Lineare Algebra II

Skript zur Vorlesung im

Sommersemester 2007
an der TU Berlin

Vorlesung: Dr. Jörg Liesen
(Raum MA 378, Email: liesen@math.tu-berlin.de)

Übungen: Dr. Michel Sortais
(Raum MA 783, Email: sortais@math.tu-berlin.de)

Skript: Prof. Volker Mehrmann

Inhaltsverzeichnis

Kapitel 1

Äquivalenzrelationen und Quotientenräume

In diesem Kapitel wollen wir uns einmal etwas abstrakter mit Äquivalenzrelationen beschäftigen. Damit legen wir die Grundlage für die Untersuchung von verschiedenen Normalformen. Zur Wiederholung:

Definition 1.1 *Sei M eine Menge. Eine Teilmenge R von $M \times M$ heißt* Äquivalenzrelation*, falls folgende Eigenschaften für beliebige $u, v, w \in M$ gelten:*

(a) Reflexivität: $(u, u) \in R$,

(b) Symmetrie: $(u, v) \in R \implies (v, u) \in R$,

(c) Transitivität: $(u, v) \in R$, $(v, w) \in R \implies (u, w) \in R$.

Falls $(u, v) \in R$, so schreiben wir $u \sim v$, u äquivalent v.

Beispiel 1.2

a) Sei $M = \mathbb{K}^{m,n}$. Es gelte $A \sim B$ genau dann wenn es invertierbare Matrizen P, Q gibt, so dass $A = PBQ$ (Äquivalenz von Matrizen).

   (i) $A \sim A$, mit $P = I_m, Q = I_n$.

   (ii) $A \sim B \Longrightarrow B \sim A$, denn $A = PBQ \Longrightarrow B = P^{-1}AQ^{-1}$.

   (iii) $A \sim B,\ B \sim C$,
   $A = PBQ,\ B = WCY,\ P, Q, W, Y$ invertierbar,
   $\Longrightarrow A = (PW)C(YQ)$.

b) Sei $\mathbb{K}$ ein Körper, V ein $\mathbb{K}$-Vektorraum und $U \subset V$ ein Untervektorraum. Setze für $v, w \in V$: $\quad v \sim w$ genau dann wenn $v - w \in U$.

(i) $v \sim v$, denn $v - v = 0 \in U$,

(ii) $v \sim w \implies v - w \in U \implies w - v = -(v - w) \in U$,

(iii) $v \sim w, w \sim z \implies v - w \in U,\ w - z \in U \implies v - z = (v - w) + (w - z) \in U$.

c) $M = \mathbb{K}^{n,n}$. Es gelte $A \sim B$ genau dann wenn es eine invertierbare Matrix P gibt, so dass $A = P^{-1}BP$ (Ähnlichkeit von Matrizen).

(i) $A \sim A$ mit $P = I$.

(ii) $A \sim B \implies A = P^{-1}BP \implies B = PAP^{-1}$.

(iii) $A \sim B,\ B \sim C \implies A = P^{-1}BP,\ B = Q^{-1}CQ,$
$A = (QP)^{-1}C(QP)$.

Definition 1.3 *Sei M eine Menge und R eine Äquivalenzrelation auf M. Eine Teilmenge $T \subset M$ heißt* Äquivalenzklasse *bezüglich R, falls gilt:*

(a) $T \neq \emptyset$,

(b) $u, v \in T \implies u \sim v$,

(c) $u \in T$, $v \in M$ *und* $u \sim v \implies v \in T$.

Lemma 1.4 *Ist R eine Äquivalenzrelation auf M, so gehört jedes $u \in M$ zu genau einer Äquivalenzklasse. Das heißt insbesondere, dass für zwei verschiedene Äquivalenzklassen T_1, T_2 gilt:*

$$T_1 \cap T_2 = \emptyset\,.$$

Insbesondere hat dann jede Äquivalenzklasse die Form $T_u = \{w \in M | u \sim w\}$ für ein $u \in M$.

Beweis: Sei $u \in M$. Dann ist $T_u = \{w \in M \mid u \sim w\}$ eine Äquivalenzklasse, denn

(a) $u \in T_u \implies T_u \neq \emptyset$,

(b) $w_1, w_2 \in T_u \implies w_1 \sim u, w_2 \sim u \overset{\text{Sym./Trans.}}{\implies} w_1 \sim w_2$,

(c) $w \in T_u$, $z \in M$, $w \sim z$,
$u \sim w, w \sim z \overset{\text{Trans.}}{\Longrightarrow} u \sim z$.

Umgekehrt folgt aus $T \neq \emptyset$, dass es $v \in T$ gibt und wegen b) und C) besteht T genau aus den Elementen die zu v äquivalent sind.

Also gehört jedes $u \in M$ zu einer Klasse. Seien T_u, T_v zwei Klassen. Falls $T_u \cap T_v \neq \emptyset$, so gibt es $w \in T_u \cap T_v$. Sei $u_1 \in T_u \Longrightarrow u_1 \sim w$. Da $w \in T_v$ und $w \sim u_1$, so folgt $u_1 \in T_v \Longrightarrow T_u \subset T_v$. Analog folgt $T_v \subset T_u \Longrightarrow T_u = T_v$. $\square$

Jede Äquivalenzrelation R auf einer Menge M liefert eine Zerlegung von M in disjunkte Äquivalenzklassen $T_i,\ i \in I$, d.h.,

$$M = \bigcup_{i \in I} T_i\,, \qquad T_i \cap T_j = \emptyset,\ i \neq j\,.$$

Dies heißt Klasseneinteilung von M.

Umgekehrt definiert jede Zerlegung einer Menge M in disjunkte Teilmengen eine Klasseneinteilung.

Definition 1.5

a) Sei R eine Äquivalenzrelation auf einer Menge M. Die Menge

$$M/R = \{T_i \mid i \in I, T_i \text{ Äquivalenzklasse bzgl. } R\}$$

heißt Quotientenmenge. *M/R ist eine Menge von Teilmengen von M.*

b) Sei T eine Äquivalenzklasse, so heißt jedes $t \in T$ ein Repräsentant *dieser Klasse.*

Zu M/R gibt es eine natürliche Abbildung

$$\Pi :\ M \to M/R,$$

die jedem $u \in M$ die Äquivalenzklasse zuordnet, zu der u gehört. Π ist offensichtlich surjektiv.

Beispiel 1.6 Sei V ein Vektorraum über dem Körper $\mathbb{K}$ und U ein Untervektorraum von V. Betrachte die Relation aus Beispiel 1.2 b).
Die Äquivalenzklassen sind alle Teilräume der Form

$$v + U, \quad v \in V.$$

Die Quotientenmenge bezüglich dieser Relation bezeichnen wir als V/U. Jede Teilmenge $v + U$ hat v als Repräsentanten.

Im folgenden werden wir Repräsentanten für die Relationen in Beispiel 1.2 a) und c) studieren.

Wir wollen aber zuert die Äquivalenzrelation für Unterräume U des Vektorraums V über $\mathbb{K}$, gegeben durch

$$u \sim v \quad \text{wenn} \quad u - v \in U$$

genauer betrachten.

Definition 1.7 *Sei V ein Vektorraum über dem Körper $\mathbb{K}$. Eine Menge $M \subset V$ heißt* affiner Unterraum *von V, falls $M = \emptyset$ oder $M = v + U$, wobei U ein Untervektorraum von V ist.*

Lemma 1.8 *Sei V ein Vektorraum über dem Körper $\mathbb{K}$.*

(a) Seien $v, \tilde{v} \in V$ und $U, \tilde{U}$ Untervektorräume von V, so gilt:

$$v + U = \tilde{v} + \tilde{U} \iff U = \tilde{U} \text{ und } v - \tilde{v} \in U.$$

(b) Sei $v \in V$, U Untervektorraum von V und $M = v + U$, so ist

$$U = \{m_1 - m_2 \mid m_1, m_2 \in M\}.$$

Beweis:

(a) „$\Longrightarrow$“ $v + U = \tilde{v} + \tilde{U}$. Da $0 \in \tilde{U}$, so folgt

$$\tilde{v} = v + u \text{ für ein } u \in U \quad \Rightarrow \quad \tilde{v} - v \in U.$$

Sei $\tilde{u} \in \tilde{U}$. Dann gibt es $u \in U$ mit

$$v + u = \tilde{v} + \tilde{u}, \text{ also } \tilde{u} = v - \tilde{v} + u \in U \quad \Rightarrow \quad \tilde{U} \subset U.$$

Umgekehrt folgt genauso $\tilde{U} \supset U$.

$$\Rightarrow \tilde{U} = U.$$

„$\Longleftarrow$“ Aus $m \in v + U$ folgt $m = v + u$ für ein $u \in U$. Sei $v - \tilde{v} = u_1 \in U$

$$\Rightarrow v = u_1 + \tilde{v} \Rightarrow m = u + u_1 + \tilde{v}.$$

Da $u + u_1 \in U$ und $U = \tilde{U}$, so folgt $m \in \tilde{v} + \tilde{U}$. Also ist $v + U \subset \tilde{v} + \tilde{U}$. Umgekehrt zeigt man analog $\tilde{v} + \tilde{U} \subset v + U$.

(b) Sei $u \in U$. Setze $m_1 = v + u \in M$ und $m_2 = v + 0 \in M$

$$\Rightarrow \; u = m_1 - m_2\,.$$

Umgekehrt sei $m_1 = v + u \in M,\ m_2 = v + \tilde{u} \in M$ mit $u,\ \tilde{u} \in U$.

$$\Rightarrow m_1 - m_2 = v + u - v - \tilde{u} = u - \tilde{u} \in U.$$

□

Man definiert für affine Räume die *Dimension* durch

$$\dim(v + U) := \dim U \qquad \text{und} \qquad \dim \emptyset = -1.$$

Geometrische Interpretation: Ist M affiner Teilraum von V, so ist das,

falls $\dim M = 0$, ein *Punkt*,
$\dim M = 1$, eine *Gerade*,
$\dim M = 2$, eine *Ebene*.

Falls $\dim M = n - 1$, $\dim V = n$, so heißt M *Hyperebene.*

Satz 1.9 *Sei $M \subset \mathbb{K}^m$ für einen Körper $\mathbb{K}$. Dann ist M ein affiner Unterraum von $\mathbb{K}^m$ genau dann, wenn M die Lösungsmenge eines linearen Gleichungssystems in m Unbekannten ist.*

Beweis: Sei M Lösungsmenge von $Ax = b$ mit $A \in \mathbb{K}^{n,m}$ und sei

$$U = \{u \in \mathbb{K}^m \mid Au = 0\}\,,$$

so haben wir schon gezeigt, dass U ein Untervektorraum ist. Damit ist

$$M = y + U \quad \text{für eine Lösung } y \text{ von } Ay = b.$$

Für die Umkehrung sei $v \in \mathbb{K}^m$ und $U \subset \mathbb{K}^m$ ein Untervektorraum, $M = v + U$.
Sei $\{u_1, \ldots, u_r\}$ Basis von U und $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$ Basis von $\mathbb{K}^m$.
Sei $\{\varphi_1, \ldots, \varphi_m\}$ duale Basis zu $\{u_1, \ldots, u_m\}$, d.h. $\varphi_i(u_j) = \delta_{ij}$. Dann gilt

$$U = \{w \in \mathbb{K}^m \mid \varphi_i(w) = 0 \;\text{ für } r + 1 \le i \le m\}, \;\text{ denn:}$$

Für $u = \sum_{j=1}^{r} \lambda_j u_j$ gilt

$$\varphi_i(u) = \sum_{j=1}^{r} \varphi_i(u_j) = 0 \text{ für } i > r.$$

Umgekehrt, sei $w \in \mathbb{K}^m$ mit $\varphi_i(w) = 0$ für $i > r$ und $w = \sum_{j=1}^{m} \lambda_j u_j$, so folgt

$$\varphi_i(w) = \lambda_i \;\Rightarrow\; \lambda_i = 0 \;\text{ für }\; i > r,$$

oder $w \in U$.

Sei $a_i \in \mathbb{K}^{1,m}$ die Matrixdarstellung von φ_i bezüglich der kanonischen Basen in $\mathbb{K}^m, \mathbb{K}$, und

$$A = \begin{bmatrix} a_{r+1} \\ \vdots \\ a_m \end{bmatrix} \in \mathbb{K}^{m-r,m},$$

so ist U Lösungsmenge von $Ax = 0$, dargestellt in der kanonischen Basis.

D.h., es gilt

$$\text{Kern}\,(A) = \bigcap_{i=r+1}^{m} \text{Kern}\,(\varphi_i) = U.$$

Setze $b = Av$, so ist $M = v + U$ Lösungsmenge von $Ax = b$.

Es bleibt der Fall $M = \emptyset$. Dann wähle

$$A = 0 \in \mathbb{K}^{1,m}\,,\; b \neq 0 \in \mathbb{K}.$$

$\square$

Definition 1.10 *Sei V Vektorraum über $\mathbb{K}$ und U Unterraum von V. Betrachte die Äquivalenzklassen V/U, d.h., die Menge aller affinen Teilräume der Form $v + U$ mit $v \in V$. Wenn man auf V/U noch die Addition $(v + U) \oplus (v' + U) = (v + v') + U$ und die sklalare Multiplikation $\lambda \otimes (v + U) = \lambda v + U$ definiert so ist V/U ein Vektorraum und heißt* Quotientenraum von V modulo U.

Wie schon vorher definiert, gibt es eine surjektive Abbildung

$$\begin{aligned} \Pi \;\; &: V \longrightarrow V/U \qquad\qquad \text{mit Kern}\,(\Pi) = U. \\ &: v \longmapsto v + U \end{aligned}$$

Es ist klar, dass Π linear ist und dass $\dim V/U = \dim V - \dim U$.

Weitere Eigenschaften des Quotientenraums.

Satz 1.11

Sei V eine Vektorraum über dem Körper $\mathbb{K}$.

*(a) Sei $\{v_1, \ldots, v_s\}$ eine Basis von V,
so bilden die verschiedenen Elemente von $\{v_1 + U, \ldots, v_s + U\}$ eine Basis von V/U.*

(b) Ist $V = U + \mathcal{L}(v_1, \ldots, v_t)$, so ist $\quad V/U = \mathcal{L}(v_1 + U, \ldots, v_t + U)$.

(c) Sind $v_1, \ldots, v_s \in V$ und sind $v_1 + U, \ldots, v_s + U$ linear unabhängig in V/U, so sind $v_1, \ldots, v_s$ linear unabhängig in V.

(d) Seien $v_1, \ldots, v_s \in V$ und seien $v_1 + U, \ldots, v_s + U$ linear unabhängig in V/U. Sind $u_1, \ldots, u_r$ linear unabhängig in U, so sind $u_1, \ldots, u_r$, $v_1, \ldots, v_s$ linear unabhängig in V.

(e) Sei $\{u_1, \ldots, u_r\}$ Basis von U, $v_1, \ldots, v_s \in V$. $\{u_1, \ldots, u_r, v_1, \ldots, v_s\}$ ist Basis von V genau dann, wenn $\{v_1 + U, \ldots, v_s + U\}$ Basis von V/U ist .

Beweis: (a) ... (d) Übung.

(e) „$\Longrightarrow$“ Wegen (b) ist $\mathcal{L}(v_1 + U, \ldots, v_s + U) = V/U$.
Da $\{u_1, \ldots, u_r, v_1, \ldots, v_s\}$ eine Basis von V bilden, gilt

$$v_i - v_j \notin U, \qquad i, j = 1, \ldots, s, \ i \neq j.$$

Damit sind alle Elemente von $\{v_1 + U, \ldots, v_s + U\}$ voneinander verschieden. Aus (a) folgt nun die lineare Unabhängigkeit von $v_1 + U, \ldots, v_s + U$.

„$\Longleftarrow$“ Nach (d) sind $u_1, \ldots, u_r, v_1, \ldots, v_s$ linear unabhängig in V. $\mathcal{L}(u_1, \ldots, u_r, v_1, \ldots, v_s) = V$, da $\dim V = r + s$. Also folgt die Behauptung.

$\square$

Kapitel 2

Eigenwerte und Eigenvektoren von Endomorphismen

In diesem Kapitel wollen wir beginnen uns mit der Normalform von Endomorphismen/Matrizen unter Ähnlichkeit zu beschäftigen. Zuvor jedoch erinnern wir uns noch einmal an die Äquivalenz von Matrizen.

Definition 2.1 *Seien $n, m \in \mathbb{N}$. Zwei Matrizen $A, B \in \mathbb{K}^{n,m}$ heißen* äquivalent, *falls es invertierbare Matrizen P, Q gibt, $P \in \mathbb{K}^{n,n}, Q \in \mathbb{K}^{m,m}$, so dass*

$$A = PBQ^{-1}.$$

Wir wissen bereits, dass die Ränge von A und B gleich sind, wenn A und B äquivalent sind. Aus Satz 4.2 wissen wir, dass A (und damit auch B) äquivalent ist zu $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, wobei $r = \text{Rang}\,(A) = \text{Rang}\,(B)$. Betrachte nun die Menge

$$M_r = \left\{ A \in \mathbb{K}^{n,m} \;\middle|\; A \text{ äquivalent zu } \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \right\} \text{ mit } 0 \leq r \leq \min(m, n).$$

Dann folgt

$$\begin{aligned} \mathbb{K}^{n,m} &= \bigcup_{r=0}^{\min(m,n)} M_r \\ &= M_0 \cup M_1 \cup \ldots \cup M_{\min(m,n)}\,. \end{aligned}$$

Die Matrix $N_r = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{K}^{n,m}$ heißt *Normalform* von $A \in M_r$ bezüglich der Relation Äquivalenz. Wir wollen nun die Normalform unter einer anderen

Äquivalenzrelation betrachten, nämlich unter der Ähnlichkeit für Matrizen in $\mathbb{K}^{n,n}$ (s. Def. 6.8). Dies ist für allgemeine Körper ein schwieriges Problem, wir werden uns daher auf $\mathbb{K} = \mathbb{C}$ beschränken.

Wir erinnern uns daran, dass ein Endomorphismus $f : V \to V$ (V Vektorraum über $\mathbb{K}$) eine lineare Abbildung (Homomorphismus) von V in sich ist. Ist $\{v_1, \ldots, v_n\}$ eine Basis von V, so besitzt f eine Matrixdarstellung bezüglich $\{v_1, \ldots, v_n\}$. Aus Lemma 11.2 folgt, dass alle Matrixdarstellungen von f ähnlich sind, denn wenn $\{v_1, \ldots, v_n\}, \{\tilde{v}_1, \ldots, \tilde{v}_n\}$ Basen von V sind mit der Basisübergangsmatrix P und wenn F die Matrixdarstellung von f bezüglich $\{v_1, \ldots, v_n\}$ ist , und Z die Matrixdarstellung von f bezüglich $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$, so folgt

$$Z = PFP^{-1}.$$

Wir können dann auch wieder $\mathbb{K}^{n,n}$ darstellen als disjunkte Vereinigung von Mengen

$$J_i = \{A \in \mathbb{K}^{n,n}, A \text{ ähnlich zu } A_i\}$$

und versuchen, eine Normalform und entsprechende Repräsentanten zu bestimmen.

Zur Erinnerung: Ähnliche Matrizen haben das gleiche charakteristische Polynom (Satz 6.9), d.h., unabhängig von der Matrixdarstellung gibt es zu einem Endomorphismus f ein *charakteristisches Polynom* $P_f(x) = P_F(x)$, wobei F eine Matrixdarstellung von f ist.

Definition 2.2 *Sei $f : V \to V$ ein Endomorphismus, V ein Vektorraum über dem Körper $\mathbb{K}$. Ein Vektor $v \in V \setminus \{0\}$ heißt* Eigenvektor *von f mit* Eigenwert *$\lambda \in \mathbb{K}$, falls $f(v) = \lambda v$.*

Zur Erinnerung: Ist $A \in \mathbb{K}^{n,n}$, so ist $v \neq 0$ Eigenvektor mit Eigenwert λ, falls $Av = \lambda v, \lambda \in \mathbb{K}$.

Jede quadratische Matrix A definiert einen Endomorphismus und es gilt, dass v, λ Eigenvektor und Eigenwert dieses Endomorphismus sind. Die Eigenwerte dieses Endomorphismus sind Nullstellen des charakteristischen Polynoms. Daraus folgt, dass wenn $\dim V = n$ ist, so hat f höchstens n Eigenwerte.

Lemma 2.3 *Sei $f : V \to V$ ein Endomorphismus des Vektorraums V und sei $\{v_1, \ldots, v_n\}$ Basis von V. Die Vektoren $v_1, \ldots, v_n$ sind Eigenvektoren von f genau dann, wenn die Matrixdarstellung von f bezüglich $\{v_1, \ldots, v_n\}$ eine Diagonalmatrix ist.*

Beweis: Seien $\lambda_i \in \mathbb{K}$ für $1 \leq i \leq n$ Eigenwerte zu Eigenvektoren v_i, also $f(v_i) = \lambda_i v_i$, für $1 \leq i \leq n$. Dann ist die Matrixdarstellung F von f bezüglich $\{v_1, \ldots, v_n\}$ die Diagonalmatrix

$$F = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Umgekehrt sei diese Matrixdarstellung F von f bezüglich $\{v_1, \ldots, v_n\}$ als Diagonalmatrix gegeben. Dann gilt $f(v_i) = \lambda_i v_i\,,\ i = 1, \ldots, n$. Da $\{v_1, \ldots, v_n\}$ eine Basis ist, sind $v_1, \ldots, v_n \neq 0$, also Eigenvektoren. □

Definition 2.4 *Sei $f : V \to V$ ein Endomorphismus eines Vektorraums V über einem Körper $\mathbb{K}$. Falls es eine Basis $\{v_1, \ldots, v_n\}$ von V aus Eigenvektoren zu f gibt, so heißt f* diagonalisierbar*. Entsprechend heißt eine Matrix $A \in \mathbb{K}^{n,n}$* diagonalisierbar*, falls sie ähnlich zu einer Diagonalmatrix ist.*

Beispiel 2.5 Die Matrix

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & -2 \\ 0 & 2 & 0 \\ 4 & 0 & 5 \end{bmatrix} \end{aligned}$$

ist diagonalisierbar, aber die Matrix

$$B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

ist nicht diagonalisierbar, denn sei

$$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix},\ P^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

so folgt

$$\begin{aligned} PBP^{-1} &= \frac{1}{ad - bc} \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \\ &= \frac{1}{ad - bc} \begin{bmatrix} -ca & a^2 \\ -c^2 & ca \end{bmatrix}. \end{aligned}$$

Diese Matrix kann nur diagonal sein, falls $a = 0$ und $c = 0$. Dann ist aber P nicht invertierbar.

Satz 2.6 *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums V über dem Körper $\mathbb{K}$. Sei v_i Eigenvektor von f mit Eigenwert $\lambda_i \in \mathbb{K}$ für $1 \leq i \leq r$. Sind die Zahlen λ_i paarweise verschieden, d.h., $\lambda_i \neq \lambda_j$ für $i \neq j$, so sind $v_1, \ldots, v_r$ linear unabhängig.*

Beweis: Wir verwenden Induktion nach r.

I.A.: $r = 1$ ist klar, da Eigenvektoren von Null verschieden sind.

I.V.: Behauptung sei richtig für $1 \leq k \leq r - 1$.

I.S.: Sei $r > 1$ und $\sum\limits_{i=1}^{r} \alpha_i v_i = 0$ mit $\alpha_1, \ldots, \alpha_r \in \mathbb{K}$.
Dann gilt

$$\begin{aligned} 0 = f(0) &= f\left(\sum_{i=1}^{r} \alpha_i v_i\right) = \sum_{i=1}^{r} \alpha_i f(v_i) \\ &= \sum_{i=1}^{r} \alpha_i \lambda_i v_i \ , \end{aligned}$$

und da

$$\begin{aligned} 0 &= \lambda_r \cdot 0 = \sum_{i=1}^{r} \alpha_i \lambda_r v_i, \quad \text{so folgt} \\ 0 &= \sum_{i=1}^{r} \alpha_i \lambda_i v_i - \sum_{i=1}^{r} \alpha_i \lambda_r v_i \\ &= \sum_{i=1}^{r-1} \alpha_i (\lambda_i - \lambda_r) v_i \ . \end{aligned}$$

Nach I.V. sind $v_1, \ldots, v_{r-1}$ linear unabhängig, also folgt $\alpha_i(\lambda_i - \lambda_r) = 0$ und damit $\alpha_i = 0$ für $i = 1, \ldots, r-1$. Dann folgt aus

$$0 = \sum_{i=1}^{r} \alpha_i v_i = \alpha_r v_r$$

auch dass $\alpha_r = 0$, da $v_r \neq 0$.

$\square$

Korollar 2.7

a) *Sei $f : V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums V über einem Körper $\mathbb{K}$, $\dim V = n$. Besitzt f die n paarweise verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_n$, so ist f diagonalisierbar.*

b) *Sei $A \in \mathbb{K}^{n,n}$. Besitzt A die n paarweise verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_n$, so ist A diagonalisierbar.*

Beispiel 2.8 Die Umkehrung von Korollar 2.7 gilt nicht, denn z. B. die Matrix λI_n hat n mal den Eigenwert λ, ist aber diagonalisierbar. Ob eine Matrix diagonalisierbar ist, hängt, da die Eigenwerte Nullstellen des charakterischen Polynoms sind, auch stark vom Körper $\mathbb{K}$ ab. Sei z.B.

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{R}^{2,2},$$

so ist $P_A(\lambda) = \lambda^2 + 1$ und es gibt keine Eigenwerte in $\mathbb{R}$. Wenn wir A aber als Element von $\mathbb{C}^{2,2}$ auffassen, so ist A ähnlich zu

$$\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Definition 2.9 *Sei V ein endlichdimensionaler Vektorraum über einem Körper $\mathbb{K}$, $\dim V = n$. Sei $f : V \to V$ ein Endomorphismus und $\lambda \in \mathbb{K}$ Eigenwert von f.*

a) *Dann heißt $E_\lambda = \mathrm{Kern}\,(\lambda\ Id - f)$ der* Eigenraum *zu λ. Die Dimension von E_λ, $g(\lambda) = \dim E_\lambda$, heißt* geometrische Vielfachheit *von λ.*

b) *Sei $P_f(x)$ das charakteristische Polynom von f. Dann ist $x - \lambda$ ein Teiler von $P_f(x)$. Mit $a(\lambda)$ bezeichnen wir die größte Zahl, so dass $(x - \lambda)^{a(\lambda)}$ ein Teiler von $P_f(x)$ ist. $a(\lambda)$ heißt* algebraische Vielfachheit *von λ.*

Eine analoge Definition und Sätze analog zu den folgenden gelten natürlich auch für quadratische Matrizen.

Lemma 2.10 *Sei V ein endlichdimensionaler Vektorraum über einem Körper $\mathbb{K}$, $\dim V = n$. Sei $f : V \to V$ ein Endomorphismus und $\lambda \in \mathbb{K}$ Eigenwert von f, so ist die algebraische Vielfachheit größer gleich der geometrischen, d.h.*

$$a(\lambda) \geq g(\lambda).$$

Beweis: Sei $r = g(\lambda)$ und $\{v_1, \ldots, v_r\}$ eine Basis von E_λ. Nach dem Basisergänzungssatz können wir $v_{r+1}, \ldots, v_n$ bestimmen, so dass $\{v_1, \ldots, v_n\}$ eine Basis von V ist. Sei B die Matrixdarstellung von f in der Basis $\{v_1, \ldots, v_n\}$. Dann ist $P_f(x) = P_B(x)$. Da $f(v_i) = \lambda v_i$ für $1 \leq i \leq r$ und $f(v_j) = \sum\limits_{i=1}^{n} b_{ij} v_i$, so hat B die Form

$$B = \begin{array}{cc} \left[\begin{array}{c|c} \lambda I_r & B_{12} \\ \hline 0 & B_{22} \end{array}\right] & \begin{array}{c} {\scriptstyle r} \\ {\scriptstyle n-r} \end{array} \\ \begin{array}{cc} {\scriptstyle r} & \quad {\scriptstyle n-r} \end{array} & \end{array} .$$

Also folgt $P_f(x) = P_B(x) = (x - \lambda)^r \cdot P_{B_{22}}(x)$. Insbesondere ist damit $a(\lambda) \geq g(\lambda) = r$. $\square$

Beispiel 2.11

$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ hat offensichtlich den Eigenwert 0 mit algebraischer Vielfachheit 2. Aber

$$\text{Kern}\,(0 \cdot I - A) = \text{Kern}\,(-A) = \text{Kern}\,(A) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

hat Dimension 1, denn $Ax = \begin{bmatrix} x_2 \\ 0 \end{bmatrix} = 0$ genau dann wenn $x_2 = 0$.

Damit erhalten wir nun **einen der wichtigsten Sätze über Eigenwerte**.

Satz 2.12 *Sei V ein endlichdimensionaler Vektorraum ($\dim V = n$) über dem Körper $\mathbb{K}$ und $f : V \to V$ ein Endomorphismus. Dann sind die folgenden Bedingungen äquivalent:*

(a) f ist diagonalisierbar.

(b) (i) $P_f(x)$ zerfällt in Linearfaktoren, d.h., $P_f(x) = \prod\limits_{i=1}^{n} (x - \lambda_i)$.
(ii) Für alle Eigenwerte λ von f gilt $a(\lambda) = g(\lambda)$.

(c) Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f, so gilt

$$V = \bigoplus_{i=1}^{r} E_{\lambda_i} = E_{\lambda_1} \dot{+} E_{\lambda_2} \dot{+} \ldots \dot{+} E_{\lambda_r} .$$

Beweis:
„$(a) \Longrightarrow (c)$“
Sei $\{v_1, \dots, v_n\}$ Basis von V aus Eigenvektoren von f. Fasse davon jeweils alle Eigenvektoren zum gleichen Eigenwert λ_i zusammen als $v_{i_1}, \dots, v_{i_l}$. Da die v_j eine Basis bilden, sind sie linear unabhängig, und damit gilt

$$E_{\lambda_i} = \mathcal{L}(v_{i_1}, \dots, v_{i_l}) \text{ und } V = E_{\lambda_1} \dot{+} \dots \dot{+} E_{\lambda_r}.$$

„$(c) \Longrightarrow (a)$“
Wähle Basen $\{v_{i_1}, \dots, v_{i_{l_i}}\}$ aus E_{λ_i}. Dann gilt $f(v_{i_j}) = \lambda_i v_{i_j}$ für $j = 1, \dots, l_i$. Es folgt, dass

$$\{v_{1_1}, \dots, v_{1_{l_1}}, v_{2_1}, \dots, v_{2_{l_2}}, \dots, v_{r_1}, \dots, v_{r_{l_r}}\}$$

eine Basis aus Eigenvektoren von V ist, da $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$, falls $i \neq j$. Denn, seien $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen

Eigenwerte von f und sei $\{v_{i_j},\ j = 1, \dots, g(\lambda_i)\}$ eine Basis von E_{λ_i}, $i = 1, \dots, r$, so sind nach Satz 2.6 die v_{i_j} linear unabhängig und die Gesamtdimension ist n.

„$(c) \Longrightarrow (b)$“
Aus (c) folgt, dass

$$\sum_{i=1}^{r} g(\lambda_i) = n,$$
$$\sum_{i=1}^{r} g(\lambda_i) \le \sum_{i=1}^{r} a(\lambda_i) = \text{grad}\,(P_f) = n\,,$$

also zerfällt P_f in Linearfaktoren und $a(\lambda_i) = g(\lambda_i), i = 1, \dots, r$.

„$(b) \Longrightarrow (c)$“ folgt trivialerweise.

$\square$

Beispiel 2.13 Sei $f = f_A$ mit $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix}$. Dann gilt

$$\begin{aligned} P_A(x) &= \det \begin{bmatrix} x & -1 & 0 \\ 1 & x-2 & 0 \\ 1 & -1 & x-1 \end{bmatrix} \\ &= (x-1)(x(x-2)+1) \\ &= (x-1)(x^2-2x+1) \\ &= (x-1)^3. \end{aligned}$$

Damit zerfällt P_A in Linearfaktoren. Wir müssen noch $g(1)$ und $a(1)$ bestimmen. Natürlich ist $a(1) = 3$.

$$B = 1 \cdot I - A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

hat Rang 1, also $\dim(\mathrm{Kern}\,(B)) = 2$ und damit $g(1) < a(1)$. f ist nicht diagonalisierbar.

Andererseits hat die Matrix

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 & 1 \\ -2 & 1 & 2 \\ -2 & -2 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \end{aligned}$$

3 verschiedene Eigenwerte und ist damit diagonalisierbar.

Satz 2.14 *Seien $A, B \in \mathbb{K}^{n,n}$. Dann gilt:*

(i) $A^\top$ hat dieselben Eigenwerte wie A.

(ii) A hat den Eigenwert 0 genau dann, wenn A singulär ist.

(iii) AB und BA haben dieselben Eigenwerte.

Beweis:

(i) $\det(A - \lambda I) = \det(A - \lambda I)^\top = \det(A^\top - \lambda I)$.

(ii) $\det(A - 0I) = \det A$.

(iii) Da $\det(AB) = \det A \cdot \det B = \det(BA)$ ist, so ist 0 Eigenwert von AB genau dann, wenn 0 Eigenwert von BA ist.

Sei $\lambda \neq 0$ ein Eigenwert von AB, d.h., $ABv = \lambda v$ für $v \neq 0$.
Setze $z = Bv$ dann gilt $Az = \lambda v \neq 0$ und damit $z \neq 0$ und $BAz = BABv = B\lambda v = \lambda Bv = \lambda z$. Also ist $z \neq 0$ Eigenvektor von BA mit Eigenwert λ und damit λ auch Eigenwert von BA.

Das gleiche Argument für einen Eigenwert $\mu \neq 0$ von BA liefert dann, dass jeder Eigenwert von BA auch Eigenwert von AB ist. Das komplettiert den Beweis.

$\square$

Zwar kann man nicht jede komplexe Matrix diagonalisieren, aber jede komplexe Matrix ist ähnlich (sogar unitär-ähnlich) zu einer Dreiecksmatrix.

Satz 2.15 ***Satz von Schur***
Sei $A \in \mathbb{C}^{n,n}$, so gibt es eine unitäre Matrix $U \in \mathbb{C}^{n,n}$ ($U^H U = I$), so dass

$$U^{-1}AU = U^H AU = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{bmatrix} \in \mathbb{C}^{n,n}$$

Beweis: Per Induktion.

I.A.: $n = 1$ klar.

I.V.: Die Behauptung sei richtig für $A \in \mathbb{C}^{n-1,n-1}$.

I.S.: Sei $A \in \mathbb{C}^{n,n}$ und sei $v \in \mathbb{C}^n \setminus \{0\}$ Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$. Dann existiert eine unitäre Matrix U_0, so dass $U_0^H v = e_1 \cdot \|v\|_2$, z.B. eine Spiegelung. Aus $Av = \lambda v$ folgt, dass

$$\underbrace{U_0^H AU_0}_{B=[b_{ij}]} \underbrace{U_0^H v}_{e_1\|v\|_2} = \lambda \underbrace{U_0^H v}_{e_1\|v\|_2}$$

und damit

$$\begin{bmatrix} b_{11} \\ \vdots \\ b_{n1} \end{bmatrix} \|v\|_2 = \lambda e_1 \|v\|_2$$

also $b_{21} = \cdots = b_{n1} = 0$ und $b_{11} = \lambda$. Also gilt $U_0^H AU_0 = B = \left[\begin{array}{c|c} \lambda & * \\ \hline 0 & B_1 \end{array}\right]$.
Nach I.V. gibt es eine unitäre Matrix U_1, so dass

$$U_1^H B_1 U_1 = B_2 \quad \text{obere Dreiecksmatrix ist.}$$

Setze $U = U_0 \left[\begin{array}{c|c} 1 & \\ \hline & U_1 \end{array}\right]$, so folgt dass

$$\begin{aligned} U^H AU &= \begin{bmatrix} 1 & 0 \\ 0 & U_1^H \end{bmatrix} B \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} \\ &= \left[\begin{array}{c|c} \lambda & * \\ \hline & B_2 \end{array}\right] \end{aligned}$$

obere Dreiecksmatrix ist.

□

Dieser Satz gilt für reelle Matrizen im allgemeinen nicht, weil es nicht unbedingt reelle Eigenwerte geben muss.

Beispiel 2.16 Für $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ gibt es keine reelle orthogonale Matrix U, so dass $U^\top AU = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}$, denn U hätte die Form

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \qquad \text{oder} \qquad \begin{bmatrix} c & s \\ s & -c \end{bmatrix},$$

mit $c = \cos\varphi, s = \sin\varphi$. Aber

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} -s & c \\ -c & -s \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

und

$$\begin{bmatrix} c & s \\ s & -c \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} c & s \\ s & -c \end{bmatrix} = \begin{bmatrix} c & s \\ s & -c \end{bmatrix} \begin{bmatrix} s & -c \\ -c & -s \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Ein anderes Argument ist, dass $\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} = \lambda^2 + 1$ Nullstellen $-i, i$ hat, aber $U^\top AU$ ist reell, also sind r_{11}, r_{22} reell und dies ist ein Widerspruch.

Kapitel 3

Nilpotente Matrizen / Endomorphismen

Für die weitere Untersuchung der Normalform unter Ähnlichkeit ist es günstig zuerst eine spezielle Klasse von Matrizen zu betrachten.

Definition 3.1 *Ein Endomorphismus $f : V \to V$ (V endlichdimensionaler Vektorraum über einem Körper $\mathbb{K}$) heißt* nilpotent, *falls $f^j = \underbrace{f \circ f \circ \ldots \circ f}_{j\text{-mal}} \equiv 0$ für ein $j \in \mathbb{N}$.*
Das kleinste j, für das $f^j \equiv 0$, heißt Nilpotenzindex *von f.*

Ist F eine Matrixdarstellung von f, so heißt F nilpotent, *falls $F^j = 0$ für $j \in \mathbb{N}$.*

Beispiel 3.2

$$F = \begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix} \in \mathbb{K}^{n,n}$$

ist nilpotent vom Nilpotenzindex n, denn

$$F^i = \begin{array}{cl} \left[\begin{array}{c|c} 0 & \begin{smallmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{smallmatrix} \\ \hline 0 & 0 \end{array}\right] & \begin{array}{l} \\ \left.\vphantom{\begin{matrix}0\\0\end{matrix}}\right\} i \end{array} \\ \underbrace{\hphantom{0000}}_{i}\hphantom{0000000} & \end{array}, \qquad F^{n-1} = \begin{bmatrix} & & 1 \\ & 0 & \\ & & \end{bmatrix}.$$

Lemma 3.3

(a) *Sei $A \in \mathbb{K}^{n,n}$ obere Dreiecksmatrix, so ist A nilpotent genau dann, wenn alle Diagonalelemente 0 sind. Der Nilpotenzindex ist höchstens n, kann aber kleiner sein.*

(b) *$A \in \mathbb{K}^{n,n}$ ist nilpotent und diagonalisierbar genau dann, wenn $A = 0$.*

(c) *Ist $A \in \mathbb{K}^{n,n}$ nilpotent, so hat A nur den Eigenwert 0.*

Beweis:

(a) Für eine Dreiecksmatrix A gilt stets

$$A^j = \begin{bmatrix} a_{11}^j & & * \\ & \ddots & \\ 0 & & a_{nn}^j \end{bmatrix}.$$

Also kann A nur nilpotent sein, falls alle $a_{ii} = 0$ sind, $i = 1, \cdots, n$. Umgekehrt, ist A strikte obere Dreiecksmatrix, so folgt per Induktion, dass A^j die Form

$$\begin{array}{cl} \left[\begin{array}{c|c} 0 & \\ \hline 0 & 0 \end{array}\right] & \begin{array}{l} \\ \} j \end{array} \\ \underbrace{\hphantom{00}}_{j}\hphantom{000} & \end{array}$$

hat, und damit $A^n = 0$.

(b) Falls $A = 0$, so ist A nilpotent vom Index 1 und diagonalisierbar.

Sei A diagonalisierbar und nilpotent, dann gibt es eine invertierbare Matrix P mit

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

und da

$$(P^{-1}AP)^j = P^{-1}(A^j)P = \begin{bmatrix} \lambda_1^j & & \\ & \ddots & \\ & & \lambda_n^j \end{bmatrix},$$

so folgt aus $A^j = 0$, dass $\begin{bmatrix} \lambda_1^j & & \\ & \ddots & \\ & & \lambda_n^j \end{bmatrix} = 0$, dass $\lambda_k = 0$ für alle $k = 1, \ldots, n$ und damit $A = P0P^{-1} = 0$.

(c) Angenommen, es gäbe einen anderen Eigenwert $\lambda \neq 0$ und zugehörigen Vektor $v \neq 0$, mit $Av = \lambda v$. Da A nilpotent ist, so gibt es ein $j \in \mathbb{N}$ mit $A^j = 0$ und $A^{j-1} \neq 0$. Damit folgt $0 = A^j v = A^{j-1}\lambda v = \lambda^j v$ also $\lambda = 0$ oder $v = 0$ und dies ist ein Widerspruch. $\square$

Wir wollen nun alle Formen nilpotenter Matrizen konstruieren?

Definition 3.4 *Eine* Partition *von n in s Teile ist eine Folge natürlicher Zahlen*

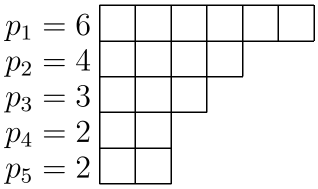
$$p_1 \geq p_2 \geq p_3 \geq \ldots \geq p_s$$

mit $\sum_{i=1}^{s} p_i = n$.

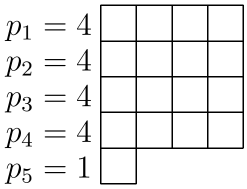
Die Darstellung erfolgt am besten durch Kästchendiagramme:

$$\begin{array}{ll} p_1 & \square\square\square\square\square\square\square \\ p_2 & \square\square\square\square\square \\ \vdots & \quad \cdots \\ p_s & \square \end{array}$$

Beispiel 3.5 $n = 17,\ s = 5.$



oder:



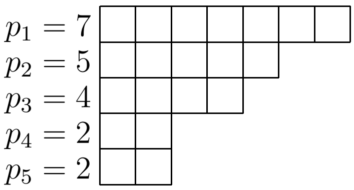
.

Es gibt sehr viele verschiedene Partitionen, z.B. für $n = 50$ gibt es 204 226 Partitionen und für $n = 10$ gibt es 42.

Definition 3.6 *Ist $p = (p_1, \ldots, p_s)$ eine Partition, so erhält man die zu p* duale *Partition, $p^* = (q_1, \ldots, q_t)$, indem man für q_j die Anzahl derjenigen p_i mit $p_i \geq j$ setzt.*

Dann sieht man sofort, dass q_j gerade die Anzahl der Kästchen in der j-ten Spalte ist und außerdem $t = p_1,\ s = q_1,\ (p^*)^* = p.$

Beispiel 3.7 $n = 20, s = 5.$



$$q_1 = 5,\ q_2 = 5,\ q_3 = 3,\ q_4 = 3,\ q_5 = 2,\ q_6 = 1,\ q_7 = 1\,.$$

$$\begin{aligned} (7)^* &= (1,1,1,1,1,1,1),\\ (7,5)^* &= (2,2,2,2,2,1,1),\\ (5,4)^* &= (2,2,2,2,1),\\ (4,2,2)^* &= (3,3,1,1). \end{aligned}$$

Zu jeder Partition p ordnen wir eine spezielle nilpotente Matrix zu:

$$N(p) = \begin{bmatrix} \boxed{\begin{smallmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{smallmatrix}} & & & \\ & \boxed{\begin{smallmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{smallmatrix}} & & \\ & & \ddots & \\ & & & \boxed{\begin{smallmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & 0 \end{smallmatrix}} \end{bmatrix} \begin{matrix} \}p_1 \\ \}p_2 \\ \vdots \\ \}p_s \end{matrix} =: \begin{bmatrix} N(p_1) & & & \\ & N(p_2) & & \\ & & \ddots & \\ & & & N(p_s) \end{bmatrix}.$$

Satz 3.8

(a) Seien $p, \tilde{p}$ Partitionen von n. Die Matrizen $N(p)$ und $N(\tilde{p})$ sind nur dann ähnlich, wenn $p = \tilde{p}$.

(b) Sei $A \in \mathbb{K}^{n,n}$ mit Nilpotenzindex m. Mit $q_i :=$ Rang (A^{i-1}) $-$ Rang (A^i) ergibt sich als $q = (q_1, \ldots, q_m)$ eine Partition von n und A ist ähnlich zu $N(p)$ mit $p = q^*$.

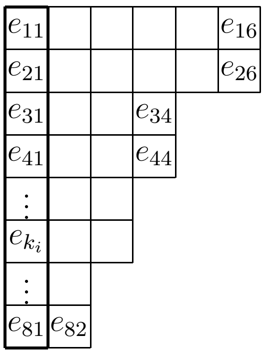
Beweis:

(a) Sei $p = (p_1, \ldots, p_s)$ eine Partition von n und $N(p)$ die zugehörige nilpotente Matrix. Da $N(p)$ Nullspalten hat, gibt es Indizes $k_1, \ldots, k_s$ mit:

$$N(p)e_{k_i} = 0, \qquad k_i = \sum_{j=1}^{i-1} p_i + 1, \quad i = 2, \ldots, s, \ k_1 = 1.$$

Ansonsten gilt für $j \notin \{k_1, \ldots, k_s\}$: $N(p)e_j = e_{j-1}$. Numeriere die Einheitsvektoren mit Doppelindizes analog zur Partition

$$\begin{array}{lll} e_{11} = e_1, & \ldots, & e_{1,p_1} = e_{p_1} \\ e_{21} = e_{p_1+1}, & \ldots, & e_{2,p_2} = e_{p_1+p_2} \\ \vdots & & \\ e_{s1} = e_{k_s}, & \ldots, & e_{s,p_s} = e_n. \end{array}$$



Dann gilt für $l \geq 0$

$$\text{Bild}\,(N(p)^l) = \mathcal{L}(\{e_{ij} \mid 1 \leq i \leq s, 1 \leq j \leq p_i - l\}),$$

und da alle e_{ij} linear unabhängig sind, erhalten wir eine Basis von Bild $(N(p)^l)$ und es gilt

$$\text{Rang}\,(N(p)^l) = n - \sum_{i=1}^{l} q_i \quad \text{ mit } \ (q_1, \ldots, q_m) = p^*\,.$$

Für $l \geq 1$ gilt dann

$$\text{Rang}\,(N(p)^{l-1}) - \text{Rang}\,(N(p)^l) = n - \sum_{i=1}^{l-1} q_i - n + \sum_{i=1}^{l} q_i = q_l.$$

Damit bestimmt $N(p)$ eindeutig die Partition $(q_1, \ldots, q_m)$ und damit auch p. Sind $p, \tilde{p}$ Partitionen von n und sind $N(p)$ und $N(\tilde{p})$ ähnlich, so gilt $\text{Rang}\,(N(p)^l) = \text{Rang}\,(N(\tilde{p})^l)$ für alle l, also $p^* = \tilde{p}^*$, und damit $p = \tilde{p}$.

(b) Wir müssen zeigen, dass es eine Matrix P gibt, so dass

$$P^{-1}AP = N(p)$$

ist, wobei p die zu q duale Partition von n ist. Sei m der Nilpotenzindex von A. Setze

$$S_i = \text{Kern}\,(A^i).$$

Wir zeigen dazu, dass gilt $\{0\} = S_0 \subset S_1 \subset \ldots \subset S_m = \mathbb{K}^n$.

Da $A^0 = I_n$, so folgt $S_0 = \{0\}$, und da $A^m = 0$, so folgt $S_m = V$. Sei $0 \leq i < m$ und $v \in S_i$, d.h. $A^i v = 0$. Dann folgt

$$A^{i+1}v = A(A^i v) = A0 = 0.$$

Also gilt $S_i \subset S_{i+1}$. Beachte, dass alle S_i Untervektorräume sind. Wir setzen

$$\begin{aligned} q_i \;=\; \dim S_i - \dim S_{i-1} &= (n - \text{Rang}\,(A^i)) - (n - \text{Rang}\,(A^{i-1})) \\ &= \text{Rang}\,(A^{i-1}) - \text{Rang}\,(A^i) \end{aligned}$$

und zeigen, dass $(q_1, \ldots, q_m)$ eine Partition von n ist.

Da die S_i Unterräume sind, können wir den Quotienten S_i/S_{i-1} bilden, d.h., wir bilden die Menge aller Äquivalenzklassen in S_i bezüglich der Relation für $u, \tilde{u} \in S_i$:

$$u \sim \tilde{u} \quad \Longleftrightarrow \quad u - \tilde{u} \in S_{i-1}.$$

Es folgt, dass S_i/S_{i-1} einen Vektorraum bildet, und dessen Dimension ist dann

$$\dim S_i/S_{i-1} = \dim S_i - \dim S_{i-1} = q_i.$$

Wir brauchen nun den folgenden Sachverhalt: Sei eine Abbildung

$$z: \ S_{i+1}/S_i \to S_i/S_{i-1}$$

definiert durch $z(x + S_i) = Ax + S_{i-1}$ für $x \in S_{i+1}$.

Wir zeigen, dass z wohldefiniert, linear und injektiv für $1 \le i \le m-1$ ist.

Da z auf Äquivalenzklassen definiert ist, müssen wir zeigen, dass z nicht von der Wahl des Repräsentanten abhängt. Seien $x, y \in S_{i+1}$ mit $x - y \in S_i$. Wir müssen zeigen, dass $Ax \in S_i$ und $Ax - Ay \in S_{i-1}$.

$$x \in S_{i+1} \implies A^{i+1}x = 0 \implies A^i(Ax) = 0 \implies Ax \in \text{Kern}\,(A^i) = S_i,$$

$$A^{i-1}(Ax - Ay) = A^i x - A^i y = A^i(x - y) = 0,$$

denn $x - y \in S_i$. Damit ist z wohldefiniert.

Seien $x, y \in S_{i+1}$ und $\lambda \in \mathbb{K}$. Dann gilt

$$\begin{aligned} z((x + S_i) + (y + S_i)) &= z(x + y + S_i) \\ &= A(x + y) + S_{i-1} \\ &= Ax + Ay + S_{i-1} \\ &= (Ax + S_{i-1}) + (Ay + S_{i-1}) \\ &= z(x + S_i) + z(y + S_i), \end{aligned}$$

$$\begin{aligned} z(\lambda(x + S_i)) &= z(\lambda x + S_i) \\ &= A\lambda x + S_{i-1} \\ &= \lambda Ax + S_{i-1} \\ &= \lambda(Ax + S_{i-1}) \\ &= \lambda z(x + S_i). \end{aligned}$$

Also ist z linear. Für die Injektivität sei $x + S_i \in \text{Kern}\,(z)$.

$$\begin{aligned} 0 &= z(x + S_i) = Ax + S_{i-1}, \qquad (0 = 0 + S_i \in S_{i+1}/S_i), \\ \Rightarrow &\ Ax \in S_{i-1} \\ \Rightarrow &\ A^{i-1}(Ax) = 0 \\ \Rightarrow &\ A^i x = 0 \\ \Rightarrow &\ x \in S_i. \end{aligned}$$

Wegen der Injektivität von z gilt

$$q_{i+1} = \dim S_{i+1}/S_i \le \dim S_i/S_{i-1} = q_i.$$

Außerdem gilt $S_{m-1} \subsetneqq S_m$, da m der Nilpotenzindex von A ist, also ist $q_m \geq 1$. Damit ist $q = (q_1, \ldots, q_m)$ eine Partition, denn

$$q_{i+1} = \dim S_{i+1}/S_i = \dim S_{i+1} - \dim S_i$$

$$\sum_{i=1}^{m} q_i = \sum_{i=1}^{m} (\dim S_i - \dim S_{i-1}) = \dim S_m - \dim S_0 = n.$$

Wir werden nun eine Basis $\{v_{ij} \mid 1 \leq j \leq m, 1 \leq i \leq q_j\}$ konstruieren, so dass

$$\begin{aligned} Av_{ij} &= v_{i,j-1} \quad \text{für } \ j > 1, \\ Av_{i1} &= 0\,. \end{aligned}$$

Dazu verwenden wir Induktion rückwärts: $j = m$.

Sei $\{v_{1,m} + S_{m-1}, \ldots, v_{q_m,m} + S_{m-1}\}$ eine beliebige Basis von S_m/S_{m-1} mit

$$v_{1,m}, \ldots, v_{q_m,m} \in S_m = \mathbb{K}^n.$$

Für $1 \leq j < m$ sei $\{v_{1,j+1} + S_j, \ldots, v_{q_{j+1},j+1} + S_j\}$ eine Basis von S_{j+1}/S_j mit

$$v_{i,j+1} \in S_{j+1}.$$

Setze nun $v_{ij} = Av_{i,j+1}$, $1 \leq i \leq q_{j+1}$.

Wir haben gezeigt, dass $v_{1,j} + S_{j-1}, \ldots, v_{q_{j+1},j} + S_{j-1}$ linear unabhängig sind, und wir ergänzen diese zu einer Basis von S_j/S_{j-1} durch Elemente

$$v_{q_{j+1}+1,j} + S_{j-1}, \ \ldots, \ v_{q_j,j} + S_{j-1}, \qquad \text{wobei } v_{q_{j+1}+1,j}, \ldots, v_{q_j,j} \in S_j.$$

Damit haben wir induktiv $v_{1j}, \ldots, v_{q_j,j} \in S_j$ konstruiert, so dass

$$\{v_{1,j} + S_{j-1}, \ \ldots, \ v_{q_j,j} + S_{j-1}\}$$

eine Basis von S_j/S_{j-1} bilden. Nach Konstruktion ist

$$\begin{aligned} Av_{i,j} &= v_{i,j-1} && \text{für } j > 1, \ \text{ und} \\ Av_{i,1} &= 0, && \text{denn } v_{i,1} \in S_1 = \text{Kern}\,(A). \end{aligned}$$

Nach Satz 12.11 (e) folgt, dass die $\{v_{ij}\}$ eine Basis bilden, und damit ist die aus $\{v_{ij}\}$ gebildete Matrix P invertierbar und mit der zu q dualen Partition folgt nach Konstruktion

$$P^{-1}AP = N(p).$$ □

Korollar 3.9 *Sei $A \in \mathbb{K}^{n,n}$, dann sind die folgenden Aussagen äquivalent:*

(a) *A ist nilpotent.*

(b) *A ist ähnlich zu $N(p)$ für eine Partition p.*

(c) $\det(\lambda I - A) = \lambda^n$.

(d) $A^n = 0$.

In der Menge aller nilpotenten Matrizen bildet Ähnlichkeit eine Äquivalenzrelation und die Repräsentanten der Äquivalenzklassen sind $N(p)$.

Beispiel 3.10 $V = \mathbb{R}^3$.

$$A = \begin{bmatrix} -1 & 1 & -2 \\ -3 & 3 & -6 \\ -1 & 1 & -2 \end{bmatrix} \qquad \boxed{\text{Rang}\,(A) = 1} \qquad \boxed{\dim(\text{Kern}\,(A)) = 2}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 3 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}}_{P} \underbrace{\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]}_{N(p)} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -2 \\ 0 & -1 & 3 \end{bmatrix}}_{P^{-1}}.$$

$A^2 = 0 \;\Rightarrow\; A$ hat den Nilpotenzindex 2.

$$\begin{aligned} S_0 &= \{0\}, \\ S_1 &= \text{Kern}\,(A) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right), \\ S_2 &= \text{Kern}\,(A^2) = \mathbb{R}^3 = V. \end{aligned}$$

Wähle Element von $V \setminus S_1 = \mathbb{R}^3 \setminus S_1$, z. B. $v_{12} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$,

also ist $\{v_{12} + S_1\}$ eine Basis von V/S_1.

$$v_{11} = Av_{12} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix},$$

$$q_2 = \dim S_2/S_1 = \dim V/S_1 = \dim V - \dim S_1 = 1\,,$$

$$q_1 = \dim S_1/S_0 = \dim S_1 - \dim S_0 = 2\,.$$

Wir ergänzen $v_{11} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ durch $v_{21} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ zu einer Basis von S_1. Wir erhalten die Partition

$$\begin{array}{l|c|c|} \cline{2-3} p_1 & v_{11} & v_{12} \\ \cline{2-3} p_2 & v_{21} \\ \cline{2-2} \end{array}$$

und es gilt

$$P = [v_{11}, v_{12}, v_{21}] = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 3 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

Wir haben damit schon für die spezielle Klasse der nilpotenten Matrizen die Äquivalenzklassen und die Normalform gefunden. Im nächsten Kapitel werden wir uns erst einmal mit speziellen Polynomen beschäftigen, bevor wir zur Charakterisierung der Normalform unter Ähnlichkeit für allgemeine quadratische Matrizen kommen.

Kapitel 4

Das charakteristische Polynom und das Minimalpolynom

Wir beginnen dieses Kapitel mit einigen Eigenschaften von Polynomen.

Wie beim Rechnen in $\mathbb{Z}$ können wir auch für Polynome den Begriff der Teilbarkeit des größten gemeinsamen Teilers (ggT)und des kleinsten gemeinsamen Vielfachen (kgV) einführen:

Definition 4.1 *Sei $\mathbb{K}$ ein Körper und seinen $p_1, \ldots, p_n \neq 0$ Polynome mit Koeffizienten in $\mathbb{K}$.*

1. *Ein Polynom $m \neq 0$, heißt* kleinstes gemeinsames Vielfaches (kgV) *von $p_1, \ldots, p_n$, wenn*

   (i) $p_i \mid m,\ i = 1, \ldots, n,$

   (ii) aus $p_i \mid m',\ i = 1, \ldots, n,$ *folgt* $m \mid m'$.

2. *Ein Polynom $d \neq 0$, heißt* größter gemeinsamer Teiler (ggT) *von $p_1, \ldots, p_n$, wenn*

   (i) $d \mid p_i,\ i = 1, \ldots, n,$

   (ii) aus $d' \mid p_i,\ i = 1, \ldots, n,$ *folgt* $d' \mid d$.

3. *Dabei bedeutet $p_i \mid m$ (p teilt m), dass es eine anderes Polynom q mit Koeffizienten in $\mathbb{K}$ gibt, so dass $m = pq$.*

4. *Ein konstantes Polynom $p(x) = a$ mit $a \in \mathbb{K} \setminus \{0\}$ heißt* Einheit.

Definition 4.2 *Ein Polynom $p \neq 0$, mit Koeffizienten in $\mathbb{K}$ heißt* irreduzibel *(oder* unzerlegbar*), wenn p nicht Einheit ist und in jeder Zerlegung des Polynoms in ein Produkt von Polynomen $p = q \cdot z$ einer der beiden Faktoren q, z Einheit ist, d.h., p hat keine echten Teiler.*

Lemma 4.3 *Sei p mit Koeffizienten in $\mathbb{K}$ ein irreduzibles Polynom, und für zwei Polynome q_1, q_2 mit Koeffizienten in $\mathbb{K}$ gelte $p \mid q_1 \cdot q_2$. Dann folgt $p \mid q_1$ oder $p \mid q_2$.*

Beweis: Sei $p \nmid q_1$. Dann ist $d = \mathrm{ggT}\,(p, q_1)$ Einheit, denn sonst wäre d ein echter Teiler von p. Dann gibt es Polynome z, z_1, so dass

$$p \cdot z + q_1 \cdot z_1 = 1.$$

Dies folgt aus dem Euklidischen Algorithmus durch Teilen mit Rest. Daraus folgt

$$p \cdot z \cdot q_2 + q_1 \cdot q_2 \cdot z_1 = q_2.$$

Da $p \mid q_1 \cdot q_2$, ist p ein Teiler der linken Seite dieser Gleichung, und somit $p \mid q_2$. $\square$

Im folgenden bezeichnen wir Polynome, welche als höchsten Koeffizienten die $1 \in \mathbb{K}$ haben als *normiert*.

Satz 4.4 *Für jedes normierte Polynom $p \neq 1$ mit Koeffizienten in $\mathbb{K}$, gibt es eine eindeutige Zerlegung in irreduzible normierte Polynome.*

Beweis: Die Existenz beweisen wir an dieser Stelle nicht, siehe Vorlesung Algebra. Für die Eindeutigkeit nehmen wir an, es gibt ein Polynom p mit nicht eindeutiger Zerlegung, d.h., es existieren irreduzible normierte Polynome $q_1, \ldots, q_r$ und $z_1, \ldots, z_s$, so dass

$$p = q_1 \cdot \ldots \cdot q_r = z_1 \cdot \ldots \cdot z_s\,. \tag{4.5}$$

Wir führen nun eine Induktion über s durch.

I.A.: $s = 1$. Das heißt $z_1 = q_1 \cdot \ldots \cdot q_r$ und daraus folgt $r = 1$ und $z_1 = q_1$, denn sonst wäre z_1 kein irreduzibles Polynom.

I.V.: Für $s = k$ sei bewiesen, dass $r = k$ und $z_i = q_i,\ i = 1, \ldots, k$.

I.S.: Aus (4.5) und Lemma 4.3 folgt, dass $q_r \mid z_i$ ist für ein $i \in \{1, \ldots, k+1\}$. Da z_i irreduzibel ist, muss deshalb $q_r = z_i$ gelten. Damit ist (4.5) äquivalent zu

$$q_1 \cdot \ldots \cdot q_{r-1} = z_1 \cdot \ldots \cdot z_{i-1} \cdot z_{i+1} \cdot \ldots \cdot z_{k+1}.$$

Nach I.V. erhalten wir die Behauptung.

□

Beispiel 4.6 *Falls $\mathbb{K} = \mathbb{R}$, so hat das Polynom*

$$p(x) = (x^2 + 1)(x + 1) = x^3 + x^2 + x + 1$$

die unzerlegbaren Faktoren $(x^2 + 1)$, $(x + 1)$.

Falls $\mathbb{K} = \mathbb{C}$, so hat $p(x)$ die unzerlegbaren Faktoren $(x - i)$, $(x + i)$, $(x + 1)$.

Wir wissen aus dem Satz von Cayley/Hamilton, dass für das charakteristische Polynom $P_A(x) = \det(xI - A)$ einer Matrix $A \in \mathbb{K}^{n,n}$ gilt, dass $P_A(A) = 0$.

Definition 4.7 *Sei $m_A(x)$ das normierte Polynom kleinsten Grades, für welches*

$$m_A(A) = 0,$$

so heißt $m_A(x)$ das Minimalpolynom *zu A.*

Es ist klar, dass

$$\text{Grad } m_A(x) \leq \text{ Grad } P_A(x).$$

Lemma 4.8 *Für das Minimalpolynom $m_A(x)$ und das charakteristische Polynom $P_A(x)$ einer Matrix A gilt: $m_A \mid P_A$.*

Beweis: Aufgrund des Divisionsalgorithmus mit Rest gibt es Polynome $r(x), s(x)$, so dass

$$P_A(x) = r(x)m_A(x) + s(x),$$

wobei entweder $s(x) = 0$ oder Grad$(s(x) <$ Grad$(m_A(x))$. Angenommen, es wäre $s(x) \neq 0$, dann wäre

$$s(A) = P_A(A) - r(A)m_A(A) = 0.$$

Das ist ein Widerspruch dazu, dass m_A Minimalpolynom ist. Also ist $s(x) = 0$ und folglich $m_A \mid P_A$. □

Bemerkung 4.9 *Analog kann man für jedes Polynom $q(x)$ mit $q(A) = 0$ zeigen, dass $m_A \mid q$.*

Lemma 4.10 *Das Minimalpolynom einer Matrix A ist eindeutig bestimmt.*

Beweis: Sei $p \neq m_A$ ein normiertes Polynom mit Koeffizienten in $\mathbb{K}$, für das $p(A) = 0$ und p habe den gleichen Grad wie m_A.
Setze $q = p - m_A$, so gilt

$$q(A) = p(A) - m_A(A) = 0.$$

Da aber p und m_A den gleichen größten Koeffizienten haben, ist der Grad von q kleiner als der von m_A und wir erhalten einen Widerspruch. □

Satz 4.11 *Sei $A \in \mathbb{K}^{n,n}$, so teilt $P_A(x)$ das Polynom m_A^n.*

Beweis: Setze

$$\begin{aligned} P_A(x) &= \sum_{i=0}^{n} \alpha_i x^i, \quad \alpha_n = 1\,, \\ m_A(x) &= \sum_{i=0}^{r} \beta_i x^i, \quad \beta_r = 1\,. \end{aligned}$$

Betrachte die Matrizenfolge

$$\begin{aligned} B_r &= I, \\ B_{r-1} &= A + \beta_{r-1} I, \\ &\vdots \\ B_1 &= A^{r-1} + \beta_{r-1} A^{r-2} + \ldots + \beta_1 I, \end{aligned}$$

d.h.,

$$\begin{aligned} B_r &= I, \\ B_{i-1} &= \beta_{i-1} I + A B_i, \qquad i = r, \ldots, 2. \end{aligned}$$

So gilt

$$\begin{aligned} -AB_1 &= -A^r - A^{r-1}\beta_{r-1} - \ldots - A\beta_1 \\ &= \beta_0 I - \underbrace{m_A(A)}_{=0} = \beta_0 I\,. \end{aligned}$$

Setze $B(x) = x^{r-1}B_r + x^{r-2}B_{r-1} + \ldots + x^1 B_2 + B_1$, so folgt

$$\begin{aligned} (xI - A)B(x) &= x^r B_r + x^{r-1}B_{r-1} + \ldots + x^2 B_2 + xB_1 \\ &\qquad - x^{r-1}AB_r - \ldots - x^2 AB_3 - xAB_2 - AB_1 \\ &= x^r I + \beta_{r-1}x^{r-1}I + \ldots + \beta_2 x^2 I + \beta_1 xI + \beta_0 I \\ &= m_A(x)I\,. \end{aligned}$$

Damit gilt

$$\det(xI - A) \cdot \det B(x) = m_A(x)^n$$

alos folgt, $P_A(x)$ teilt $m_A(x)^n$. □

Korollar 4.12 *Sei $A \in \mathbb{K}^{n,n}$, so haben $P_A(x)$ und $m_A(x)$ dieselben irreduziblen Faktoren.*

Beweis: Sei $z(x)$ irreduzibler Faktor von $m_A(x)$. Da m_A das charakteristische Polynom P_A teilt und z Teiler von m_A ist, folgt, dass z auch P_A teilt.
Sei $z(x)$ irreduzibler Faktor von $P_A(x)$, so gilt, dass $z(x)$ Teiler von $(m_A)^n$ ist. Also folgt aus der Irreduzibilität und Satz 4.4, dass $z(x)$ auch m_A teilt. □

Korollar 4.13 *Das Minimalpolynom der Matrix*

$$J(\lambda, n) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} = \lambda I_n + N(p) \in \mathbb{K}^{n,n},\ p = (n),$$

ist gleich dem charakteristischen Polynom und hat die Form

$$P_{J(\lambda,n)}(x) = m_{J(\lambda,n)}(x) = (x - \lambda)^n.$$

Beweis: Das charakteristische Polynom hat per Definition diese Form.
$m_{J(\lambda,n)}(x)$ teilt $P_{J(\lambda,n)}(x)$, also $m_{J(\lambda,n)}(x) = (x-\lambda)^j$ (wegen Satz 4.4).
Nach Definition des Minimalpolynoms gilt

$$\begin{aligned} &0 = m_{J(\lambda,n)}(J(\lambda,n)), \qquad \text{aber} \\ &m_{J(\lambda,n)}(J(\lambda,n)) = (J(\lambda,n) - \lambda I)^j = N(n)^j \end{aligned}$$

Da aber $N(n)$ nilpotent vom Index n ist, so folgt $j = n$. □

Definition 4.14 *Seien $A_1, \ldots, A_k$ quadratische Matrizen mit $A_i \in \mathbb{K}^{p_i,p_i}$. Wir bezeichnen mit $\bigoplus_{i=1}^{k} A_i$ die Matrix*

$$\bigoplus_{i=1}^{k} A_i := \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix}.$$

Dies heißt direkte Summe *von Matrizen (und entspricht natürlich einer direkten Summe von Vektorräumen).*
Eine Matrix der Form

$$A = \bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il}) =: \bigoplus_{i=1}^{k} J(\lambda_i, (p_i))$$

mit den Partitionen $p_i = (p_{i1}, \ldots, p_{i,s_i})$, $i = 1, \ldots, k$, heißt Matrix in Jordan'scher Normalform.

Beispiel 4.15

$$A = \left[\begin{array}{cc|ccc} J(1,5) & & & & \\ & J(1,3) & & & \\ \hline & & J(2,3) & & \\ & & & J(2,2) & \\ & & & & J(2,1) \end{array}\right] = J(1,(5,3)) \bigoplus J(2,(3,2,1))$$

$$= \left[\begin{array}{cc|ccc} \boxed{\begin{matrix}1&1&&&\\&1&1&&\\&&1&1&\\&&&1&1\\&&&&1\end{matrix}} & & & & \\ & \boxed{\begin{matrix}1&1&\\&1&1\\&&1\end{matrix}} & & & \\ \hline & & \boxed{\begin{matrix}2&1&\\&2&1\\&&2\end{matrix}} & & \\ & & & \boxed{\begin{matrix}2&1\\&2\end{matrix}} & \\ & & & & \boxed{2} \end{array}\right],$$

$$A \;=\; \begin{bmatrix} J(1,1) & & \\ & J(2,1) & \\ & & J(3,1) \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}.$$

Lemma 4.16 *Seien $A_1, \ldots, A_k$, mit $A_i \in \mathbb{K}^{k_i,k_i}$, $i = 1, \ldots, k$, mit den Minimalpolynomen $m_{A_i}(x)$. Dann ist für $A = \bigoplus_{i=1}^{k} A_i$ das Polynom m_A das normierte kleinste gemeinsame Vielfache von*

$$m_{A_1}(x), \ldots, m_{A_k}(x).$$

Beweis: Per Induktion nach k.

I.A.: $k = 1$ ist klar.
$k = 2:$

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \qquad m_A(A) = \begin{bmatrix} m_A(A_1) & 0 \\ 0 & m_A(A_2) \end{bmatrix} = 0$$

$\Longrightarrow$ $m_A(A_1) = 0$ und $m_A(A_2) = 0$.
$\Longrightarrow$ m_{A_1} teilt m_A und m_{A_2} teilt m_A (nach Bemerkung 4.9).
$\Longrightarrow$ m_A ist gemeinsames Vielfaches von m_{A_1} und m_{A_2}.
Sei q ein anderes gemeinsames Vielfaches von m_{A_1}, m_{A_2}.

$$q(A) = \begin{bmatrix} q(A_1) & \\ & q(A_2) \end{bmatrix} = 0\,.$$

m_A ist Minimalpolynom $\implies$ m_A teilt q (aufgrund der Bemerkung 4.9). Also ist m_A das normierte kleinste gemeinsame Vielfache von m_{A_1} und m_{A_2}.

I.V.: Sei die Behauptung richtig für $k = l$ Blöcke.

I.S.: Sei $k = l + 1$, so können wir A schreiben als

$$A = \begin{bmatrix} B_1 & \\ & A_{l+1} \end{bmatrix}, \quad \text{mit} \quad B_1 = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_l \end{bmatrix}.$$

Nach I.V. ist m_{B_1} normiertes kgV von $m_{A_1}, \ldots, m_{A_l}$ und nach I.A. ist m_A normiertes kgV von m_{B_1} und $m_{A_{l+1}}$, und damit normiertes kgV von $m_{A_1}, \ldots, m_{A_{l+1}}$.

$\square$

Beispiel 4.17

$$A = \left[\begin{array}{cc|ccc|c|cc} 1 & 1 & & & & & & \\ & 1 & & & & & & \\ \hline & & 1 & 1 & & & & \\ & & & 1 & 1 & & & \\ & & & & 1 & & & \\ \hline & & & & & 1 & & \\ \hline & & & & & & 1 & 1 \\ & & & & & & & 1 \end{array}\right], \qquad P_A(x) = (x-1)^8,$$

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad P_{A_1} = (x-1)^2, \quad m_{A_1} = (x-1)^2,$$

$$\text{denn} \ \ (A_1 - I_2)^1 \neq 0, \quad \text{aber} \ (A_1 - I_2)^2 = 0.$$

$$A_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_{A_2} = (x-1)^3, \quad m_{A_2} = (x-1)^3,$$

$$A_3 = \begin{bmatrix} 1 \end{bmatrix}, \qquad P_{A_3} = (x-1), \quad m_{A_3} = (x-1),$$

$$A_4 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad P_{A_4} = (x-1)^2, \quad m_{A_4} = (x-1)^2.$$

Normiertes kleinstes gemeinsames Vielfaches von $(x-1)^2, (x-1)^3, (x-1), (x-1)^2$ ist $(x-1)^3$ und das ist tatsächlich das Minimalpolynom, da

$$\begin{array}{rcl} (A-I_8)^1 & \neq & 0\,, \\ (A-I_8)^2 & \neq & 0\,, \\ (A-I_8)^3 & = & 0 \end{array} .$$

Satz 4.18 *Sei* $A = \bigoplus_{i=1}^{k} A_i$*, mit* $A_i = \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il})$ *eine Matrix in Jordan'scher Normalform, wobei* $p_i = (p_{i,1}, \ldots, p_{i,s_i})$ *für alle* i *eine Partition ist und* λ_i*,* $i = 1, \ldots, k$*, paarweise verschieden sind.*
Dann ist

$$m_A = \prod_{i=1}^{k} (x-\lambda_i)^{p_{i,1}}.$$

Beweis: Nach Lemma 4.16 ist m_A kleinstes gemeinsames Vielfaches von $m_{A_1}, \ldots, m_{A_k}$. Jedes m_{A_i} ist kleinstes gemeinsames Vielfaches von

$$q_{i,1} = m_{J(\lambda_i, p_{i,1})},\ \ldots,\ q_{i,s_i} = m_{J(\lambda_i, p_{i,s_i})}.$$

Da p_i eine Partition ist, gilt $p_{i,1} \geq p_{i,2} \geq \ldots \geq p_{i,s_i}$.

$$q_{i,j} = m_{J(\lambda_i, p_{i,j})} = (x-\lambda_i)^{p_{i,j}}.$$

$\Longrightarrow$ Kleinstes gemeinsames Vielfaches von $q_{i,1},\ \ldots,\ q_{i,s_i}$ ist $q_{i,1}$.

Also ist m_A kleinstes gemeinsames Vielfaches von $q_{1,1},\ \ldots,\ q_{k,1}$, und da alle λ_i paarweise verschieden sind, folgt

$$m_A = \prod_{i=1}^{k} (x-\lambda_i)^{p_{i,1}}.$$

□

Kapitel 5

Die Jordan'sche Normalform

Wir wollen nun einen der zentralen Sätze der Linearen Algebra beweisen, nämlich die Klassifikation der Äquivalenzklassen unter Ähnlichkeit.

Satz 5.1

(a) Sei $A \in \mathbb{K}^{n,n}$, so ist A ähnlich zu einer Jordan'schen Normalform genau dann, wenn m_A in Linearfaktoren, d.h. in Polynome vom Grad 1, zerfällt.

(b) Seien A und B zwei Jordan'sche Normalformen,

$$\begin{aligned} A &= \bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il}) & \qquad B &= \bigoplus_{i=1}^{r} \bigoplus_{l=1}^{t_i} J(\mu_i, c_{il}) \\ &= \bigoplus_{i=1}^{k} J(\lambda_i, (p_i)), & &= \bigoplus_{i=1}^{r} J(\mu_i, (c_i)), \end{aligned}$$

mit Partionen $p_i = (p_{i,1}, \ldots, p_{i,s_i})$, $c_i = (c_{i,1}, \ldots, c_{i,t_i})$.

A und B sind ähnlich genau dann, wenn

$$\begin{aligned} &r = k, \\ &\{\lambda_1, \ldots, \lambda_k\} = \{\mu_1, \ldots, \mu_k\}, \\ &p_i = (p_{i1}, \ldots, p_{i,s_i}) = (c_{i1}, \ldots, c_{i,t_i}) = c_i, \qquad i = 1, \ldots, k. \end{aligned}$$

Zum Beweis dieses Satzes brauchen wir einige Hilfssätze.

Lemma 5.2 *Sei $A \in \mathbb{K}^{n,n}$, und für $\lambda \in \mathbb{K}$ sei $q(x) = (x - \lambda)^m$, so dass $q(A) = 0$. Dann ist A ähnlich zu einer Matrix*

$$\bigoplus_{l=1}^{t} J(\lambda, p_l),$$

wobei $(p_1, \ldots, p_t)$ *eine Partition von* n *ist und alle* $p_i \leq m$.

Beweis: Sei $B = A - \lambda I$.

$$0 = q(A) = (A - \lambda I)^m = B^m.$$

B ist nilpotent mit Nilpotenzindex $\leq m$. Nach Satz 3.8 gibt es dann eine invertierbare Matrix P, so dass

$$P^{-1}(A - \lambda I)P = N(p)$$

für eine Partition $p = (p_1, \ldots, p_t)$ von n mit $p_i \leq m$ für alle i. Dann folgt

$$P^{-1}AP \;=\; \lambda I_n + N(p) \;=\; \bigoplus_{l=1}^{t} J(\lambda, p_l)$$

□

Definition 5.3 *Sei* V *ein Vektorraum über einem Körper* $\mathbb{K}$ *und sei* $f : V \to V$ *linear. Ein Unterraum* U *von* V *heißt* invarianter Unterraum *von* V *zu* f, *falls* $f(U) \subset U$, *d.h.,*

$$f(u) \in U \qquad \textit{für alle } u \in U.$$

Beispiel 5.4 Sei $A \in \mathbb{K}^{n,n}$, λ Eigenwert von A und v Eigenvektor zu λ, so ist $\mathcal{L}(v)$ invarianter Unterraum zu A, denn

$$A(\mu v) = (\lambda\mu)v \in \mathcal{L}(v), \quad \text{für alle } \mu \in \mathbb{K}.$$

Lemma 5.5 *Sei* V *endlichdimensionaler Vektorraum über einem Körper* $\mathbb{K}$ *und sei* $a : V \to V$ *linear. Sei* $q(x)$ *ein Polynom mit Koeffizienten in* $\mathbb{K}$ *und* $q(a) = 0$, *d.h., für jede Matrixdarstellung* A *von* a *gilt* $q(A) = 0$. *Sei* $q(x) = q_1(x) \cdot \ldots \cdot q_r(x)$, *wobei alle* $q_i(x)$ *paarweise teilerfremd sind.*

Sei $V_i = \text{Kern}\,(q_i(a))$, *so ist* V_i *invarianter Unterraum zu* a *und*

$$V = \bigoplus_{i=1}^{r} V_i.$$

Beweis: Setze für $j = 1, \ldots, r$

$$z_j = \prod_{i \neq j} q_i, \ \text{ d.h.}, \ q = z_j \cdot q_j.$$

Dann gilt, dass ggT $(z_1, \ldots, z_r) = w$ eine Einheit ist, denn: Es gilt $w \mid z_1$, also existiert ein $i \neq 1$, so dass $w \mid q_i$. Es gilt auch $w \mid z_i$, also existiert ein $j \neq i$, so dass $w \mid q_j$. Da q_i und q_j teilerfremd sind, ist w Einheit. Dann gibt es wieder auf Grund des euklidischen Algorithmus Polynome $y_1, \ldots, y_r$ mit Koeffizienten in $\mathbb{K}$, so dass

$$\sum_{j=1}^{r} z_j y_j = 1\,,$$

denn 1 ist größter gemeinsamer Teiler von $z_1, \ldots, z_r$.

Also folgt

$$\sum_{j=1}^{r} z_j(a) y_j(a) = 1.$$

Sei $v \in V$ und $v_j = z_j(a)y_j(a)(v)$. Dann gilt

$$\begin{aligned} q_j(a)(v_j) &= q_j(a)z_j(a)y_j(a)(v) \\ &= q(a)y_j(a)(v) = 0, \qquad \text{da } q(a) = 0, \end{aligned}$$

also ist $v_j \in \operatorname{Kern} q_j(a) = V_j$.

$$\begin{aligned} v &= \underbrace{\left(\sum_{j=1}^{r} z_j(a)y_j(a)\right)}_{1}(v) \\ &= \sum_{j=1}^{r} (z_j(a)y_j(a)(v)) = \sum_{j=1}^{r} v_j. \end{aligned}$$

Wir haben damit schon gezeigt, dass jedes $v \in V$ als $\sum\limits_{j=1}^{r} v_j$ mit $v_j \in V_j$ geschrieben werden kann. Wir müssen noch zeigen, dass diese Darstellung eindeutig ist. Sei

$$v = \sum_{j=1}^{r} b_j = \sum_{j=1}^{r} v_j, \qquad v_j, b_j \in V_j.$$

Dann gilt

$$
\begin{aligned}
v_j &= z_j(a)y_j(a)(v) \\
&= z_j(a)y_j(a)\left(\sum_{i=1}^{r} b_i\right) \\
&= z_j(a)y_j(a)(b_j), \quad \text{da } b_i \in \operatorname{Kern}(q_i) \ \text{ und für } i \neq j:\ q_i \ \text{ Faktor von } z_j \\
&= \underbrace{\left(\sum_{i=1}^{r}(z_i(a)y_i(a))\right)}_{1}(b_j) = b_j\,.
\end{aligned}
$$

Also gilt: $V = \bigoplus_{j=1}^{r} V_j$.

Wir müssen dann nur noch zeigen, dass V_i invarianter Unterraum ist.
Sei $v \in V_i$. Da $q_i(a)\cdot a = a \cdot q_i(a)$, so folgt

$$
q_i(a)(a(v)) = (a\cdot q_i(a))(v) = a(0) = 0 \quad \Rightarrow \quad a(v) \in V_i\,.
$$

□

Beweis: (Satz 5.1).

(a) „$\Longleftarrow$“. Sei m_A das Minimalpolynom von A und zerfalle in Linearfaktoren,

$$
m_A = \prod_{i=1}^{r}(x-\lambda_i)^{m_i}, \qquad \text{mit } \lambda_i \neq \lambda_j \ \text{ für } \ i \neq j\,.
$$

Setze $\mu_i(x) = (x-\lambda_i)^{m_i}$, dann gilt $m_A = \prod_{i=1}^{r} \mu_i$, und die μ_i sind paarweise teilerfremd.
Setze $V_i = \operatorname{Kern}(\mu_i(f_A))$, wobei $f_A(x) = Ax$.
Nach Lemma 5.5 sind $V_1, \dots, V_r$ invariante Unterräume zu f_A und

$$
V = \bigoplus_{i=1}^{r} V_i.
$$

Sei $f_i : V_i \to V_i$ die lineare Abbildung, die durch $f_i(v) = f_A(v)$ für $v \in V_i$ definiert ist, und A_i eine Matrixdarstellung von f_i bzgl. einer Basis von V_i. So ist

$$
\begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{bmatrix}
$$

eine Matrixdarstellung von f_A bzgl. einer Basis von V.

Es gilt $\mu_i(f_i) = 0$. Dann folgt mit Lemma 5.2, dass

$$A_i \quad \text{ähnlich zu} \quad \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il})$$

und damit folgt diese Richtung.

(a) „$\Longrightarrow$". A sei ähnlich zu Jordan'scher Normalform. Dann folgt, dass P_A in Linearfaktoren zerfällt also zerfällt auch m_A in Linearfaktoren, denn $m_A \mid P_A$ (Lemma 4.8).

(b) „$\Longleftarrow$" ist klar, denn dann ist $A = B$.

(b) „$\Longrightarrow$" A ähnlich B, und B in Jordan'scher Normalform.

Wir wissen aus der Änlichkeit, dass die charakteristischen Polynome gleich sind, also

$$k = r\,,$$
$$\{\lambda_1, \ldots, \lambda_k\} = \{\mu_1, \ldots, \mu_k\}\,,$$
$$\sum_{l=1}^{s_i} p_{il} = \sum_{l=1}^{t_i} c_{il}\,, \qquad i = 1, \ldots, k\,.$$

Nun folgt aber gerade aus der Tatsache, dass wir eine direkte Summe haben,

$$\begin{array}{rcl} q_{il} & = & \text{Rang}\,(A - \lambda_i I)^{l-1} - \text{Rang}\,(A - \lambda_i I)^l \\ d_{il} & = & \text{Rang}\,(B - \lambda_i I)^{l-1} - \text{Rang}\,(B - \lambda_i I)^l \end{array},$$

wobei q die duale Partition zu p und d die duale Partition zu c ist. Aber da A und B ähnlich sind, so sind diese Ränge und damit die Partitionen q_i und d_i alle gleich. Folglich sind auch die Partitionen p_i und c_i alle gleich. □

Korollar 5.6

(a) Jedes $A \in \mathbb{C}^{n,n}$ ist ähnlich zu einer Jordan'schen Normalform.

(b) Hat $A \in \mathbb{R}^{n,n}$ nur reelle Eigenwerte, so ist A ähnlich zu einer Jordan'schen Normalform.

(c) Ist $A \in \mathbb{K}^{n,n}$ diagonalisierbar, so ist A ähnlich zu einer Jordan'schen Normalform.

(d) *Sei $A \in \mathbb{C}^{n,n}$, so ist $A = D + N$ mit D diagonalisierbar, N nilpotent und $DN = ND$.*

(e) *A ist ähnlich zu $A^\top$.*

Beweis:

(a) Da jedes Polynom über $\mathbb{C}$ in Linearfaktoren zerfällt (siehe Vorlesung Algebra), so folgt die Behauptung aus Satz 5.1(a).

(b) Falls alle Eigenwerte von A reell sind, so zerfällt m_A als Teiler von P_A in Linearfaktoren, also folgt die Behauptung aus Satz 5.1(a).

(c) A ist ähnlich zu einer Diagonalmatrix. Diese ist eine Jordan'sche Normalform.

(d) Nach (a) ist A ähnlich zu einer Jordan'schen Normalform, d.h.,

$$\begin{aligned}
PAP^{-1} &= \bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} J(\lambda_i, p_{il}) \\
&= \bigoplus_{i=1}^{k} (\lambda_i I_{n_i}) + \bigoplus_{i=1}^{k} \left(\bigoplus_{l=1}^{s_i} N(p_{il}) \right), \\
D &= P^{-1} \left(\bigoplus_{i=1}^{k} \lambda_i I_{n_i} \right) P, \\
N &= P^{-1} \left(\bigoplus_{i=1}^{k} \bigoplus_{l=1}^{s_i} N(p_{il}) \right) P,
\end{aligned}$$

(natürlich in der gleichen Ordnung).

Es bleibt zu zeigen, dass $DN = ND$. (Übung)

(e) Aus der Jordan'schen Normalform folgt die Existenz einer invertierbaren Matrix P, so dass $PAP^{-1} = D + N$ mit D diagonal und N strikte obere Dreiecksmatrix. Dann folgt $P^{-\top} A^\top P^\top = D^\top + N^\top = D + N^\top$. Damit haben A und $A^\top$ das gleiche charakteristische Polynom und damit sind sie ähnlich.

$\square$

Beispiel 5.7 Betrachte

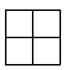
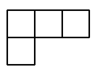
$$
\begin{aligned}
A &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & -1 & 2 \end{bmatrix}, \\
P_A &= \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda - 2 \end{bmatrix} \cdot \det \begin{bmatrix} \lambda & -1 \\ 1 & \lambda - 2 \end{bmatrix} \\
&= (\lambda^2 - 2\lambda + 1)^2 = (\lambda - 1)^4 , \\
(A - I)^1 &= \left[\begin{array}{cc|cc} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ \hline -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{array}\right], \\
(A - I)^2 &= \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} = 0 , \\
\text{also ist} m_A &= (\lambda - 1)^2 .
\end{aligned}
$$

Rang $(A - I_4)^0 = 4$,
Rang $(A - I_4)^1 = 2$.

Alle Eigenwerte sind reell. Nach Korollar 5.6(b) ist A ähnlich zu Jordan'scher Normalform.

$$
\begin{aligned}
n = \sum_{l=1}^{s} p_i &= 4 \\
p_1 + p_2 &= 4
\end{aligned}
$$

Es bleiben als Partitionen nur

$p = (2, 2)$  oder $p = (3, 1)$  .

Die duale Partition zu p ist $q = p^*$, und diese erhalten wir aus

$$\begin{array}{rcl}
q_j & = & \mathrm{Rang}\,(A - 1\,I_4)^{j-1} - \mathrm{Rang}\,(A - 1\,I_4)^{j}\,, \\
q_1 & = & 4 - 2 = 2\,, \\
q_2 & = & 2 - 0 = 2\,, \\
 & & \\
\Rightarrow & & q = p^* = (2,2), \\
\Rightarrow & & p = (2,2)\,.
\end{array}$$

Also

$$P^{-1}AP = \left[\begin{array}{cc|cc} 1 & 1 & & \\ & 1 & & \\ \hline & & 1 & 1 \\ & & & 1 \end{array}\right],$$

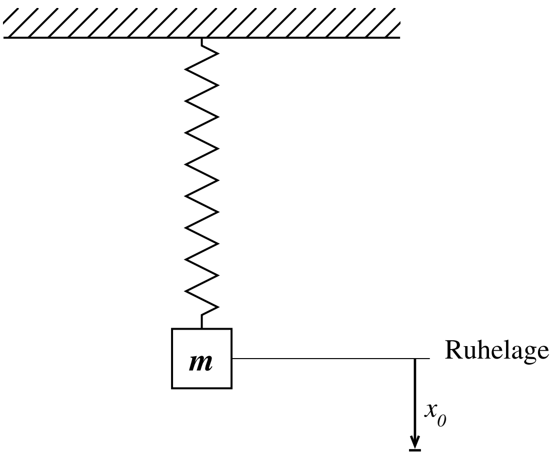
$$P = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Kapitel 6

Lineare Differentialgleichungen

Im Kapitel 16 haben wir die Jordan'sche Normalform einer Matrix kennengelernt. Wir wollen diese Ergebnisse nun anwenden, um damit die Lösungstheorie für lineare Differentialgleichungen zu bekommen.

Beispiel 6.1 *Schwingungsgleichung*
Ein Gewichtstück mit Masse m sei an einer Schraubenfeder mit Federkonstante μ aufgehängt. Zu Anfang wird das Gewicht um die Strecke x_0 ausgelenkt.



Außerdem erhalte es eine Anfangsgeschwindigkeit v_0. Die Bewegungsgleichung lautet (mit dem Hook'schen Gesetz):

$$x'' = \frac{d^2x}{dt^2} = -\frac{\mu}{m}x$$

mit Anfangsbedingungen

$$x(0) = x_0, \quad x'(0) = v_0 \,.$$

Wir führen als eine neue Variable die Geschwindigkeit ein:

$$v = x', \quad (v' = x'')$$

und erhalten das System von gewöhnlichen Differentialgleichungen 1. Ordnung

$$\begin{bmatrix} x \\ v \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ \frac{-\mu}{m} & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}. \tag{6.2}$$

Anfangsbedingungen : $\begin{bmatrix} x \\ v \end{bmatrix}(0) = \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}.$

Wir möchten eine allgemeine Theorie zur Lösung von linearen Systemen gewöhnlicher Differentialgleichungen

$$y' = Ay + f, \quad y(0) = y_0, \qquad A \in \mathbb{C}^{n,n}, \; y_0 \in \mathbb{C}^n, \tag{6.3}$$

wobei

$$y, f : [0, a] \to \mathbb{C}^n.$$

Satz 6.4 *Die Lösungen des homogenen Differentialgleichungssystems*

$$y' = Ay, \qquad y : [0, a] \to \mathbb{C}^n, \quad A \in \mathbb{C}^{n,n}, \tag{6.5}$$

bilden einen Untervektorraum $\mathcal{L}_A$ des (unendlich dimensionalen) Vektorraums der über dem Intervall $[0, a]$ stetig differenzierbaren Funktionen

$$\mathcal{C}^1\left([0, a], \mathbb{C}^n\right).$$

Beweis: Seien $y_1, \ldots, y_k$ Lösungen von (6.5), so folgt für alle $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$,

$$\left(\sum_{i=1}^{k} \alpha_i y_i\right)' = \sum_{i=1}^{k} \alpha_i y_i' = \sum_{i=1}^{k} \alpha_i A y_i = A\left(\sum_{i=1}^{k} \alpha_i y_i\right).$$

Also ist jede Linearkombination von Lösungen eine Lösung und die anderen Gesetze lassen sich leicht überprüfen, insbesondere ist die Funktion $y = 0$ das Nullelement. □

Satz 6.6 *Jede Lösung des inhomogenen Systems*

$$y' = Ay + f \tag{6.7}$$

ist von der Form

$$y = y_1 + y_2\,, \tag{6.8}$$

wobei y_1 eine spezielle Lösung von (6.7) ist und y_2 eine Lösung des homogenen Systems (6.5).

Beweis: Seien y, y_1 Lösungen von (6.7), so gilt dass

$$\begin{aligned} (y-y_1)' &= Ay - Ay_1 + f - f \\ &= A(y-y_1)\,, \end{aligned}$$

also ist $y - y_1$ Lösung des homogenen Systems. □

Wie erhalten wir nun die Lösungen von (6.3)?

Da $A \in \mathbb{C}^{n,n}$, so gibt es nach Satz 5.1 bzw. Folgerung 5.6 eine nichtsinguläre Matrix T, so dass TAT^{-1} in Jordan'scher Normalform ist. Es gilt

$$Ty' = (Ty)' = TAT^{-1}Ty + Tf, \tag{6.9}$$

$$Ty(0) = (Ty)(0) = Ty_0 =: z_0. \tag{6.10}$$

Setze $Ty =: z,\ Tf = g$, so folgt

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}' = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_k \end{bmatrix}, \tag{6.11}$$

wobei $J_i = \lambda_i I + N_i\,, \qquad N_i = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$ und λ_i Eigenwert von A.

Wir erhalten lauter kleine Systeme

$$z_i' = J_i z_i + g_i, \qquad i = 1, \dots, k. \tag{6.12}$$

Schauen wir uns so ein System im Detail an.

Sei $z_i = \begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,l_i} \end{bmatrix}, \quad g_i = \begin{bmatrix} g_{i,1} \\ \vdots \\ g_{i,l_i} \end{bmatrix},$ so gilt

$$\begin{bmatrix} z_{i,1}'(x) \\ \vdots \\ z_{i,l_i}'(x) \end{bmatrix} = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \begin{bmatrix} z_{i,1}(x) \\ \vdots \\ z_{i,l_i}(x) \end{bmatrix} + \begin{bmatrix} g_{i,1}(x) \\ \vdots \\ g_{i,l_i}(x) \end{bmatrix}, \tag{6.13}$$

$$\begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,l_i} \end{bmatrix}(0) = \begin{bmatrix} z_{i,1}^0 \\ \vdots \\ z_{i,l_i}^0 \end{bmatrix}.$$

Dieses System können wir durch Rückwärtseinsetzen lösen, sofern wir eine Gleichung der Form

$$w'(x) = \lambda w(x) + \gamma(x), \quad w(0) = w_0 \tag{6.14}$$

lösen können.

Multipliziere beide Seiten von (6.14) mit $e^{-\lambda x} \neq 0$, so gilt

$$e^{-\lambda x} w'(x) = \lambda e^{-\lambda x} w(x) + e^{-\lambda x} \gamma(x), \tag{6.15}$$

und da

$$\frac{d}{dx}\left(e^{-\lambda x} w(x)\right) \;=\; \frac{d}{dx}(e^{-\lambda x}) w(x) + e^{-\lambda x} w'(x) \;=\; -\lambda e^{-\lambda x} w(x) + e^{-\lambda x} w'(x),$$

so folgt äquivalent

$$\begin{array}{rcl} e^{-\lambda x}(w' - \lambda w) & = & \frac{d}{dx}(e^{-\lambda x} w) \;=\; e^{-\lambda x}\gamma \\ \iff \qquad e^{-\lambda x} w & = & \int\limits_0^x (e^{-\lambda t}\gamma) dt + c \\ \iff \qquad\qquad w & = & e^{\lambda x} \int\limits_0^x (e^{-\lambda t}\gamma) dt + e^{\lambda x} c \end{array}$$

$$w(0) = w_0 \implies w_0 = e^0 \left(\int\limits_0^0 e^{-\lambda t} \gamma dt \right) + ce^0 \implies \boxed{c = w_0}.$$

Satz 6.16 *Die eindeutige Lösung von (6.14) ist*

$$w = e^{\lambda x} \int\limits_0^x (e^{-\lambda t} \gamma) dt + e^{\lambda x} w_0 .$$

Beweis: Die Eindeutigkeit und Lösbarkeit folgt aus der obigen Konstruktion. □

Wir erhalten also rekursiv aus (6.13)

$$z_{i,l_i} = e^{\lambda_i x} \int\limits_0^x \left(e^{-\lambda_i t} g_{i,l_i}(t) \right) dt + z^0_{i,l_i} e^{\lambda_i x} \tag{6.17}$$

und

$$\begin{aligned} z_{i,j} \quad = \quad & e^{\lambda_i x} \int\limits_0^x e^{-\lambda_i t} \left(g_{i,j}(t) + z_{i,j+1}(t) \right) dt + z^0_{i,j} e^{\lambda_i x}, \\ & j = l_i - 1, \ldots, 1. \end{aligned} \tag{6.18}$$

Korollar 6.19 *Das Differentialgleichungssystem (6.3) hat eine eindeutige Lösung.*

Beweis: System (6.3) hat eine eindeutige Lösung $\iff$ (6.9), (6.10) hat eindeutige Lösung $\iff$ alle Systeme (6.13), $i = 1, \ldots, k$, haben eindeutige Lösung. □

Wir können also die Existenz und Eindeutigkeit nachweisen und auch die Lösung mit Hilfe der Jordan'schen Normalform von A bestimmen.

Um die Lösungsdarstellung zu vereinfachen, führen wir jetzt für $A \in \mathbb{C}^{n,n}$ die Exponentialfunktion für Matrizen ein.

Definition 6.20 *Die* <u>Exponentialfunktion</u> *von $A \in \mathbb{C}^{n,n}$ ist definiert durch*

$$e^A \;=\; \sum_{i=0}^{\infty} \frac{1}{i!} A^i \;=\; \lim_{k \to \infty} \sum_{i=0}^{k} \frac{1}{i!} A^i, \qquad (e^0 \;=\; I).$$

Lemma 6.21 *Es gilt für $A, B \in \mathbb{C}^{n,n}$:*

(a) $\dfrac{d}{dx}(e^{Ax}) = Ae^{Ax}$.

(b) $(e^A)^{-1} = e^{-A}$.

(c) Falls $AB = BA$*, so ist* $e^{A+B} = e^A \cdot e^B$.

Beweis: Mit formaler Potenzreihe. $\square$

Satz 6.22 *Die Lösung des Systems von linearen Differentialgleichungen*

$$y' = Ay + f, \qquad y(0) = y^0$$

ist gegeben durch

$$y = e^{Ax} \int\limits_0^x e^{-At} f dt + e^{Ax} y^0. \tag{6.23}$$

Beweis: Da wir die Eindeutigkeit schon nachgewiesen haben, brauchen wir nur zu zeigen, dass (6.23) eine Lösung darstellt.

$$\begin{aligned}
y' &= \frac{d}{dx}\left(e^{Ax}\int\limits_0^x e^{-At} f dt + e^{Ax} y^0\right) \\
&= \left(\frac{d}{dx} e^{Ax}\right)\int\limits_0^x e^{-At} f dt + e^{Ax}\left(\frac{d}{dx}\int\limits_0^x e^{-At} f dt\right) + \left(\frac{d}{dx} e^{Ax}\right) y^0 \\
&= A \underbrace{\left[e^{Ax}\int\limits_0^x e^{-At} f dt + e^{Ax} y^0\right]}_{y} + f\,.
\end{aligned}$$

Weiter gilt

$$\begin{aligned}
y(0) &= e^0 \int\limits_0^0 e^{-At} f dt + e^0 y^0 \\
&= y^0.
\end{aligned}$$

Das heißt, die Anfangsbedingung ist auch erfüllt. $\square$

Satz 6.24 *Sei $A \in \mathbb{C}^{n,n}$, so hat die Lösungsmenge des homogenen Systems $y' = Ay$ die Dimension n.*

Beweis: Die gleiche Konstruktion wie bei der Lösung von $w' = \lambda w + f$ ergibt

$$y = \underbrace{e^{Ax} \int_0^x e^{-At} 0 \, dt}_{0} + e^{Ax} c \,,$$

wobei c ein beliebiger Vektor aus $\mathbb{C}^n$ ist. □

Beispiel 6.25 Betrachte das System aus (6.2):

$$\begin{bmatrix} x \\ v \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}, \qquad \begin{bmatrix} x \\ v \end{bmatrix}(0) = \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}.$$

Lösung:

$$\begin{bmatrix} x \\ v \end{bmatrix} = e^{\begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix} t} \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}.$$

Da $\begin{bmatrix} 0 & 1 \\ -\frac{\mu}{m} & 0 \end{bmatrix}$ die Eigenwerte $\lambda_{1,2} = \pm i\sqrt{\frac{\mu}{m}}$ hat und die Eigenvektoren dazu sich aus

$$\begin{bmatrix} \mp i\sqrt{\frac{\mu}{m}} & 1 \\ -\frac{\mu}{m} & \mp i\sqrt{\frac{\mu}{m}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \qquad \text{ergeben,}$$

so ist $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -i\sqrt{\frac{\mu}{m}} \end{bmatrix}$ Eigenvektor zu $-i\sqrt{\dfrac{\mu}{m}}$

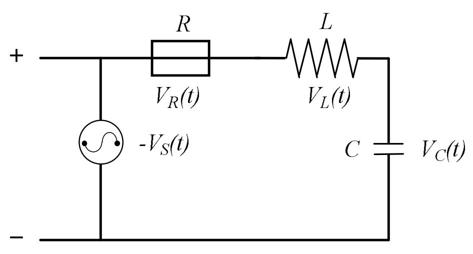
und $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ +i\sqrt{\frac{\mu}{m}} \end{bmatrix}$ Eigenvektor zu $+i\sqrt{\dfrac{\mu}{m}}$.

Kapitel 7

Lineare differentiell–algebraische Gleichungen

Heutzutage spielen bei der Modellierung von technischen Systemen neben Differentialgleichungen, sogenannte differentiell–algebraische Gleichungen eine besondere Rolle.

Beispiel 7.1 Betrachte den einfachen Schaltkreis



V_S ist die Spannung, R Widerstand, L Induktivität und C Kapazität. Die Spannungsabfälle an Widerstand, Spule und Kondensator sind V_R, V_L, V_C. $I(t)$ ist die Stromstärke. Die Anwendung der Kirchhoff'schen Gesetze ergibt das System

$$\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{I}(t) \\ \dot{V}_L(t) \\ \dot{V}_C(t) \\ \dot{V}_R(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} I(t) \\ V_L(t) \\ V_C(t) \\ V_R(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -V_S(t) \end{bmatrix}.$$

Dies ist ein lineares System von differentiellen und algebraischen Gleichungen.

Ein allgemeines System dieser Art hat die Form

$$E\dot{x} = Ax + f, \qquad E, A \in \mathbb{K}^{n,m}, \quad \begin{array}{l} x \in \mathcal{C}^1\left([t_0, t_1], \mathbb{K}^m\right), \\ f \in \mathcal{C}\left([t_0, t_1], \mathbb{K}^n\right). \end{array} \tag{7.2}$$

Wie können wir nun so ein System lösen? Ähnlich wie im Fall der linearen Differentialgleichungen können wir das System von links mit nichtsingulären Matrizen $P \in \mathbb{K}^{n,n}$ multiplizieren und außerdem $x = Qy$ setzen, wobei $Q \in \mathbb{K}^{m,m}$ nichtsingulär.

Definition 7.3 *Zwei Paare von Matrizen* (E_1, A_1), (E_2, A_2), $\quad E_i, A_i \in \mathbb{K}^{n,m}$, $i = 1, 2$, *heißen* äquivalent, *wenn es nichtsinguläre Matrizen* $P \in \mathbb{K}^{n,n}, Q \in \mathbb{K}^{m,m}$ *gibt, so dass*

$$(E_1, A_1) = (PE_2Q, PA_2Q) \tag{7.4}$$

Beispiel 7.5 $m = n, E_1 = E_2 = I_n, P = Q^{-1}$ ergibt die Ähnlichkeit von A_1 und A_2. Wir erwarten also eine Verallgemeinerung der Jordan'schen Normalform.

Wir wollen hier nur einen Spezialfall betrachten, nämlich sogenannte *reguläre Paare.*

Definition 7.6 *Ein komplexes Matrizenpaar* (E, A), $\quad E, A \in \mathbb{C}^{n,m}$, *heißt* regulär, *wenn* $n = m$ *und ein* $\lambda \in \mathbb{C}$ *existiert, so dass*

$$\det(\lambda E - A) \neq 0.$$

Falls (E, A) *regulär, so heißen die Nullstellen von* $\det(\lambda E - A)$ verallgemeinerte *Eigenwerte von* (E, A).

Beispiel 7.7 (I, A) ist immer regulär, denn für alle λ, die nicht Eigenwert von A sind, gilt $\det(\lambda I - A) \neq 0$. Verallgemeinerte Eigenwerte von (I, A) sind die Eigenwerte von A.

Beispiel 7.8 Für den Schaltkreis aus Beispiel 7.1 gilt

$$\det\left(\lambda\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}\right)$$

$$= \det\begin{bmatrix} \lambda L & -1 & 0 & 0 \\ -1 & 0 & \lambda C & 0 \\ R & 0 & 0 & -1 \\ 0 & -1 & -1 & -1 \end{bmatrix} = \lambda L \det\begin{bmatrix} 0 & \lambda C & 0 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix} + \det\begin{bmatrix} -1 & \lambda C & 0 \\ R & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$= \lambda^2 LC + (-1)\det\begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix} - (\lambda C)\det\begin{bmatrix} R & -1 \\ 0 & -1 \end{bmatrix} = \lambda^2 LC + \lambda RC + 1\,.$$

Es gibt also nur zwei verallgemeinerte Eigenwerte $\lambda_{1,2} = -\dfrac{1}{2L}\left(R \pm \sqrt{R^2 - \dfrac{4L}{C}}\right)$, obwohl die Matrizen 4×4 sind.

Beispiel 7.9

$$E = \begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & & \\ & 0 & \\ & & 1 \end{bmatrix}$$

so folgt

$$\det(\lambda E - A) = \begin{bmatrix} -1 & \lambda & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & -1 \end{bmatrix} = 0 \qquad \text{für alle } \lambda,$$

also ist das Paar (E, A) *nicht regulär.*

Lemma 7.10 *Jedes Matrizenpaar, welches äquivalent zu einem regulären Paar ist, ist ebenfalls regulär. Äquivalente Paare haben dieselben verallgemeinerten Eigenwerte.*

Beweis: Sei $E_2 = PE_1Q,\ A_2 = PA_1Q$, mit P, Q nicht singulär. Dann gilt

$$\begin{aligned} \det(\lambda E_2 - A_2) &= \det(\lambda PE_1Q - PA_1Q) \\ &= \det P \cdot \det(\lambda E_1 - A_1) \cdot \det Q\,, \end{aligned}$$

also gilt

$$\det(\lambda E_2 - A_2) = 0$$

genau dann wenn

$$\iff \quad \det(\lambda E_1 - A_1) = 0.$$

$\square$

Satz 7.11 Weierstraß-Normalform
Seien $E, A \in \mathbb{C}^{n,n}$ und sei (E, A) regulär. Dann gibt es nichtsinguläre Matrizen $P, Q \in \mathbb{C}^{n,n}$, so dass

$$PEQ = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \quad PAQ = \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix},$$

wobei J eine Matrix in Jordan'scher Normalform ist und N eine nilpotente Matrix in Jordan'scher Normalform. Dabei sind N, J (bis auf Permutation der Jordanblöcke zu verschiedenen Eigenwerten) eindeutig bestimmt.

Beweis: Da (E, A) regulär ist, so gibt es $\lambda_0 \in \mathbb{C}$ mit $\det(\lambda_0 E - A) \neq 0$. Dann ist (E, A) äquivalent zu

$$\begin{aligned}
& ((\lambda_0 E - A)^{-1} E, (\lambda_0 E - A)^{-1} A) \\
= \quad & ((\lambda_0 E - A)^{-1} E, (\lambda_0 E - A)^{-1}(-\lambda_0 E + A + \lambda_0 E)) \\
= \quad & ((\lambda_0 E - A)^{-1} E, -I + \lambda_0 (\lambda_0 E - A)^{-1} E) \\
=: \quad & (\tilde{E}, -I + \lambda_0 \tilde{E})
\end{aligned}$$

Sei P_1 nichtsingulär, so dass

$$P_1^{-1} \tilde{E} P_1 = P_1^{-1} \left((\lambda_0 E - A)^{-1} E \right) P_1 = \tilde{J}$$

in Jordan'scher Normalform ist (diese existiert nach Folgerung 5.6). Dann ist $(\tilde{E}, -I + \lambda_0 \tilde{E})$ äquivalent zu $\left(P_1^{-1} \tilde{E} P_1, P_1^{-1}(-I + \lambda_0 \tilde{E}) P_1 \right) = (\tilde{J}, -I + \lambda_0 \tilde{J})$.

Teile $\tilde{J}$ auf als

$$\begin{bmatrix} \tilde{J}_1 & 0 \\ 0 & \tilde{N} \end{bmatrix}, \quad \text{mit } \tilde{J}_1 \text{ nichtsingulär und } \tilde{N} \text{ nilpotent.}$$

Wir erhalten also

$$(\tilde{J}, -I + \lambda_0 \tilde{J}) = \left(\begin{bmatrix} \tilde{J}_1 & 0 \\ 0 & \tilde{N} \end{bmatrix}, \left[\begin{array}{c|c} -I + \lambda_0 \tilde{J}_1 & 0 \\ \hline 0 & -I + \lambda_0 \tilde{N} \end{array}\right] \right).$$

Die Matrix $-I + \lambda_0 \tilde{N}$ ist nichtsingulär, da $\tilde{N}$ strikte obere Dreiecksmatrix ist, also folgt, dass $(\tilde{J}, -I + \lambda_0 \tilde{J})$ äquivalent zu

$$\left(\begin{bmatrix} \tilde{J}_1^{-1} & 0 \\ 0 & (-I + \lambda_0 \tilde{N})^{-1} \end{bmatrix} \begin{bmatrix} \tilde{J}_1 & 0 \\ 0 & \tilde{N} \end{bmatrix}, \begin{bmatrix} \tilde{J}_1^{-1} & 0 \\ 0 & (-I + \lambda_0 \tilde{N})^{-1} \end{bmatrix} \begin{bmatrix} -I + \lambda_0 \tilde{J}_1 & 0 \\ 0 & -I + \lambda_0 \tilde{N} \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} I & 0 \\ 0 & \hat{N} \end{bmatrix}, \begin{bmatrix} \hat{J} & 0 \\ 0 & I \end{bmatrix} \right)$$

ist. Seien P_2, P_3 nichtsinguläre Matrizen, so dass $J = P_2 \hat{J} P_2^{-1}$ in Jordan'scher Normalform und $N = P_3 \hat{N} P_3^{-1}$ nilpotent in Jordan'scher Normalform ist. Dann folgt, dass

$$\begin{bmatrix} P_2 & 0 \\ 0 & P_3 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \hat{N} \end{bmatrix} \begin{bmatrix} P_2^{-1} & 0 \\ 0 & P_3^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}$$

und

$$\begin{bmatrix} P_2 & 0 \\ 0 & P_3 \end{bmatrix} \begin{bmatrix} \hat{J} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_2^{-1} & 0 \\ 0 & P_3^{-1} \end{bmatrix} = \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}.$$

Wir haben gezeigt, dass jedes reguläre Paar äquivalent zu einem Paar der Form

$$\left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right)$$

ist, und müssen jetzt noch die Eindeutigkeit zeigen.

Seien

$$\left(\begin{bmatrix} I & 0 \\ 0 & N_1 \end{bmatrix}, \begin{bmatrix} J_1 & 0 \\ 0 & I \end{bmatrix} \right), \qquad \left(\begin{bmatrix} I & 0 \\ 0 & N_2 \end{bmatrix}, \begin{bmatrix} J_2 & 0 \\ 0 & I \end{bmatrix} \right)$$

zwei Normalformen von (E, A), d.h.

$$P_1EQ_1 = \begin{bmatrix} I & 0 \\ 0 & N_1 \end{bmatrix}, \quad P_1AQ_1 = \begin{bmatrix} J_1 & 0 \\ 0 & I \end{bmatrix},$$
$$P_2EQ_2 = \begin{bmatrix} I & 0 \\ 0 & N_2 \end{bmatrix}, \quad P_2AQ_2 = \begin{bmatrix} J_2 & 0 \\ 0 & I \end{bmatrix}.$$

$$P_1P_2^{-1} \begin{bmatrix} I & 0 \\ 0 & N_2 \end{bmatrix} Q_2^{-1}Q_1 = \begin{bmatrix} I & 0 \\ 0 & N_1 \end{bmatrix},$$
$$P_1P_2^{-1} \begin{bmatrix} J_2 & 0 \\ 0 & I \end{bmatrix} Q_2^{-1}Q_1 = \begin{bmatrix} J_1 & 0 \\ 0 & I \end{bmatrix}.$$

Setze $P_1P_2^{-1} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$, $Q_1^{-1}Q_2 = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$, so folgt $P_{11} = Q_{11},\ P_{12}N_2 = Q_{12},\ P_{21} = N_1Q_{21},\ P_{22}N_2 = N_1Q_{22}$ und $P_{11}J_2 = J_1Q_{11},\ P_{12} = J_1Q_{12},\ P_{21}J_2 = Q_{21},\ P_{22} = Q_{22}$. Damit gilt, dass $P_{21} = N_1P_{21}J_2 = N_1^2P_{21}J_2^2 = \ldots = N_1^lP_{21}J_2^l = 0$, fr ein l. Dies impliziert, dass $P_{11} = Q_{11},\ P_{22} = Q_{22}$ nichtsingulär sind und das $P_{11}J_2P_{11}^{-1} = J_1$. Damit folgt $J_1 = J_2$, bis auf Permutation verschiedener Jordanblöcke und analog folgt aus $P_{22}N_2P_{22}^{-1} = N_1$, dass $N_1 = N_2$, denn die Jordan'sche Normalform ist (bis auf Permutationen) eindeutig. □

Beispiel 7.12 *Betrachte das Paar (E, A) aus dem Schaltungsbeispiel. Es gilt*

$$\left(\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right)$$
$$\sim \left(\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & -1 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right)$$
$$\sim \left(\begin{bmatrix} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -R & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right)$$

$$
\begin{aligned}
&\sim \left(\left[\begin{array}{ccc|c} L & 0 & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{ccc|c} -R & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \right) \\
&\sim \left(\left[\begin{array}{ccc|c} L & 0 & 0 & 0 \\ 0 & C & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{ccc|c} -R & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \right) \\
&\sim \left(\left[\begin{array}{cc|c|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{cc|c|c} -\frac{R}{L} & -\frac{1}{L} & 0 & 0 \\ \frac{1}{C} & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \right) \\
&\sim \left(\left[\begin{array}{cc|c|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{cc|c|c} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \right).
\end{aligned}
$$

Es folgt, dass

$$
N = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \end{array} \right], \quad J = \left[\begin{array}{cc} \lambda_1 & 0 \\ 0 & \lambda_2 \end{array} \right].
$$

Definition 7.13 *Sei (E, A) ein reguläres Paar mit Weierstraß–Normalform*

$$
\left(\left[\begin{array}{cc} I & 0 \\ 0 & N \end{array} \right], \left[\begin{array}{cc} J & 0 \\ 0 & I \end{array} \right] \right),
$$

dann heißt der Nilpotenzindex ν von N der Index von (E, A). *Wir schreiben $\nu = \mathrm{Ind}\,(E, A)$).*

Die verallgemeinerten Eigenwerte sind die Eigenwerte von J und dazu kommt noch der Eigenwert ∞ von $(\lambda N - I)$, denn

$$
\det(N - \mu I) = 0
$$

hat nur die Nullstelle 0 und das charakteristische Polynom $\mu^{\mathrm{Ind}\,(N,I)}$.

$$
\det(N - \mu I) = \mu \det \left(\frac{1}{\mu} N - I \right).
$$

Für „$\lambda = \frac{1}{\mu} = \infty$“, d.h. ($\mu = \frac{1}{\lambda} = 0$), ergibt sich die entsprechende Jordanstruktur von N.

Nun können wir auch die differentiell-algebraische Gleichung

$$E\dot{x} = Ax + f, \qquad x(t_0) = x^0,$$

mit (E, A) regulär, lösen. Seien dazu P, Q die Matrizen, die (E, A) auf Weierstraß–Normalform bringen. Dann gilt mit $x = Qy$

$$PEQ\dot{y} = PAQy + Pf, \quad y(t_0) = Q^{-1}x^0 =: y^0.$$

Teile $g := Pf$ und y entsprechend der Aufteilung von (PEQ, PAQ) auf, d.h.

$$\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \tag{7.14}$$

$$y_1(t_0) = y_1^0, \quad y_2(t_0) = y_2^0.$$

Dann ist y_1 die Lösung von

$$\dot{y}_1 = Jy_1 + g_1, \quad y_1(t_0) = y_1^0. \tag{7.15}$$

Für die Existenz, Eindeutigkeit und Lösungsdarstellung siehe letztes Kapitel. Der Teil y_2 ist die Lösung von

$$N\dot{y}_2 = y_2 + g_2 \tag{7.16}$$

mit Anfangsbedingung

$$y_2(t_0) = y_2^0. \tag{7.17}$$

Satz 7.18 *Sei $g_2 \in \mathcal{C}^\nu\left([t_0, t_1], \mathbb{C}^l\right)$, wobei ν der Nilpotenzindex der nilpotenten Matrix $N \in \mathbb{C}^{l,l}$ ist, so hat die Lösung von (7.16) die Form*

$$y_2 = -\sum_{i=0}^{\nu-1} N^i g_2^{(i)}. \tag{7.19}$$

Die Lösung ist eindeutig, falls keine Anfangsbedingung (7.17) gegeben ist. Ist $y_2(t_0) = y_2^0$ als Anfangsbedingung gegeben, so existiert die Lösung und ist eindeutig genau dann, wenn

$$y_2^0 = -\sum_{i=0}^{\nu-1} N^i g_2^{(i)}(t_0). \tag{7.20}$$

Beweis: Dass jede Funktion der Form (7.19) die Gleichung (7.16) löst, folgt durch Einsetzen.

$$\begin{aligned}
N\dot{y}_2 &= N \cdot \left(-\sum_{i=0}^{\nu-1} N^i g_2^{(i+1)}\right) \\
&= -\sum_{i=0}^{\nu-1} N^{i+1} g_2^{(i+1)} \\
&= -\sum_{i=1}^{\nu} N^i g_2^{(i)} \\
&= -\sum_{i=1}^{\nu-1} N^i g_2^{(i)} \qquad (N^\nu = 0) \\
&= \underbrace{-\sum_{i=0}^{\nu-1} N^i g_2^{(i)}}_{y_2} + \underbrace{N^0 g_2^{(0)}}_{g_2}.
\end{aligned}$$

Dass jede Lösung die Form (7.19) hat, folgt durch Lösung der einzelnen Teilblöcke der Jordan–Form von N.

$$N = \begin{bmatrix} N_1 & & \\ & \ddots & \\ & & N_k \end{bmatrix}, \quad N_i = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \in \mathbb{C}^{l_i, l_i}$$

$$N_i \dot{z}_i = z_i + h_i.$$

Der Rest ist klar. Da (7.19) schon eindeutig ist, muss die Anfangsbedingung erfüllt sein oder die Lösung existiert nicht. □

Bemerkung 7.21 *Im Gegensatz zu gewöhnlichen Differentialgleichungen können wir Anfangsbedingungen nicht beliebig wählen.*
*Außerdem muss die Inhomogenität genügend oft (*Ind (E, A)*-mal) differenzierbar sein.*

Kapitel 8

Bilinearformen

Ein wichtiges Hilfsmittel in vielen Teilen der Mathematik ist der Begriff der Bilinearform, den wir nun einführen wollen.

Definition 8.1 *Sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine Abbildung*

$$\beta : V \times V \to \mathbb{K}$$

heißt Bilinearform*, falls folgende Bedingungen gelten:*

(i) $\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w)$,

(ii) $\beta(\lambda v, w) = \lambda\beta(v, w)$,

(iii) $\beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2)$,

(iv) $\beta(v, \lambda w) = \lambda\beta(v, w)$

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

Die Bilinearform heißt symmetrisch*, falls*

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

Die Abbildung β heißt Hermite'sche Form*, falls (i), (ii) und (v) erfüllt sind.*

(v) $\beta(v, w) = \overline{\beta(w, v)}, \quad \text{für alle } v, w \in V.$

Beispiel 8.2

(a) Das in Kap. 9 eingeführte Skalarprodukt in $\mathbb{R}^n$

$$\langle u, v\rangle = v^\top u$$

ist eine symmetrische Bilinearform. Das analoge komplexe Skalarprodukt in $\mathbb{C}^n$ ist eine Hermite'sche Form.

(b) Sei $V = \mathcal{C}([a, b], \mathbb{R})$. So bildet die Abbildung

$$\beta(f, g) = \int_a^b f \cdot g dx$$

eine symmetrische Bilinearform.

Lemma 8.3

(i) Sei V ein Vektorraum über $\mathbb{R}$ und $\beta : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. β definiert ein Skalarprodukt in V genau dann, wenn $\beta(v, v) > 0$ für $v \neq 0$.

(ii) Sei V ein Vektorraum über $\mathbb{C}$ und $\beta : V \times V \to \mathbb{C}$ eine Hermite'sche Form. β definiert ein komplexes Skalarprodukt in V genau dann, wenn $\beta(v, v) > 0$ für $v \neq 0$.

Beweis: Aus den Eigenschaften der symmetrischen Bilinearform (Hermite'schen Form) folgen sofort Bedingungen (1) und (2) des Skalarprodukts und die dritte Bedingung „positive Definitheit“ ist eben genau

$$\beta(v, v) > 0, \qquad v \neq 0.$$

□

Zu jeder Bilinearform können wir Matrixdarstellungen konstruieren. Sei V ein endlichdimensionaler Vektorraum über $\mathbb{K}$ und $\beta : V \times V \to \mathbb{K}$ eine Bilinearform. Sei $\{v_1, \cdots, v_n\}$ eine Basis von V. Dann heißt die Matrix $B \in \mathbb{K}^{n,n}$, mit den Einträgen

$$b_{ij} = \beta(v_j, v_i), \quad i, j = 1, \dots, n,$$

die Matrixdarstellung von β bezüglich der Basis $\{v_1, \dots, v_n\}$.

Falls β eine symmetrische Bilinearform ist, so ist die Matrixdarstellung B bezüglich jeder Basis symmetrisch, d.h., $B = B^\top$. Falls B eine Hermite'sche

Form ist, so ist die Matrixdarstellung B bezüglich jeder Basis Hermite'sch, d.h., $B = B^H = \overline{B}^\top$.

Beachte, auch im komplexen Fall gibt es symmetrische Bilinearformen, die nicht Hermite'sch sind (komplex symmetrische Matrizen).

Beispiel 8.4 Sei

$$V = \mathbb{C}^3 \, , D = \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix} .$$

Dann ist

$$\beta_1(v,w) = w^\top Dv = v_1w_1 + v_2w_2 - v_3w_3$$

ist eine symmetrische Bilinearform, denn

$$\beta_1(w,v) \;=\; w_1v_1 + w_2v_2 - w_3v_3 \;=\; \beta_1(v,w)\,.$$

Dagegen ist $\quad \beta_2(v,w) = w^H Dv \quad$ eine Hermite'sche Form.

Für die Basis $\{e_1, e_2, e_3\}$ ergibt sich die Matrixdarstellung $B = D$ und für

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ c \\ -\overline{s} \end{bmatrix}, \begin{bmatrix} 0 \\ s \\ \overline{c} \end{bmatrix} \right\} \qquad \text{mit} \qquad |c|^2 + |s|^2 = 1$$

ergibt sich

$$\begin{aligned} B &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & -\overline{s} & \overline{c} \end{bmatrix}^H D \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & -\overline{s} & \overline{c} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \overline{c} & -s \\ 0 & \overline{s} & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & \overline{s} & -\overline{c} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & |c|^2 - |s|^2 & 2\overline{c}s \\ 0 & 2\overline{s}c & |s|^2 - |c|^2 \end{bmatrix} . \end{aligned}$$

Lemma 8.5 *Seien* $\{v_1, \cdots, v_n\}$, $\{\tilde{v}_1, \cdots, \tilde{v}_n\}$ *Basen eines Vektorraumes* V *über* $\mathbb{R}$ *und sei* $P = [p_{i,j}]$ *die Basisübergangsmatrix zwischen diesen Basen, d.h.,*

$$v_i = \sum_{j=1}^{n} p_{j,i} v_j$$

. *Ist* $\beta : V \times V \to \mathbb{R}$ *eine Bilinearform und* B *die Matrixdarstellung von* β *bezüglich* $\{v_1, \cdots, v_n\}$, *so ist*

$$\tilde{B} = P^\top B P$$

die Matrixdarstellung von β *bezüglich* $\{\tilde{v}_1, \cdots, \tilde{v}_n\}$.

Beweis: Falls

$$\begin{aligned} v &= \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \tilde{\lambda}_i \tilde{v}_i, \\ w &= \sum_{i=1}^{n} \delta_i v_i = \sum_{i=1}^{n} \tilde{\delta}_i \tilde{v}_i \end{aligned}$$

und

$$\Lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \tilde{\Lambda} = \begin{bmatrix} \tilde{\lambda}_1 \\ \vdots \\ \tilde{\lambda}_n \end{bmatrix}, \quad \Delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix}, \quad \tilde{\Delta} = \begin{bmatrix} \tilde{\delta}_1 \\ \vdots \\ \tilde{\delta}_n \end{bmatrix},$$

so gilt $\Lambda = P\tilde{\Lambda}$, $\Delta = P\tilde{\Delta}$ und

$$\begin{aligned} \beta(v, w) &= \beta\left(\sum_{i=1}^{n} \lambda_i v_i, \sum_{j=1}^{n} \delta_j v_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \delta_j \beta(v_i, v_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \delta_j b_{j,i} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}^\top B \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix} = (P\tilde{\Lambda})^\top B (P\tilde{\Delta}) \\ &= \tilde{\Lambda}^\top \underbrace{P^\top B P}_{\tilde{B}} \tilde{\Delta} = \sum_{ij=1}^{n} \tilde{\lambda}_i \tilde{\delta}_j \tilde{b}_{j,i} = \sum_{i=1}^{n} \tilde{\lambda}_i \tilde{\delta}_j \beta(\tilde{v}_i, \tilde{v}_j) \\ &= \beta(\tilde{v}, \tilde{w}) . \end{aligned}$$

□

Definition 8.6 *Seien $A, B \in \mathbb{R}^{n,n}$, so heißen A und B* kongruent*, falls es eine nichtsinguläre Matrix P gibt, so dass $B = P^\top AP$.*
Analog, falls $A, B \in \mathbb{C}^{n,n}$, so heißen A und B kongruent*, falls es eine nichtsinguläre Matrix P gibt, so dass $B = P^H AP$.*

Zur Erinnerung: Ein reeller Vektorraum V mit Skalarprodukt (positiv definiter symmetrischer Bilinearform) heißt *Euklidischer Raum* (V, β) oder $(V, \langle\cdot,\cdot\rangle)$.

Analog nennen wir einen komplexen Vektorraum mit Skalarprodukt (positiv definiter Hermite'scher Form) *unitären Raum* (V, β) oder $(V, \langle\cdot,\cdot\rangle)$.

Definition 8.7 *Sei V ein euklidischer (unitärer) Raum. Ein Endomorphismus $f : V \to V$ heißt orthogonal (unitär), falls*

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

für alle $v, w \in V$.

Korollar 8.8 *Sei V ein euklidischer (unitärer) Raum mit Orthonormalbasis $\{v_1, \cdots, v_n\}$. Sei $f : V \to V$ ein orthogonaler (unitärer) Endomorphismus. Dann gelten folgende Aussagen:*

(i) Die Matrixdarstellung von f bezüglich der Basis $\{v_1, \cdots, v_n\}$ ist eine orthogonale (unitäre) Matrix. (Die Umkehrung, dass jede orthogonale (unitäre) Matrix einen orthogonalen (unitären) Endomorphismus definiert, gilt natürlich auch.)

(ii) f ist injektiv und f^{-1} ist ebenfalls orthogonal (unitär).

(iii) Ist λ ein Eigenwert von f, so ist $|\lambda| = 1$.

(iv) Ist $\|\cdot\|$ die durch $\langle\cdot,\cdot\rangle$ induzierte Norm, d.h., $\| v \| = \sqrt{\langle v, v\rangle}$ so gilt $\| f(v) \| = \| v \|$.

Beweis:

(i) Sei $F = [f_{ij}]$ die Matrixdarstellung von f bezüglich einer Basis $\{v_1, \ldots, v_n\}$, d.h.,

$$f(v_j) = \sum_{i=1}^{n} f_{ij} v_i \quad j = 1, \ldots, n\,.$$

Falls $v = \sum_{i=1}^{n} \lambda_i v_i,\ w = \sum_{i=1}^{n} \gamma_i v_i$, und

$$\Lambda \ = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \Delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix},$$

so folgt

$$\langle v, w\rangle = \sum_{i,j=1}^{n} \lambda_i \gamma_j \langle v_i, v_j\rangle \ = \ \sum_{i=1}^{n} \lambda_i \gamma_i = \Gamma^\top \Lambda$$

und damit

$$\begin{aligned} \langle f(v), f(w)\rangle \ &= \ \sum_{i,j=1}^{n} \lambda_i \gamma_j \langle f(v_i), f(v_j)\rangle \\ &= \ (F\Gamma)^\top (F\Lambda) = \Gamma^\top F^\top F \Lambda. \end{aligned}$$

Falls f orthogonal ist, so folgt $\langle v, w\rangle = \langle f(v), f(w)\rangle \quad$ für alle $v, w \in V$, und damit:

$$\Gamma^\top F^\top F \Lambda = \Gamma^\top \Lambda, \qquad \text{für alle } \Lambda, \Gamma \in \mathbb{K}^n$$

insbesondere damit $F^\top F = I$.

Die Umkehrung ist trivialerweise erfüllt.

(iii) Sei v ein Eigenvektor von f, d.h., $f(v) = \lambda v$.
Dann folgt, da f (als invertierbare Matrix F) nicht den Eigenwert 0 hat, dass

$$\begin{aligned} 0 \neq \langle v, v\rangle \ &= \ \langle f(v), f(v)\rangle = \langle \lambda v, \lambda v\rangle \\ &= \ \begin{cases} \lambda^2 \langle v, v\rangle\,, & \text{falls} \quad \mathbb{K} = \mathbb{R} \ \text{ und } \lambda \in \mathbb{R} \\ \lambda\overline{\lambda} \langle v, v\rangle\,, & \text{falls} \quad \mathbb{K} = \mathbb{C} \ \text{ oder } \lambda \notin \mathbb{R}\,, \end{cases}, \end{aligned}$$

also

$$1 = \begin{cases} \lambda^2, & \text{falls} \quad \mathbb{K} = \mathbb{R} \ \text{ und } \lambda \in \mathbb{R} \\ |\lambda|^2, & \text{falls} \quad \mathbb{K} = \mathbb{C} \ \text{ oder } \lambda \notin \mathbb{R}\,. \end{cases}$$

(Falls $\mathbb{K} = \mathbb{R}$ und $\lambda \notin \mathbb{R}$, so muss das komplexe Skalarprodukt verwendet werden.)
Es folgt $\lambda \in \{+1, -1\}$, falls $\mathbb{K} = \mathbb{R}$ und $\lambda \in \mathbb{R}$, und $|\lambda| = 1$, falls $\mathbb{K} = \mathbb{C}$ oder $\lambda \notin \mathbb{R}$.

(iv)

$$\begin{aligned}\|v\| &= \sqrt{\langle v,v\rangle} = \sqrt{\langle f(v),f(v)\rangle} \\ &= \|f(v)\|\end{aligned}$$

□

Lemma 8.9 *Sei V ein euklidischer Vektorraum und $f : V \to V$ eine beliebige Abbildung mit $f(0) = 0$ und*

$$\|v-w\| \;=\; \|f(v)-f(w)\| \;=\; \|f(v-w)\| \qquad \textit{für alle } v,w \in V\,,$$

so ist f ein orthogonaler Endomorphismus.

Beweis: Es gilt für alle $v, w \in V$

$$\langle v,w\rangle = \frac{1}{2}\left(\langle v,v\rangle + \langle w,w\rangle - \langle v-w,v-w\rangle\right)$$

Analog

$$\begin{aligned}\langle f(v),f(w)\rangle &= \frac{1}{2}\left(\langle f(v),f(v)\rangle + \langle f(w),f(w)\rangle - \langle f(v-w),f(v-w)\rangle\right) \\ &= \frac{1}{2}\left(\langle f(v),f(v)\rangle + \langle f(w),f(w)\rangle - \langle v-w,v-w\rangle\right)\end{aligned}$$

mit $w = 0$ ergibt sich

$$\begin{aligned}0 \;=\; \langle f(v),0\rangle &= \frac{1}{2}\left(\langle f(v),f(v)\rangle + \langle 0,0\rangle - \langle v,v\rangle\right) \\ \Longrightarrow \qquad & \\ \langle f(v),f(w)\rangle &= \frac{1}{2}\left((v,v) + \langle w,w\rangle - \langle v-w,v-w\rangle\right) \\ &= \langle v,w\rangle\,.\end{aligned}$$

Wir müssen nur noch die Linearität von f zeigen, d.h.,

$$f(\lambda_1 v_1 + \lambda_2 v_2) \;=\; \lambda_1 f(v_1) + \lambda_2 f(v_2), \qquad \forall\, v_1,v_2 \in V,\; \lambda_1,\lambda_2 \in \mathbb{K}.$$

$$
\begin{aligned}
&\|f(\lambda_1 v_1 + \lambda_2 v_2) - \lambda_1 f(v_1) - \lambda_2 f(v_2)\|^2 \\
&\quad = \langle f(\lambda_1 v_1 + \lambda_1 v_2), f(\lambda_1 v_1 + \lambda_2 v_2)\rangle + \lambda_1^2 \langle f(v_1), f(v_1)\rangle + \lambda_2^2 \langle f(v_2), f(v_2)\rangle \\
&\qquad - 2\lambda_1 \langle f(v_1), f(\lambda_1 v_1 + \lambda_2 v_2)\rangle - 2\lambda_2 \langle f(v_2), f(\lambda_1 v_1 + \lambda_2 v_2)\rangle \\
&\qquad + 2\lambda_1\lambda_2 \langle f(v_1), f(v_2)\rangle \\
&\quad = \langle \lambda_1 v_1 + \lambda_2 v_2, \lambda_1 v_1 + \lambda_2 v_2\rangle + \lambda_1^2 \langle v_1, v_1\rangle + \lambda_2^2 \langle v_2, v_2\rangle \\
&\qquad - 2\langle \lambda_1 v_1 + \lambda_2 v_2, \lambda_1 v_1 + \lambda_2 v_2\rangle \\
&\qquad + 2\lambda_1\lambda_2 \langle v_1, v_2\rangle \\
&\quad = 0.
\end{aligned}
$$

Damit folgt die Behauptung. $\square$

Wir haben bereits gesehen, dass alle Eigenwerte von orthogonalen bzw. unitären Matrizen den Betrag 1 haben und wir hatten bereits den Satz von Schur gezeigt. Im reellen erhallten wir eine reelle Version des Satzes von Schur.

Satz 8.10 ***Reelle Schur-Form*** *(Reelle Version von Satz 2.15)*

Sei $A \in \mathbb{R}^{n,n}$, dann gibt es $U \in \mathcal{O}(n) = \{A \in \mathbb{R}^{n,n} \mid A \text{ orthogonal }\}$, so dass

$$
U^{-1}AU = \begin{bmatrix} R_{11} & \cdots & R_{1m} \\ & \ddots & \vdots \\ & & R_{mm} \end{bmatrix}
$$

wobei die Diagonalblöcke R_{ii} entweder 1×1 oder 2×2 sind, und falls sie 2×2 sind, so haben sie ein Paar komplex konjugierter Eigenwerte.

Beweis: Da A reell ist, treten alle nicht reellen Eigenwerte in komplex konjugierten Paaren auf. Falls alle Eigenwerte reell sind, folgt der Beweis genau wie bei Satz 2.15, nur reell. Ansonsten verwenden wir vollständige Induktion:

I.A.: $n = 1, n = 2 \quad \Rightarrow \quad$ klar.

I.V.: Die Behauptung sei richtig für $A \in \mathbb{R}^{n-1,n-1}$

I.S.: Falls $\lambda = a + ib,\ a, b \in \mathbb{R},$ Eigenwert von A mit Eigenvektor $x + iy$, so gilt

$$
A(x + iy) = (a + ib)(x + iy)
$$

und nach Konjugation

$$A(x - iy) = (a - ib)(x - iy), \qquad (x + iy)^\top (x - iy) = 1.$$

Es gilt dann durch Addition und Subtraktion dieser Gleichungen

$$A[x, y] = [x, y] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

und damit ist $[x, y]$ ein reeller invarianter Unterraum.

Mit der QR-Zerlegung (Gram-Schmidt, Folgerung 10.13) gibt es U_1 orthogonal, so dass

$$U_1^\top [x, y] = \left[\begin{array}{cc} x_{11} & y_{11} \\ 0 & y_{21} \\ \hline 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{array}\right], \qquad \text{mit} \quad \begin{bmatrix} x_{11} & y_{11} \\ 0 & y_{21} \end{bmatrix} \quad \text{invertierbar,}$$

und damit

$$U_1^\top A U_1 U_1^\top [x, y] = U_1^\top [x, y] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

und es folgt, dass

$$U_1^\top A U_1 = \left[\begin{array}{cc|c} a_{11} & a_{12} & \multirow{2}{*}{$*$} \\ a_{21} & a_{12} & \\ \hline 0 & 0 & \\ \vdots & \vdots & A_2 \\ 0 & 0 & \end{array}\right]$$

und dass $R_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{12} \end{bmatrix}$ ähnlich ist zu $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

Per Induktionsvoraussetzung gibt es $\tilde{U}_2$, so dass

$$U_2^\top A_2 U_2 = \begin{bmatrix} R_{22} & \dots & R_{2m} \\ & \ddots & \vdots \\ & & R_{mm} \end{bmatrix}$$

von der gewünschten Form ist.

Mit $U = U_1 \begin{bmatrix} I_2 & 0 \\ 0 & U_2 \end{bmatrix}$ folgt die Behauptung.

$\square$

Wir kommen nun zum zentralen Satz für orthogonale Matrizen.

Satz 8.11

(i) Sei $A \in \mathcal{U}(n) = \{A \in \mathbb{C}^{n,n} | A \text{ unitär }\}$. Dann gibt es eine unitäre Matrix P, so dass

$$P^{-1}AP = P^H AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

(ii) Sei $A \in \mathcal{O}(n) := \{A \in \mathbb{R}^{n,n} | A \text{ orthogonal }\}$. Dann gibt es eine orthogonale Matrix P, so dass

$$P^{-1}AP = P^\top AP = \begin{bmatrix} I_p & & & & \\ & -I_q & & & \\ & & D_1 & & \\ & & & \ddots & \\ & & & & D_s \end{bmatrix}$$

mit $D_i = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \in \mathbb{R}^{2,2}, \quad c^2 + s^2 = 1.$

Beweis:

(i) Nach dem Satz von Schur (Satz 2.15) gibt es eine unitäre Matrix $P \in \mathbb{C}^{n,n}$, so dass

$$P^{-1}AP = P^H AP = R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{bmatrix}.$$

Da P^{-1}, A, P unitär sind, folgt da $\mathcal{U}(n)$ eine Gruppe ist, dass auch R unitär ist. Damit

$$R^H R = I_n, \qquad \text{d.h.,} \quad \delta_{ik} = \sum_{j=1}^{n} \overline{r}_{ji} r_{jk}.$$

Es gilt dann, dass $|r_{11}|^2 = 1$ und damit $r_{11} \neq 0$. Weiterhin folgt $\overline{r}_{11} r_{1j} = 0$, für $j = 2, \ldots, n$, und damit $r_{1,j} = 0$, für $j = 2, \ldots, n$. Per Induktion folgt dann die Behauptung.

Um Teil (ii) zu beweisen, verwenden wir analog die reelle Schurform aus 8.10. $\square$

Kapitel 9

Selbstadjungierte Endomorphismen

In diesem Kapitel wollen wir uns mit Hemite'schen und symmetrischen Matrizen, sowie selbstadjungierten Endomorphismen beschäftigen. Im folgenden sei V ein euklidischer oder ein unitärer Raum.

Definition 9.1 *Seien $f, g : V \to V$ Endomorphismen. Dann heißt g* adjungiert *zu f, falls*

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

Lemma 9.2 *Sei V ein endlichdimensionaler euklidischer (unitärer) Raum und $f : V \to V$ ein Endomorphismus. Dann gibt es einen eindeutigen adjungierten Endomorphismus f^{ad} zu f, und es gilt $(f^{ad})^{ad} = f$.*

Beweis: Wir zeigen nur den reellen Fall, der Beweis im komplexen Fall ist eine Übungsaufgabe.

Sei $V^* = L(V, \mathbb{R})$ der Dualraum zu V, (siehe Kapitel 11) und sei:

$$\begin{aligned} \varphi : \ & V \to V^* \\ & v \mapsto \langle v, \cdot \rangle. \end{aligned}$$

Dann ist φ ist ein Vektorraumisomorphismus (Satz 11.7).

Sei $f^* : V^* \to V^*$ die zu f duale Abbildung

$$f^*(g) = gf, \quad \text{d.h.,} \quad (f^*(g))(v) = g(f(v)), \quad \text{für alle } v \in V, \ g \in V^*,$$

und definiere

$$f^{ad} : V \to V \qquad \text{durch} \quad f^{ad}(v) = (\varphi^{-1} f^* \varphi)(v).$$

f^{ad} ist natürlich ein Endomorphismus und es gilt

$$\begin{aligned} \langle f^{ad}(v), w \rangle &= \varphi\left(f^{ad}(v)\right)(w) \\ &= \left(f^*(\varphi(v))\right)(w) \\ &= \varphi(v)\left(f(w)\right) = \langle v, f(w) \rangle. \end{aligned}$$

Wir müssen noch die Eindeutigkeit zeigen. Sei $g : V \to V$ ein Endomorphismus mit

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

dann gilt

$$\varphi\left(g(v)\right) = \varphi\left(f^{ad}(v)\right).$$

Da φ injektiv ist, folgt

$$g(v) = f^{ad}(v))$$

und damit

$$\langle (f^{ad})^{ad}(v), w \rangle \;=\; \langle v, f^{ad}(w), v \rangle = \langle w, f(v) \rangle = \langle f(v), w \rangle.$$

Daraus erhalten wir $f(v) = (f^{ad})^{ad}(v)$ für alle $v \in V$ und damit $f = (f^{ad})^{ad}$. $\square$

Dieses Lemma gilt auch in unendlichdimensionalen Vektorräumen.

Beispiel 9.3

(a) *Sei $V = \mathbb{R}^3$ und*

$$f : \begin{array}{l} V \to V \\ v \mapsto Fv \end{array} \qquad \textit{mit} \qquad F = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 3 & 0 \end{bmatrix}$$

Dann ist

$$\langle g(v), w\rangle \;=\; \langle v, Fw\rangle \;=\; (Fw)^\top v \;=\; w^\top F^\top v \;=\; \langle F^\top v, w\rangle$$

und damit

$$g = f^{ad} \;:\; \begin{array}{l} V \to V \\ v \to F^\top v. \end{array}$$

(b) *Sei* $V = \{f \in \mathcal{C}^\infty([a,b],\mathbb{R}), f(a) = f(b) = 0\}$ *und betrachte das Skalarprodukt* $\langle f, g\rangle = \int_a^b f \cdot g\,dx$ *. Sei*

$$D: \begin{array}{l} V \to V \\ f \mapsto f' \end{array}$$

und sei z *so dass*

$$\langle z(f), g\rangle = \langle f, D(g)\rangle \text{ für alle } g \in V^*.$$

Dann folgt

$$\int_a^b z(f) \cdot g\,dx \;=\; \int_a^b f \cdot g' dx \;=\; \underbrace{f \cdot g|_a^b}_{=0} - \int_a^b f' g\,dx \;=\; -\int_a^b f' g\,dx$$

und damit

$$z(f) \;=\; -f' \;=\; D^{ad}(f).$$

Definition 9.4 *Ein Endomorphismus* $f : V \to V$ *heißt* selbstadjungiert*, falls*

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

für alle $v, w \in V$*.*

Lemma 9.5 *Sei* V *ein endlichdimensionaler euklidischer (unitärer) Raum mit Orthonormalbasis* $\{v_1, \ldots, v_n\}$*. Sei* $f : V \to V$ *ein Endomorphismus und* F *die Matrixdarstellung von* f *bzgl.* $\{v_1, \ldots, v_n\}$*.*

i) Falls V *euklidisch ist, so ist* f *selbstadjungiert genau dann, wenn* F *symmetrisch ist.*

ii) Falls V unitär ist, so ist f selbstadjungiert genau dann, wenn F hermite'sch ist.

Beweis:

i) Sei $v = \sum_{i=1}^{n} \lambda_i v_i, \; w = \sum_{i=1}^{n} \gamma_i v_i, \qquad \Lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \; \Gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}$. Dann gilt

$$\begin{aligned} \langle f(v), w \rangle &= \Gamma^\top F \Lambda \\ \langle v, f(w) \rangle &= (F\Gamma)^\top \Lambda \; = \; \Gamma^\top F^\top \Lambda \end{aligned}$$

und damit

$$\Gamma^\top F^\top \Lambda = \Gamma^\top F \Lambda, \qquad \text{für alle } \; \Lambda, \Gamma \in \mathbb{R}^n.$$

Also folgt $F^\top = F$.

ii) Übungsaufgabe.

□

Wir haben also den folgenden Zusammenhang

<u>Euklidischer Raum</u>	<u>Unitärer Raum</u>
• symmetrische Bilinearform	• Hermite'sche Form
• selbstadjungierter Endomorphismus	• selbstadjungierter Endomorphismus
• symmetrische Matrix	• Hermite'sche Matrix

Satz 9.6

(i) Falls $A \in \mathbb{R}^{n,n}$ symmetrisch, so gibt es $Q \in \mathcal{O}(n)$, so dass

$$Q^{-1} A Q = Q^\top A Q = D \in \mathbb{R}^{n,n}$$

diagonal ist. Insbesondere ist damit A diagonalisierbar und alle Eigenwerte sind reell.

(ii) *Falls $A \in \mathbb{C}^{n,n}$ Hermite'sch, so gibt es $Q \in \mathcal{U}(n)$, so dass*

$$Q^{-1}AQ = Q^H AQ = D \in \mathbb{R}^{n,n}$$

diagonal ist. Insbesondere ist damit A diagonalisierbar und alle Eigenwerte sind reell.

Beweis:

(i) Nach Satz 8.10 (reelle Schur-Form) gibt es $Q \in \mathcal{O}(n)$, so dass

$$Q^{-1}AQ = Q^\top AQ = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ & \ddots & & \\ & & \ddots & \\ & & & R_{m,m} \end{bmatrix} = R$$

Da $R^\top = (Q^\top AQ)^\top = Q^\top A^\top Q = Q^\top AQ = R$, so folgt $R_{i,j} = 0$ für $j > i$. Nach Satz 8.10 sind die Blöcke R_{ii} (1×1) oder (2×2) reell symmetrisch. Da aber im (2×2)-Fall $R_{ii} = \begin{bmatrix} a_i & b_i \\ b_i & c_i \end{bmatrix}$ das charakteristische Polynom

$$\lambda^2 - (a_i + c_i)\lambda + (a_i c_i) - b_i^2 = 0$$

die Nullstellen $\dfrac{a_i + c_i \pm \sqrt{(a_i - c_i)^2 + b_i^2}}{2}$ hat und diese beide reell sind, gibt es keine (2×2)-Blöcke. Damit folgt die Behauptung.

(ii) Nach dem Satz von Schur gibt es $Q \in \mathcal{U}(n)$, so dass

$$Q^{-1}AQ = Q^H AQ = \begin{bmatrix} r_{11} & & r_{1n} \\ & \ddots & \\ & & r_{nn} \end{bmatrix}$$

Da $(Q^H AQ)^H = Q^H A^H Q = Q^H AQ$, so folgt

$r_{ij} = 0$, für $j > i$ und $r_{ii} = \overline{r_{ii}}$ für alle i. Damit sind alle r_{ii} reell.

□

Beispiel 9.7 Betrachte

$$A = \begin{bmatrix} 2 & -1 & \frac{3}{2}\sqrt{2}\,i \\ -1 & 2 & -\frac{3}{2}\sqrt{2}\,i \\ -\frac{3}{2}\sqrt{2}\,i & \frac{3}{2}\sqrt{2}\,i & 3 \end{bmatrix}$$

mit dem charakteristischen Polynom

$$\begin{aligned} P_A(\lambda) &= \lambda^3 - 7\lambda^2 + \lambda(6+6+4-\frac{9}{2}-\frac{9}{2}-1) + (-12-\frac{9}{2}-\frac{9}{2}+9+3+9) \\ &= \lambda^3 - 7\lambda^2 + 6\lambda \;=\; (\lambda-1)\lambda(\lambda-6). \end{aligned}$$

Die Eigenwerte sind $\lambda_{1,2,3} = 1, 0, 6$. Da wir die Eigenwerte kennen, können wir den Ansatz im Beweis des Satzes von Schur verwenden. Bilde

$$(\lambda_1 I - A) = \begin{bmatrix} -1 & 1 & -\frac{3}{2}\sqrt{2}i \\ 1 & -1 & \frac{3}{2}\sqrt{2}i \\ \frac{3}{2}\sqrt{2}i & -\frac{3}{2}\sqrt{2}i & -2 \end{bmatrix}.$$

Mit

$$Q_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad Q_1^H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

ergibt sich

$$Q_1^H(\lambda_1 I - A)Q_1 \;=\; \begin{bmatrix} 0 & 0 & 0 \\ \frac{2}{\sqrt{2}} & -\frac{2}{\sqrt{2}} & 3\sqrt{2}i \\ \frac{3}{2}\sqrt{2}i & -\frac{3}{2}\sqrt{2}i & -2 \end{bmatrix} Q_1 \;=\; \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 3i \\ 0 & -3i & -2 \end{bmatrix}$$

also

$$Q_1^H A Q_1 = \lambda_1 I - \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 3i \\ 0 & -3i & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -3i \\ 0 & 3i & 3 \end{bmatrix}.$$

Dann fahren wir mit der kleineren Matrix analog fort.

Satz 9.8 *Sei V endlichdimensionaler euklidischer oder unitärer Raum. Seien f_1, f_2 selbstadjungierte Endomorphismen von $V \to V$. Dann sind die folgenden Aussagen äquivalent.*

(i) Es gibt eine Orthonormalbasis $\{v_1, \ldots, v_n\}$ von V, so dass $\{v_1, \ldots, v_n\}$ Eigenvektoren für f_1 und f_2 sind, d.h. f_1 und f_2 sind simultan diagonalisierbar.

(ii) $f_1 \circ f_2 = f_2 \circ f_1$

(In Matrixterminologie: Zwei reelle symmetrische (Hermite'sche) Matrizen F_1, F_2 sind genau dann simultan diagonalisierbar, wenn $F_1F_2 = F_2F_1$.)

Beweis:
(i) $\Longrightarrow$ (ii)
Sei v_i Eigenvektor von f_1 zum Eigenwert λ_i^1 und von f_2 zum Eigenwert λ_i^2. Dann gilt

$$(f_1 \circ f_2)(v_i) = f_1\left(f_2(v_i)\right) = \lambda_i^1\lambda_i^2 v_i = \lambda_i^2\lambda_i^1 v_i = (f_2 \circ f_1)(v_i) \quad \text{für alle } i = 1, \ldots, n$$

und damit folgt die Behauptung.

(ii) $\Longrightarrow$ (i)
Da f_1, f_2 selbstadjungierte Endomorphismen sind, so gibt es Matrixdarstellungen F_1, F_2 und eine orthogonale (unitäre) Matrix P_1, so dass

$$P_1^H F_1 P_1 = D_1, \qquad \text{mit} D_1 = \operatorname{diag}(\lambda_1 I_{j_1}, \ldots, \lambda_k I_{j_k}) \text{ und } \lambda_i \neq \lambda_j \text{ für } i \neq j,$$

Dann gilt

$$\begin{array}{rclcl} P_1^H F_1 F_2 P_1 & = & P_1^H F_1 P_1 P_1^H F_2 P_1 & = & D_1 P_1^H F_2 P_1 \\ & = & P_1^H F_2 F_1 P_1 & = & P_1^H F_2 P_1 D_1. \end{array}$$

Setze $P_1^H F_2 P_1 = \begin{bmatrix} F_{11} & \cdots & F_{1k} \\ \vdots & \ddots & \vdots \\ F_{k1} & \cdots & F_{kk} \end{bmatrix}$, analog zu D_1 partitioniert. Es folgt, dass

$$\lambda_i F_{ij} = F_{ij}\lambda_j \qquad \text{für alle } i, j.$$

Wegen $\lambda_i \neq \lambda_j$ für $i \neq j$ gilt $F_{ij} = 0$ für $i \neq j$, also

$$P_1^H F_2 P_1 = \begin{bmatrix} F_{11} & & \\ & \ddots & \\ & & F_{kk} \end{bmatrix}.$$

Setze $\tilde{P}_1 = \operatorname{diag} \begin{bmatrix} \Pi_1 & & \\ & \ddots & \\ & & \Pi_k \end{bmatrix}$, Π_i orthogonal (unitär), so dass $\Pi_i^H F_{ii} \Pi_i = \Delta_i$ diagonal ist. Dann folgt, dass

$$\begin{aligned} \tilde{P}_1^H P_1^H F_1 P_1 \tilde{P}_1 &= \begin{bmatrix} \lambda_1 I_{j_1} & & \\ & \ddots & \\ & & \lambda_k I_{j_k} \end{bmatrix} \qquad \text{und} \\ \tilde{P}_1^H P_1^H F_2 P_1 \tilde{P}_1 &= \begin{bmatrix} \Delta_1 & & \\ & \ddots & \\ & & \Delta_k \end{bmatrix} \end{aligned}$$

diagonal sind. □

Definition 9.9 *Sei V ein endlichdimensionaler Vektorraum und $f : V \to V$ ein Endomorphismus. Dann heißt f* normal, *falls*

$$f^{ad} \circ f = f \circ f^{ad}.$$

In Matrixterminologie in $\mathbb{C}^{n,n}$ (oder $\mathbb{R}^{n,n}$) heißt F normal, *falls*

$$F^H F = F F^H \qquad (F^\top F = F F^\top).$$

Beispiel 9.10 Sei

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{so ist} \quad A^\top A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$\text{und} \quad A A^\top = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

also ist A nicht normal. Für

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

gilt $A^\top A = A^2 = A A^\top$ und damit ist A normal.

Satz 9.11 *Sei V ein endlichdimensionaler Vektorraum, so gilt* *(i) Ist $f : V \to V$ ein orthogonaler (unitärer) Endomorphismus, so ist f normal.*

(ii) Ist $f : V \to V$ selbstadjungiert, so ist f normal.

Beweis:

(i) Für alle $v, w \in V$ gilt

$$\langle f(v), f(w)\rangle = \langle v, w\rangle \qquad \text{und} \qquad \langle f^{ad}(v), w\rangle = \langle v, f(w)\rangle\,.$$

Da V endlichdimensional ist, so ist f invertierbar und f^{-1} orthogonal (unitär). Also gilt für alle $v, w \in V$, dass

$$\langle v, f(w)\rangle = \langle f^{-1}(v), w\rangle = \langle f^{ad}(v), w\rangle$$

und damit $f^{-1} = f^{ad}$ und somit

$$f \circ f^{ad} = f^{ad} \circ f.$$

(ii) Falls $f = f^{ad}$, so folgt $f \circ f^{ad} = f \circ f = f^{ad} \circ f$.

□

Normale Matrizen (Endomorphismen) bilden also eine Obermenge der orthogonalen und symmetrischen (bzw. unitären und Hermite'schen) Matrizen (Endomorphismen). Die Eigenschaft der Diagonalisierbarkeit wird dabei auch vererbt.

Lemma 9.12 *Sei V ein endlichdimensionaler Vektorraum und $f : V \to V$ ein normaler Endomorphismus. Dann gilt*

(i) $\mathrm{Kern}\,(f) = \mathrm{Kern}\,(f^{ad})$.

(ii) Für alle Eigenwerte λ von f gilt $E(\lambda, f) = E(\bar{\lambda}, f^{ad})$.

(Zur Erinnerung: $E(\lambda, f)$ ist der Eigenraum von f zum Eigenwert λ).

Beweis:

(i) Für alle $v \in V$ gilt

$$\begin{aligned}\langle f(v), f(v)\rangle &= \langle (f^{ad} \circ f)(v), v\rangle \\ &= \langle (f \circ f^{ad})v, v\rangle = \langle f^{ad}(v), f^{ad}(v)\rangle.\end{aligned}$$

Aus der Definitheit von $\langle\cdot,\cdot\rangle$ folgt die Behauptung.

(ii) Sei $h = f - \lambda \cdot id_V$, so folgt

$$\begin{aligned}
\langle h^{ad}(v), w\rangle &= \langle v, h(w)\rangle \\
&= \langle v, f(w)\rangle - \langle v, \lambda w\rangle \\
&= \langle f^{ad}(v), w\rangle - \bar{\lambda}\langle v, w\rangle \\
&= \langle (f^{ad} - \bar{\lambda} \cdot id_V(v), w\rangle\,.
\end{aligned}$$

und damit $h^{ad} = f^{ad} - \bar{\lambda} \cdot id_V$.

(Beachte, dass bei komplexen Eigenwerten auch das komplexe Skalarprodukt verwendet werden muss.) Durch Nachrechnen ergibt sich sofort, dass auch h normal ist, und damit folgt

$$\begin{aligned}
E(\lambda, f) &= \text{Kern}\,(f - \lambda \cdot id_V) \overset{(i)}{=} \text{Kern}\,(f^{ad} - \bar{\lambda} \cdot id_V) \\
&= E(\bar{\lambda}, f^{ad}).
\end{aligned}$$

□

Damit können wir nun zeigen, dass die durch unitäre Ähnlichkeit diagonalisierbaren Endomorphismen genau die normalen sind.

Satz 9.13 *Sei V ein endlichdimensionaler Vektorraum und $f : V \to V$ ein Endomorphismus. Dann sind die folgenden Aussagen äquivalent.*

(i) Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f.

(ii) f ist normal.

(In Matrixterminologie: A ist unitär diagonalisierbar genau dann, wenn A normal ist.)

Beweis: (i) $\Longrightarrow$ (ii) Sei $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis von V aus Eigenvektoren von f. Es gilt dann

$$\langle f^{ad}(v_i), v_j\rangle = \langle v_i, f(v_j)\rangle = \langle v_i, \lambda v_j\rangle = \bar{\lambda}\langle v_i, v_j\rangle = \begin{cases} \bar{\lambda}, & i = j \\ 0, & i \neq j \end{cases}$$

und damit

$$f^{ad}(v_i) = \sum_{j=1}^{n}\langle f^{ad}(v_i), v_j\rangle v_j = \bar{\lambda} v_i \text{ für alle } i, j = 1, \ldots, n.$$

Daraus folgt, dass

$$(f \circ f^{ad})(v_i) = \lambda\bar{\lambda}v_i = (f^{ad} \circ f)(v_i) \qquad \text{für alle } i = 1, \ldots, n$$

also $f \circ f^{ad} = f^{ad} \circ f$.

(ii) $\Longrightarrow$ (i) Mit Induktion:

I.A.: $n = 1$ ist klar.

I.V.: Behauptung sei richtig für $n - 1$.

I.S.: Sei v_1 mit $\|v_1\|_2 = 1$ ein Eigenvektor von f zum Eigenwert λ und $W = \mathcal{L}(v_1)^\perp$. Sei $w \in W$, so gilt

$$\langle v_1, f(w)\rangle = \langle f^{ad}(v_1), w\rangle \overset{\text{Lemma 9.12}}{=} \langle \bar{\lambda}v_1, w\rangle = \bar{\lambda}\langle v_1, w\rangle = 0$$

und damit $f(W) \subset W$. Weiter gilt, dass $f|_W$ natürlich auch normal ist. Da $\dim(W) = \dim(V) - 1$, so gibt es nach Induktionsvoraussetzung eine Orthonormalbasis $\{v_2, \ldots, v_n\}$ aus Eigenvektoren von $f|_W$. Dann ist $\{v_1, \ldots, v_n\}$ Orthonormalbasis von V aus Eigenvektoren von f.

In Matrizenterminologie ist der Beweis eine Folgerung des Satzes von Schur und Satz 9.8.

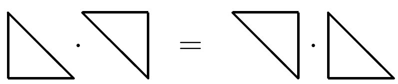
(i) $\Longrightarrow$ (ii) Falls A unitär diagonalisierbar, so gilt $P^HAP = D$,

$$\begin{aligned} P^HA^HAP &= P^HA^HPP^HAP \\ &= D^HD = DD^H = P^HAA^HP \end{aligned}$$

also folgt $A^HA = AA^H$.

(ii) $\Longrightarrow$ (i) Sei $A^HA = AA^H$.
Nach dem Satz von Schur gibt es eine unitäre Matrix P, so dass $P^HAP = \Delta = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{bmatrix}$ obere Dreiecksmatrix ist. Dann folgt

$$(P^HAP)^H = P^HA^HP = \Delta^H$$



$$\Delta^H \cdot \Delta = \Delta \cdot \Delta^H.$$

Wenn wir das komponentenweise betrachten erhalten wir

$$|r_{11}|^2 = |r_{11}|^2 + \sum_{j=2}^{n} |r_{1j}|^2$$

und damit sofort

$$\sum_{j=2}^{n} |r_{1j}|^2 = 0.$$

Also gilt

$$\Delta = \left[\begin{array}{c|ccc} r_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \ddots & \\ 0 & & & \end{array}\right]$$

Der Rest des Beweises folgt mit Induktion.

□

Für normale Matrizen liefert also der Satz von Schur die (unitäre) Diagonalisierung. Zusammenfassung:

A Hermite'sch, $A = A^H$ ONB aus Eigenvektoren Eigenwerte reell	A reell symmetrisch, $A = A^\top$ ONB aus reellen Eigenvektoren Eigenwerte reell
A unitär, $A^H A = I$ ONB aus Eigenvektoren Eigenwerte $\|\lambda\| = 1$	A orthogonal, $A^\top A = I$ ONB aus i.a. komplexen Eigenvektoren Eigenwerte $\|\lambda\| = 1$
A schiefhermite'sch, $A = -A^H$ ONB aus Eigenvektoren Eigenwerte rein imaginär	A reell schiefsymmetrisch, $A = -A^\top$ ONB aus komplexen Eigenvektoren Eigenwerte rein imaginär

Kapitel 10

Die Singulärwertzerlegung

Wir haben bisher verschiedene Transformationen kennengelernt,die $A \in \mathbb{K}^{n,m}$, auf eine Normalform transformieren.

	Transformation		Normalform
	$A \to PAQ$	P, Q nichtsingulär	Treppen NF $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$
$n = m$	$A \to Q^{-1}AQ$	Q nichtsingulär	Jordan'sche NF.
$n = m$	$A \to Q^H AQ$	Q unitär	Schur–Form

Es stellt sich die Frage, was sich ergibt, wenn wir $A \to PAQ$ betrachten, P, Q unitär (orthogonal).

Um die Betrachtungen elegant zu machen, führen wir zuerst mal ein paar neue Normen ein:

Sei $A = [a_{i,j}] \in \mathbb{C}^{n,m}(\mathbb{R}^{n,m})$.

$$\|A\|_F := \left[\sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|^2\right]^{\frac{1}{2}} = (\operatorname{tr}(A^H A))^{\frac{1}{2}} \tag{10.1}$$

heißt *Frobeniusnorm* von A. Dies ist die euklidische Norm in $\mathbb{C}^{n \cdot m}$ $(\mathbb{R}^{n \cdot m}))$ und

$$\|A\|_2 := \sup_{\substack{x \neq 0 \\ x \in \mathbb{C}^m}} \frac{\|Ax\|_2}{\|x\|_2} \tag{10.2}$$

heißt *Spektralnorm* von A.

Lemma 10.3 *Für $A \in \mathbb{C}^{n,m}, P \in \mathbb{C}^{n,n}, Q \in \mathbb{C}^{m,m}$, mit P, Q unitär gilt:*

$$\begin{aligned} \|A\|_F &= \|PAQ\|_F, \\ \|A\|_2 &= \|PAQ\|_2. \end{aligned}$$

(Analog für $A \in \mathbb{R}^{n,m}, P \in \mathbb{R}^{n,n}, Q \in \mathbb{R}^{m,m}$, P, Q orthogonal).

Beweis:

$$\|\cdot\|_F : \quad \left(\mathrm{tr}\left((PAQ)^H(PAQ)\right)\right)^{\frac{1}{2}} = \left(\mathrm{tr}\,(Q^H A^H AQ)\right)^{\frac{1}{2}} = \left(\mathrm{tr}\,(A^H A)\right)^{\frac{1}{2}} .$$

$$\begin{aligned} \|\cdot\|_2 : \quad \sup_{\substack{x \neq 0 \\ x \in \mathbb{C}^n}} \frac{\|Ax\|_2}{\|x\|_2} &= \sup_{x \neq 0} \frac{\sqrt{\langle Ax, Ax\rangle}}{\sqrt{\langle x, x\rangle}}, \\ \sup_{x \neq 0} \frac{\|PAQx\|_2}{\|x\|_2} &= \sup_{x \neq 0} \frac{\langle PAQx, PAQx\rangle^{\frac{1}{2}}}{\langle x, x\rangle^{\frac{1}{2}}} = \sup_{x \neq 0} \frac{\langle AQx, AQx\rangle^{\frac{1}{2}}}{\langle Qx, Qx\rangle^{\frac{1}{2}}} \\ &= \sup_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2}, \quad \text{mit } \; y = Qx. \end{aligned}$$

$\square$

Satz 10.4 *Sei $A \in \mathbb{C}^{n,m}$ $(\mathbb{R}^{n,m})$, $n \geq m$, so gibt es P, Q unitär (orthogonal), $P \in \mathbb{C}^{n,n}$, $Q \in \mathbb{C}^{m,m}$ $(P \in \mathbb{R}^{n,n}$, $Q \in \mathbb{R}^{m,m})$, so dass*

$$P^H AQ = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & & \\ & 0 & \\ & & \end{bmatrix} \qquad \textit{mit } \; \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0 \,.$$

Beweis: Falls $A = 0$ so ist der Beweis trivial. Sei nun $\sigma = \|A\|_2 \neq 0$ und seien $x \in \mathbb{C}^m, y \in \mathbb{C}^n (x \in \mathbb{R}^m, y \in \mathbb{R}^n)$, mit $\|x\|_2 = \|y\|_2 = 1$ und $Ax = \sigma y$. Wähle dazu als x den Vektor mit $\|x\|_2 = 1$, so dass

$$\sigma = \sup_{\substack{x \neq 0 \\ x \in \mathbb{C}^m}} \frac{\|Ax\|_2}{\|x\|_2}$$

und setze $y = Ax\sigma^{-1}$. Dann gilt ja $\|Ax\|_2 = \sigma\|x\|_2$ und damit $\|y\|_2 = 1$.

Seien $P = [y,\ P_1] \in \mathbb{C}^{n,n}, Q = [x,\ Q_1] \in \mathbb{C}^{m,m}$ unitär (diese gibt es mit dem Basisergänzungssatz und der Gram-Schmidt Orthonormalisierung), so folgt

$$A_1 = P^H A Q \quad = \quad \begin{array}{cc} \left[\begin{array}{cc} \sigma & w^H \\ 0 & A_2 \end{array}\right] & \begin{array}{c} {\scriptstyle 1} \\ {\scriptstyle n-1} \end{array} \\ \begin{array}{cc} {\scriptstyle 1} & {\scriptstyle m-1} \end{array} & \end{array} .$$

Da

$$\left\| A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \sigma^2 + w^H w \\ A_2 w \end{bmatrix} \right\|_2^2 \geq (\sigma^2 + w^H w)^2 ,$$

so folgt

$$\|A_1\|_2 = \sup_{\substack{z \neq 0 \\ z \in \mathbb{C}^m}} \frac{\|A_1 z\|_2}{\|z\|_2}$$

und damit

$$\|A_1\|_2^2 \geq \sigma^2 + w^H w.$$

Aber da $\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2$, so folgt $w^H w = 0$ und damit $w = 0$.

Der Rest folgt mit Induktion. □

Definition 10.5 *Sei $A \in \mathbb{C}^{n,m}(\mathbb{R}^{n,m}),\ n \geq m$. Die Zerlegung in*

$$A = P \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & 0 & \end{bmatrix} Q^H,$$

mit P, Q unitär (orthogonal), heißt Singulärwertzerlegung von A, *die σ_i heißen* Singulärwerte *und die Spalten von* $P\,(Q)$ Linkssingulärvektoren *(Rechtssingulärvektoren). (Matlab: SVD) (Für $m \geq n$ gilt alles mit der Singulärwertzerlegung von A^H).*

Satz 10.6 *Sei* $A \in \mathbb{C}^{n,m}(\mathbb{R}^{n,m})$, $n \geq m$, *und sei mit* $P = [p_1, \ldots, p_n]$, $Q = [q_1, \ldots, q_m]$

$$A = P \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & 0 & \end{bmatrix} Q^H$$

mit $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_m = 0$ *die Singulärwertzerlegung von* A. *Dann gilt*

(i) $r = \text{Rang}\,(A)$,

(ii) $\text{Kern}\,(A) = \text{span}\,\{q_{r+1}, \ldots, q_m\}$,

(iii) $\text{Bild}\,(A) = \text{span}\,\{p_1, \ldots, p_r\}$,

(iv) $\|A\|_2 = \sigma_1$,

(v) $\|A\|_F{}^2 = \sigma_1^2 + \ldots + \sigma_m^2$.

Beweis:

(i) Die Behauptung folgt aus den Sätzen 11.5 und 4.2.

(ii) Die Spalten von Q bilden eine Orthonormalbasis von $\mathbb{C}^m$,

$$A[q_{r+1}, \ldots, q_m] = 0 \quad \Longrightarrow \quad \{q_{r+1}, \ldots, q_m\} = \text{Kern}\,(A)$$

(iii) Für das Bild gilt die Betrachtung analog.

(iv) Nach Konstruktion

$$\left\|P^H A Q\right\|_2 = \left\| \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & 0 & \end{bmatrix} \right\|_2 = \sup_{\substack{x \in \mathbb{C}^n \\ x \neq 0}} \frac{\|[\sigma_i x_i]\|_2}{\|x\|_2} = \sigma_1 \ \text{ mit } x = e_1.$$

(v) $\|A\|_F{}^2 = \left\|P^H A Q\right\|_F{}^2 = \sigma_1^2 + \ldots + \sigma_m^2$.

$\square$

Bemerkung: Die Singulärwerte sind die Wurzeln der Eigenwerte von $A^H A$ bzw. AA^H.

Mit der Singulärwertzerlegung erhalten wir das Analogon zur Normalform unter Äquivalenz. Allerdings bietet die Singulärwertzerlegung neben der Tatsache, dass man sie numerisch sehr gut ausrechnen kann, auch noch weitere wichtige Eigenschaften.
Wir kommen nochmal auf das allgemeine lineare Gleichungssystem $Ax = b$ mit

$$A \in \mathbb{C}^{n,m}\ , b \in \mathbb{C}^n, x \in \mathbb{C}^m \quad (A \in \mathbb{R}^{n,m}, b \in \mathbb{R}^n, x \in \mathbb{R}^m)\,.$$

zurück. Wir setzen jetzt aber nicht voraus, dass $Rang(A) = m$ und $b \in Bild(A)$. Damit ist $Ax = b$ aber nicht mehr unbedingt lösbar. Wir lösen statt dessen

$$\|Ax - b\| = \ \min\ , \tag{10.7}$$

und fordern außerdem, dass

$$\|x\| = \ \min\ , \tag{10.8}$$

damit die Lösung eindeutig wird.

Satz 10.9 *Betrachte die Minimierungsaufgabe $\|Ax - b\|_2$ =min unter der Zusatzforderung $\|x\|_2$ =min. Sei Rang $A = r$. Diese Aufgabe hat eine eindeutige Lösung, die mittels der Singulärwertzerlegung von $A = P\Sigma Q^H = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^H$, mit Σ_r nichtsingulär, wie folgt angegeben werden kann:*

$$\hat{x} = Q \begin{bmatrix} \Sigma_r^{-1}[I_r\ 0]P^H b \\ 0 \end{bmatrix}. \tag{10.10}$$

Das Residuum ist gegeben durch

$$\|A\hat{x} - b\|_2 = \left\|[0\ I_{n-r}]P^H b\right\|_2.$$

Beweis:
$\|Ax - b\|_2 = \min$ und $\|x\|_2 = \min \quad \Leftrightarrow \quad \left\|P^H AQQ^H x - P^H b\right\|_2 = \min$ und $\left\|Q^H x\right\|_2 = \min$. Das Gleichungssystem $P^H AQQ^H x = P^H b$ hat die Form

$$\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} [\,I_r,\, 0\,]P^H b \\ [0,\, I_{n-r}]P^H b \end{bmatrix}, \tag{10.11}$$

mit

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} [I_r, \ \ 0]Q^H x \\ [0, \ I_{m-r}]Q^H x \end{bmatrix}$$

Dessen Minimallösung lautet

$$\begin{aligned} y_1 &= \Sigma_r^{-1} b_1 = \Sigma_r^{-1}[I_r \ 0]P^H b, \\ y_2 &= 0. \end{aligned}$$

Also gilt

$$\hat{x} = Q \begin{bmatrix} \Sigma_r^{-1}[I_r \ 0]P^H b \\ 0 \end{bmatrix}.$$

Damit haben wir eine Lösung. Und da die Minimallösung von (10.11) eindeutig ist, ist auch die Lösung (10.10) eindeutig. Für das Residuum gilt:

$$\|A\hat{x} - b\|_2 = \left\|P^H AQQ^H \hat{x} - P^H b\right\|_2 = \|b\|_2 = \left\|[0 \, I_{n-r}]P^H b\right\|_2$$

□

Beispiel 10.12 Sei

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

Die Singulärwertzerlegung von

$$A = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^H.$$

ergibt sich mit

$$P^H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P^H A = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad P^H b = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 3 \end{bmatrix},$$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \Sigma = P^H AQ = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad y = Q^H x, \qquad .$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 3 \end{bmatrix}$$

Also folgt $y_1 = \dfrac{3}{2\sqrt{2}}$ und $y_2 = 0$ und damit

$$x = Qy = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{3}{2\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Das Residuum ist

$$\|Ax - b\|_2 = \left\| \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -3 \end{bmatrix} \right\|_2 = \sqrt{\frac{1}{4} + \frac{1}{4} + 9} = \sqrt{\frac{19}{2}}.$$

Definition 10.13 *Sei $A \in \mathbb{C}^{n,m}$ $(\mathbb{R}^{n,m})$ und sei $A = P \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} Q^H = P\Sigma Q^H$ mit $\Sigma_r = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$ die Singulärwertzerlegung von A. Die Matrix*

$$A^+ = Q\Sigma^+ P^H \qquad \textit{mit} \qquad \Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m,n}$$

heißt Moore–Penrose–Inverse von A.

Satz 10.14 *Sei $A \in \mathbb{C}^{n,m}$ $(\mathbb{R}^{n,m})$.*

(i) Falls Rang $(A) = m$*, so gilt* $A^+ = (A^H A)^{-1} A^H$.

(ii) Falls $n = m$ und A nichtsingulär, so gilt $A^+ = A^{-1}$.

(iii) A^+ ist die eindeutige Matrix X, welche die vier Moore–Penrose-Gleichungen erfüllt:

(a) $AXA = A$,

(b) $XAX = X$,

(c) $(AX)^H = AX$,

(d) $(XA)^H = XA$.

Beweis: Sei $A = P\Sigma Q^H$ mit $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ wobei Σ_r invertierbar ist.

(i) Da Rang $(A) = m$, d.h., $PAQ^H = \begin{bmatrix} \Sigma_m \\ 0 \end{bmatrix}$, so folgt $A^+ = Q\Sigma^+P^H = Q[\Sigma_m^{-1}\, 0]P^H$. Also gilt

$$\begin{aligned}(A^HA)^{-1}A^H &= \left(Q[\Sigma_m\ 0]P^HP\begin{bmatrix} \Sigma_m \\ 0 \end{bmatrix}Q^H\right)^{-1} \cdot Q[\Sigma_m\ 0]P^H \\ &= Q\Sigma_m^{-2}Q^HQ[\Sigma_m\ 0]P^H = Q[\Sigma_m^{-1}\, 0]P^H.\end{aligned}$$

(ii) Klar.

(iii) Sei $Y = Q^HXP$. Damit erhalten wir:

$$P^HAQ \quad Q^HXP \quad P^HAQ \quad = \quad P^HAQ,$$

(a) Aus $\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ folgt $\Sigma_rY_{11}\Sigma_r = \Sigma_r$ und damit $Y_{11} = \Sigma_r^{-1}$.

(b) Aus $\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$, folgt

$$\begin{aligned}Y_{11}\Sigma_rY_{11} &= Y_{11}\,, \\ Y_{21}\Sigma_rY_{11} &= Y_{21}\,, \\ Y_{11}\Sigma_rY_{12} &= Y_{12}\,, \\ Y_{21}\Sigma_rY_{12} &= Y_{22}\,.\end{aligned}$$

(c) Wenn $\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$ Hermite'sch ist, so gilt $\Sigma_rY_{12} = 0,\ \Sigma_rY_{11} = Y_{11}^H\Sigma_r^H$.

(d) Wenn $\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ Hermite'sch ist, so folgt $Y_{21}\Sigma_r = 0,\ (Y_{11}\Sigma_r)^H = Y_{11}\Sigma_r$.

Alle 4 Gleichungen zusammen ergeben Aus (c) folgt $Y_{12} = 0$ und aus (d) folgt $Y_{21} = 0$. Dann folgt aus (b), dass $Y_{22} = 0$, und wir erhalten aus (a) dann $Y_{11} = \Sigma_r^{-1}$. Da die Lösung eindeutig ist folgt die Behauptung. □

Kapitel 11

Die Hauptachsentransformation

Nachdem wir mit der Singulärwertzerlegung bereits eines der wichtigsten praktischen Werkzeuge der linearen Algebra kennengelernt haben, welches uns wichtige Eigenschaften wie Rang, Kern, Bild einer Matrix berechnen läßt, kommen wir nun zu weiteren wichtigen Transformationen. Bisher haben wir Äquivalenztransformationen PAQ mit P, Q nichtsingulär, orthogonal oder unitär betrachtet und $P^{-1}AP$ mit P nichtsingulär, orthogonal oder unitär. Im Fall P orthogonal ist das $P^\top AP$.
Was passiert nun aber, wenn P nicht orthogonal ist und wir $P^\top AP$ betrachten? (Im Komplexen betrachten wir natürlich analog $P^H AP$.) Wir haben schon gesehen, dass dies die Form ist, die ein Basiswechsel bei einer Bilinearform hat, siehe Lemma 8.5.

Wir haben außerdem gesehen, dass wir zu jeder Bilinearform eine Matrixdarstellung finden. Sei $\mathbb{K}$ ein beliebiger Körper, V ein Vektorraum und $\{v_1, \ldots, v_n\}$ eine Basis von V. Ist $\alpha : V \times V \to \mathbb{R}$ eine Bilinearform, so ist $A = [a_{ij}]$ mit $a_{ij} = \alpha(v_j, v_i)$, $i, j = 1, \ldots, n$, die Matrixdarstellung von α bezüglich der Basis $\{v_1, \ldots, v_n\}$,

Umgekehrt sei $A = [a_{ij}] \in \mathbb{R}^{n,n}$, so definieren wir auf V eine Bilinearform $\alpha : V \times V \to \mathbb{R}$ wie folgt: Seien $v, v' \in V$ und sei

$$v = \sum_{i=1}^{n} b_i v_i, \; v' = \sum_{i=1}^{n} c_i v_i \quad \text{und} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \; c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

so setze $\alpha(v, v') := c^\top Ab = \sum_{i,j=1}^{n} c_i a_{ij} b_j$.

Beachte: diese Bijektion zwischen Matrizen und Bilinearformen hängt natürlich von der Basis ab. Wir hatten schon in Lemma 8.5 gesehen, dass für zwei Basen

$\{v_1,\ldots,v_n\},\{v_1',\ldots,v_n'\}$ von V mit Basisübergangsmatrix P, die Matrixdarstellung von einer Bilinearform $\alpha : V\times V\to\mathbb{K}$ bezüglich dieser beiden Basen durch Kongruenztransformation gegeben ist.

Falls $v,w\in V$, $v=\sum\limits_{i=1}^{n} b_iv_i=\sum\limits_{i=1}^{n} b_i'v_i'$ und $w=\sum\limits_{i=1}^{n} c_iv_i=\sum\limits_{i=1}^{n} c_i'v_i'$, so ist

$$\begin{aligned} b := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} &= P\begin{bmatrix} b_1' \\ \vdots \\ b_n' \end{bmatrix} =: Pb' \\ c := \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} &= P\begin{bmatrix} c_1' \\ \vdots \\ c_n' \end{bmatrix} =: Pc' \end{aligned}$$

und es ergibt sich

$$\alpha(v,w)=c^\top Ab=(Pc')^\top APb'=c'^\top P^\top APb'\,.$$

(Analoge Betrachten gelten im komplexen.)

Wir hatten auch bereits gesehen, dass die Matrixdarstellung von α symmetrisch ist genau dann, wenn α symmetrisch ist, und dass es in diesem Fall eine Orthonormalbasis aus Eigenvektoren gibt.

Beispiel 11.1 *Sei* $V=\mathbb{R}^2,\{v_1,v_2\}=\{e_1,e_2\}$, und

$$A=\begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix},$$

Dann ist

$$\begin{aligned} \alpha\left(\begin{bmatrix} b_1 \\ b_2 \end{bmatrix},\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}\right) &= \alpha(b_1e_1+b_2e_2,c_1e_1+c_2e_2) \;=\; c^\top Ab \\ &= 4b_1c_1+b_2c_1+b_1c_2+2b_2c_2. \end{aligned}$$

A hat das charakteristische Polynom $\quad P_A(\lambda)=\lambda^2-6\lambda+7$ und die Eigenwerte:

$$\lambda_1,\lambda_2=3\pm\sqrt{9-7}=3\pm\sqrt{2}.$$

Die Eigenvektoren ergeben sich aus

$$A - \lambda_1 I = \begin{bmatrix} 4-3+\sqrt{2} & 1 \\ 1 & 2-3+\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1+\sqrt{2} & 1 \\ 1 & -1+\sqrt{2} \end{bmatrix}.$$

$$Q^\top = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad c = \frac{1+\sqrt{2}}{\sqrt{(1+\sqrt{2})^2+1^2}}, \quad s = \frac{1}{\sqrt{(1+\sqrt{2})^2+1^2}},$$

$$\begin{aligned}
Q^\top(A-\lambda_1 I)Q &= \begin{bmatrix} c(1+\sqrt{2})+s & c+s(-1+\sqrt{2}) \\ -s(1+\sqrt{2})+c & -s+c(-1+\sqrt{2}) \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \\
&= \begin{bmatrix} \frac{(1+\sqrt{2})^2+1}{\sqrt{(1+\sqrt{2})^2+1}} & \frac{1+\sqrt{2}-1+\sqrt{2}}{\sqrt{(1+\sqrt{2})^2+1}} \\ 0 & \frac{-1+2-1}{\sqrt{(1+\sqrt{2})^2+1}} \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \\
&= \frac{1}{\sqrt{(1+\sqrt{2})^2+1}} \begin{bmatrix} (1+\sqrt{2})^2+1 & 2\sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \\
&= \frac{1}{\left(\sqrt{(1+\sqrt{2})^2+1}\right)^2} \begin{bmatrix} ((1+\sqrt{2})^2+1)(1+\sqrt{2})+\sqrt{2} & (1+\sqrt{2})2\sqrt{2}-1-(1+\sqrt{2})^2 \\ 0 & 0 \end{bmatrix} \\
&= \frac{1}{(1+\sqrt{2})^2+1} \begin{bmatrix} (4+2\sqrt{2})(1+\sqrt{2})+2\sqrt{2} & 2\sqrt{2}+4-1-1-2\sqrt{2}-2 \\ 0 & 0 \end{bmatrix} \\
&= \frac{1}{(1+\sqrt{2})^2+1} \begin{bmatrix} 8+8\sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

Damit folgt

$$Q^\top A Q = \begin{bmatrix} 3+\sqrt{2} & 0 \\ 0 & 3-\sqrt{2} \end{bmatrix}.$$

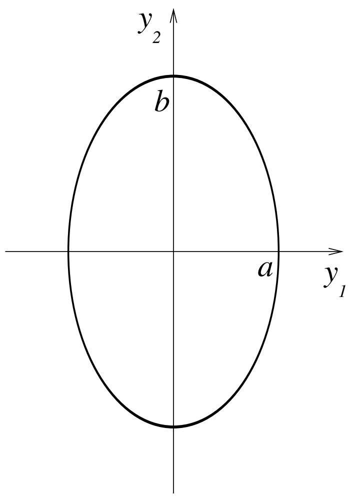
Die Menge

$$E = \left\{ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \mid x^\top A x - 1 = 0 \right\} = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \mid 4x_1^2 + 2x_1 x_2 + 2x_2^2 = 1 \right\}$$

beschreibt eine Ellipse.

Wenn wir zu dem Koordinatensystem übergehen, welches durch die Orthonormalbasis $\left\{ \begin{bmatrix} c \\ s \end{bmatrix}, \begin{bmatrix} -s \\ c \end{bmatrix} \right\}$ gegeben ist, so ist

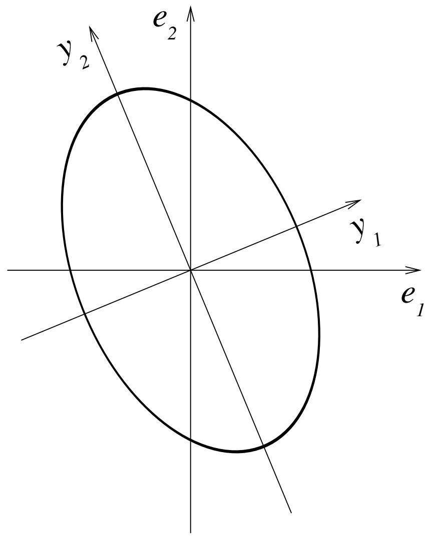
$$\begin{aligned} E &= \left\{ y = Q^\top x \in \mathbb{R}^2 \;\middle|\; y^\top \begin{bmatrix} 3+\sqrt{2} & 0 \\ 0 & 3-\sqrt{2} \end{bmatrix} y = 1 \right\} \\ &= \left\{ y \in \mathbb{R}^2 \mid y_1^2(3+\sqrt{2}) + y_2^2(3-\sqrt{2}) = 1 \right\}, \quad \text{d.h., die Ellipse} \end{aligned}$$



$$E = \left\{ y \in \mathbb{R}^2 \;\middle|\; \frac{y_1^2}{a^2} + \frac{y_2^2}{b^2} = 1 \right\}, \quad \text{mit}$$

$$a = \sqrt{\frac{1}{3+\sqrt{2}}},\; b = \sqrt{\frac{1}{3-\sqrt{2}}}.$$

Im kanonischen Koordinatensystem mit der Basis $\{e_1, e_2\}$ ist das also die Ellipse



Definition 11.2 *Sei V ein euklidischer Raum und $\alpha : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann heißt α* positiv definit*, falls $\alpha(v,v) > 0$, für alle $v \in V$, $v \neq 0$, und* positiv semidefinit*, falls $\alpha(v,v) \geq 0$ für alle $v \in V$.*

Satz 11.3 *Sei V ein endlichdimensionaler euklidischer Raum und $\alpha : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform.*

(i) α ist positiv definit genau dann, wenn es eine Basis $\{v_1, \ldots, v_n\}$ von V gibt, so dass die Matrixdarstellung von α bezüglich $\{v_1, \ldots, v_n\}$ nur positive reelle Eigenwerte hat.

(ii) α ist positiv semidefinit genau dann, wenn es eine Basis $\{v_1, \ldots, v_n\}$ von V gibt, so dass die Matrixdarstellung von α bezüglich $\{v_1, \ldots, v_n\}$ nur nichtnegative reelle Eigenwerte hat.

(iii) Es gibt eine Orthonormalbasis $\{v_1, \ldots, v_n\}$ von V, so dass die Matrixdarstellung von α bezüglich dieser Basis diagonal ist.

Beweis: Wir beweisen (i) und (ii) zusammen. Sei α positiv (semi)definit und $A = [a_{ij}]$ die Matrixdarstellung von α bezüglich $\{v_1, \ldots, v_n\}$, d.h., $a_{ij} = \alpha(v_j, v_i)$. Da A reell symmetrisch ist folgt, dass es eine Orthonormalbasis aus Eigenvektoren von A gibt, gegeben durch die Spalten einer orthogonalen Matrix $P = [p_1, \ldots, p_n]$, und bezüglich der Basis $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$ die sich aus $\{v_1, \ldots, v_n\}$ mit Hilfe der Übergangsmatrix P ergibt, hat α die Matrixdarstellung

$$P^\top A P = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Ist w der Koordinatenvektor von $z \in V$ in der Basis $\{\tilde{v}_1, \ldots, \tilde{v}_n\}$, dann gilt

$$\alpha(z, z) = w^\top P^\top A P w = \sum_{i=1}^{n} w_i^2 \lambda_i.$$

Falls es $\lambda_i \leq 0$ ($\lambda_i < 0$) gibt, so betrachte $w = e_i$ und es gilt $\alpha(z, z) = \lambda_i \leq 0$ (< 0) und das ist ein Widerspruch zur positiven Definitheit (Semidefinitheit).

Für die andere Richtung: Ist x der Koordinatenvektor von $z \in V$ in der Basis $\{v_1, \ldots, v_n\}$, dann ergibt sich aus

$$\alpha(z, z) = x^\top A x = x^\top P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^\top x = \sum_{i=1}^{n} x_i^2 \lambda_i > 0 \ (\geq 0)$$

sofort, dass α positiv definit (semidefinit) ist.

(iii) ist klar. Man nimmt einfach eine Matrixdarstellung bezüglich irgendeiner Basis und erhält A. Mit der Orthonormalbasis aus Eigenvektoren ergibt sich die Behauptung. $\square$

Beachte: Wir haben sogar bewiesen, dass die Matrixdarstellung einer positiv (semi-)definiten, symmetrischen Bilinearform bezüglich jeder Basis von V nur positive (nichtnegative) reelle Eigenwerte hat.

Wir kommen nun zu der Frage, welches die Invarianten unter der Transformation $P^\top AP$, angewandt auf symmetrische Matrizen A, sind. Es ist klar, dass für orthogonales (unitäres)P dies die Eigenwerte sind, aber für nichtorthogonale (nichtunitäre) P sind dies nur die Anzahlen der positiven, negativen oder Null-Eigenwerte.

Satz 11.4 (Trägheitssatz von Sylvester)
Sei V ein endlichdimensionaler euklidischer Raum und $\alpha : V \times V \longrightarrow \mathbb{R}$ eine symmetrische Bilinearform. Seien $\{v_1, \ldots, v_n\}, \{v'_1, \ldots, v'_n\}$ Basen von V und A, A' Matrixdarstellungen von α bezüglich dieser beiden Basen.
Seien (π, ν, ω) und (π', ν', ω') die Anzahl der positiven, negativen oder Null-Eigenwerte von A bzw. A'. Dann gilt $\pi = \pi', \nu = \nu', \omega = \omega'$ und $\mathrm{Rang}\,(A) = \mathrm{Rang}\,(A')$.

Beweis: Es gibt Q, Q' orthogonal, so dass

$$Q^\top AQ = D,\ Q'^\top A'Q' = D'$$

und die Diagonalelemente von D, D' seien geordnet als

$$D = \begin{bmatrix} D_1 & & \\ & D_2 & \\ & & 0 \end{bmatrix},\ D' = \begin{bmatrix} D'_1 & & \\ & D'_2 & \\ & & 0 \end{bmatrix},$$

wobei D_1 und D'_1 die positiven Eigenwerte von A und A' enthalten und D_2 und D'_2 die negativen.

Da dies mit einer Permutation immer möglich ist, können wir gleich Q, Q' so wählen, dass dies gilt. Wir wissen, dass $A' = P^\top AP$, also gilt

$$D' = \tilde{P}^\top D\tilde{P} \quad \text{mit} \quad \tilde{P} = Q^\top PQ'$$

und außerdem $\mathrm{Rang}\,(A) = \mathrm{Rang}\,(D) = \mathrm{Rang}\,(D') = \mathrm{Rang}\,(A')$.

Da A' und A symmetrisch sind und mit der Singulärwertzerlegung Rang $(A) = \pi + \nu$ und Rang $(A') = \pi' + \nu'$ ist, reicht es also zu zeigen, dass $\pi = \pi'$, dann gilt auch $\nu = \nu'$.

Wir erhalten aus der Diagonalisierung Räume V_π, V_ν, V_ω bzw. $V'_{\pi'}, V'_{\nu'}, V'_{\omega'}$, die aus den Spalten der orthogonalen Matrizen Q, Q' gebildet werden. Wir zeigen nun, dass

$$V_\pi \cap (V'_{\nu'} \oplus V'_{\omega'}) = \{0\}.$$

Dazu sei v aus $V_\pi \cap (V'_{\nu'} \oplus V'_{\omega'})$, so gilt $\alpha(v,v) > 0$ und $\alpha(v,v) \leq 0$, also $v = 0$.

Mit der Dimensionsformel (Satz 8.26) folgt

$$\dim V_\pi + \dim V'_{\nu'} + \dim V'_{\omega'} \leq n.$$

Aber da

$$\dim V'_{\pi'} + \dim V'_{\nu'} + \dim V'_{\omega'} = n,$$

so folgt, dass $\dim V_\pi \leq \dim V'_{\pi'}$ und damit$\pi \leq \pi'$.

Mit $V'_{\pi'} \cap [V_\nu \oplus V_\omega]$ folgt $\pi' \leq \pi$.

Also gilt $\pi' = \pi$ und dann $\nu' = \nu$. □

Definition 11.5 *Das zu einer symmetrischen Bilinearform α gehörende Tripel*

$$(\pi, \nu, \omega)$$

heißt Trägheitsindex von α.

Dieser Begriff stammt ursprünglich aus der Mechanik (Beschreibung von Trägheitsmomenten). Ein direktes Korollar aus Satz 11.4 ist die Normalform unter Kongruenz.

Korollar 11.6 *Sei V ein endlichdimensionaler euklidischer Raum der Dimension n und α eine symmetrische Bilinearform von V, so gibt es eine Basis $\{v_1, \ldots, v_n\}$ von V, so dass die Matrixdarstellung von α bezüglich dieser Basis die Form*

$$\begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0 \end{bmatrix} \in \mathbb{R}^{n,n}.$$

hat.

Beweis: Wir wissen schon, dass π, ν, ω Invarianten unter Kongruenz sind und dass es P orthogonal gibt, so dass

$$P^\top AP = \begin{bmatrix} D_1 & & \\ & D_2 & \\ & & 0 \end{bmatrix} \begin{matrix} \}\pi \\ \}\nu \\ \}\omega \end{matrix}. \quad \text{Setze} \quad \tilde{P} = P \cdot \begin{bmatrix} D_1^{-\frac{1}{2}} & & \\ & (-D_2)^{-\frac{1}{2}} & \\ & & I_\omega \end{bmatrix},$$

wobei für $D = \operatorname{diag}(d_1, \ldots, d_n)$

$$D^{\frac{1}{2}} = \operatorname{diag}(d_1^{\frac{1}{2}}, \ldots, d_n^{\frac{1}{2}}).$$

Damit folgt

$$\tilde{P}^\top A\tilde{P} \;=\; \begin{bmatrix} D_1^{-\frac{1}{2}} D_1 D_1^{-\frac{1}{2}} & & \\ & (-D_2)^{-\frac{1}{2}} D_2 (-D_2)^{-\frac{1}{2}} & \\ & & 0 \end{bmatrix} = \begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0 \end{bmatrix}.$$

□

Man beachte, dass die Transformationsmatrix nicht eindeutig ist, jedoch die Form $\begin{bmatrix} I_\pi & & \\ & -I_\nu & \\ & & 0 \end{bmatrix}$.

Definition 11.7 *Sei $A \in \mathbb{R}^{n,n}$ symmetrisch, so heißt A* positiv definit positiv semidefinit*), falls*

$$x^\top Ax > 0 \; (\geq 0), \qquad \textit{für alle } \; x \in \mathbb{R}^n \setminus \{0\}.$$

A heißt negativ definit negativ semidefinit*), falls*

$$x^\top Ax < 0 \; (\leq 0), \qquad \textit{für alle } x \in \mathbb{R}^n \setminus \{0\}.$$

Korollar 11.8 *Sei $A \in \mathbb{R}^{n,n}$ symmetrisch.*

(a) Die folgenden Aussagen sind äquivalent.

(i) A ist positiv definit.

(ii) Alle Eigenwerte von A sind positiv.

(iii) Der Trägheitsindex von A ist $(n, 0, 0)$.

(b) Die folgenden Aussagen sind äquivalent.

(i) A *ist positiv semidefinit.*

(ii) Alle Eigenwerte von A *sind nichtnegativ.*

(iii) Der Trägheitsindex von A *ist* $(\pi, 0, \omega)$.

Beweis: Der Beweis ist eine Übungsaufgabe. $\square$

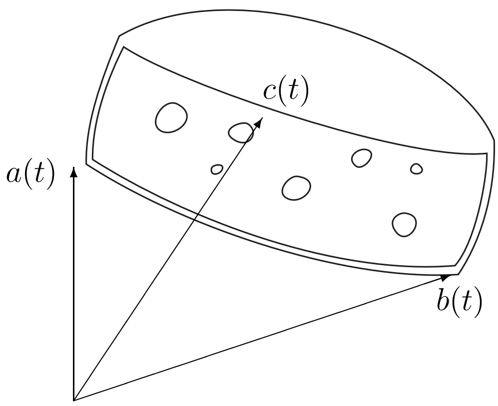
Kapitel 12

Bewegung starrer Körper

Als Beispiel wollen wir nun die Bewegung eines starren Körpers betrachten.
Markiere einen seiner Punkte und beschreibe dessen Bahn durch die Funktion $a(t)(t \geq 0)$. Da der Körper starr ist, ändern die anderen Punkte ihren Abstand zu a nicht, so dass für alle Punkte $x(t)$ gilt

$$x(t) - a(t) = Q(t)\,(x(0) - a(0))$$

mit einer t-abhängigen orthogonalen Matrix Q.



Mit der Kenntnis der Bahnbewegung zweier weiterer Punkte $b(t), c(t)$ (die zu $a(t)$ in allgemeiner Lage liegen, d.h. die entsprechenden Vektoren in $\mathbb{R}63$ sind linear unabhängig) kann man dann $Q(t)$ bestimmen.

$$\begin{aligned}
x_1(t) &= b(t) - a(t) = Q(t)\,(b(0) - a(0))\,, \\
x_2(t) &= c(t) - a(t) = Q(t)\,(c(0) - a(0))\,, \\
x_3(t) &= (b(t) - a(t)) \times (c(t) - a(t)) \\
&= (Q(t)\,(b(0) - a(0))) \times (Q(t)\,(c(0) - a(0)))\,.
\end{aligned}$$

Die Vektoren sind linear unabhängig.

Da Q orthogonal und $x_3(0) \perp x_1(0), x_2(0)$, gilt auch

$$Q(t)x_3(0) \perp Q(t)x_1(0), Q(t)x_2(0) \implies Q(t)x_3(0) = \lambda(Q(t)x_1(0) \times Q(t)x_2(0)).$$

Da sich die Orientierung des Körpers nicht ändern kann, gilt $\lambda > 0$ und

$$\begin{aligned}
\|(Qx_1) \times (Qx_2)\|_2^{\,2} &= \|Qx_1\|_2^{\,2}\|Qx_2\|_2^{\,2} - \langle Qx_1, Qx_2\rangle^2 \\
&= \|x_1\|_2^{\,2}\|x_2\|_2^{\,2} - \langle x_1, x_2\rangle^2 = \|x_3\|_2^{\,2} \\
&= \|Qx_3\|_2^{\,2}
\end{aligned}$$

so folgt $\lambda = 1$. Damit haben wir gezeigt, dass auch für x_3 die Bewegung durch $x_3(t) = Q(t)x_3(0)$ beschrieben wird.

Da $\det[Qx_1, Qx_2, Qx_3] = \det Q \cdot \det[x_1, x_2, x_3]$, aber $\det Q \in \{+1, -1\}$, bedeutet die Erhaltung der Orientierung also: $\det Q = +1$.

$$Q(t) \in \mathcal{SO}(3) = \{Q \in \mathcal{O}(3) | \det Q = 1\}\,.$$

Man beachte, dass $a(0), a(t)$ nicht eindeutig bestimmt sind, aber $Q(t)$. Für die Geschwindigkeit $\dot{x} = \frac{d}{dt}x(t)$ eines Punktes $x(t)$ gilt

$$\dot{x}(t) = \dot{a}(t) + \dot{Q}(t)\,(x(0) - a(0))\,.$$

Was wissen wir über die Ableitung einer orthogonalen Matrix?

$$\begin{array}{lllll}
Q^\top Q = I & \Longrightarrow & \dot{Q}^\top Q + Q^\top \dot{Q} = 0 & \Longrightarrow & Q^\top \dot{Q} \text{ ist schiefsymmetrisch,} \\
QQ^\top = I & \Longrightarrow & \dot{Q}Q^\top + Q\dot{Q}^\top = 0 & \Longrightarrow & \dot{Q}Q^\top \text{ ist schiefsymmetrisch.}
\end{array}$$

Nun gilt aber

$$\dot{Q}\,(x(0) - a(0)) = \dot{Q}Q^\top Q\,(x(0) - a(0))\,,$$

also

$$\dot{x} = \dot{a} + \dot{Q}Q^\top\,(Q(x(0) - a(0))) \;=\; \dot{a} + \dot{Q}Q^\top\,(x(t) - a(t))\,.$$

Die Matrix $\dot{Q}Q^\top$ ist schiefsymmetrisch und kann daher geschrieben werden als

$$\dot{Q}Q^\top = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix}, \qquad \text{so ist}$$

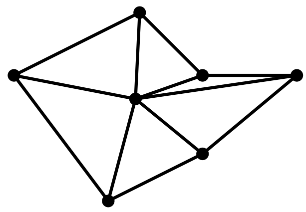
$$\dot{Q}Q^\top\,(x(t)-a(t)) = \underbrace{\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}}_{\omega(t)} \times (x(t)-a(t)) \qquad \text{(Vektorprodukt in } \mathbb{R}^3\text{).}$$

$$\Longrightarrow \quad \dot{x} = \dot{a} + \omega(t) \times (x(t)-a(t))\,.$$

Die Momentanbewegung eines starren Körpers ist zusammengesetzt aus der Translation des Punktes $a(t)$ mit der Geschwindigkeit $\dot{a}(t)$ und einer Rotation um die durch $\omega(t)$ definierte Achse durch $a(t)$ mit der Winkelgeschwindigkeit $\|\omega(t)\|_2$.

> Beachte: $\omega(t)$ ist eindeutig, $\dot{a}$ und die Zerlegung von $\dot{x}$ in $\dot{a}$ und ω sind nicht eindeutig.

Wir wählen nun als $a(t)$ die Bahn des Schwerpunktes eines Körpers, der aus starr verbundenen Einzelmassen besteht.



$$a(t) = \sum_{i=1}^{N} \frac{m_i}{m} y_i(t)$$

wobei y_i die Orte einzelner Massenpunkte sind, m_i ihre Massen und $m = \sum\limits_{i=1}^{N} m_i$ die Gesamtmasse.
(Im nichtdiskreten Fall sind die Summen durch Volumenintegrale zu ersetzen).

Setze $v = \|\dot{a}\|_2, v_i = \|\dot{y}_i\|_2,\ r_i = y_i - a\,,\ \varrho_i = \|r_i\|_2$.

Dann ist die kinetische Energie des Körpers gegeben durch:

$$\begin{aligned} T &= \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2 = \sum_{i=1}^{N} \frac{1}{2} m_i \|\dot{a} + \omega \times r_i\|_2{}^2 \\ &= \sum_{i=1}^{N} \left(\frac{m_i}{2} v^2 + m_i \langle \dot{a}, \omega \times r_i \rangle + \frac{1}{2} m_i \,\|\omega \times r_i\|_2^2 \right) \\ &= \frac{m}{2} v^2 + \underbrace{\langle \dot{a}, \omega \times \sum_{i=1}^{N} m_i r_i \rangle}_{0} + \sum_{i=1}^{N} \frac{1}{2} m_i \left(\|\omega\|_2^2\, \varrho_i^2 - \langle \omega, r_i \rangle^2 \right) \end{aligned}$$

Nun ist aber $\sum\limits_{i=1}^{N} m_i r_i = 0$ und mit $r_i = \begin{bmatrix} r_{i1} \\ r_{i2} \\ r_{i3} \end{bmatrix}$, $\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}$ folgt

$$\|\omega\|_2{}^2 \varrho_i^2 - \langle \omega, r_i \rangle^2 = (\omega_1^2 + \omega_2^2 + \omega_3^2)\varrho_i^2 - (\omega_1 r_{i1} + \omega_2 r_{i2} + \omega_3 r_{i3})^2$$

$$= \sum_{k,l=1}^{3} \omega_k \omega_l (\varrho_i^2 \delta_{kl} - r_{ik} r_{il}).$$

Setze $\Theta = [\Theta_{kl}] \in \mathbb{R}^{3,3}$ mit

$$\Theta_{k,l} = \sum_{i=1}^{N} m_i \{\varrho_i^2 \delta_{kl} - r_{ik} r_{il}\},$$

so folgt

$$T = \frac{m}{2} v^2 + \frac{1}{2} \omega^\top \Theta \omega.$$

Dann nennt man $\frac{m}{2} v^2$ die *Translationsenergie*, und $\frac{1}{2} \omega^\top \Theta$ die *Rotationsenergie.* Die Normierung der Bilinearform Θ als

$$\Theta_e = \frac{\omega^\top \Theta \omega}{\|\omega\|_2^{\,2}}$$

heißt das auf die Achse $\dfrac{\omega}{\|\omega\|_2}$ bezogene *Trägheitsmoment* des Körpers.

Θ ist symmetrisch und (falls nicht alle Massenpunkte auf einer Geraden liegen) positiv definit. Die Gleichung $\omega^\top \Theta \omega = 1$ beschreibt ein Ellipsoid, das Trägheitsellipsoid, das die bei fester Rotationsenergie zu den verschiedenen Richtungen gehörenden Winkelgeschwindigkeiten veranschaulicht. Die Hauptachsen dieses Ellipsoids sind Achsen, um die der Körper reine Rotationsbewegungen ausführen kann.
Dazu kommt dann noch die Kreiselbewegung, die behandeln wir hier nicht.

Kapitel 13

Quadriken

Wir wollen nun einen Bezug zur Geometrie herstellen und damit die Klassifikation von geometrischen Objekten im $\mathbb{R}^2$ und $\mathbb{R}^3$ vornehmen. Sei $\mathbb{K}$ ein Körper. Ein quadratisches Polynom in n Variablen ist ein Ausdruck der Form

$$P(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} a_{ij} x_i x_j + \sum_{1 \le i \le n} b_i x_i + c, \tag{13.1}$$

in dem nicht alle a_{ij} verschwinden. P ist eine *nichtlineare* Abbildung

$$P : \mathbb{K}^n \to \mathbb{K}.$$

Definition 13.2 *Eine Teilmenge $Q \subset \mathbb{K}^n$ heißt* Quadrik *oder* Hyperfläche zweiter Ordnung, *falls es ein quadratisches Polynom gibt, so dass*

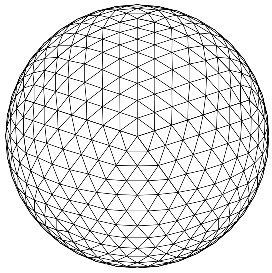
$$Q = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n \;\middle|\; P(x_1, \ldots, x_n) = 0 \right\}.$$

Beispiel 13.3

(1) $\mathbb{K} = \mathbb{R},\ n = 3$

$$P(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1,$$

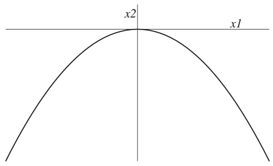
$$Q = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; x_1^2 + x_2^2 + x_3^2 - 1 = 0 \right\}$$



ist gerade die Oberfläche einer Kugel mit Mittelpunkt $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ und Radius 1.

(2) $\mathbb{K} = \mathbb{R},\ n = 2$

$$P(x_1, x_2) = x_1^2 + 2x_2$$
$$Q = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; x_1^2 + 2x_2 = 0 \right\}$$

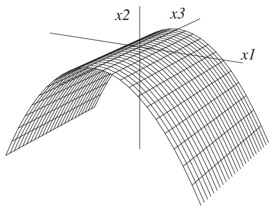


Parabel

(3) $\mathbb{K} = \mathbb{R},\ n = 3$

$$P(x_1, x_2, x_3) = x_1^2 + 2x_2$$

$$Q = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1^2 + 2x_2 = 0 \right\}$$



Parabelzylinder

Im folgenden betrachten wir nur Körper, in denen $1 + 1 \neq 0$ ist.
Sei $P(x_1, \ldots, x_n)$ quadratisches Polynom wie in (13.1). Wir wollen nun alle Quadriken mit Hilfe von Matrizen beschreiben. Dazu konstruieren wir aus $A = [a_{ij}] \in \mathbb{K}^{n,n}, b \in \mathbb{K}^n, c \in \mathbb{K}$ die folgende Matrix

$$\hat{A} \;=\; \begin{bmatrix} c & \frac{b_1}{2} & \frac{b_2}{2} & \cdots & \frac{b_n}{2} \\ \frac{b_1}{2} & a_{11} & \frac{a_{12}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{b_2}{2} & \frac{a_{12}}{2} & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \frac{a_{n-1,n}}{2} \\ \frac{b_n}{2} & \frac{a_{1n}}{2} & \cdots & \frac{a_{n-1,n}}{2} & a_{nn} \end{bmatrix} = [\hat{a}_{ij}]. \tag{13.4}$$

$\hat{A}$ ist symmetrisch und enthält alle Koeffizienten des Polynoms. Weiter gilt, dass $x \in Q$ genau dann wenn

$$\hat{x}^\top \hat{A}\hat{x} = 0 \quad \text{für } \hat{x} = \begin{bmatrix} 1 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Durch diese Erweiterung gehören also zu den Punkten aus Q diejenigen erweiterten Vektoren, für die die über $\hat{A}$ definierte Bilinearform

$$\hat{\alpha} : \mathbb{K}^{n+1} \times \mathbb{K}^{n+1} \to \mathbb{K}$$

ergibt $\hat{\alpha}(\hat{x}, \hat{x}) = \hat{x}^\top \hat{A}\hat{x} = 0$.

Wir wollen nun spezielle Abbildungen betrachten, die Abstände erhalten.

Sei V ein endlichdimensionaler euklidischer Vektorraum mit einer Abstandsfunktion

$$d(v, w) = \|v - w\|_2 .$$

Eine Abbildung $f : V \to V$, für die gilt $d(v, w) = d(f(v), f(w))$ für alle $v, w \in V$ heißt *abstandserhaltend* (wird manchmal auch *Kongruenzabbildung* genannt).

Betrachte nun eine lineare Abbildung

$$\begin{aligned} g : \ & V \to V \\ & v \mapsto g(v) = f(v) - f(0) \end{aligned}$$

mit f abstandserhaltend.

Es gilt natürlich, dass g wieder abstandserhaltend ist und $g(0) = 0$.

Also folgt aus Lemma 8.9, dass g ein orthogonaler Endomorphismus ist. Es gibt also zu jeder abstandserhaltenden Funktion $f : V \to V$ einen orthogonalen Endomorphismus g, so dass

$$f(v) = a + g(v) \quad \text{für alle } v \in V, \qquad (a = f(0)).$$

Umgekehrt gilt natürlich sofort, dass alle Abbildungen $v \mapsto a + g(v)$, mit $a \in V$ und g orthogonal, abstandserhaltend sind.

Was ist die Matrixdarstellung von f bzw. g in dem Raum der Vektoren $\begin{bmatrix} 1 \\ x \end{bmatrix}$, mit $x \in V$?
Von g ist das natürlich eine orthogonale Matrix und von f eine Matrix der Form

$$\hat{G} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_1 & & & \\ \vdots & & G & \\ a_n & & & \end{array}\right] \tag{13.5}$$

wobei G die (orthogonale) Matrixdarstellung von g ist und $a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$.

Lemma 13.6 *Ist Q eine Quadrik in $\mathbb{R}^n$, beschrieben durch die Matrix $\hat{A}$ und f eine abstandserhaltende Abbildung mit der Matrixdarstellung*

$$\hat{G} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_1 & & & \\ \vdots & & G & \\ a_n & & & \end{array}\right],$$

so ist $f(Q)$ eine Quadrik, beschrieben durch die Matrix

$$\hat{G}^{-\top} \hat{A} \hat{G}^{-1}. \tag{13.7}$$

Wir sehen, dass dies eine (spezielle) Kongruenztransformation mit $\hat{G}^{-1}$ ist.

Beweis: Sei $y = f(x)$, $\hat{y} = \hat{G}\hat{x}$ mit

$$\hat{y} = \begin{bmatrix} 1 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}, \hat{x} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} \Longrightarrow \hat{x} = \hat{G}^{-1}\hat{y}.$$

(Aus der Gruppeneigenschaft folgt, dass $\hat{G}$ invertierbar ist.) Damit haben wir, dass $y \in f(Q)$ genau dann wenn $x \in Q$, genau dann wenn $\hat{x}^\top \hat{A} \hat{x} = 0$ und ist dies gilt genau dann wenn $\hat{y}^\top \hat{G}^{-\top} \hat{A} \hat{G}^{-1} \hat{y} = 0$. Also wird $f(Q)$ gerade durch die

Matrix $\hat{G}^{-\top}\hat{A}\hat{G}^{-1}$ beschrieben und ist damit eine Quadrik, denn $\hat{G}^{-1}$ hat die gleiche Form wie $\hat{G}$. $\square$

Wir wollen noch einmal anschauen, was

$$\hat{y}^\top\hat{G}^{-\top}\hat{A}\hat{G}^{-1}\hat{y} = 0 \tag{13.8}$$

ist:

$$\hat{y} = \begin{bmatrix} 1 \\ y \end{bmatrix}, \quad \hat{G} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_1 & & & \\ \vdots & & G & \\ a_n & & & \end{array}\right] =: \begin{bmatrix} 1 & 0 \\ a & G \end{bmatrix}, \quad \hat{G}^{-1} = \begin{bmatrix} 1 & 0 \\ -G^{-1}a & G^{-1} \end{bmatrix},$$

$$\hat{A} = \begin{bmatrix} c & \frac{b_1}{2} & \cdots & \frac{b_n}{2} \\ \frac{b_1}{2} & a_{11} & & \frac{a_{1n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{b_n}{2} & \frac{a_{n1}}{2} & \cdots & a_{nn} \end{bmatrix} =: \begin{bmatrix} c & \frac{b^\top}{2} \\ \frac{b}{2} & \tilde{A} \end{bmatrix}$$

$$\hat{G}^{-1}\begin{bmatrix} 1 \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -G^{-1}a + G^{-1}y \end{bmatrix} = \begin{bmatrix} 1 \\ G^{-1}(y-a) \end{bmatrix}$$

$$\hat{y}^\top\hat{G}^{-\top}\hat{A}\hat{G}^{-1}\hat{y} = \begin{bmatrix} 1 \\ G^{-1}(y-a) \end{bmatrix}^\top \begin{bmatrix} c & \frac{b^\top}{2} \\ \frac{b}{2} & \tilde{A} \end{bmatrix} \begin{bmatrix} 1 \\ G^{-1}(y-a) \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ G^{-1}(y-a) \end{bmatrix}^\top \begin{bmatrix} c + \frac{b^\top}{2}G^{-1}(y-a) \\ \frac{b}{2} + \tilde{A}G^{-1}(y-a) \end{bmatrix}$$

$$= c + \frac{b^\top}{2}(G^{-1}(y-a)) + (G^{-1}(y-a))^\top\frac{b}{2} + (G^{-1}(y-a))^\top\tilde{A}(G^{-1}(y-a))$$

$$= c + \left(\frac{(y-a)}{2}^\top G^{-\top}b\right) + \left(\frac{(y-a)}{2}^\top G^{-\top}b\right)^\top + (y-a)^\top(G^{-\top}\tilde{A}G^{-1})(y-a).$$

Wir erhalten also, dass $\tilde{A}$ durch eine orthogonale Kongruenztransformation mit G^{-1} transformiert wird.

Satz 13.9 (Klassifikation der Quadriken)
Sei Q eine Quadrik in $\mathbb{R}^n$. Dann gibt es eine abstandserhaltende Abbildung f, natürliche Zahlen $\tilde{\pi}, \tilde{\nu}$ und reelle Zahlen $\beta_i > 0, 1 \leq i \leq \tilde{\pi}+\tilde{\nu}$, so dass die Quadrik $f(Q)$ durch eine der folgenden Gleichungen beschrieben wird:

$$\sum_{i=1}^{\tilde{\pi}} \frac{x_i^2}{\beta_i^2} - \sum_{i=\tilde{\pi}+1}^{\tilde{\pi}+\tilde{\nu}} \frac{x_i^2}{\beta_i^2} = 0, \qquad \tilde{\pi} \geq \tilde{\nu}, \tag{13.10}$$

$$\sum_{i=1}^{\tilde{\pi}} \frac{x_i^2}{\beta_i^2} - \sum_{i=\tilde{\pi}+1}^{\tilde{\pi}+\tilde{\nu}} \frac{x_i^2}{\beta_i^2} = 1, \tag{13.11}$$

$$\sum_{i=1}^{\tilde{\pi}} \frac{x_i^2}{\beta_i^2} - \sum_{i=\tilde{\pi}+1}^{\tilde{\pi}+\tilde{\nu}} \frac{x_i^2}{\beta_i^2} = x_{\tilde{\pi}+\tilde{\nu}+1}, \qquad \tilde{\pi}+\tilde{\nu} < n. \tag{13.12}$$

Beweis: Sei $Q = \left\{x \in \mathbb{R}^n \mid P(x_1, \ldots, x_n) = \hat{x}^\top \hat{A} \hat{x} = 0\right\}$, wobei $\hat{x}, \hat{A}$ wie in (13.4) gebildet sind,

$$\hat{A} = \begin{bmatrix} c & \frac{b^\top}{2} \\ \frac{b}{2} & \tilde{A} \end{bmatrix}, \qquad \text{und} \qquad \tilde{A} = \tilde{A}^\top.$$

1. Schritt: Diagonalisierung von $\tilde{A}$. Da $\tilde{A}$ reell symmetrisch ist, so gibt es nach dem Trägheitssatz von Sylvester eine orthogonale Matrix P, so dass

$$P^\top \tilde{A} P = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_\pi & & & & & & \\ & & & \lambda_{\pi+1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_{\pi+\nu} & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix}$$

mit $\lambda_1, \ldots, \lambda_\pi > 0$, $\lambda_{\pi+1} \ldots \lambda_{\pi+\nu} < 0$. Setze $\hat{P} = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array}\right]$, so gilt

$$A_1 = \hat{P}^\top \hat{A} \hat{P} = \left[\begin{array}{c|c} c & \frac{b^\top}{2} P \\ \hline P^\top \frac{b}{2} & P^\top \tilde{A} P \end{array}\right] =: \begin{bmatrix} c & \gamma_1 & \cdots & \gamma_n \\ \gamma_1 & \lambda_1 & & \\ \vdots & & \ddots & \\ \gamma_n & & & \lambda_n \end{bmatrix}$$

2. Schritt: Verschiebung des Nullpunktes

$$\text{Sei } T = \begin{bmatrix} 1 & & & & & & \\ -\dfrac{\gamma_1}{\lambda_1} & 1 & & & & & \\ \vdots & & \ddots & & & & \\ -\dfrac{\gamma_{\pi+\nu}}{\lambda_{\pi+\nu}} & & & 1 & & & \\ 0 & & & & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix},$$

$$A_2 = T^\top A_1 T = \left[\begin{array}{c|ccc|c} \tilde{c} & 0 & \cdots & 0 & \gamma_{\pi+\nu+1} \cdots \gamma_n \\ \hline 0 & \lambda_1 & & & \\ \vdots & & \ddots & & 0 \\ 0 & & & \lambda_{\pi+\nu} & \\ \hline \gamma_{\pi+\nu+1} & & & & \\ \vdots & & 0 & & 0 \\ \gamma_n & & & & \end{array}\right].$$

3. Schritt: Jetzt unterscheiden wir 3 Fälle.

(a) $\tilde{c} = 0$ und $\gamma_{\pi+\nu+1} = \ldots = \gamma_n = 0$.
Mit der Setzung $\beta_i = \dfrac{1}{\sqrt{|\lambda_i|}}$ erhalten wir dann Rang $\tilde{A}$ = Rang $\hat{A}$ und Teil (13.10).

(b) $\tilde{c} \neq 0$ und $\gamma_{\pi+\nu+1} = \ldots = \gamma_n = 0$, also Rang $\hat{A}$ = Rang $\tilde{A} + 1$.
Mit $\beta_i = \sqrt{\dfrac{|\tilde{c}|}{|\lambda_i|}}$ erhalten wir (13.11).

(c) $\tilde{c} \neq 0$ und es gibt $\gamma_j \neq 0,\ \pi + \nu < j \leq n \implies$ Rang $\hat{A}$ = Rang $\tilde{A} + 2$.
Setze $\gamma = \begin{bmatrix} \gamma_{\pi+\nu+1} \\ \vdots \\ \gamma_n \end{bmatrix}$ und $\varphi_1 = \dfrac{\gamma}{\|\gamma\|_2}$.
Ergänze φ_1 durch $\varphi_2, \ldots, \varphi_{n-\pi-\nu}$ zu einer Orthonormalbasis von $\mathbb{R}^{n-\pi-\nu}$,

setze

$$\tilde{\tilde{c}} = -\frac{\tilde{c}}{2\|\gamma\|_2} \qquad \text{und} \qquad V = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & I_{\pi+\nu} & 0 \\ \hline \tilde{\tilde{c}}\varphi_1 & 0 & \tilde{V} \end{array}\right]$$

mit $\tilde{V} = [\varphi_1 \varphi_2 \ldots \varphi_{n-\pi-\nu}]$. Dann gilt

$$V^\top A_2 V = \left[\begin{array}{c|ccc|cccc} 0 & 0 & \cdots & 0 & \|\gamma\|_2 & 0 & \cdots & 0 \\ \hline 0 & \lambda_1 & & & & & & \\ \vdots & & \ddots & & & & 0 & \\ 0 & & & \lambda_{\pi+\nu} & & & & \\ \hline \|\gamma\|_2 & & & & & & & \\ 0 & & & & & & & \\ \vdots & & 0 & & & & 0 & \\ 0 & & & & & & & \end{array}\right]$$

und mit $\beta_i = \sqrt{\dfrac{2\|\gamma\|_2}{|\lambda_i|}}$, $i = 1, \ldots, \pi + \nu$ erhalten wir (13.12).

Beachte, je nach Vorzeichen von $\tilde{c}$ wechseln die Rollen von π und ν. Ausserdem haben wir in den Gleichungen noch eine Skalierung durchgeführt, so dass die rechte Seite in (13.11) und (13.12) 1 bzw. $x_\pi + \nu + 1$ ist. □

Beispiel 13.13

$$p(x) = x_1^2 + 9x_2^2 - 6x_1x_2 + 20x_1 - 4x_2 - 10$$

$$\Longrightarrow \tilde{A} = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}, \quad \hat{A} = \begin{bmatrix} -10 & \frac{20}{2} & -\frac{4}{2} \\ \frac{20}{2} & 1 & -3 \\ -\frac{4}{2} & -3 & 9 \end{bmatrix} = \begin{bmatrix} -10 & 10 & -2 \\ 10 & 1 & -3 \\ -2 & -3 & 9 \end{bmatrix}$$

Rang $\tilde{A} = 1$, Rang $\hat{A} = 3$, da invertierbar. Also haben wir den Fall (13.12).

(1) Diagonalisierung von $\tilde{A}$:

$$P_{\tilde{A}}(\lambda) = \lambda^2 - 10\lambda, \quad \Longrightarrow \quad \text{Eigenwerte } 0, 10$$

$$P = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}, \text{ so dass } P^\top(\tilde{A} - 0 \cdot I)P = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}$$

$$s = \frac{-3}{\sqrt{3^2+1}} = -\frac{3}{\sqrt{10}}, \quad c = \frac{1}{\sqrt{10}}$$

$$P^\top \tilde{A} P = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & P^\top \end{bmatrix} \hat{A} \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix} = \begin{bmatrix} -10 & \frac{8}{5}\sqrt{10} & \frac{14}{5}\sqrt{10} \\ \frac{8}{5}\sqrt{10} & 10 & 0 \\ \frac{14}{5}\sqrt{10} & 0 & 0 \end{bmatrix}$$

(2) Verschiebung des Nullpunktes

$$T = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline -\frac{8\sqrt{10}}{5\cdot 10} & 1 & 0 \\ 0 & 0 & 1 \end{array}\right], \qquad A_2 = T^\top A_1 T = \begin{bmatrix} -\frac{314}{25} & 0 & \frac{14}{5}\sqrt{10} \\ 0 & 10 & 0 \\ \frac{14}{5}\sqrt{10} & 0 & 0 \end{bmatrix}$$

$$\gamma = \left[\frac{14}{5}\sqrt{10}\right] \in \mathbb{R}^1, \quad \varphi_1 = [1], \quad \tilde{\tilde{c}} = \frac{314}{25 \cdot 2 \cdot \|\gamma\|_2} = \frac{157}{700}\sqrt{10} \qquad .$$

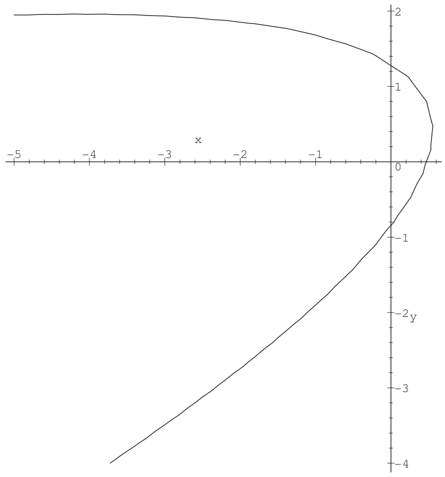
$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{157}{700}\sqrt{10} & 0 & 1 \end{bmatrix}, \qquad V^\top A_2 V = \begin{bmatrix} 0 & 0 & \frac{14}{5}\sqrt{10} \\ 0 & 10 & 0 \\ \frac{14}{5}\sqrt{10} & 0 & 0 \end{bmatrix}$$

Setze $\beta_1 = \sqrt{\dfrac{2 \cdot 14\sqrt{10}}{5 \cdot 10}}$. Dann erhalten wir die transformierte Gleichung

$$\frac{y_1^2}{\beta_1^2} = y_2 \qquad \text{oder} \qquad \frac{5}{28}\sqrt{10}y_1^2 - y_2 = 0, \quad \text{wobei}$$

$$\hat{x} = P \cdot T \cdot V\, \hat{y} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{359}{700} & \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \\ \frac{493}{700} & -\frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 1 \\ y_1 \\ y_2 \end{bmatrix}, \qquad \text{bzw.}$$

$$\hat{y} = V^{-1} \cdot T^{-1} \cdot \hat{P}^{-1}\, \hat{x} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{4}{25}\sqrt{10} & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ -\frac{157}{700}\sqrt{10} & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}.$$

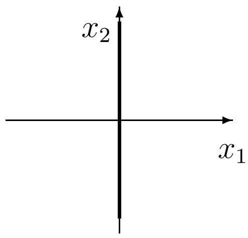


$$p(x) = x_1^2 + 9x_2^2 - 6x_1x_2 + 20x_1 - 4x_2 - 10 = 0$$

Wir können damit alle Quadriken klassifizieren. Im $\mathbb{R}^2$ erhalten wir die folgenden Möglichkeiten:

Tabelle 13.14 Quadriken in $\mathbb{R}^2$

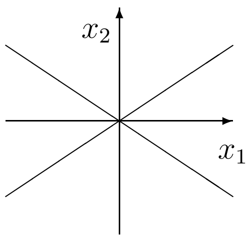
(13.10) $\nu = 0, \quad \pi = 1, \qquad x_1^2 = 0$



Gerade

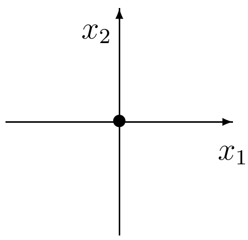
$\nu = 1, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} - x_2^2 = 0$

$x_1 = \pm\beta_1 x_2$



zwei sich schneidende Geraden

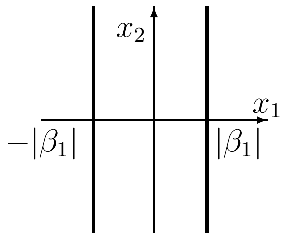
$\nu = 0, \quad \pi = 2, \qquad \dfrac{x_1^2}{\beta_1^2} + x_2^2 = 0$



Punkt

(13.11) $\nu = 1, \quad \pi = 0, \qquad \dfrac{-x_1^2}{\beta_1^2} = 1$ $\qquad \emptyset$

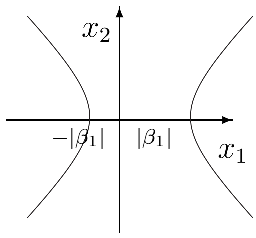
$\nu = 0, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} = 1$



zwei parallele Geraden

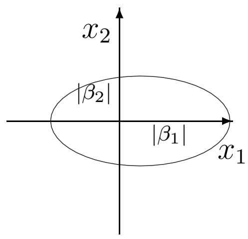
$\nu = 2, \quad \pi = 0, \qquad \dfrac{-x_1^2}{\beta_1^2} - \dfrac{x_2^2}{\beta_2^2} = 1$ $\qquad \emptyset$

$\nu = 1, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} - \dfrac{x_2^2}{\beta_2^2} = 1$



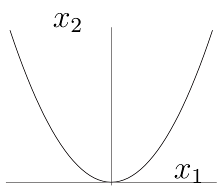
Hyperbel

$\nu = 0, \quad \pi = 2, \qquad \dfrac{x_1^2}{\beta_1^2} + \dfrac{x_2^2}{\beta_2^2} = 1$



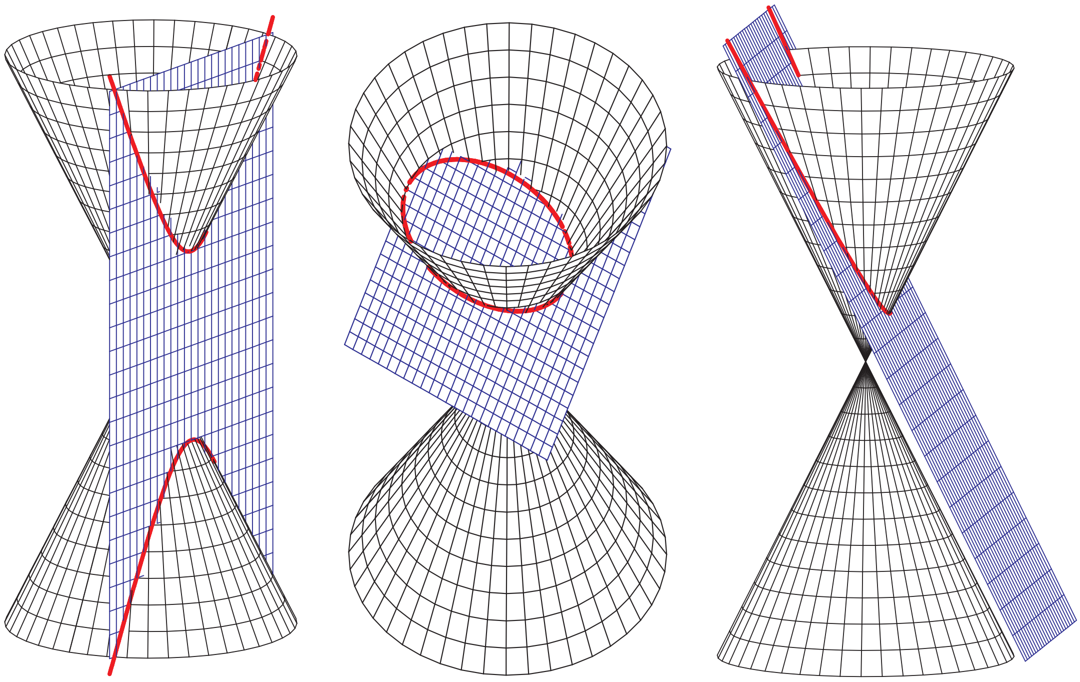
Ellipse

(13.12) $\nu = 0, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} = x_2$



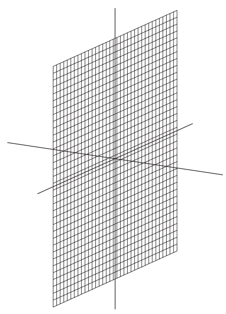
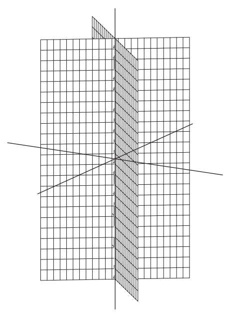
Parabel

Quadriken sind Schnitte von Ebenen mit einem doppelten Kreiskegel, sie werden daher auch Kegelschnitte genannt. Der besondere Fall der parallelen Geraden gehört zu einer degenerierten Situation wo der Doppelkegel in einen Zylinder entartet.

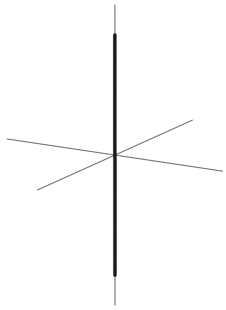


Hyperbel Ellipse Parabel

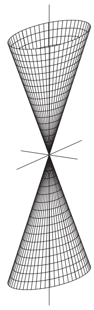
Tabelle 13.15 Quadriken in $\mathbb{R}^3$

(13.10)	$\nu = 0, \quad \pi = 1,$	$x_1^2 = 0$		eine Ebene
	$\nu = 1, \quad \pi = 1,$	$\dfrac{x_1^2}{\beta_1^2} - x_2^2 = 0$		zwei sich schneidende Ebenen

$\nu = 0, \quad \pi = 2, \qquad \dfrac{x_1^2}{\beta_1^2} + x_2^2 = 0$ — eine Gerade



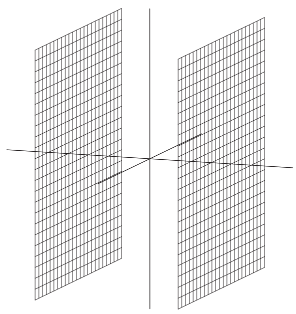
$\nu = 1, \quad \pi = 2, \qquad \dfrac{x_1^2}{\beta_1^2} + \dfrac{x_2^2}{\beta_2^2} - x_3^2 = 0$ — Ellipsenkegel



$\nu = 0, \quad \pi = 3, \qquad \dfrac{x_1^2}{\beta_1^2} + \dfrac{x_2^2}{\beta_2^2} + x_3^2 = 0$ — Punkt

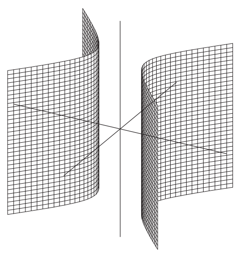
$$\nu = 1, \quad \pi = 0, \qquad \frac{-x_1^2}{\beta_1^2} = 1 \qquad \emptyset \tag{13.11}$$

$\nu = 0, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} = 1$ — zwei parallele Ebenen

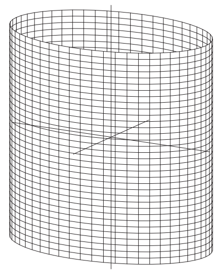


$\nu = 2, \quad \pi = 0, \qquad \dfrac{-x_1^2}{\beta_1^2} - \dfrac{x_2^2}{\beta_2^2} = 1$ — $\emptyset$

$\nu = 1, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} - \dfrac{x_2^2}{\beta_2^2} = 1$ — Hyperbelzylinder

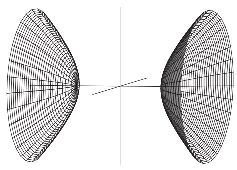


$\nu = 0, \quad \pi = 2, \qquad \dfrac{x_1^2}{\beta_1^2} + \dfrac{x_2^2}{\beta_1^2} = 1$ Ellipsenzylinder

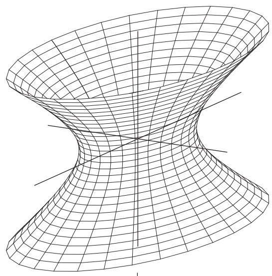


$\nu = 3, \quad \pi = 0, \qquad \dfrac{-x_1^2}{\beta_1^2} - \dfrac{x_2^2}{\beta_2^2} - \dfrac{x_3^2}{\beta_3^2} = 1$ $\emptyset$

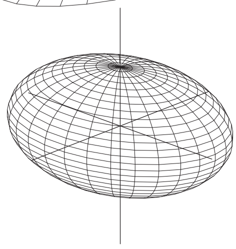
$\nu = 2, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} - \dfrac{x_2^2}{\beta_2^2} - \dfrac{x_3^2}{\beta_3^2} = 1$ zweischaliges Hyperboloid



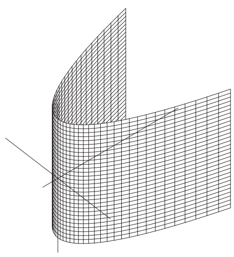
$\nu = 1, \quad \pi = 2, \qquad \dfrac{x_1^2}{\beta_1^2} + \dfrac{x_2^2}{\beta_2^2} - \dfrac{x_3^2}{\beta_3^2} = 1$ einschaliges Hyperboloid



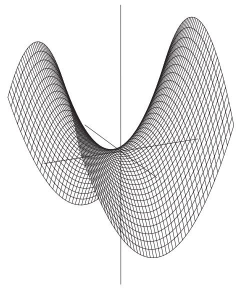
$\nu = 0, \quad \pi = 3, \qquad \dfrac{x_1^2}{\beta_1^2} + \dfrac{x_2^2}{\beta_2^2} + \dfrac{x_3^2}{\beta_3^2} = 1$ Ellipsoid



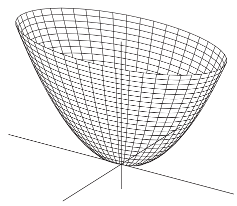
(13.12) $\nu = 0, \quad \pi = 1, \qquad \dfrac{x_1^2}{\beta_1^2} = x_2$ parabolischer Zylinder



$\nu = 1, \quad \pi = 1, \qquad \frac{x_1^2}{\beta_1^2} - \frac{x_2^2}{\beta_2^2} = x_3$ hyperbolisches Paraboloid



$\nu = 0, \quad \pi = 2, \qquad \frac{x_1^2}{\beta_1^2} + \frac{x_2^2}{\beta_2^2} = x_3$ elliptisches Paraboloid



Definition 13.16 *Sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung $f : V \to V$ heißt* affin linear*, falls es $a \in V$ und lineare Abbildung $g : V \to V$ gibt, so dass $f(v) = a + g(v)$.*
Zwei Quadriken Q_1, Q_2 heißen affin äquivalent*, wenn es eine bijektive affine Abbildung $f : V \to V$ gibt mit $f(Q_1) = Q_2$.*

Beachte, dass damit eine abstandserhaltenden Abbildung affin linear ist mit einem speziellen (orthogonalen) g.

Korollar 13.17 *Jede Quadrik in $\mathbb{R}^n$ ist affin äquivalent zu einer Quadrik in $\mathbb{R}^n$, die durch eine der folgenden Gleichungen gegeben ist:*

$$\sum_{i=1}^{\pi} x_i^2 - \sum_{i=\pi+1}^{\pi+\nu} x_i^2 = 0 \tag{13.18}$$

$$\sum_{i=1}^{\pi} x_i^2 - \sum_{i=\pi+1}^{\pi+\nu} x_i^2 = 1 \tag{13.19}$$

$$\sum_{i=1}^{\pi} x_i^2 - \sum_{i=\pi+1}^{\pi+\nu} x_i^2 = x_{\pi+\nu+1} \tag{13.20}$$

Beweis: Wir können natürlich annehmen, dass wir schon eine Beschreibung der Quadrik Q in der Form (13.10 - 13.12) haben. Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ die lineare Abbildung mit der Matrixdarstellung

$$\begin{bmatrix} \beta_1^{-1} & & & & & \\ & \ddots & & & & \\ & & \beta_{\pi+\nu}^{-1} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \quad \text{bzgl.} \quad \{e_1, \dots, e_n\},$$

so ist f natürlich bijektiv und wir haben sofort, dass $f(Q)$ einer der Gleichungen (13.18 - 13.20) genügt, denn sei zum Beispiel (13.10) die Form von Q, so gilt

$$\begin{aligned} y = f(x) \quad \text{mit} \quad & y_i = \frac{1}{\beta_i} x_i \quad i = 1, \dots, \pi + \nu \\ & y_i = x_i \qquad i = \pi + \nu + 1, \dots, n \end{aligned}$$

$$\begin{aligned} y \in f(Q) \iff x \in Q &\iff \sum_{i=1}^{\pi} \frac{x_i^2}{\beta_i^2} - \sum_{i=\pi+1}^{\pi+\nu} \frac{x_i^2}{\beta_i^2} = 0 \\ &\iff \sum_{i=1}^{\pi} y_i^2 - \sum_{i=\pi+1}^{\pi+\nu} y_i^2 = 0. \end{aligned}$$

Die anderen Fälle sind analog. □

Kapitel 14

Positiv definite Matrizen (Bilinearformen)

Wir haben bereits gesehen, dass symmetrisch positiv definite Matrizen (Bilinearformen) eine wichtige Rolle spielen (Skalarprodukte, Quadriken usw.), und dass diese dadurch charakterisiert sind, dass alle Eigenwerte reell und positiv sind. Jetzt wollen wir noch einige weitere wichtige Eigenschaften kennenlernen.
Zuerst noch ein paar Bemerkungen zum komplexen Fall.

Definition 14.1 *Sei V ein unitärer Raum und $\beta : V \times V \to \mathbb{C}$ eine Hermite'sche Form, so heißt β* positiv definit (positiv semidefinit) *genau dann, wenn $\beta(v, v) > 0$ ($\beta(v, v) \geq 0$) für alle $v \in V \setminus \{0\}$.*

Beachte, dass damit auch $\beta(v, v) \in \mathbb{R}$ gilt.

Korollar 14.2 *Sei V ein unitärer Raum und $\beta : V \times V \to \mathbb{C}$ eine Hermite'sche Form. Dann sind folgende Aussagen äquivalent:*

(i) β ist positiv definit.

(ii) Jede Matrixdarstellung B von β (bezüglich einer Basis $\{v_1, \dots, v_n\}$ von V) hat nur positive reelle Eigenwerte.

(iii) Für jede Matrixdarstellung B von β gibt es eine nichtsinguläre Matrix T, so dass

$$T^{-H} B T^{-1} = I\,.$$

(iv) Für jede Matrixdarstellung B von β gibt es eine nichtsinguläre untere Dreiecksmatrix L, so dass $B = LL^H$ (Cholesky-Zerlegung).

(v) *Für jede Matrixdarstellung B von β gibt es eine Hermite'sche Matrix A mit beliebigem Trägheitsindex $(\pi, \nu, 0)$, so dass $A^2 = B$.*

Beweis: Der Beweis, dass (i) $\iff$ (ii), gilt analog wie im reellen Fall.
„(ii) $\implies$ (iii)“: Nach dem Satz von Schur gibt es eine unitäre Matrix Q, so dass $B = Q^H DQ$ mit $D = \mathrm{diag}(d_1, \dots, d_n), d_i > 0$. Setze

$$D^{\frac{1}{2}} = \begin{bmatrix} d_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & d_n^{\frac{1}{2}} \end{bmatrix},$$

so folgt mit $T = D^{\frac{1}{2}}Q$, dass T invertierbar und

$$B = T^H T \qquad \implies \qquad T^{-H} B T^{-1} = I.$$

„(iii) $\implies$ (iv)“: Nach Satz **??** mit Folgerung **??** aus Teil I gibt es Q unitär, so dass $T = QR$ mit einer oberen Dreiecksmatrix R. Es folgt

$$B = T^H T = R^H Q^H Q R = R^H R \quad \text{mit} \quad L = R^H.$$

„(iv) $\implies$ (iii)“: trivial

„(iii) $\implies$ (ii)“: Übungsaufgabe

„(ii) $\implies$ (v)“: $B = Q^H DQ, \;\; D = \mathrm{diag}(d_1, \dots, d_n)$ mit $d_i > 0, \; i = 1, \dots, n$.
Setze $A = Q^H D^{\frac{1}{2}} \begin{bmatrix} I_\pi & \\ & -I_\nu \end{bmatrix} Q$, so folgt, dass A Hermite'sch ist und

$$A^2 = Q^H D^{\frac{1}{2}} \begin{bmatrix} I_\pi & \\ & -I_\nu \end{bmatrix} Q Q^H D^{\frac{1}{2}} \begin{bmatrix} I_\pi & \\ & -I_\nu \end{bmatrix} Q = Q^H D^{\frac{1}{2}} D^{\frac{1}{2}} Q = Q^H DQ.$$

„(v) $\implies$ (ii)“: Bilde $A = Q^H DQ$ mit $D = \mathrm{diag}(d_1, \dots, d_n)$ und $d_i \neq 0$, reell, $i = 1, \dots, n$, so folgt $B = A^2 = Q^H D^2 Q$, und damit hat B nur positive Eigenwerte. □

Um die Cholesky-Zerlegung einer Hermite'schen, positiv definiten Matrix $A \in \mathbb{C}^{n,n}$ zu berechnen, betrachten wir die Gleichung

$$A = LL^H = \begin{bmatrix} l_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} \overline{l_{11}} & \cdots & \overline{l_{n1}} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overline{l_{nn}} \end{bmatrix}, \quad L = [l_{ij}], \quad l_{ij} = 0 \;\; \text{für } j > i.$$

So gilt (für Zeile 1 von A)

$$\begin{aligned} a_{11} &= l_{11}\overline{l_{11}} \implies l_{11} = \sqrt{a_{11}} = \overline{l_{11}}, \qquad \text{(wir können } l_{11} \text{ so wählen)} \\ a_{1j} &= l_{11}\overline{l_{j1}} \implies l_{j1} = \frac{\overline{a_{1j}}}{l_{11}}, \qquad j = 2, \dots, n. \end{aligned} \tag{14.3}$$

Analog (für Zeile i)

$$\begin{aligned} a_{ii} &= \sum_{j=1}^{i} l_{ij}\overline{l_{ij}} \implies l_{ii} = \sqrt{a_{ii} - \sum_{j=1}^{i-1} |l_{ij}|^2} = \overline{l_{ii}}, \quad \text{(wir können } l_{ii} \text{ so wählen)} \\ a_{ij} &= \sum_{k=1}^{n} l_{ik}\overline{l_{jk}} = \sum_{k=1}^{i} l_{ik}\overline{l_{jk}} = \sum_{k=1}^{i-1} l_{ik}\overline{l_{jk}} + l_{ii}\overline{l_{ji}} \\ \implies & \quad l_{ji} = \frac{1}{l_{ii}}\overline{\left(a_{ij} - \sum_{k=1}^{i-1} l_{ik}\overline{l_{jk}}\right)} \quad \text{für } \; j > i. \end{aligned} \tag{14.4}$$

Beispiel 14.5

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

$$l_{11} = \sqrt{a_{11}} = \sqrt{2}, \qquad l_{j1} = \frac{\overline{a_{1j}}}{l_{11}} \implies l_{21} = \frac{\overline{a_{12}}}{l_{11}} = \frac{1}{\sqrt{2}}, \quad l_{31} = 0,$$

$$l_{22} = \sqrt{a_{22} - \sum_{j=1}^{1} |l_{2j}|^2} = \sqrt{a_{22} - |l_{21}|^2} = \sqrt{2 - \frac{1}{2}} = \sqrt{\frac{3}{2}},$$

$$l_{32} = \frac{1}{l_{22}}\left(a_{23} - \sum_{k=1}^{1} l_{2k}\overline{l_{3k}}\right) = \frac{a_{23} - l_{21}\overline{l_{31}}}{l_{22}} = \sqrt{\frac{2}{3}}(1-0) = \sqrt{\frac{2}{3}},$$

$$l_{33} = \sqrt{a_{33} - |l_{31}|^2 - |l_{32}|^2} = \sqrt{2 - \frac{2}{3}} = \sqrt{\frac{4}{3}},$$

$$L = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} & 0 \\ 0 & \sqrt{\frac{2}{3}} & \sqrt{\frac{4}{3}} \end{bmatrix}, \quad L^H = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \sqrt{\frac{3}{2}} & \sqrt{\frac{2}{3}} \\ 0 & 0 & \sqrt{\frac{4}{3}} \end{bmatrix}.$$

Probe:

$$LL^H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Es gibt noch viele weitere wichtige Eigenschaften Hermite'scher (positiv definiter) Matrizen. Um diese wichtigen Sätze zu betrachten, ordnen wir die Eigenwerte im folgenden.

Falls $\lambda_1(A), \ldots, \lambda_n(A) \in \mathbb{R}$ die Eigenwerte von A sind (gezählt mit ihrer Multiplizität), so gelte

$$\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A).$$

Dann gilt der

Satz 14.6 (Satz von Courant–Fischer)
Sei $A \in \mathbb{C}^{n,n}$ Hermite'sch, so gilt für $k = 0, \ldots, n-1$

$$\lambda_{k+1}(A) = \min_{\substack{\dim(S)=n-k \\ S\subset\mathbb{C}^n}} \max_{0\neq x\in S} \frac{x^H A x}{x^H x}.$$

Beweis: Betrachte

$$\max x^H A x \quad \text{mit den Nebenbedingungen} \quad \|x\|_2 = 1 \text{ und } p_i^H x = 0,\ i = 1, \ldots, k,$$

wobei $p_i \neq 0$ irgendwelche linear unabhängige Vektoren sind. Nach Satz 9.6 gibt es eine unitäre Matrix Q, so dass

$$Q^H A Q = D, \qquad \text{mit } D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Wir setzen $y = Q^H x$, so folgt $x^H A x = x^H Q Q^H A Q Q^H x = y^H D y$. Setzen wir noch $q_i = Q^H p_i$, können wir also stattdessen betrachten:

$$\max \sum_{i=1}^{n} \overline{y}_i \lambda_i y_i \quad \text{mit den Nebenbedingungen} \quad y^H y = 1 \text{ und } q_i^H y = 0,\ i = 1, \ldots, k.$$

Für beliebige Vektoren p_i und damit q_i erhalten wir damit $k+1 \leq n$ Gleichungen für die Komponenten von y.
Für jede Wahl y der Form

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{k+1} \\ 0 \end{bmatrix} \in \mathbb{R}^n, \quad \text{so dass} \quad \begin{bmatrix} & q_1^H & \\ & \vdots & \\ & q_k^H & \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_k \\ y_{k+1} \\ 0 \end{bmatrix} = 0,$$

können wir y so normieren, dass $\|y\|_2 = 1$, und erhalten die zu maximierende Funktion

$$\sum_{i=1}^{k+1} |y_i|^2 \lambda_i \;\geq\; \lambda_{k+1} \sum_{i=1}^{k+1} |y_i|^2 \;=\; \lambda_{k+1}.$$

Also $\max x^H A x \geq \lambda_{k+1}$ für alle Wahlen p_i und damit auch

$$\min_{\substack{\dim(S)=n-k \\ S \subset \mathbb{C}^n}} \max_{0 \neq x \in S} \frac{x^H A x}{x^H x} \geq \lambda_{k+1}.$$

Für die spezielle Wahl $p_i = Q e_i$, $i = 1, \ldots, k$, ergibt sich $q_i = e_i$ und damit $y_1, \ldots, y_k = 0$. Also

$$x^H A x = y^H \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} y = \sum_{i=k+1}^{n} \lambda_i |y_i|^2 \;\leq\; \lambda_{k+1} \sum_{i=k+1}^{n} |y_i|^2 = \lambda_{k+1}$$

$$\Longrightarrow \max \frac{x^H A x}{x^H x} \leq \lambda_{k+1} \Longrightarrow \min \max \frac{x^H A x}{x^H x} \leq \lambda_{k+1}.$$

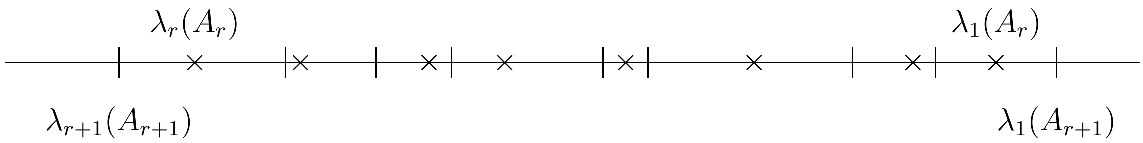
Also folgt die Behauptung. $\square$

Damit können wir die Eigenwerte über eine Optimierung bekommen. Dieser Satz hat viele wichtige Folgerungen. Wohl die wichtigste ist der Cauchy'sche Trennungssatz.

Korollar 14.7 *Sei A_r die führende $r \times r$ - Hauptabschnittsmatrix von $A = [a_{ij}] \in \mathbb{C}^{n,n}$,*

$A = A^H$*, d.h.* $A_r = \begin{bmatrix} a_{11} \cdots a_{1r} \\ \vdots \quad\quad \vdots \\ a_{r1} \cdots a_{rr} \end{bmatrix}$*, so gilt für* $r = 1, \ldots, n-1$

$$\lambda_{r+1}(A_{r+1}) \leq \lambda_r(A_r) \leq \lambda_r(A_{r+1}) \leq \ldots \leq \lambda_2(A_{r+1}) \leq \lambda_1(A_r) \leq \lambda_1(A_{r+1}).$$



Beweis: Mit Satz 14.6 gilt für $k = 0, \ldots, r-1$

$$\lambda_{k+1}(A_r) = \min_{\substack{\dim(S)=n-k \\ S \subset \mathbb{C}^r}} \max_{0 \neq x \in S} \frac{x^H A_r x}{x^H x} = \min_{\substack{\dim(S)=n-k \\ S \subset \mathbb{C}^r}} \max_{0 \neq x \in S} \frac{\begin{bmatrix} x \\ 0 \end{bmatrix}^H A_{r+1} \begin{bmatrix} x \\ 0 \end{bmatrix}}{\begin{bmatrix} x \\ 0 \end{bmatrix}^H \begin{bmatrix} x \\ 0 \end{bmatrix}}$$

$$\leq \min_{\substack{\dim(S)=n-k \\ S \subset \mathbb{C}^{r+1}}} \max_{0 \neq x \in S} \frac{x^H A_{r+1} x}{x^H x} = \lambda_{k+1}(A_{r+1})$$

Andererseits gilt: Die Menge der Werte $x^H A_r x,\ x \in \mathbb{C}^r$, ist gleich der Menge der Werte $x^H A_{r+1} x,\ x \in \mathbb{C}^{r+1},\ x_{r+1} = 0$. Also

$$\lambda_{k+1}(A_r) = \min_{\substack{S \subset \mathbb{C}^{r+1} \\ \dim(S)=n-k}} \max_{\substack{0 \neq x \in S \\ x_{r+1}=0}} \frac{x^H A_{r+1} x}{x^H x}.$$

Es gibt eine Menge von Vektoren $p_1, \ldots, p_k$, für die dieses Minimum angenommen wird, d.h. $\lambda_{k+1}(A_r)$ ist das Maximum von $x^H A_{r+1} x$, wobei $x^H x = 1,\ x_{r+1} = 0$, $p_i^H x = 0$, $i = 1, \ldots, k$.

Aber $\lambda_{k+2}(A_{r+1})$ ist das Minimum über alle Maxima für jede Wahl von x_{r+1} und p_i, $i = 1, \ldots, k$,

$$\Longrightarrow \quad \lambda_{k+2}(A_{r+1}) \leq \lambda_{k+1}(A_r), \quad k = 0, \ldots, r-1.$$

$\square$

Ähnliche Sätze spielen eine große Rolle bei der Störungstheorie von Matrizen (Operatoren), die bei der Modellbildung und in vielen Bereichen der Numerischen Mathematik auftreten.