der Diagonalen) ist.

Man sagt für (b) wieder, dass A ***orthogonal*** *bzw.* ***unitär diagonalisierbar*** *ist.*

Beweis. Es sei $f = f_A\colon \mathbb{K}^n \to \mathbb{K}^n$, $x \mapsto Ax$ die zu A gehörige lineare Abbildung, so dass umgekehrt also A die Abbildungsmatrix von f bezüglich der Standardbasis von $\mathbb{K}^n$ ist. Wir versehen $\mathbb{K}^n$ mit dem Standardskalarprodukt.

Da die Standardbasis eine Orthonormalbasis für das Standardskalarprodukt ist, ist A nach Bemerkung 22.24 (a) genau dann normal mit in Linearfaktoren zerfallendem charakteristischen Polynom, wenn dies für f gilt. Nach dem Spektralsatz 22.27 ist dies wiederum genau dann der Fall, wenn $\mathbb{K}^n$ eine Orthonormalbasis $(x_1,\dots,x_n)$ aus Eigenvektoren von f bzw. A besitzt. Mit $T = (\,x_1 \mid \cdots \mid x_n\,)$ und $\lambda_1,\dots,\lambda_n$ den Eigenwerten zu $x_1,\dots,x_n$ ist diese Aussage nach Folgerung 19.29 und 22.5 (b) aber genau äquivalent dazu, dass T orthogonal bzw. unitär ist und $T^{-1}AT$ die Diagonalmatrix mit Einträgen $\lambda_1,\dots,\lambda_n$ ist. □

Beispiel 22.29. Die für uns wichtigsten Spezialfälle von Folgerung 22.28 sind:

(a) $\mathbb{K} = \mathbb{C}$ und A ist hermitesch oder unitär: Dann ist A nach Bemerkung 22.25 (b) normal, und wegen des Fundamentalsatzes der Algebra aus Satz 6.11 zerfällt χ_A in Linearfaktoren. Also ist A dann nach Folgerung 22.28 unitär diagonalisierbar.

(b) $\mathbb{K} = \mathbb{R}$ und A ist symmetrisch (also selbstadjungiert): Wie in (a) ist A auch dann wieder normal. Außerdem können wir A auch als komplexe hermitesche Matrix auffassen; das charakteristische Polynom zerfällt dann zumindest über $\mathbb{C}$ wieder in Linearfaktoren, d. h. es ist $\chi_A(t) = (t-\lambda_1)\cdots(t-\lambda_n)$ mit den Eigenwerten $\lambda_1,\dots,\lambda_n \in \mathbb{C}$ von A. Diese Eigenwerte sind nach Lemma 22.22 (b) aber reell, und damit zerfällt χ_A sogar über $\mathbb{R}$ in Linearfaktoren. Wir erhalten so den sicher wichtigsten Fall des Spektralsatzes:

$$\boxed{\text{Symmetrische reelle Matrizen sind stets orthogonal diagonalisierbar.}}$$

(c) $\mathbb{K} = \mathbb{R}$ und A ist orthogonal: Dann muss χ_A nicht notwendig in Linearfaktoren zerfallen — wie z. B. bei einer Drehung in $\mathbb{R}^2$ wie in Beispiel 19.12 (c). Nach Folgerung 19.39 ist A dann also auch nicht notwendig diagonalisierbar.

Beispiel 22.30. Die Berechnung der Matrix T in Folgerung 22.28 bzw. der Orthonormalbasis in Satz 22.27 ist mit unserem Vorwissen nicht mehr weiter kompliziert: Wir berechnen mit dem Verfahren von Gram-Schmidt aus Satz 21.31 Orthonormalbasen aller Eigenräume — besonders einfach ist dies natürlich für alle eindimensionalen Eigenräume, weil wir für diese nur jeweils einen normierten Eigenvektor benötigen. Nach Satz 22.26 (c) erhalten wir auf diese Art eine orthonormale Familie, und wegen der aus dem Spektralsatz folgenden Diagonalisierbarkeit sogar eine Orthonormalbasis. Als konkretes Beispiel sei

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \in \mathrm{Mat}(2\times 2,\mathbb{R}).$$

Die Matrix A ist offensichtlich reell symmetrisch und damit nach Beispiel 22.29 (b) orthogonal diagonalisierbar. Wie üblich (siehe Beispiel 19.15) berechnen wir dazu zuerst die Eigenwerte und -räume von A. Die Rechnung zeigt, dass A zwei Eigenwerte besitzt, und zwar

- $\lambda_1 = 3$ mit Eigenraum $\mathrm{Eig}(A,3) = \mathrm{Lin}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, also normiertem Eigenvektor $\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, und
- $\lambda_1 = 1$ mit Eigenraum $\mathrm{Eig}(A,1) = \mathrm{Lin}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, also normiertem Eigenvektor $\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

(wie man durch Rückwärtseinsetzen auch leicht bestätigen kann). Beachte, dass wir dabei für die Bestimmung der Orthonormalbasen das Standardskalarprodukt nehmen müssen und *nicht* die durch A bestimmte Bilinearform $b(x,y) = x^{\mathsf{T}}Ay$. Dies wird aus dem Beweis von Folgerung 22.28 deutlich; man sieht es aber auch schon daran, dass wir nach Bemerkung 22.2 (b) für eine orthogonale Matrix ja Spaltenvektoren brauchen, die *bezüglich des Standardskalarprodukts* orthonormal sind.

Schreiben wir diese Vektoren nun wie üblich als Spalten in eine Matrix, so erhalten wir also die orthogonale Matrix

$$T = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \in \mathrm{O}(2)$$

(wie man leicht sieht, bilden die Spalten in der Tat eine Orthonormalbasis von $\mathbb{R}^2$ bezüglich des Standardskalarprodukts), und es gilt dann

$$T^{-1}AT = T^{\mathsf{T}}AT = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix},$$

wie man auch leicht durch direkte Matrixmultiplikation (bzw. Inversenbildung) überprüfen kann.

Aufgabe 22.31 (Wurzelziehen für Matrizen)**.** Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine symmetrische Matrix. Man zeige:

(a) Ist A positiv semidefinit, so gibt es eine eindeutig bestimmte symmetrische positiv semidefinite Matrix $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ mit $B^2 = A$. Man nennt sie die *Wurzel* aus A.

(b) Ist A nicht positiv semidefinit, so kann es zwar eine Matrix $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ mit $B^2 = A$ geben, aber keine symmetrische.

Aufgabe 22.32. Es seien $n \in \mathbb{N}_{>0}$ und $A, B \in \mathrm{Mat}(n \times n, \mathbb{R})$ zwei symmetrische Matrizen. Man zeige:

(a) Die Menge

$$M_A := \left\{ \frac{x^{\mathsf{T}}Ax}{x^{\mathsf{T}}x} : x \in \mathbb{R}^n \backslash \{0\} \right\} \quad \subset \mathbb{R}$$

hat ein Maximum und Minimum, und zwar den größten bzw. kleinsten Eigenwert von A.

(b) Liegen alle Eigenwerte von A und B in den Intervallen $[a_1, a_2]$ bzw. $[b_1, b_2]$, so liegen alle Eigenwerte von $A + B$ im Intervall $[a_1 + b_1, a_2 + b_2]$.

Aufgabe 22.33. Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$. Man zeige:

(a) Ist A symmetrisch und nilpotent (siehe Aufgabe 20.25), so ist $A = 0$.

(b) Ist A diagonalisierbar und $A^{\mathsf{T}} = -A$, so ist $A = 0$.

Aufgabe 22.34. Es sei $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine symmetrische Matrix. Zeige, dass es eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ gibt mit $A^3 + A = B$.

In seiner ursprünglichen Form macht der Spektralsatz 22.27 eine Aussage über Endomorphismen. Wir wollen nun sehen, dass man ihn jedoch ebenso gewinnbringend auf Bilinearformen bzw. Sesquilinearformen anwenden kann — denn obwohl sich deren Gramsche Matrizen nach Bemerkung 21.8 zunächst einmal anders transformieren als die Abbildungsmatrizen von Endomorphismen (nämlich mit $A \mapsto \overline{T}^{\mathsf{T}}AT$ statt mit $A \mapsto T^{-1}AT$), stimmen diese beiden Transformationsregeln für eine orthogonale bzw. unitäre Matrix T ja überein.

Als Erstes wollen wir eine in der Praxis besonders wichtige Charakterisierung positiv definiter Matrizen (bzw. Bilinear- oder Sesquilinearformen) mit Hilfe von Eigenwerten angeben. Wir stellen sie im Folgenden noch einmal mit dem uns bereits aus Satz 21.42 bekannten Hurwitz-Kriterium zusammen und erweitern diese Aussagen auch auf negativ definite und indefinite Matrizen, da wir dies für die spätere Anwendung auf Extremwertuntersuchungen benötigen werden (siehe Satz 26.20).

Satz 22.35 (Kriterien für die Definitheit von Matrizen)**.** *Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine symmetrische (für $\mathbb{K} = \mathbb{R}$) bzw. hermitesche (für $\mathbb{K} = \mathbb{C}$) Matrix. Ferner sei $A_k \in \mathrm{Mat}(k \times k, \mathbb{K})$ für $k = 1, \dots, n$ die Matrix, die aus den ersten k Zeilen und Spalten von A besteht. Dann gilt:*

(a) ***(Eigenwertkriterium)***

A ist genau dann positiv definit / negativ definit / positiv semidefinit / negativ semidefinit, wenn $\lambda > 0$ / $\lambda < 0$ / $\lambda \geq 0$ / $\lambda \leq 0$ für jeden (nach Lemma 22.22 (b) reellen) Eigenwert λ von A gilt.

A ist genau dann indefinit, wenn A mindestens einen positiven und einen negativen Eigenwert besitzt.

(b) ***(Hurwitz-Kriterium)***

A ist genau dann positiv definit, wenn $\det A_k > 0$ *für alle k.*

A ist genau dann negativ definit, wenn $\det A_k > 0$ *für alle geraden und* $\det A_k < 0$ *für alle ungeraden k.*

Ist $\det A \neq 0$*, so ist A genau dann positiv bzw. negativ semidefinit, wenn A positiv bzw. negativ definit ist.*

Ist $\det A \neq 0$*, so ist A genau dann indefinit, wenn A weder positiv noch negativ definit ist, also wenn die Vorzeichenfolge von* $\det A_k$ *für* $k = 1,\dots,n$ *weder* $(+,+,+,\dots)$ *noch* $(-,+,-,+,\dots)$ *ist.*

Beweis.

(a) Nach der Folgerung 22.28 aus dem Spektralsatz gibt es eine orthogonale bzw. unitäre Matrix T mit $\overline{T}^{\mathsf{T}} A T = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$, wobei $\lambda_1,\dots,\lambda_n$ die Eigenwerte von A sind. Mit der invertierbaren Transformation $y := \overline{T}^{\mathsf{T}} x = T^{-1}x$, d. h. $x = Ty$, ist A also genau dann positiv definit, d. h. $\overline{x}^{\mathsf{T}} A x > 0$ für alle $x \in \mathbb{K}^n \backslash \{0\}$, wenn

$$\overline{y}^{\mathsf{T}}\overline{T}^{\mathsf{T}} A T y > 0 \quad \Leftrightarrow \quad \overline{y}^{\mathsf{T}} \operatorname{diag}(\lambda_1,\dots,\lambda_n) y > 0 \quad \Leftrightarrow \quad \lambda_1 |y_1|^2 + \cdots + \lambda_n |y_n|^2 > 0$$

für alle $y \neq 0$ gilt. Dies ist offensichtlich der Fall, wenn alle Eigenwerte $\lambda_1,\dots,\lambda_n$ positiv sind; umgekehrt folgt aus dieser Ungleichung durch Einsetzen des Einheitsvektors e_i aber auch sofort $\lambda_i > 0$ für alle i.

Die Aussagen zur negativen Definitheit und Semidefinitheit zeigt man analog. Die Matrix A ist indefinit, wenn sie weder negativ noch positiv semidefinit ist, also wenn sie mindestens einen positiven und einen negativen Eigenwert hat.

(b) Der Fall der positiven Definitheit ist genau Satz 21.42. Weiterhin ist A genau dann negativ definit, wenn $\overline{x}^{\mathsf{T}} A x < 0$ und damit $\overline{x}^{\mathsf{T}} (-A) x > 0$ für alle $x \in \mathbb{K}^n \backslash \{0\}$ gilt, also genau dann wenn $-A$ positiv definit ist. Anwenden von Satz 21.42 auf $-A$ ergibt in diesem Fall also wegen $\det(-A_k) = (-1)^k \det A_k$ die Behauptung.

Ist nun $\det A \neq 0$, so ist A invertierbar, d. h. es ist $\operatorname{Eig}(A,0) = \operatorname{Ker} A = \{0\}$ und damit 0 kein Eigenwert von A. Nach (a) ist A damit genau dann positiv bzw. negativ semidefinit, wenn A positiv bzw. negativ definit ist, und indefinit, wenn dies beides nicht der Fall ist. □

Bemerkung 22.36. Zur Bestimmung der Definitheitseigenschaften einer symmetrischen bzw. hermiteschen Matrix A mit Satz 22.35 ist das Hurwitz-Kriterium oftmals geeigneter, da die Berechnung von Determinanten einfacher ist als die aller Eigenwerte. Es liefert aber nicht in jedem Fall eine Entscheidung: Ist $\det A = 0$, so lässt sich mit dem Hurwitz-Kriterium in der Regel keine allgemeine Aussage treffen. Dies zeigt das Beispiel der drei Matrizen

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

die nach Satz 22.35 (a) positiv semidefinit, negativ semidefinit bzw. indefinit sind, für die aber alle Untermatrizen der ersten $k = 1,2,3$ Zeilen und Spalten die Determinante 0 haben.

Aufgabe 22.37. Es sei $A = (a_{i,j})_{i,j} \in \operatorname{Mat}(n \times n, \mathbb{R})$ eine symmetrische reelle Matrix mit

$$a_{i,i} > \sum_{j \neq i} |a_{i,j}|$$

für alle $i = 1,\dots,n$. Beweise, dass A dann positiv definit ist — symmetrische Matrizen, deren Diagonaleinträge im Vergleich zu den anderen „groß genug" sind, sind also positiv definit.

Als weitere Anwendung des Spektralsatzes auf symmetrische Bilinearformen bzw. Sesquilinearformen können wir den folgenden „Normalformensatz" zeigen, der die Aussage von Bemerkung 21.33 auf den nicht notwendig positiv definiten Fall verallgemeinert:

Satz 22.38 (**Trägheitssatz von Sylvester**). *Es seien V ein endlich erzeugter $\mathbb{K}$-Vektorraum und b eine symmetrische Bilinearform (für $\mathbb{K} = \mathbb{R}$) bzw. hermitesche Sesquilinearform (für $\mathbb{K} = \mathbb{C}$). Dann gibt es eine Basis B von V, bezüglich der die Gramsche Matrix von b die einfache Form*

$$A_b^B = \operatorname{diag}(\underbrace{1,\dots,1}_{k},\underbrace{-1,\dots,-1}_{l},0,\dots,0) = \left(\begin{array}{c|c|c} E_k & & 0 \\ \hline & -E_l & \\ \hline 0 & & 0 \end{array}\right)$$

hat. Dabei ist die Anzahl k bzw. l der Einträge 1 *bzw.* -1 *auf der Diagonalen durch b bereits eindeutig bestimmt, und zwar ist k bzw. l gleich*

(a) *der maximalen Dimension eines Unterraums von V, auf dem b positiv bzw. negativ definit ist; und*

(b) *der (mit Vielfachheiten gezählten) Anzahl der positiven bzw. negativen Eigenwerte einer beliebigen Gramschen Matrix zu b.*

Beweis. Wir teilen den Beweis in zwei Teile:

Teil 1: Als Erstes zeigen wir die Eindeutigkeit von k und l und den Ausdruck aus (a). Dazu betrachten wir eine beliebige Basis $B = (x_1,\dots,x_n)$ von V, für die $A_b^B = (b(x_i,x_j))_{i,j}$ die im Satz angegebene Form hat. Dann ist $U_+ := \operatorname{Lin}(x_1,\dots,x_k)$ sicher ein k-dimensionaler Unterraum von V, auf dem b positiv definit ist, denn die Gramsche Matrix der Einschränkung von b auf U_+ ist ja gerade die positiv definite Matrix E_k. Genauso sieht man natürlich, dass $U_- := \operatorname{Lin}(x_{k+1},\dots,x_n)$ ein $(n-k)$-dimensionaler Unterraum ist, auf dem b negativ semidefinit ist.

Ist nun andererseits $U \le V$ ein beliebiger Unterraum, auf dem b positiv definit ist, so ist mit dem eben gefundenen U_- sicher $U \cap U_- = \{0\}$, denn für ein $x \in U \cap U_-$ mit $x \neq 0$ ergäbe sich aus $b(x,x) > 0$ wegen $x \in U$ und $b(x,x) \le 0$ wegen $x \in U_-$ sofort ein Widerspruch. Mit der Dimensionsformel aus Folgerung 15.30 erhalten wir also

$$n \ge \dim(U + U_-) = \dim U + \dim U_- = \dim U + n - k$$

und damit $\dim U \le k$. Also ist k wirklich die maximale Dimension eines Unterraums von V, auf dem b positiv definit ist. Analog zeigt man natürlich die entsprechende Aussage für l. Wir haben damit also den Ausdruck für k und l in (a) bewiesen, und somit insbesondere auch die Eindeutigkeit der im Satz angegebenen Matrixdarstellung.

Teil 2: Wir zeigen nun (konstruktiv) die Existenz einer Basis B wie in der Behauptung und dabei den Ausdruck aus (b). Dazu sei zunächst B' eine beliebige Basis von V. Nach Lemma 21.11 bzw. Konstruktion 21.18 ist mit b auch die Gramsche Matrix $A := A_b^{B'}$ symmetrisch bzw. hermitesch. Wir gehen nun in zwei Schritten vor:

- Schritt 1: Drehung auf Diagonalform.

  Aus dem Spektralsatz folgt wie in Beispiel 22.29 die Existenz einer orthogonalen bzw. unitären Matrix T, so dass

  $$T^{-1} A T = \overline{T}^{\mathsf{T}} A T = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$$

  eine Diagonalmatrix mit den (nach Lemma 22.22 (b) reellen) Eigenwerten $\lambda_1,\dots,\lambda_n$ von A ist. Dabei wählen wir die Reihenfolge der Spalten von T und damit der Diagonaleinträge der Matrix so, dass die Eigenwerte $\lambda_1,\dots,\lambda_k$ positiv, $\lambda_{k+1},\dots,\lambda_{k+l}$ negativ, und $\lambda_{k+l+1},\dots,\lambda_n$ gleich 0 sind.

- Schritt 2: Koordinatenstreckung auf Normalform.

  Wir setzen nun

  $$S = \overline{S}^{\mathsf{T}} = \operatorname{diag}\left(\frac{1}{\sqrt{|\lambda_1|}},\dots,\frac{1}{\sqrt{|\lambda_{k+l}|}},1,\dots,1\right) \in \mathrm{GL}(n,\mathbb{K}).$$

Dann ist
$$\overline{TS}^{\mathsf{T}} A\,TS = \overline{S}^{\mathsf{T}} \overline{T}^{\mathsf{T}} A\,TS = \overline{S}^{\mathsf{T}} \cdot \operatorname{diag}(\lambda_1,\dots,\lambda_n) \cdot S = \operatorname{diag}(1,\dots,1,-1,\dots,-1,0,\dots,0).$$

Nach Satz 16.32 (b) gibt es nun aber eine Basis B, so dass die Basiswechselmatrix $A^{B,B'}$ gleich TS ist. Mit der Transformationsregel für Bilinearformen aus Lemma 21.7 (bzw. für Sesquilinearformen aus Konstruktion 21.18) hat die Matrix A_b^B dann die gewünschte Form, wobei k und l wie in (b)
54 sind. □

Bemerkung 22.39 (Trägheitssatz in Matrixform)**.** Natürlich gibt es auch vom Trägheitssatz 22.38 eine Matrixform: Für jede reelle symmetrische bzw. komplexe hermitesche Matrix $A \in \operatorname{Mat}(n \times n, \mathbb{K})$ gibt es eine invertierbare Matrix T, so dass $\overline{T}^{\mathsf{T}} A\,T$ die in Satz 22.38 angegebene Gestalt hat. Darüber hinaus sind auch hier dann die Anzahlen k und l der Diagonaleinträge 1 bzw. -1 eindeutig bestimmt und gleich der maximalen Dimension eines Unterraums von $\mathbb{K}^n$, auf dem A positiv bzw. negativ definit ist, sowie gleich der Anzahl der positiven bzw. negativen Eigenwerte von $\overline{S}^{\mathsf{T}} A\,S$ für eine beliebige invertierbare Matrix $S \in \operatorname{GL}(n,\mathbb{K})$. Mit den uns bekannten Entsprechungen zwischen Sesquilinearformen und Matrizen ergeben sich diese Aussagen unmittelbar aus Satz 22.38 angewendet auf die Sesquilinearform $b(x,y) = \overline{x}^{\mathsf{T}} A\,y$ auf $\mathbb{K}^n$.

Insbesondere bedeutet dies, dass die Matrizen A und $\overline{S}^{\mathsf{T}} A\,S$ für alle $S \in \operatorname{GL}(n,\mathbb{K})$ die gleiche Anzahl positiver (und analog negativer) Eigenwerte haben — obwohl die Eigenwerte selbst ja nicht übereinstimmen, da A und $\overline{S}^{\mathsf{T}} A\,S$ im Allgemeinen nicht ähnlich zueinander sind. Diese Aussage, die aus unserem Satz 22.38 folgt, wird in der Literatur auch oft Trägheitssatz von Sylvester genannt. Aus ihr leitet sich auch der Name „Trägheitssatz" ab: Der Satz zeigt, dass sich die Anzahlen der positiven und negativen Eigenwerte einer hermiteschen Matrix unter Transformationen der Form $A \mapsto \overline{S}^{\mathsf{T}} A\,S$ nicht ändern, sich also in diesem Sinne „träge" verhalten.

Die in Teil 2 vom Beweis des Trägheitssatzes 22.38 konstruierte Transformation, um eine gegebene reelle symmetrische Bilinearform in die dort angegebene Normalform zu bringen, hat auch eine einfache geometrische Bedeutung. Der Einfachheit halber betrachten wir diese zunächst im positiv definiten Fall, also wenn alle Diagonaleinträge der transformierten Matrix gleich 1 sind.

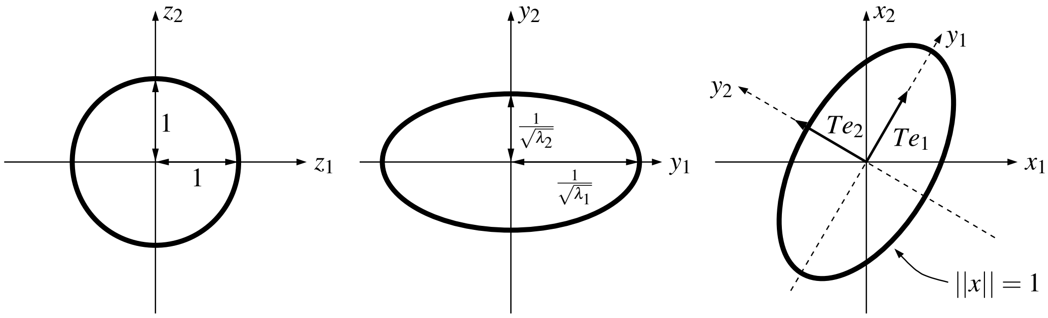
Beispiel 22.40 (Hauptachsentransformation). Es sei b ein Skalarprodukt auf $\mathbb{R}^n$ mit Gramscher Matrix $A \in \operatorname{Mat}(n \times n, \mathbb{R})$. Um dieses Skalarprodukt zu visualisieren, können wir die nach Definition 21.13 zugehörige Norm $||x|| = \sqrt{b(x,x)} = \sqrt{x^{\mathsf{T}} A x}$ betrachten und die Menge
$$K := \{x \in \mathbb{R}^n : ||x|| = 1\} = \{x \in \mathbb{R}^n : x^{\mathsf{T}} A x = 1\}$$
aller Punkte mit Norm 1 zeichnen, also den „Rand der Einheitskugel" bezüglich b. Um K geometrisch zu beschreiben, führen wir die zwei Schritte in Teil 2 des Beweises von Satz 22.38 durch:

- Drehung auf Diagonalform: Wir finden nach dem Spektralsatz eine orthogonale Matrix T, so dass $T^{\mathsf{T}} A\,T = \operatorname{diag}(\lambda_1,\dots,\lambda_n) =: D$ eine Diagonalmatrix ist. Als Eigenwerte von A sind diese $\lambda_1,\dots,\lambda_n$ nach Satz 22.35 (a) positiv.

  Wir machen nun die Koordinatentransformation $y := T^{-1}x = T^{\mathsf{T}} x$, also $x = Ty$, die wir uns wegen $T \in \mathrm{O}(n)$ als Drehung (bzw. Spiegelung) in $\mathbb{R}^n$ vorstellen können. Mit diesen neuen Koordinaten ist
  $$x^{\mathsf{T}} A x = 1 \quad \Leftrightarrow \quad y^{\mathsf{T}} T^{\mathsf{T}} A\,T y = 1 \quad \Leftrightarrow \quad y^{\mathsf{T}} D y = 1 \quad \Leftrightarrow \quad \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 = 1.$$

- Koordinatenstreckung auf Normalform: Wegen $\lambda_i > 0$ für alle i können wir nun noch die weitere Koordinatentransformation $z_i := \sqrt{\lambda_i}\, y_i$, also $y_i = \frac{1}{\sqrt{\lambda_i}} z_i$ für $i = 1,\dots,n$ durchführen, die einer Streckung der Koordinatenachsen (mit unterschiedlichen Streckfaktoren) entspricht. In diesen neuen Koordinaten ist nun einfach
  $$x^{\mathsf{T}} A x = 1 \quad \Leftrightarrow \quad \lambda_1 y_1^2 + \dots + \lambda_n y_n^2 = 1 \quad \Leftrightarrow \quad z_1^2 + \dots + z_n^2 = 1,$$
  d. h. hier bekommen wir nun den Rand der „gewöhnlichen Einheitskugel".

Unsere ursprüngliche Menge K ensteht also aus dem Rand der normalen Einheitskugel (in den Koordinaten z_i), indem wir zuerst die einzelnen Koordinaten strecken (Übergang von den z_i zu den y_i) und das resultierende Ellipsoid dann im Raum drehen (Übergang von den y_i zu den x_i). Die gesuchte Menge K ist also wie im folgenden Bild für $n = 2$ dargestellt ein im Raum gedrehtes Ellipsoid.

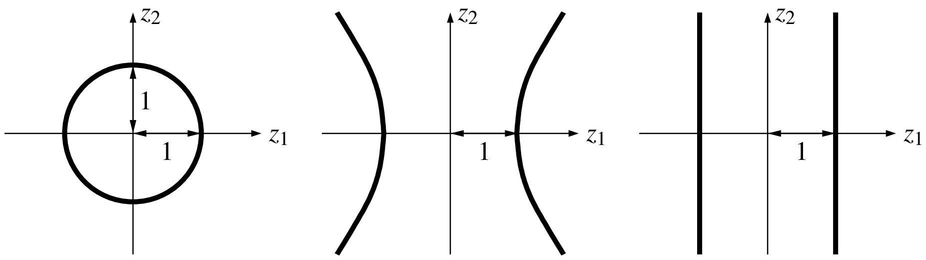


Wie im Bild angedeutet sind die charakteristischen Merkmale dieses Ellipsoids:

- Seine Radien sind gerade $\frac{1}{\sqrt{\lambda_i}}$, denn dies sind die y_i-Werte, die $z_i = 1$ entsprechen;
- Seine Symmetrieachsen werden aufgespannt von den x-Vektoren, die im y-Koordinatensystem den Einheitsvektoren entsprechen — wegen $x = Ty$ also von den Spalten $Te_1, \dots, Te_n$ von T und damit genau von den Eigenvektoren von A.

Die Symmetrieachsen eines Ellipsoids werden auch als seine **Hauptachsen** bezeichnet — daher der Name Hauptachsentransformation.

Bemerkung 22.41. Untersuchen wir die Menge $K = \{x \in \mathbb{R}^n : x^\mathsf{T} A x = 1\}$ wie in Beispiel 22.40 für eine symmetrische Matrix A, die nicht mehr notwendig positiv definit ist, so können wir immer noch das Verfahren aus Teil 2 des Beweises von Satz 22.38 anwenden, erhalten jedoch am Ende in den Koordinaten z_i eine quadratische Gleichung, deren Koeffizienten 1, -1 und 0 sein können (je nachdem, wie viele Eigenwerte von A positiv, negativ bzw. 0 sind). Bis auf Permutation dieser Variablen erhalten wir so z. B. für $n = 2$ als Möglichkeiten für ein nicht-leeres K wie im folgenden Bild neben einem Kreis (bei zwei positiven Eigenwerten) eine *Hyperbel* (bei einem positiven und einem negativen Eigenwert) und ein *Geradenpaar* (bei einem positiven Eigenwert und einem Eigenwert 0):



$z_1^2 + z_2^2 = 1$ Kreis | $z_1^2 - z_2^2 = 1$ Hyperbel | $z_1^2 = 1$ Geradenpaar

Wie in Beispiel 22.40 ist die ursprüngliche Menge K dann eine in den Koordinatenrichtungen gestreckte und anschließend gedrehte Variante dieser Bilder. Hat A keinen positiven Eigenwert, so ist $K = \emptyset$, da eine Summe von Quadraten mit nicht-positiven Vorfaktoren niemals 1 ergeben kann. Für $n > 2$ gibt es natürlich entsprechend mehr qualitativ verschiedene Möglichkeiten für K.

Aufgabe 22.42. Die symmetrische reelle Matrix $A = \begin{pmatrix} 4 & -2 & -2 \\ -2 & 7 & 4 \\ -2 & 4 & 7 \end{pmatrix}$ hat genau die beiden Eigenwerte 3 und 12. Da alle Eigenwerte positiv sind, ist A nach Satz 22.35 (b) also positiv definit und bestimmt somit ein Skalarprodukt auf $\mathbb{R}^3$.

(a) Berechne eine orthogonale Matrix T, so dass $T^{-1}AT$ eine Diagonalmatrix ist.

(b) Bestimme mit einer Hauptachsentransformation die Punkte der Menge $\{x \in \mathbb{R}^3 : x^{\mathsf{T}}Ax = 3\}$, die vom Ursprung (bezüglich des Standardskalarprodukts) den kleinsten Abstand haben.

Aufgabe 22.43. Man zeige:

(a) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ normal, so ist A genau dann hermitesch, wenn alle Eigenwerte von A reell sind.

(b) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch mit $x^{\mathsf{T}}Ax \le x^{\mathsf{T}}Bx$ für alle $x \in \mathbb{R}^n$, so hat B mindestens so viele positive Eigenwerte wie A (mit ihrer jeweiligen algebraischen Vielfachheit gezählt).

Aufgabe 22.44. Es sei b eine Bilinearform auf einem endlich-dimensionalen Vektorraum V über einem Körper K, in dem $1+1 \neq 0$ gilt. Wir nennen b *antisymmetrisch*, wenn $b(x,y) = -b(y,x)$ für alle $x, y \in V$ gilt. Eine quadratische Matrix $A \in \mathrm{Mat}(n \times n, K)$ heißt *antisymmetrisch*, wenn $A^{\mathsf{T}} = -A$.

(a) Zeige, dass b genau dann antisymmetrisch ist, wenn A_b^B für eine beliebige Basis B von V antisymmetrisch ist.

(b) Man zeige mit Induktion über $n = \dim V$: Ist b antisymmetrisch, so ist A_b^B für eine geeignete Basis B von V von der Form

$$A_b^B = \begin{pmatrix} \boxed{I} & & & 0 \\ & \ddots & & \\ & & \boxed{I} & \\ 0 & & & 0 \end{pmatrix}$$

mit $I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, ist also eine Blockdiagonalmatrix mit einer gewissen Anzahl k von Blöcken I (mit $0 \le 2k \le n$) und $n-2k$ anschließenden Nullzeilen und -spalten.

(c) Zeige, dass die Determinante jeder ganzzahligen antisymmetrischen Matrix eine Quadratzahl ist.

22.D Die Singulärwertzerlegung

Als Abschluss der linearen Algebra wollen wir nun noch sehen, dass die Eigenwerttheorie und der Spektralsatz überraschenderweise sogar auf beliebige lineare Abbildungen mit unterschiedlichem Start- und Zielraum angewendet werden können — und damit eine Brücke zurück zu Kapitel 16 schlagen, in dem wir derartige Abbildungen untersucht haben.

Dazu erinnern wir uns an unseren ersten Normalformensatz 16.43: Zu einem Morphismus $f : V \to W$ zwischen endlich-dimensionalen Vektorräumen V und W gibt es stets Basen B und C von V bzw. W, so dass die zugehörige Abbildungsmatrix die besonders einfache Form

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

hat. Wir suchen nun nach einer analogen Aussage, bei der V und W Vektorräume mit Skalarprodukt sind und wir nur *Orthonormalbasen* B und C zulassen wollen. Da uns dies mehr einschränkt, erwarten wir natürlich, dass die zugehörige Normalform dann nicht mehr ganz so einfach sein wird. Wir werden allerdings sehen, dass es genügt, statt der Einsen auf der „Diagonalen“ der Abbildungsmatrix beliebige positive reelle Zahlen zuzulassen. Daher führen wir für derartige Matrizen zunächst eine einfache Notation ein:

Notation 22.45 (Nichtquadratische Diagonalmatrizen)**.** Im Folgenden wollen wir auch eine nicht notwendig quadratische Matrix $D = (d_{i,j})_{i,j} \in \mathrm{Mat}(m \times n, K)$ als *Diagonalmatrix* bezeichnen, wenn

$d_{i,j} = 0$ für alle $i \neq j$. Analog zu Definition 19.28 schreiben wir für $\lambda_1, \ldots, \lambda_r \in K$ mit $r \leq \min(m,n)$

$$D = \operatorname{diag}(\lambda_1, \ldots, \lambda_r) := \left(\begin{array}{ccc|c} \lambda_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \lambda_r & \\ \hline & 0 & & 0 \end{array}\right) \in \operatorname{Mat}(m \times n, K)$$

(wobei die Größe der Matrix aus dieser Schreibweise nicht ersichtlich ist und aus dem Zusammenhang klar sein muss). Oft verlangen wir dabei, dass $\lambda_1, \ldots, \lambda_r$ ungleich 0 sind, in diesem Fall ist dann natürlich $r = \operatorname{rk} D$.

Beachte, dass auch diese Diagonalmatrizen wie erwartet transponiert und multipliziert werden können: Ist $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_r) \in \operatorname{Mat}(m \times n, K)$, so ist z. B. auch

$$D^{\mathsf{T}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_r) \in \operatorname{Mat}(n \times m, K) \quad \text{und} \quad D^{\mathsf{T}} D = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_r^2) \in \operatorname{Mat}(n \times n, K).$$

Damit können wir nun den angekündigten Normalformensatz beweisen, der in der Literatur unter dem Namen *Singulärwertzerlegung* bekannt ist. Für den Beweis ist es praktisch, zunächst die Matrixform dieses Satzes zu betrachten.

Satz 22.46 (Singulärwertzerlegung)**.** *Zu jeder Matrix $A \in \operatorname{Mat}(m \times n, \mathbb{K})$ gibt es orthogonale bzw. unitäre Matrizen $S \in \operatorname{Mat}(m \times m, \mathbb{K})$ und $T \in \operatorname{Mat}(n \times n, \mathbb{K})$, so dass*

$$S^{-1} A T = \overline{S}^{\mathsf{T}} A T = \operatorname{diag}(\lambda_1, \ldots, \lambda_r) =: D \quad \in \operatorname{Mat}(m \times n, \mathbb{K})$$

eine Diagonalmatrix mit reellen positiven $\lambda_1, \ldots, \lambda_r$ und $r = \operatorname{rk} A$ ist.

*Die $\lambda_1, \ldots, \lambda_r$ sind dabei bis auf die Reihenfolge eindeutig bestimmt und heißen die **Singulärwerte** von A. Man nennt die Produktdarstellung $A = SDT^{-1} = SD\overline{T}^{\mathsf{T}}$ die **Singulärwertzerlegung** von A.*

Beweis. Wir zeigen zunächst die Eindeutigkeit der Singulärwerte: Sind S und T beliebige orthogonale bzw. unitäre Matrizen der passenden Größen, so dass $\overline{S}^{\mathsf{T}} A T = \operatorname{diag}(\lambda_1, \ldots, \lambda_r) =: D$ eine Diagonalmatrix mit reellen positiven $\lambda_1, \ldots, \lambda_r$ ist, so folgt durch Transponieren und Konjugieren auch $\overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} S = \overline{D}^{\mathsf{T}} = D^{\mathsf{T}}$, und damit

$$T^{-1} \overline{A}^{\mathsf{T}} A T = \overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} A T = \overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} S \overline{S}^{\mathsf{T}} A T = D^{\mathsf{T}} D = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_r^2) \quad \in \operatorname{Mat}(n \times n, \mathbb{K}).$$

Die Matrix $\overline{A}^{\mathsf{T}} A$ ist also ähnlich zur Diagonalmatrix $D^{\mathsf{T}} D$ und hat damit dieselben Eigenwerte. Die $\lambda_1^2, \ldots, \lambda_r^2$ (und damit auch $\lambda_1, \ldots, \lambda_r$) sind daher bis auf ihre Reihenfolge eindeutig bestimmt, nämlich als die Eigenwerte von $\overline{A}^{\mathsf{T}} A$ ungleich 0.

Der Existenzbeweis der Singulärwertzerlegung ist konstruktiv und erfolgt in zwei Schritten:

(a) Bestimmung von T.

Die Matrix $\overline{A}^{\mathsf{T}} A \in \operatorname{Mat}(n \times n, \mathbb{K})$ ist symmetrisch bzw. hermitesch (da $\overline{\overline{A}^{\mathsf{T}} A}^{\mathsf{T}} = \overline{A}^{\mathsf{T}} A$) und positiv semidefinit (da $\overline{x}^{\mathsf{T}} \overline{A}^{\mathsf{T}} A x = \overline{Ax}^{\mathsf{T}} A x = ||Ax||^2 \geq 0$ für alle $x \in \mathbb{K}^n$, wobei $|| \cdot ||$ die Norm des Standardskalarprodukts ist). Nach dem Spektralsatz gibt es wie in Beispiel 22.29 also eine orthogonale bzw. unitäre Matrix T, so dass $\overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} A T$ eine Diagonalmatrix ist. Ihre Diagonaleinträge, also die Eigenwerte von $\overline{A}^{\mathsf{T}} A$, liegen nach Satz 22.35 (a) in $\mathbb{R}_{\geq 0}$. Wir können (nach geeigneter Anordnung der Spalten von T) also

$$\overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} A T = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_r^2)$$

mit reellen $\lambda_1, \ldots, \lambda_r > 0$ für ein $r \leq n$ schreiben.

(b) Bestimmung von S.

Für alle $i \leq r$ setzen wir

$$s_i := \frac{1}{\lambda_i} A T e_i \quad \in \mathbb{K}^m.$$

Diese Vektoren sind bezüglich des Standardskalarprodukts orthonormal, denn für alle $i, j \leq r$ gilt

$$\overline{s_i}^{\mathsf{T}} s_j = \frac{1}{\lambda_i \lambda_j} \overline{e_i}^{\mathsf{T}} \overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} A T e_j \overset{(a)}{=} \frac{1}{\lambda_i \lambda_j} \overline{e_i}^{\mathsf{T}} \cdot \operatorname{diag}(\lambda_1^2, \dots, \lambda_r^2) \cdot e_j = \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Wir können sie nach Satz 21.31 also zu einer Orthonormalbasis von $\mathbb{K}^m$ ergänzen (insbesondere ist damit auch $r \leq m$) und daraus die orthogonale bzw. unitäre Matrix $S := (s_1 \,|\, \cdots \,|\, s_m)$ bilden.

Wir setzen nun $D := \operatorname{diag}(\lambda_1, \dots, \lambda_r) \in \operatorname{Mat}(m \times n, \mathbb{K})$ und behaupten, dass damit wie gewünscht $\overline{S}^{\mathsf{T}} A T = D$, also $AT = SD \in \operatorname{Mat}(m \times n, \mathbb{K})$ gilt. Die Gleichheit dieser Matrizen überprüfen wir spaltenweise, d. h. wir zeigen $A T e_i = S D e_i$ für alle $i = 1, \dots, n$:

- Für $i \leq r$ gilt $A T e_i = s_i \lambda_i = S D e_i$ nach (b).
- Für $i > r$ ist

$$||A T e_i||^2 = \overline{e_i}^{\mathsf{T}} \overline{T}^{\mathsf{T}} \overline{A}^{\mathsf{T}} A T e_i \overset{(a)}{=} \overline{e_i}^{\mathsf{T}} \cdot \operatorname{diag}(\lambda_1^2, \dots, \lambda_r^2) \cdot e_i = 0$$

  und damit $A T e_i = 0 = S D e_i$, da die i-te Spalte $D e_i$ von D eine Nullspalte ist.

Mit $\overline{S}^{\mathsf{T}} A T = D$ sind A und D nun insbesondere auch äquivalent, so dass $r = \operatorname{rk} D = \operatorname{rk} A$ nach Lemma 16.42 gilt. Damit ist alles gezeigt. □

Bemerkung 22.47 (Singulärwertzerlegung für Morphismen)**.** Wie am Anfang dieses Abschnitts erwähnt gibt es natürlich auch eine Variante der Singulärwertzerlegung für Morphismen: Ist $f : V \to W$ eine lineare Abbildung zwischen zwei endlich-dimensionalen Vektorräumen V und W mit Skalarprodukt, so gibt es Orthonormalbasen B und C von V bzw. W, so dass die Abbildungsmatrix $A_f^{B,C}$ diagonal mit reellen nicht-negativen Diagonaleinträgen ist. Dies ergibt sich unmittelbar aus Satz 22.46 angewendet auf eine Abbildungsmatrix von f zu beliebigen Orthonormalbasen.

Beispiel 22.48.

(a) Ist $A \in \operatorname{Mat}(n \times n, \mathbb{K})$ quadratisch, symmetrisch bzw. hermitesch und positiv semidefinit, so finden wir nach dem Spektralsatz wie in Beispiel 22.29 bereits eine orthogonale bzw. unitäre Matrix T, so dass $\overline{T}^{\mathsf{T}} A T$ eine Diagonalmatrix ist, deren Einträge nach Satz 22.35 (a) in $\mathbb{R}_{\geq 0}$ liegen. Wir können in Satz 22.46 dann also $S = T$ wählen, und die Singulärwerte von A sind genau die positiven Eigenwerte von A.

(b) Wir wollen eine Singulärwertzerlegung der Matrix

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \in \operatorname{Mat}(3 \times 2, \mathbb{R})$$

finden. Um wie im Beweis von Satz 22.46 zunächst T zu bestimmen, müssen wir die symmetrische, positiv semidefinite Matrix

$$A^{\mathsf{T}} A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \in \operatorname{Mat}(2 \times 2, \mathbb{R})$$

nach dem Spektralsatz orthogonal diagonalisieren. Dies haben wir bereits in Beispiel 22.30 getan: Mit

$$T = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \in \mathrm{O}(2) \quad \text{ist} \quad T^{\mathsf{T}} A^{\mathsf{T}} A T = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} =: \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}.$$

Die Singulärwerte von A sind also $\lambda_1 = \sqrt{3}$ und $\lambda_2 = \sqrt{1} = 1$. Um nun auch S zu bestimmen, setzen wir

$$s_1 := \frac{1}{\lambda_1} A\,Te_1 = \frac{1}{\sqrt{3}} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$$\text{und} \qquad s_2 := \frac{1}{\lambda_2} A\,Te_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Beachte, dass diese beiden Vektoren in der Tat wie in der Konstruktion von Satz 22.46 bezüglich des Standardskalarprodukts orthonormal sind. Wir ergänzen sie leicht (z. B. mit dem Gram-Schmidtschen Orthonormalisierungsverfahren aus Satz 21.31, falls man das Ergebnis nicht bereits sieht) mit

$$s_3 := \frac{1}{\sqrt{3}} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

zu einer Orthonormalbasis von $\mathbb{R}^3$. Nach Satz 22.46 gilt dann mit der orthogonalen Matrix $S := (s_1 \,|\, s_2 \,|\, s_3) \in \mathrm{O}(3)$ also

$$S^{\mathsf{T}} A\,T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

55

Beispiel 22.49 (Approximation von Matrizen)**.** Eine interessante praktische Anwendung der Singulärwertzerlegung liegt im Bereich der Approximation von Matrizen durch Matrizen von kleinem Rang. Dazu betrachten wir einmal eine (große) reelle Matrix $A = (a_{i,k})_{i,k} \in \mathrm{Mat}(m \times n, \mathbb{R})$ in ihrer Singulärwertzerlegung $A = SDT^{\mathsf{T}}$, so dass für alle $i = 1, \dots, m$ und $k = 1, \dots, n$ nach Definition 16.5 der Matrixmultiplikation also

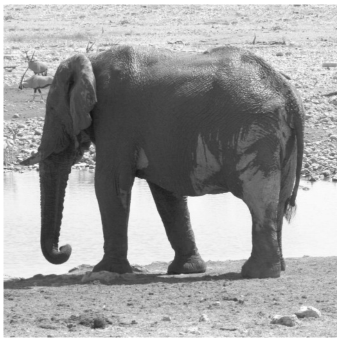
$$a_{i,k} = \sum_{j=1}^{r} s_{i,j}\, \lambda_j\, t_{k,j} \qquad (*)$$

mit $S = (s_{i,j})_{i,j} \in \mathrm{O}(m)$, $D = \mathrm{diag}(\lambda_1, \dots, \lambda_r) \in \mathrm{Mat}(m \times n, \mathbb{R})$ und $T = (t_{j,k})_{j,k} \in \mathrm{O}(n)$ gilt, wobei wir die Singulärwerte so anordnen, dass $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ (mit $r = \mathrm{rk}\,A$). Beachte dabei, dass in $(*)$ nur die ersten r Spalten von S und T benötigt werden, da in der Summe stets $j \leq r$ ist. Wir können A also exakt rekonstruieren, wenn wir nur diese Teile von S und T sowie die Diagonaleinträge von D kennen, was insgesamt

$$rm + rn + r = r(m+n+1)$$

reelle Zahlen sind. Wenn der Rang r von A klein ist, können dies deutlich weniger Zahlen sein als wenn wir uns die mn Einträge von A direkt merken würden. Matrizen von kleinem Rang lassen sich so also z. B. in einem Computer sehr platzsparend abspeichern.

Für beliebige Matrizen, deren Rang in der Regel nicht klein ist, hilft dies natürlich erst einmal nicht weiter. Ist es aber akzeptabel, die Werte in der Matrix nur näherungsweise abzuspeichern — wie etwa in dem Foto rechts, das mit seinen Helligkeitswerten eine reelle Matrix der Größe 1000×1000 darstellt — so können wir sehr einfach eine Näherung des Ausdrucks $(*)$ bilden, indem wir ein $r' \leq r$ wählen und in der Summe nur die größten (also „wichtigsten“) r' Singulärwerte $\lambda_1, \dots, \lambda_{r'}$ berücksichtigen, d. h. A durch die Matrix



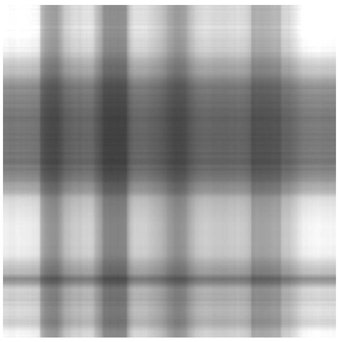
$$A' = (a'_{i,k})_{i,k} \in \mathrm{Mat}(m \times n, \mathbb{R}) \quad \text{mit} \quad a'_{i,k} = \sum_{j=1}^{r'} s_{i,j}\, \lambda_j\, t_{k,j}$$

vom Rang r' approximieren, die sich dann wieder wie oben platzsparend abspeichern lässt. Wir können den Fehler, den wir dabei machen, auch genau berechnen: In der Norm zum gewöhnlichen Skalarprodukt auf $\mathrm{Mat}(m \times n, \mathbb{R}) = \mathbb{R}^{mn}$ ist

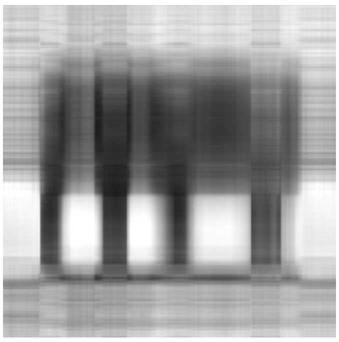
$$||A - A'||^2 = \sum_{i=1}^{m} \sum_{k=1}^{n} (a_{i,k} - a'_{i,k})^2 = \sum_{i=1}^{m} \sum_{k=1}^{n} \sum_{j>r'} \sum_{l>r'} s_{i,j}\, \lambda_j\, t_{k,j}\, s_{i,l}\, \lambda_l\, t_{k,l} = \sum_{j>r'} \lambda_j^2,$$

weil sowohl $\sum_{i=1}^{m} s_{i,j}\, s_{i,l}$ als auch $\sum_{k=1}^{n} t_{k,j}\, t_{k,l}$ nach Bemerkung 22.2 (b) wegen der Orthogonalität von S und T gleich 1 für $j = l$ und 0 für $j \neq l$ ist. Da die weggelassenen Singulärwerte λ_j für $j > r'$ ja die kleinsten sind, ist diese Näherung also wirklich recht gut. In der Tat kann man beweisen, dass sie im Sinne dieser Norm die beste Approximation von A durch eine Matrix vom Rang höchstens r' ist.

Für das obige Foto sind drei dieser Näherungen für (sehr) kleine Werte von r' unten dargestellt. Die ursprüngliche Matrix $A \in \mathrm{Mat}(1000 \times 1000, \mathbb{R})$ hat wie erwartet maximalen Rang $r = 1000$, ihre Singulärwerte fallen allerdings sehr schnell ab: Schon der erste im mittleren Näherungsfoto weggelassene Singulärwert λ_4 ist nur noch etwa $\frac{1}{10}$-mal so groß wie λ_1, und nahezu die Hälfte aller Singulärwerte ist kleiner als $\frac{1}{1000} \cdot \lambda_1$, so dass deren Vernachlässigung praktisch nicht mehr erkennbar ist. Das rechte Näherungsfoto belegt weniger als 5 % des Speicherplatzes der ursprünglichen Matrix A. Beachte auch, dass man dem linken Näherungsfoto „ansieht“, dass es Rang 1 hat: Alle Spalten der Matrix sind ein Vielfaches desselben Vektors, haben also von oben nach unten die gleiche Helligkeitsverteilung und sind nur insgesamt heller oder dunkler — wodurch das deutlich sichtbare Streifenmuster entsteht.



$r' = 1$



$r' = 3$

$r' = 20$

Aufgabe 22.50. Bestimme eine Singulärwertzerlegung für die reelle Matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}$.