

5. Euklidische und unitäre Räume (und selbstadjungierte, orthogonale, unitäre, normale Endomorphismen).

5.1. Reelle symmetrische Matrizen sind diagonalisierbar.

Satz: *Reelle symmetrische Matrizen sind diagonalisierbar.*

Bevor wir dies beweisen, wollen wir den Satz kommentieren. Ein typisches Beispiel für die Aussage ist die Matrix, die die Fibonacci-Zahlen erzeugt: $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, dies ist eine reelle symmetrische Matrix, also diagonalisierbar.

Andererseits sollte daran erinnert werden, dass es reelle quadratische Matrizen gibt, die nicht symmetrisch, aber diagonalisierbar sind, zum Beispiel $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$.

Nun betrachten wir symmetrische Matrizen über anderen Körpern, zum Beispiel: Die Matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in M(2 \times 2, \mathbb{F}_2)$ ist symmetrisch, aber nicht diagonalisierbar (sie ist nilpotent, aber nicht die Null-Matrix).

Oder auch: Die Matrix $A = \begin{bmatrix} 0 & i \\ i & 2 \end{bmatrix} \in M(2 \times 2, \mathbb{C})$ ist symmetrisch, aber nicht diagonalisierbar. (Das charakteristische Polynom ist $T^2 - 2T + 1 = (T - 1)^2$. Wäre sie diagonalisierbar, so wäre sie ähnlich zur Diagonalmatrix E_2 , also gleich E_2). Noch ein Beispiel: die komplexe Matrix $\begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$ ist nilpotent

Beweis des Satzes: Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch mit $n \geq 1$. Als erstes soll gezeigt werden:

(1) *Es gibt einen reellen Eigenwert.*

Beweis: Wir fassen A als komplexe Matrix auf.

Es sei an folgendes erinnert: zu jeder komplexen Zahl $z = x + iy$ mit $x, y \in \mathbb{R}$ wird mit $\bar{z} = x - iy$ die zu z konjugiert komplexe Zahl bezeichnet. Genau dann ist $z \in \mathbb{R}$, wenn $\bar{z} = z$ gilt.

Ist $v \in \mathbb{C}^n$, so setzt man $\|v\| = \sqrt{v^t \bar{v}}$, dies ist immer eine nicht-negative reelle Zahl und für $v \neq 0$ sogar eine positiv. Denn ist $v = (v_1, \dots, v_n)$ und $v_s = x_s + iy_s$ mit $x_s, y_s \in \mathbb{R}$, so ist

$$v^t \bar{v} = \sum_s v_s \cdot \bar{v}_s = \sum_s (x_s + iy_s)(x_s - iy_s) = \sum_s x_s^2 + y_s^2.$$

Jedes komplexe Polynom vom Grad mindestens 1 hat eine Nullstelle, also hat $\chi_A \in \mathbb{C}[T]$ eine Nullstelle, etwa λ . Es gibt also einen Vektor $v \neq 0$ mit komplexen Koordinaten mit $Av = \lambda v$. Wir rechnen:

$$\begin{aligned} \lambda \|v\|^2 &= \lambda v^t \bar{v} = (Av)^t \bar{v} = v^t A^t \bar{v} \\ &= v^t A \bar{v} \\ &= v^t \overline{A v} \\ &= v^t \overline{\lambda v} = v^t (\bar{\lambda}) \bar{v} = \bar{\lambda} v^t \bar{v} = \bar{\lambda} \|v\|^2 \end{aligned}$$

(dabei ist $A^t = A$, da A symmetrisch ist und $\overline{A} = A$, da A reelle Koeffizienten hat). Also $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.

Nun eine Erinnerung an LA I. Ist $v \in \mathbb{R}^n$, so bezeichnet man mit v^\perp den folgenden Unterraum:

$$v^\perp = \{w \in \mathbb{R}^n \mid v^t \cdot w = 0\}$$

(Statt $v^t \cdot w$ schreibt man oft auch $\langle v, w \rangle$: man betrachtet den \mathbb{R}^n als euklidischen Raum mit dem kanonischen inneren Produkt.)

Wichtig ist: Es gilt

$$L(v) \oplus v^\perp = \mathbb{R}^n.$$

(siehe 3.1.11 (b)).

(2) Ist v ein Eigenvektor zu A , so ist v^\perp ein A -invarianter Unterraum.

Beweis: Sei $w \in v^\perp$. Wir zeigen, dass Aw auch zu v^\perp gehört.

$$v^t \cdot Aw = v^t A \cdot w = v^t A^t \cdot w = (Av)^t \cdot w = (\lambda v)^t w = \lambda v^t w = \lambda \cdot 0 = 0.$$

Der eigentliche Beweis des Satzes wird mit Induktion geführt. Ist $n = 1$, so ist nichts zu zeigen. Sei also $n \geq 2$. Wegen (1) gibt es einen Eigenwert λ (in \mathbb{R}), also einen Eigenvektor v zum Eigenwert λ . Ersetze v durch $v_1 = \frac{1}{\|v\|}v$, dann ist auch v_1 ein Eigenvektor mit Eigenwert λ und zusätzlich ist dieser Eigenvektor normiert. Wir wählen eine Orthonormalbasis w_2, \dots, w_n von $v_1^\perp = v^\perp$. Die Matrizendarstellung von l_A bezüglich dieser neuen Basis ist wieder eine symmetrische Matrix und hat die Form

$$\begin{bmatrix} \lambda & & & & 0 \\ & \vdots & & & \\ & & \ddots & & \\ & & & A' & \\ & 0 & & & \end{bmatrix}$$

Nach Induktion wissen wir, dass es zur symmetrischen Matrix A' eine Orthonormalbasis aus Eigenvektoren gibt, etwa v_2, \dots, v_n . Dann ist v_1, \dots, v_n eine Orthonormalbasis von \mathbb{R}^n und alle diese Vektoren v_i sind Eigenvektoren für die Matrix A .

Wir haben also die folgende Verschärfung gezeigt:

Satz (zweite Formulierung). *Ist A eine reelle symmetrische Matrix, so gibt es eine Orthonormalbasis des \mathbb{R}^n aus Eigenvektoren.*

Bilden wir aus diesen Basis-Vektoren eine Matrix S , so gilt also:

- $S^{-1}AS$ ist Diagonalmatrix,
- S ist orthogonale Matrix, d.h. $S^{-1} = S^t$. (LA I,3.1.19)

Man kann demnach die Aussage des Satzes auch so formulieren:

Satz (dritte Formulierung). *Ist A eine symmetrisch reelle Matrix, so gibt es eine orthogonale Matrix S so dass $S^{-1}AS$ eine Diagonalmatrix ist oder auch: so dass S^tAS eine Diagonalmatrix ist (für eine orthogonale Matrix ist ja $S^{-1} = S^t$).*

Betrachten wir noch einmal das Beispiel der Matrix, die die Fibonacci-Zahlen erzeugt, symmetrisch: $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Berechnet man Eigenvektoren zu den beiden Eigenwerten, so sieht man, dass diese senkrecht aufeinander stehen: Wie wir nun wissen, muss das so sein: *Eigenvektoren einer symmetrischen Matrix mit verschiedenen Eigenwerten sind orthogonal.* — Das bedeutet aber auch: sobald wir **einen** Eigenvektor zur Fibonacci-Matrix berechnet haben, etwa $[x_1 \ x_2]^t$, wissen wir (ohne jede weitere Rechnung), dass auch $[x_2 \ -x_1]^t$ ein Eigenvektor ist.

Erinnerung: Sei $(V, \langle -, - \rangle)$ euklidischer Vektorraum. Eine lineare Abbildung $f: V \rightarrow V$ heißt *selbstadjungiert*, wenn für alle $v, w \in V$ gilt:

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

Eine symmetrische Matrix liefert einen selbstadjungierten Endomorphismus bezüglich des kanonischen inneren Produkts auf \mathbb{R} . Und umgekehrt gilt: Ist V euklidischer Vektorraum mit einer Orthonormalbasis \mathcal{B} , und ist $f: V \rightarrow V$ selbstadjungiert, so ist die Matrizarstellung $M_{\mathcal{B}}^{\mathcal{B}}(f)$ eine symmetrische Matrix.

Offensichtlich gilt auch: *Ist f selbstadjungiert, und U f -invarianter Unterraum von V , so ist auch $f|_U$ selbstadjungiert.*

Dies wird beim Induktionsbeweis gebraucht.

Satz (vierte Formulierung). *Ist V endlich-dimensionaler euklidischer Vektorraum und $f: V \rightarrow V$ ein selbstadjungierter Endomorphismus, so gibt es eine Orthonormalbasis aus Eigenvektoren.*

Beweis. Wir gehen noch einmal den Beweis durch:

(1) *Die Abbildung f hat einen Eigenwert.* Um dies zu zeigen, ordnen wir f eine symmetrische Matrix zu: Wähle eine Orthonormalbasis \mathcal{B} . Dann ist $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ eine symmetrische Matrix. Wir fassen A als komplexe Matrix auf, als komplexe Matrix hat A einen (komplexen) Eigenwert. Man zeigt (siehe oben): jeder (komplexe) Eigenwert von A ist reell. (Man könnte hier auch die Rechnung im Abschnitt 5.3 verwenden: dort wird gezeigt: jeder Eigenwert einer hermiteschen Matrix ist reell, und reelle symmetrische Matrizen sind hermitesch.)

(2) Da f einen reellen Eigenwert besitzt, gibt es dazu einen Eigenvektor $v \in V$. Behauptung: *Mit $L(v)$ ist auch v^{\perp} f -invariant.*

Beweis:

$$\langle v, Aw \rangle = \langle Av, w \rangle = \langle \lambda v, w \rangle = \lambda \langle v, w \rangle = \lambda \cdot 0 = 0.$$

Nun der Induktionsschritt: Die Einschränkung von f auf v^{\perp} ist wieder selbstadjungiert, also gibt es in v^{\perp} eine Orthonormalbasis aus Eigenvektoren. Zusammen mit $\frac{1}{\|v\|}v$ ist dies dann eine Orthonormalbasis von V aus Eigenvektoren.

Folgerung. Die einzige Matrix in $M(n \times n, \mathbb{R})$, die symmetrisch und nilpotent ist, ist die Nullmatrix.

Beweis: Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch und nilpotent. Da A symmetrisch ist, gibt es eine orthogonale Matrix P , sodass $D = P^{-1}AP$ eine Diagonalmatrix ist. Mit A ist auch $P^{-1}AP$ symmetrisch. Andererseits ist mit A auch $P^{-1}AP$ nilpotent. Es ist also D eine symmetrische nilpotente Diagonalmatrix. Also ist D die Nullmatrix. Dann ist aber auch $A = PDP^{-1} = 0$.

Man sieht hier: der Anfang des Beweises ist wie der für Matrizen, danach ist aber alles einfacher: wir brauchen keine Orthonormalbasis von v^\perp zu betrachten um zu sehen, dass mit der Einschränkung von f auf v^\perp arbeiten zu können — die Einschränkung einer selbstadjungierten Abbildung auf einen f -invarianten Unterraum ist offensichtlich wieder selbstadjungiert.

Hier noch einmal der **Algorithmus**, wie man zur reellen symmetrischen Matrix C eine orthogonale Matrix P findet, sodass $P^{-1}CP$ eine Diagonalmatrix ist:

- Bestimme das charakteristische Polynom χ_C der Matrix C .
[Hier ist also eine Determinante zu berechnen.]
- Bestimme die Nullstellen $\gamma_1, \dots, \gamma_t$ von χ_C (auf diese Weise erhält man die Eigenwerte von C).
[Dies ist meist der schwierigste Schritt. Die Theorie besagt, dass χ_C Produkt von Linearfaktoren ist — sie zu berechnen kann aber schwierig sein.]
- Zu jedem γ_i bestimmt man eine Basis v_{i1}, \dots, v_{i,j_i} von $\text{Eig}(C, \gamma_i)$.
[Hier handelt es sich um das Lösen eines linearen Gleichungssystems, nämlich dem homogenen Gleichungssystem mit Koeffizientenmatrix $C - \gamma_i E_n$. Dafür gibt es zum Beispiel den Gauß-Algorithmus.]
- Orthonormiere die Folge v_{i1}, \dots, v_{i,j_i} , man erhält auf diese Weise eine Orthonormalbasis w_{i1}, \dots, w_{i,j_i} von $\text{Eig}(C, \gamma_i)$.
[Hierfür gibt es das Gram-Schmidt-Verfahren.]
- Die Matrix P hat als Spalten die Vektoren

$$w_{11}, \dots, w_{1,j_1}, w_{21}, \dots, w_{2,j_2}, \dots, w_{t1}, \dots, w_{t,j_t}.$$

Beachte: Diese Folge ist eine Orthonormalbasis von \mathbb{R}^n , daher ist P orthogonale Matrix. Will man zusätzlich haben, dass $\det P = 1$ gilt, so muss man gegebenenfalls zwei Spalten vertauschen (dies geht immer, wenn $n \geq 2$ ist; im Fall $n = 1$ ist nichts zu zeigen).

Das Wichtigste in Kürze:

- Jede reelle symmetrische $(n \times n)$ -Matrix C ist diagonalisierbar.
- Die Eigenräume $\text{Eig}(C, \lambda)$ sind paarweise orthogonal (bezüglich des kanonischen inneren Produkts von \mathbb{R}^n).

5.2. Unitäre Vektorräume.

Definition: Sei V ein \mathbb{C} -Vektorraum. Eine Abbildung $\langle -, - \rangle: V \times V \rightarrow \mathbb{C}$ heißt *Sesquilinearform* (sesquilinear = anderthalb-linear), wenn die folgenden beiden Bedingungen für alle $v, v_1, v_2 \in V$ und $\lambda_1, \lambda_2 \in \mathbb{C}$ gelten:

- (1) $\langle \lambda_1 v_1 + \lambda_2 v_2, v \rangle = \lambda_1 \langle v_1, v \rangle + \lambda_2 \langle v_2, v \rangle$
- (2) $\langle v, \lambda_1 v_1 + \lambda_2 v_2 \rangle = \overline{\lambda_1} \langle v, v_1 \rangle + \overline{\lambda_2} \langle v, v_2 \rangle$

Die erste Bedingung besagt Linearität in der ersten Variablen. In der zweiten Variablen wird immer noch vorausgesetzt, dass die Form additiv ist, statt aber des üblichen "Herausziehen" eines Skalars in der zweiten Variablen wird hier verlangt, dass der Skalar nach dem Herausziehen zu konjugieren ist.

Eine *hermite'sche Form* auf V ist eine Abbildung $\langle -, - \rangle: V \times V \rightarrow \mathbb{C}$ mit den folgenden beiden Eigenschaften:

- (1) Seien $v_1, v_2, w \in V$ und $\alpha_1, \alpha_2 \in \mathbb{C}$. Dann ist

$$\langle \alpha_1 v_1 + \alpha_2 v_2, w \rangle = \alpha_1 \langle v_1, w \rangle + \alpha_2 \langle v_2, w \rangle.$$

- (2) Für $v, w \in V$ gilt

$$\langle w, v \rangle = \overline{\langle v, w \rangle}.$$

Eine hermite'sche Form ist offensichtlich sesquilinear, denn

$$\begin{aligned} \langle v, \lambda_1 v_1 + \lambda_2 v_2 \rangle &= \overline{\langle \lambda_1 v_1 + \lambda_2 v_2, v \rangle} \\ &= \overline{\lambda_1 \langle v_1, v \rangle + \lambda_2 \langle v_2, v \rangle} \\ &= \overline{\lambda_1} \overline{\langle v_1, v \rangle} + \overline{\lambda_2} \overline{\langle v_2, v \rangle} \\ &= \overline{\lambda_1} \langle v, v_1 \rangle + \overline{\lambda_2} \langle v, v_2 \rangle \end{aligned}$$

Typisches Beispiel: Sei $H \in M(n \times n, \mathbb{C})$. Man nennt H *hermite'sche Matrix*, falls gilt $\overline{H}^t = H$ (d.h.: Es ist $H = (h_{ij})$ mit $h_{ji} = \overline{h_{ij}}$ für alle i, j). Setze

$$\langle v, w \rangle = v^t \cdot H \cdot \overline{w}.$$

Dann ist $\langle -, - \rangle$ eine hermite'sche Form auf \mathbb{C}^n .

Wörterbuch: Hermitesche Form, hermitesche Matrix.

- (a) Genau dann ist die Form $\langle v, w \rangle = v^t A \overline{w}$ hermitesch, wenn A hermitesch ist.
- (b) Ist $(V, \langle -, - \rangle)$ hermitesch, und ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis, so ist die Matrix

$$M_{\mathcal{B}}(\langle -, - \rangle) = (\langle v_r, v_s \rangle)_{rs}$$

hermitesch.

Wichtig. Ist $\langle -, - \rangle$ eine hermite'sche Form auf V , so ist $\langle v, v \rangle \in \mathbb{R}$ für alle $v \in V$.

(Denn die Vertauschungsregel besagt gerade: $\langle v, v \rangle = \overline{\langle v, v \rangle}$. Für eine Zahl $c \in \mathbb{C}$ gilt aber: Ist $c = \bar{c}$, so ist $c \in \mathbb{R}$.)

Man nennt die hermite'sche Form $\langle -, - \rangle$ *positiv definit* (oder auch ein *inneres Produkt*), falls für alle $0 \neq v \in V$ gilt: $\langle v, v \rangle > 0$. (Beachte: Für eine beliebige komplexe Zahl c würde die Ungleichung $c > 0$ keinen Sinn machen! Hier wird gebraucht, dass $c \in \mathbb{R}$ gilt.) Einen komplexen Vektorraum mit einem inneren Produkt nennt man einen *unitären Raum*.

Der kanonische unitäre Vektorraum \mathbb{C}^n : Hier nimmt man als hermitesche Matrix H die Einheitsmatrix, arbeitet also mit

$$\langle v, w \rangle = v^t \bar{w}.$$

Wie wir schon gesehen haben, ist diese hermitesche Form positiv definit, also ein inneres Produkt; man nennt dies das *kanonische innere Produkt* auf \mathbb{C}^n .

Sei nun $(V, \langle -, - \rangle)$ ein unitärer Raum.

Norm. Ist $v \in V$, so nennt man $\|v\| = \sqrt{\langle v, v \rangle}$ die *Norm* von v (beachte: das Wurzelzeichen dürfen wir hier verwenden, da $\langle v, v \rangle$ eine **nicht-negative reelle** Zahl ist). Man nennt v *normiert*, falls $\|v\| = 1$ gilt.

Natürlich kann man jeden von Null verschiedenen Vektor v normieren: multipliziere ihn mit $\frac{1}{\|v\|}$, denn es ist für λ in \mathbb{C}

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda \bar{\lambda} \langle v, v \rangle} = \sqrt{\lambda \bar{\lambda}} \sqrt{\langle v, v \rangle} = |\lambda| \cdot \|v\|,$$

Also gilt für $\lambda = \frac{1}{\|v\|}$

$$\left\| \frac{1}{\|v\|} v \right\| = \left| \frac{1}{\|v\|} \right| \cdot \|v\| = \frac{1}{\|v\|} \cdot \|v\| = 1.$$

Orthogonalität. Man nennt Vektoren v, w *orthogonal* und schreibt $v \perp w$, falls $\langle v, w \rangle = 0$ gilt (und dies gilt genau dann, wenn $\langle w, v \rangle = 0$ gilt).

Wie in euklidischen Vektorräumen nennt man eine Basis eine *Orthogonalbasis*, wenn die Vektoren paarweise orthogonal sind.

Man nennt eine Basis *Orthonormalbasis*, falls dies eine Orthogonalbasis ist und alle Vektoren normiert sind.

Ist X eine Teilmenge von V , so nennt man

$$X^\perp = \{v \in V \mid \langle v, x \rangle = 0\} = \{v \in V \mid \langle x, v \rangle = 0\}$$

das *orthogonale Komplement* zu X (oder zu $L(X)$). Das orthogonale Komplement ist ein Unterraum von V und es gilt

$$L(X) \oplus X^\perp = V$$

(Beweis analog zum Beweis für euklidische Vektorräume).

Wie im Fall eines euklidischen Vektorraums kann man auch in unitären Räumen eine gegebene linear unabhängige Folge von Vektoren orthonormalisieren:

Orthonormalisierung: Gram-Schmidt-Verfahren. Sei U ein Unterraum von V . Sei u_1, \dots, u_m eine Orthonormalbasis von U . Ist $v \in V$, so setze

$$v' = v - \sum_{i=1}^m \langle v, u_i \rangle u_i.$$

Es ist $v' \in U^\perp$. Ist $v \in U$, so ist $v' = 0$. Ist $v \notin U$, so sind die Vektoren u_1, \dots, u_m, v' linear unabhängig. Daraus folgt: *Ist V ein endlich-dimensionaler unitärer Vektorraum und U ein Unterraum von V , so läßt sich jede Orthonormalbasis von U zu einer Orthonormalbasis von V fortsetzen. Insbesondere gilt: Jeder endlich-dimensionale unitäre Vektorraum besitzt eine Orthonormalbasis.*

5.3. Selbstadjungierte Endomorphismen eines unitären Vektorraums.

Definition: Sei $(V, \langle -, - \rangle)$ unitärer Raum, sei $f: V \rightarrow V$ linear. Man nennt f selbstadjungiert, wenn gilt:

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

Wörterbuch: Selbstadjungierter Endomorphismus, hermitesche Matrix.

Man nennt eine Matrix $A \in M(n \times n, \mathbb{C})$ *hermitesch*, falls $\bar{A} = A^t$ gilt. Behauptung:

Ist A hermitesch, so ist l_A selbstadjungiert. Umgekehrt gilt: Die Matrix-Darstellung eines selbstadjungierten Endomorphismus bezüglich einer Orthonormalbasis ist eine hermitesche Matrix.

Die Abbildung $l_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ ist selbstadjungiert, denn

$$\langle l_A v, w \rangle = \langle Av, w \rangle = (Av)^t \bar{w} = v^t A^t \bar{w} = v^t \bar{A} \bar{w} = v^t \overline{Aw} = \langle v, Aw \rangle = \langle v, l_A w \rangle.$$

Satz. *Sei W unitärer Vektorraum. Sei $f: W \rightarrow W$ selbstadjungierter Endomorphismus. Dann gilt: Jeder Eigenwert ist reell. Ist W endlich-dimensional, so besitzt W eine Orthonormalbasis aus Eigenvektoren.*

Wir brauchen: Ist v Eigenvektor zu f , so ist v^\perp f -invariant.

Allgemein gilt: Ist U f -invariant, so ist U^\perp f -invariant (falls f selbstadjungiert ist).

Satz. *Ist A hermitesche Matrix, so sind alle Eigenwerte reell und \mathbb{C}^n besitzt eine Basis aus Eigenvektoren, und zwar sogar eine Orthonormalbasis bezüglich des kanonischen inneren Produkts.*

5.4. Unitäre Endomorphismen und unitäre Matrizen.

Lemma. Die folgenden Eigenschaften sind für eine Matrix $A \in M(n \times n, \mathbb{C})$ äquivalent:

- (i) A ist invertierbar und $A^{-1} = \overline{A}^t$.
- (ii) $\overline{A}^t A = E_n$.
- (iii) $A \overline{A}^t = E_n$.
- (ii') Die Spalten von A bilden eine Orthonormalbasis des kanonischen unitären Raums \mathbb{C}^n .
- (iii') Die Zeilen von A (transponiert) bilden eine Orthonormalbasis des kanonischen unitären Raums \mathbb{C}^n .

Falls diese Eigenschaften gelten, nennt man die Matrix A unitär.

Ist A eine unitäre Matrix, so ist $|\det A| = 1$. (Denn es ist $\det \overline{A} = \overline{\det A}$, also $1 = \det A^{-1} A = \det(\overline{A}^t \cdot A) = \overline{d} \cdot d$ mit $d = \det A$).

Beispiele:

(a) Die unitären 2×2 -Matrizen. Als erste Spalte nimmt man einen beliebigen normierten Vektor in \mathbb{C}^2 . Also: man nimmt einen Vektor $v \neq 0$ in \mathbb{C}^2 und normiert ihn, bildet also $v_1 = \frac{1}{\|v\|} v$. Sei die erste Spalte $[a \ c]^t$. Dann kann man als zweite Spalte $[-\overline{c} \ \overline{a}]^t$ nehmen, oder allgemeiner: $\omega[-\overline{c} \ \overline{a}]^t$, wobei $\omega \in \mathbb{C}$ den Betrag 1 haben muss. Auf diese Weise erhält man alle unitären (2×2) -Matrizen. (Die erste Spalte v_1 haben wir ja so allgemein wie nur möglich gewählt. Die zweite Spalte muss ein Vektor in v_1^\perp sein, dies ist ein eindimensionaler Vektorraum, und er enthält für $v_1 = [a \ c]^t$ den Vektor $w = [-\overline{c} \ \overline{a}]^t$. Da w schon normiert ist, erhält man alle normierten Vektoren in v_1^\perp indem man w mit einer beliebigen komplexen Zahl ω vom Betrag 1 multipliziert.)

Beispiel. Wir beginnen mit $[1 \ 2]^t$, die Norm ist $\sqrt{5}$, also $v_1 = \frac{1}{\sqrt{5}}[1 \ 2]^t$, also erhält man als erstes

$$A = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}, \quad \text{oder} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2i \\ 2 & i \end{bmatrix}, \quad \text{oder ganz allgemein} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2\omega \\ 2 & \omega \end{bmatrix},$$

dabei ist $\omega \in \mathbb{C}$ mit $|\omega| = 1$, also etwa $\omega = \frac{1}{2}(1 + \sqrt{3}i)$.

Beginnen wir mit $v = [1 + 2i, 3 - 4i]^t$, so ist die Norm $\|v\| = \sqrt{30}$, also nimmt man als erste Spalte $\frac{1}{\sqrt{30}}[1 + 2i, 3 - 4i]^t$. Eine unitäre Matrix mit dieser ersten Spalte ist

$$\frac{1}{\sqrt{30}} \begin{bmatrix} 1 + 2i & -3 - 4i \\ 3 - 4i & 1 - 2i \end{bmatrix},$$

weitere erhält man, in dem man die zweite Spalte mit einer komplexen Zahl ω mit $|\omega| = 1$ multipliziert.

(b) Die unitären 3×3 -Matrizen: Wieder nimmt man als erste Spalte einen beliebigen normierten Vektor v_1 , dann aber muss man etwas arbeiten: Als zweite Spalte braucht man einen zu v_1 orthogonalen Vektor, der auch wieder zu normieren ist. Es bietet sich hier an, einen beliebigen Vektor in $\mathbb{C}^3 \setminus L(v_1)$ zu nehmen, und das Gram-Schmidt-sche Orthonormalisierungsverfahren anzuwenden, usw.

(1) Ist V ein n -dimensionaler unitärer Raum und sind zwei Orthonormalbasen \mathcal{A} und \mathcal{B} gegeben, so sind die zugehörigen Basiswechselmatrizen unitäre Matrizen.

Ist $\mathcal{A} = (v_1, \dots, v_n)$ eine Basis des K -Vektorraums V , so gibt es einen Vektorraum-Isomorphismus $\Phi_{\mathcal{A}}: K^n \rightarrow V$, der den kanonischen i -ten Basisvektor von K^n auf den Vektor v_i abbildet. Sind nun \mathcal{A} und \mathcal{B} zwei Basen von V , so sind die linearen Abbildungen

$$\Phi_{\mathcal{A}}^{-1}\Phi_{\mathcal{B}}, \quad \text{und} \quad \Phi_{\mathcal{B}}^{-1}\Phi_{\mathcal{A}}$$

durch invertierbare Matrizen in $M(n \times n, K)$ gegeben, man nennt sie die Basiswechselmatrizen.

Ist $f: V \rightarrow V$ ein Endomorphismus des unitären Vektorraums V , so nennt man f *unitär*, falls gilt

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V.$$

(2) Sei A eine komplexe $(n \times n)$ -Matrix. Genau dann ist A unitär, wenn die Abbildung $l_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ ein unitärer Endomorphismus (bezüglich des kanonischen inneren Produkts auf \mathbb{C}^n) ist.

(3) Ist V ein unitärer Vektorraum mit Orthonormalbasis $\mathcal{B} = (v_1, \dots, v_n)$, und ist $f: V \rightarrow V$ unitärer Endomorphismus, so ist die Matrizendarstellung $M_{\mathcal{B}}^{\mathcal{B}}(f)$ von f bezüglich dieser Basis eine unitäre Matrix.

Hauptsatz über unitäre Endomorphismen. Sei V ein endlich-dimensionaler unitärer Vektorraum und $f: V \rightarrow V$ ein unitärer Endomorphismus. Dann gibt es eine Orthonormalbasis von V , so daß die zugehörige Matrizendarstellung von f eine Diagonalmatrix ist, deren Diagonalkoeffizienten den Betrag 1 haben.

Umformulierung: **Hauptsatzes über unitäre Matrizen.** Ist A eine unitäre Matrix, so gibt es eine unitäre Matrix P , so daß $P^{-1}AP$ eine Diagonalmatrix ist, deren Diagonalkoeffizienten den Betrag 1 haben.

Der Beweis des Hauptsatzes verwendet die folgenden beiden Teilargumente:

Eigenwerte. Sei f unitärer Endomorphismus des unitären Vektorraums V . Ist λ ein Eigenwert von f , so ist $|\lambda| = 1$. Insbesondere ist die Abbildung f injektiv. Ist V endlich-dimensional, so ist f bijektiv und das charakteristische Polynom von f hat die Form $(T - \gamma_1) \cdots (T - \lambda_n)$ mit $|\lambda_i| = 1$ für $1 \leq i \leq n$.

Beweis: Sei $0 \neq v \in V$ mit $f(v) = \lambda v$. Dann ist

$$\lambda \bar{\lambda} \langle v, v \rangle = \langle \lambda v, \lambda v \rangle = \langle f(v), f(v) \rangle = \langle v, v \rangle.$$

Da $\langle v, v \rangle \neq 0$ folgt $\lambda \bar{\lambda} = 1$, also auch $|\lambda| = \sqrt{\lambda \bar{\lambda}} = 1$.

Orthogonales Komplement. Sei f unitärer Endomorphismus des unitären Vektorraums V . Ist U ein endlich-dimensionaler Unterraum von V , der f -invariant ist, so ist auch U^\perp f -invariant.

Beweis: Sei $v \in U^\perp$. Wir wollen zeigen, dass $f(v) \in U^\perp$ gilt, also dass $\langle f(v), u \rangle = 0$ für alle $u \in U$ gilt. Da die Einschränkung von f auf U bijektiv ist, gibt es $u' \in U$ mit $u = f(u')$. Nun sehen wir:

$$\langle f(v), u \rangle = \langle f(v), f(u') \rangle = \langle v, u' \rangle = 0.$$

5.5. Orthogonale Endomorphismen, orthogonale Matrizen.

Wir nehmen nun wieder als Grundkörper den Körper \mathbb{R} der reellen Zahlen. Es soll hier an eine Vielzahl von Definition und elementarer Aussagen erinnert werden, die im Rahmen der Vorlesung LA I behandelt wurden.

Erinnerung. $A \in M(n \times n, \mathbb{R})$ heißt orthogonal, wenn gilt $A^{-1} = A^t$.

Die folgenden Eigenschaften sind äquivalent: A orthogonal, $A^t A = E_n$, $AA^t = E_n$, die Spalten von A bilden eine Orthonormalbasis, die Zeilen von A bilden (transponiert) eine Orthonormalbasis.

Ein Endomorphismus f des euklidischen Vektorraums $(V, \langle -, - \rangle)$ heißt orthogonal, wenn

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

für alle v, w gilt.

Genau dann ist eine lineare Abbildung $f: V \rightarrow V$ f orthogonal, wenn $\|f(v)\| = \|v\|$ für alle $v \in V$ gilt, wenn also f "normerhaltend" ist. (Siehe LA I, 3.1.17. Klar ist, daß ein orthogonaler Endomorphismus normerhaltend ist. Umgekehrt kann man das innere Produkt mit Hilfe der Norm von Vektoren berechnen.)

Typische Beispiele orthogonaler Endomorphismen des kanonischen euklid'schen Vektorraums \mathbb{R}^2 sind die Drehungen $l_{D(t)}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$D(t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$$

und $t \in \mathbb{R}$; wegen der Periodizität von Sinus und Cosinus reicht es, $0 \leq t < 2\pi$ zu betrachten. Man nennt A die *Drehmatrix zum Winkel t* . Ist $t = 0$, so ist $A = E_2$ die Einheitsmatrix, ist $t = \pi$, so ist $A = -E_2$ ebenfalls eine Skalarmatrix (dies ist gerade die Punktspiegelung am Ursprung).

Weitere Beispiele orthogonaler Endomorphismen des kanonischen euklid'schen Vektorraums \mathbb{R}^2 sind die Spiegelungen $l_{S(t)}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$D(t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$$

und $t \in \mathbb{R}$ und $0 \leq t < 2\pi$. Wie wir sehen werden ist jeder orthogonale Endomorphismus des kanonischen euklidischen Raums \mathbb{R}^2 entweder eine Drehung oder eine Spiegelung (und nicht beides).

Wörterbuch: orthogonaler Endomorphismus, orthogonale Matrix.

(a). Sei $A \in M(n \times n, \mathbb{R})$ Genau dann ist die Abbildung f_A orthogonal bezüglich des kanonischen inneren Produkts auf \mathbb{R}^n , wenn A orthogonale Matrix ist.

(b) Ist $(V, \langle -, - \rangle)$ euklidischer Vektorraum und ist $f: V \rightarrow V$ orthogonale Abbildung, und \mathcal{B} eine Orthonormalbasis von V , so ist die Matrixdarstellung von f eine orthogonale Matrix.

Hauptsatz über orthogonale Endomorphismen. Sei V ein endlich-dimensionaler euklid'scher Vektorraum. Sei $f: V \rightarrow V$ ein orthogonaler Endomorphismus von V . Dann besitzt V eine Orthonormalbasis, so daß f bezüglich dieser Basis eine Matrixdarstellung der Form $\bigoplus_{i=1}^m A_i$ hat, wobei die Matrizen A_i (1×1) -Matrizen der Form $[1]$ oder $[-1]$ oder (2×2) -Drehmatrizen zu Winkeln $0 < t_i < \pi$ sind.

Umformulierung: **Hauptsatzes über orthogonale Matrizen.** Ist A eine orthogonale Matrix, so gibt es eine orthogonale Matrix P mit $P^{-1}AP = \bigoplus_{i=1}^m C_i$, wobei die Matrizen C_i (1×1) -Matrizen $[1]$ oder $[-1]$ oder (2×2) -Drehmatrizen zu Winkeln $0 < t_i < \pi$ sind.

Bevor wir mit dem Beweis beginnen, ist an den Fundamentalsatz der Algebra zu erinnern.

Nachtrag zum Fundamentalsatz der Algebra. Der Fundamentalsatz der Algebra impliziert, daß der Grad eines jeden irreduziblen Polynoms $h \in \mathbb{R}[T]$ 1 oder 2 ist. Eine Folgerung daraus ist:

Lemma. Sei $V \neq 0$ ein endlich-dimensionaler reeller Vektorraum. Sei $f: V \rightarrow V$ ein Endomorphismus. Besitzt f keine Eigenvektoren, so gibt es in V einen zweidimensionalen f -invarianten Unterraum.

Beweis: Sei $\chi_f(T)$ das charakteristische Polynom von f , zerlege $\chi_f = h_t \cdots h_1$ mit normierten irreduziblen Polynomen h_i . Nach dem Fundamentalsatz der Algebra haben diese Polynome h_i den Grad 1 oder 2 (und Grad 1 kann hier gar nicht auftreten, da wir voraussetzen, daß f keine Eigenvektoren besitzt). Wähle $v \neq 0$ in V . Bilde $v_0 = v$, $v_i = h_i(f)(v_{i-1})$ für $1 \leq i \leq t$. Es ist $v_0 \neq 0$, aber $v_n = h_t(f) \cdots h_1(f)(v) = \chi_f(f)(v) = 0$ (nach dem Satz von Cayley-Hamilton). Also gibt es ein $1 \leq i \leq n$ mit $v_{i-1} \neq 0$, und $v_i = 0$. Sei $w = v_{i-1}$. Betrachte den Unterraum U , der von w und $f(w)$ erzeugt wird. Wir zeigen, daß U f -invariant ist. Sei $h_i(T) = T^2 + cT + d$ mit reellen Zahlen c, d . Nach Konstruktion ist

$$0 = h_i(f)(w) = (f^2 + cf + d \cdot 1)(w) = f^2(w) + cf(w) + dw,$$

und demnach ist $f^2(w) = -cf(w) - dw$ eine Linearkombination von $f(w)$ und w , also wieder ein Element von U . Andererseits ist U auf jeden Fall von Null verschieden, denn $w \neq 0$. Wäre U eindimensional, so hätten wir einen eindimensionalen f -invarianten Unterraum, also Eigenvektoren. Da wir dies ausgeschlossen haben, muß U zweidimensional sein.

Nun zum Beweis des Hauptsatzes über orthogonale Endomorphismen. Wir beginnen mit drei Vorbemerkungen:

- (1) *Ist γ Eigenwert von f , so ist $\gamma = 1$ oder -1 .* Beweis: Sei v zugehöriger Eigenvektor, also $v \neq 0$ und $f(v) = \gamma \cdot v$. Es ist

$$\|v\| = \|f(v)\| = \|\gamma \cdot v\| = |\gamma| \cdot \|v\|.$$

Wegen $\|v\| \neq 0$ folgt $1 = |\gamma|$, also $\gamma = 1$ oder $\gamma = -1$.

- (2) Insbesondere sehen wir, daß ein orthogonaler Endomorphismus injektiv ist (denn 0 ist kein Eigenwert). Ist also f ein orthogonaler Endomorphismus eines endlich-dimensionalen Vektorraums (und das wird hier vorausgesetzt), so ist f auch surjektiv.
- (3) *Ist U ein f -invarianter Unterraum von V , so ist auch U^\perp f -invariant.* Beweis: Sei $v \in U^\perp$. Zu zeigen ist: $f(v) \in U^\perp$. Sei also $u \in U$. Da die Einschränkung von f auf U ein orthogonaler Endomorphismus von U ist, ist diese Einschränkung surjektiv: es gibt also $u' \in U$ mit $f(u') = u$. Also

$$\langle f(v), u \rangle = \langle f(v), f(u') \rangle = \langle v, u' \rangle = 0.$$

Nun der eigentliche Beweis, mit Induktion nach der Dimension $\dim V = n$.

Wir betrachten zuerst den Fall, daß V keine f -invarianten Unterräume außer 0 und V besitzt (dies schließt insbesondere den Fall $n = 1$ ein). Nach dem Fundamentalsatz der Algebra ist demnach $n = 1$ oder $n = 2$.

Ist $n = 1$, so ist jeder von Null verschiedene Vektor $v \in V$ ein Eigenvektor, die zugehörige Matrizendarstellung von f ist $[1]$ oder $[-1]$, weil 1 und -1 die einzigen möglichen Eigenwerte sind.

Sei nun $n = 2$. Da wir voraussetzen, daß V keine f -invarianten Unterräume hat, kann f keine Eigenvektoren besitzen. Sei v_1, v_2 eine beliebige Orthonormalbasis von V . Sei $f(v_1) = av_1 + bv_2$ mit $a, b \in \mathbb{R}$. Wegen

$$1 = \|v_1\| = \|f(v_1)\| = \|av_1 + bv_2\| = \sqrt{a^2 + b^2}$$

folgt also $a^2 + b^2 = 1$. Da v_1, v_2 orthogonal sind, müssen auch die Vektoren $f(v_1)$, $f(v_2)$ orthogonal sein. Und natürlich ist mit v_2 auch $f(v_2)$ normiert. In V gibt es nur zwei normierte Vektoren, die zu $f(v_1) = av_1 + bv_2$ orthogonal sind, nämlich $-bv_1 + av_2$ und $bv_1 - av_2$. Ist $f(v_2) = bv_1 - av_2$, so rechnet man sofort nach, daß das charakteristische Polynom von f das Polynom $(T + a)(T - a) - b^2 = T^2 - 1 = (T - 1)(T + 1)$ ist, aber dies impliziert, daß f Eigenvektoren besitzt, unmöglich. Also ist $f(v_2) = -bv_1 + av_2$, und demnach ist die darstellende Matrix eine Drehmatrix. Nun ist

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} D(t) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = D(-t) = D(2\pi - t),$$

und natürlich ist die Matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ orthogonal, also braucht nur die Drehmatrizen $D(t_i)$ mit $0 < t_i < \pi$.

Wir betrachten nun den Fall, daß V einen f -invarianten Unterraum U besitzt mit $0 \subset U \subset V$. Dann ist nach (3) auch U^\perp f -invariant. Betrachten wir die Einschränkung von f auf U , so gibt es (nach Induktion) eine Orthonormalbasis u_1, \dots, u_m bezüglich derer $f|_U$ die angegebene Form hat; die darstellende Matrix sei mit B bezeichnet. Betrachten wir entsprechend die Einschränkung von f auf U^\perp , so gibt es (nach Induktion) eine Orthonormalbasis u_{m+1}, \dots, u_n bezüglich derer $f|_{U^\perp}$ die angegebene Form hat; hier sei die Matrix mit C bezeichnet. Bezüglich der Basis u_1, \dots, u_n hat f die Matrizen-Darstellung

$$A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix},$$

und mit B und C hat auch A die gewünschte Form.

Orthogonale Matrizen sind unitäre Matrizen: *Die orthogonalen Matrizen sind gerade die unitären Matrizen mit reellen Koeffizienten!*

Jede orthogonale Matrix A ist auch eine unitäre Matrix (jede reelle Matrix kann natürlich als komplexe Matrix aufgefaßt werden). Wie wir wissen, gibt es eine orthogonale Matrix P mit $P^{-1}AP = \bigoplus_{i=1}^m C_i$, wobei die Matrizen C_i (1×1) -Matrizen $[1]$ oder $[-1]$ oder (2×2) -Drehmatrizen $D(\alpha_i)$ zu Winkeln $0 < \alpha_i < \pi$ oder $\pi < \alpha_i < 2\pi$ sind.

Die Drehmatrix $D(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$ ist (als komplexe Matrix) ähnlich zur Diagonalmatrix mit den Diagonalkoeffizienten $\cos(\alpha) + i \sin(\alpha)$ und $\cos(\alpha) - i \sin(\alpha)$; es ist nämlich

$$\underbrace{\begin{bmatrix} \frac{-i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix}}_{\overline{P}^t} \cdot \underbrace{\begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}}_{D(\alpha)} \cdot \underbrace{\begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}}_P = \begin{bmatrix} \cos \alpha + i \sin \alpha & 0 \\ 0 & \cos \alpha - i \sin \alpha \end{bmatrix}$$

dabei ist $P = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{bmatrix}$ eine unitäre Matrix.

5.6. Normale Matrizen.

Der Grundkörper sei \mathbb{C} . Wir haben gesehen, dass es zu hermiteschen Matrizen wie auch zu unitären Matrizen eine Orthonormalbasis aus Eigenvektoren gibt, dabei sind für hermitesche Matrizen die Eigenwerte reell, bei unitären Matrizen haben sie den Betrag 1. Man kann nun fragen, wie man die Matrizen $A \in M(n \times n, \mathbb{C})$ charakterisieren kann, für die es eine Orthonormalbasis aus Eigenvektoren gibt, ohne Einschränkung an die Eigenwerte.

Sei $A \in M(n \times n, \mathbb{C})$. Wir nennen \overline{A}^t die zu A adjungierte Matrix. Man nennt A normal, wenn $A\overline{A}^t = \overline{A}^t A$ gilt, wenn also A mit seiner adjungierten Matrix vertauscht.

Beispiele: Jede hermitesche Matrix A ist normal, denn für eine hermitesche Matrix ist $\overline{A}^t = A$ (und A vertauscht mit sich selbst). Jede unitäre Matrix A ist normal, denn für eine unitäre Matrix A ist $\overline{A}^t = A^{-1}$ (und A vertauscht mit A^{-1}).

Hauptsatz über normale Matrizen. Sei $A \in M(n \times n, \mathbb{C})$. Genau dann ist A normal, wenn es eine Orthonormalbasis aus Eigenvektoren gibt (im kanonischen unitären Raum \mathbb{C}^n).

Der Beweis ist ähnlich zu den Beweisen der Hauptsätze für hermitesche und unitäre Matrizen.

Natürlich gibt es eine entsprechende Begriffsbildung und Behauptung auch für Endomorphismen. Sei $(V, \langle -, - \rangle)$ ein endlich-dimensionaler unitärer Raum mit Orthonormalbasis \mathcal{B} . Ist $f: V \rightarrow V$ ein Endomorphismus mit Matrizendarstellung $M_{\mathcal{B}}^{\mathcal{B}}(f) = A$, so definiert man den zu f adjungierten Endomorphismus f^* durch $M_{\mathcal{B}}^{\mathcal{B}}(f^*) = \overline{A}^t$. Dann gilt für alle $v, w \in V$

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle$$

(durch diese Gleichung ist f^* für einen gegebenen Endomorphismus f eindeutig bestimmt, daher nimmt man oft diese Gleichung zur Definition des zu f adjungierten Endomorphismus). Man nennt f normal, falls $f \circ f^* = f^* \circ f$ gilt (falls also f mit dem zu f adjungierten Endomorphismus f^* vertauscht).

Ein Endomorphismus eines endlich-dimensionalen unitären Vektorraums hat genau dann eine Orthogonalbasis aus Eigenvektoren, wenn er normal ist.

5.7. Iwasawa-Zerlegung (oder QR-Zerlegung).

Satz. Jede invertierbare Matrix $A \in M(n \times n, \mathbb{R})$ lässt sich eindeutig als Produkt

$$A = QR$$

mit einer orthogonalen Matrix A und einer oberen Dreiecksmatrix R mit positiven Diagonalkoeffizienten schreiben.

Satz. Jede invertierbare Matrix $A \in M(n \times n, \mathbb{C})$ lässt sich eindeutig als Produkt

$$A = QR$$

mit einer unitären Matrix A und einer oberen Dreiecksmatrix R mit positiven reellen Diagonalkoeffizienten schreiben.

Beweis der Existenz einer derartigen Zerlegung, im Fall \mathbb{C} . Sei v_1, \dots, v_n die Folge der Spalten von A , dies ist also eine Basis von \mathbb{C} . Wir wissen, dass wir jede linear unabhängige Folge von Vektoren orthonormalisieren können: induktiv bilden wir eine Linearkombination

$$u'_s = v_s + \sum_{r < s} \lambda'_{rs} v_r \in L(v_1, \dots, v_{s-1})^\perp,$$

die dann noch zu normalisieren ist: Sei

$$\lambda_{ss} = 1/||u'_s|| \quad \text{und} \quad \lambda_{rs} = \lambda'_{rs}/||u'_s||.$$

Setze

$$u_s = \lambda_{ss}v_s + \sum_{r<s} \lambda_{rs}v_r,$$

dieser Vektor entsteht also aus u'_s durch normieren und gehört wie u'_s zu $L(v_1, \dots, v_{s-1})^\perp$. Nach Konstruktion ist

$$L(v_1, \dots, v_s)^\perp = L(u_1, \dots, v_s)^\perp,$$

für alle s . Wir sehen also: Die Folge u_1, \dots, u_n ist eine Orthonormalbasis des kanonischen unitären Raums \mathbb{C}^n .

Sei $S = (\lambda_{rs})_{rs}$ und Q die Matrix mit Spaltenfolge u_1, \dots, u_n . Dann gilt nach Konstruktion

$$AS = Q,$$

und S ist eine obere Dreiecksmatrix mit positiven reellen Diagonalkoeffizienten.

Setzen wir $R = S^{-1}$, so ist auch R eine obere Dreiecksmatrix mit positiven reellen Diagonalkoeffizienten und es gilt

$$A = QR.$$

Ist $A \in M(n \times n, \mathbb{R})$, so liefert dieses Verfahren **reelle** Matrizen Q, R , also eine orthogonale Matrix Q und eine reelle obere Dreiecksmatrix R mit positiven Diagonalkoeffizienten.

Eindeutigkeit, im Fall $A \in M(n \times n, \mathbb{C})$: Sei $Q_1R_1 = A = Q_2R_2$ mit unitären Matrizen Q_1, Q_2 und mit oberen Dreiecksmatrizen R_1, R_2 mit positiven reellen Diagonalkoeffizienten. Aus $Q_1R_1 = Q_2R_2$ folgt

$$(Q_2)^{-1}Q_1 = R_2(R_1)^{-1},$$

wir nennen diese Matrix B . Die Matrix B ist, wie man links sieht, ein Produkt zweier unitärer Matrizen, also unitär. Wie man rechts sieht, ist B eine obere Dreiecksmatrix mit positiven Diagonalkoeffizienten. Rechts sieht man also, dass alle Eigenwerte von B positive reelle Zahlen sind; Eigenwerte einer unitären Matrix haben aber den Betrag 1. Es folgt, dass alle Eigenwerte von B gleich 1 sind. Eine unitäre Matrix mit einzigem Eigenwert 1 ist ähnlich zur Matrix E_n und somit sogar gleich E_n . Aus $B = E_n$ folgt aber $Q_1 = Q_2$ und $R_1 = R_2$.

Wie haben hier verwendet: Die Menge $\mathcal{U}(n)$ der unitären $(n \times n)$ -Matrizen, wie auch die Menge $\mathcal{R}(n, \mathbb{C})$ der oberen Dreiecksmatrizen in $M(n \times n, \mathbb{C})$ mit positiven reellen Diagonalkoeffizienten sind Untergruppen von $GL(n, \mathbb{C})$ (also abgeschlossen unter Produkt- und Inversenbildung).

Es sollte noch darauf hingewiesen werden, dass man oft die Matrix R noch weiter faktorisiert: man schreibt

$$R = DU$$

dabei ist D eine Diagonalmatrix mit positiven reellen Diagonalkoeffizienten und U eine obere Dreiecksmatrix mit nur Einsen als Diagonalkoeffizienten (man nennt eine solche Matrix eine *unipotente* obere Dreiecksmatrix).

Entsprechend zur QR-Zerlegung gibt es auch QL-, RQ-, LQ-Zerlegungen, auch diese sind alle eindeutig, dabei steht jeweils Q für orthogonal oder unitär, R für obere Dreiecksmatrix mit positiven reellen Diagonalkoeffizienten und L für untere Dreiecksmatrix mit positiven reellen Diagonalkoeffizienten: Die QL-Zerlegung erhält man, indem man die Folge der Spaltenvektoren v_1, \dots, v_n der Matrix A in der umgekehrten Reihenfolge orthonormalisiert (also man beginnt mit v_n , betrachtet danach v_{n-1} , und so weiter). Wendet man den Existenzsatz für die QR-Zerlegung auf A^t an so erhält man eine Zerlegung $A^t = QR$; transponieren liefert $A = R^t Q^t$, dabei ist dann R^t eine untere Dreiecksmatrix mit positiven reellen Diagonalkoeffizienten und Q^t ist wieder orthogonal oder unitär.

5.8. Hauptachsen.

Sei K ein Körper. Sei $A \in M(n \times n, K)$ eine symmetrische Matrix. Durch

$$q_A(x) = x^t A x$$

für $x \in K^n$ wird auf K^n eine quadratische Form definiert. Ist die Charakteristik von K ungleich 2, so erhält man auf diese Weise auch alle quadratischen Formen (siehe LA I, 3.2.1).

Sei nun $K = \mathbb{R}$, sei $A \in M(n \times n, \mathbb{R})$ symmetrisch. Der Hauptsatz über symmetrische reelle Matrizen liefert eine orthogonale Matrix P , sodass $P^{-1}AP$ eine Diagonalmatrix ist. Wegen $P^{-1} = P^t$ wissen wir also: P^tAP ist eine Diagonalmatrix. Wir wählen ein neues Koordinatensystem, und zwar nehmen wir die Spalten von P als neue Basisvektoren (da P orthogonal ist, ist dies eine Orthonormalbasis). Bezüglich dieser neuen Basis wird die quadratische Form q_A durch die Matrix P^tAP beschrieben (LA I, 3.2.3). Ist P^tAP die Diagonalmatrix mit den Diagonalkoeffizienten d_1, \dots, d_n , so gilt also:

Satz. *Zu jeder quadratischen Form auf \mathbb{R}^n gibt es eine Orthonormalbasis, sodass sie bezüglich dieses neuen Koordinatensystems durch eine quadratische Form $\sum_i d_i T_i^2$ (ohne gemischte Terme) beschrieben wird.*

Man nennt ein derartiges Koordinatensystem *Hauptachsen* für die gegebene quadratische Form.

Erinnerung. (LA I, 3.2.11). Ist die quadratische Form $\sum_i d_i T_i^2$ gegeben, so nennt man die Anzahl der positiven Koeffizienten d_i den *Positivitätsindex* r , die Anzahl der negativen Koeffizienten d_i den *Negativitätsindex* s , und die Anzahl der Koeffizienten $d_i = 0$ den *Entartungsindex* t . Das Tripel (r, s, t) (oder auch das Paar (s, t)) nennt man die *Signatur* der quadratischen Form.