

Lineare Algebra 2

[MAT.153UF]

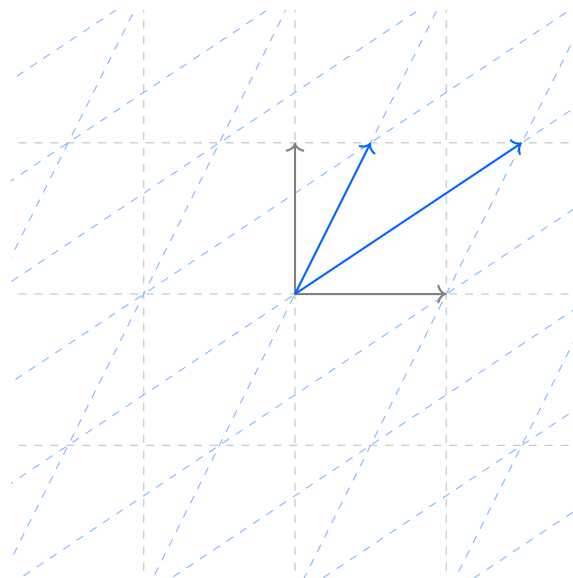
gelesen von: Franz Lehner, Assoc.Prof. Dipl.-Ing. Dr.
am Institut für Diskrete Mathematik
Technische Universität Graz

Verfasst von: Moritz Mossböck

11820925

moritz.mossboeck@student.tugraz.at

Sommersemester 2022



Inhaltsverzeichnis

7	Determinanten	4
7.1	Berechnung von Determinanten	13
7.2	Anwendungen	16
8	Innere Produkte	18
8.1	Orthogonalität	29
8.2	Orthonormalisierungsverfahren	37
8.3	Nachtrag zu Kapitel 5	44
9	Polynome und Algebra	46
9.1	Division mit Rest	49
10	Eigenwerte und Eigenvektoren	52
11	Die Jordansche Normalform	60
11.1	Anwendungen	69
12	Normale Matrizen	74
12.1	Anwendungen	77
13	Eigenwertabschätzungen	82
Anhang A:	Verzeichnisse	88
	Abbildungsverzeichnis	88
	Verzeichnis der Sätze	89
	Verzeichnis der Definitionen	90
	Verzeichnis der Lemmata	90
	Verzeichnis der Korollare	91

Ankündigungen

Die Vorlesung wird Montags (08:05-08:50) auf der KFU und Mittwochs (08:15-09:00, 09:10-09:55) auf der TU abgehalten. Das Konversatorium wird Montags von 10:00 bis 11:00 im HS 12.11 (KFU) abgehalten. Weiters gibt es ein Tutorium (siehe Website).

Wiederholung zu Linearen Abbildungen

Eine lineare Abbildung $f: V \rightarrow W$ ist ein Element des Vektorraums $\text{Hom}(V, W)$.

1. $\text{Hom}(\mathbb{K}^n, \mathbb{K}^m) \simeq \mathbb{K}^{m \times n}$, sprich jede lineare Abbildung $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ hat die Form $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, wobei $\mathbf{A} \in \mathbb{K}^{m \times n}$
2. $\dim(V) = n \Rightarrow V \simeq \mathbb{K}^n$

$$V \xrightarrow{f} W \Rightarrow \mathbb{K}^n \xrightarrow{\mathbf{A}} \mathbb{K}^m$$

Sei $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V , dann definieren wir Φ_B die Koordinatenabbildung in den \mathbb{K}^n :

$$\mathbf{x} = \sum_{k=1}^n \mathbf{b}_k x_k \mapsto \sum_{k=1}^n \mathbf{e}_k x_k$$

Sei analog Φ_C die Koordinatenabbildung $\Phi_C: W \rightarrow \mathbb{K}^m$, wir suchen für $f \in \text{Hom}(V, W)$ eine Matrix \mathbf{A} , sodass:

$$f(\mathbf{x}) = (\Phi_C^{-1} \circ f_{\mathbf{A}} \circ \Phi_B)(\mathbf{x})$$

Dabei ist Φ_C^{-1} gegeben durch:

$$\Phi_C^{-1}: \mathbb{K}^m \rightarrow W \quad \sum_{k=1}^m \mathbf{e}_k x_k \mapsto \sum_{k=1}^m \mathbf{c}_k x_k$$

Die Matrix, die der linearen Abbildung: $(\Phi_C \circ f \circ \Phi_B^{-1}): \mathbb{K}^n \rightarrow \mathbb{K}^m$ entspricht, heißt die Matrixdarstellung der linearen Abbildung f bezüglich der Basen B und C . Wir bezeichnen diese Matrix $\mathbf{A} = \Phi_C^B(f)$.

Wie finden wir diese Matrix? Die Idee ist: Für $f_{\mathbf{A}} \in \text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ ist $\mathbf{s}_i(\mathbf{A}) = f_{\mathbf{A}}(\mathbf{e}_i) = f_{\mathbf{A}}(\Phi_B(\mathbf{b}_i))$.

Lemma 6.1

$$\forall \mathbf{v} \in V: \Phi_C(f(\mathbf{v})) = \Phi_C^B(f)\Phi_B(\mathbf{v})$$

Beweis.

$$\Phi_C(f(\mathbf{v})) = \Phi_C \circ f(\mathbf{v}) = f_{\mathbf{A}} \circ \Phi_B(\mathbf{v}) = \mathbf{A}\Phi_B(\mathbf{v})$$

□

Satz 6.1

Die Spaltenvektoren der Matrixdarstellung $\Phi_C^B(f)$ sind die Koordinatenvektoren der Bilder der Basisvektoren bezüglich der Basis C .

$$\mathbf{A} = \Phi_C^B(f) = [\Phi_C(f(\mathbf{b}_1)) \quad \Phi_C(f(\mathbf{b}_2)) \quad \dots \quad \Phi_C(f(\mathbf{b}_n))] \in \mathbb{K}^{m \times n}$$

Satz 6.2

Die Abbildung:

$$\Phi_C^B: \text{Hom}(V, W) \rightarrow \mathbb{K}^{m \times n}$$

ist ein Isomorphismus, sprich:

$$\Phi_C^B(\lambda f + \mu g) = \lambda \Phi_C^B(f) + \mu \Phi_C^B(g)$$

und Φ_C^B ist bijektiv.

Beweis. Übung

□

Satz 6.3

Seien V , W und Z Vektorräume:

$$V \xrightarrow{f} W \xrightarrow{g} Z$$

Seien $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V , $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ eine Basis von W und $D = (\mathbf{d}_1, \dots, \mathbf{d}_p)$ eine Basis von Z :

$$\Phi_D^B(g \circ f) = \Phi_D^C(g) \Phi_C^B(f)$$

Beweis. Zu zeigen ist:

$$\Phi_D(g \circ f(\mathbf{v})) = \Phi_D^B(g \circ f) \Phi_B(\mathbf{v}) = \Phi_D^C(g) \Phi_C^B(f) \Phi_B(\mathbf{v})$$

$$\Phi_D((g \circ f)(\mathbf{v})) = \Phi_D(g(f(\mathbf{v}))) = \Phi_D^C(g) \Phi_C(f(\mathbf{v})) = \Phi_D^C(g) \Phi_C^B(f) \Phi_B(\mathbf{v})$$

□

Was passiert bei einem Basiswechsel?

Abbildung 1: Kommutatives Diagramm zur Matrixdarstellung eines Basiswechsels

$$\Phi_{B'}(\mathbf{v}) = \Phi_{B'}^B(\text{id}) \Phi_B(\mathbf{v})$$

Satz 6.4

Sei $f \in \text{Hom}(V, W)$ und $B \subseteq V, C \subseteq W$ Basen:

- i $\text{rg}(\Phi_C^B(f)) = \dim(\text{im}(f))$
- ii $V \simeq W \Leftrightarrow \exists \Phi_C^B(f)^{-1}$ und $\Phi_B^C(f^{-1}) = (\Phi_C^B(f))^{-1}$

Beweis. Zu i):

$$\begin{aligned} \text{im}(f) &= L(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)) \subseteq W \\ &\stackrel{\Phi_C}{\simeq} L(\Phi_C(f(\mathbf{b}_1)), \dots, \Phi_C(f(\mathbf{b}_n))) \end{aligned}$$

Zu ii): Wenn f ein Isomorphismus ist, dann $\dim(V) = \dim(W) = n$ und $\text{im}(f) = W$, womit $\dim(\text{im}(f)) = n$:

$$\Rightarrow \text{rg}(\Phi_C^B(f)) = n$$

Dann ist $\Phi_C^B(f)$ regulär. Wir zeigen $\Phi_B^C(f^{-1})\Phi_C^B(f) = \mathbf{I}$ und $\Phi_C^B(f)\Phi_B^C(f^{-1}) = \mathbf{I}$:

$$\Phi_B^C(f^{-1})\Phi_C^B(f) = \Phi_B^B(f^{-1} \circ f) = \Phi_B^B(\text{id}) = \mathbf{I}$$

Da $\Phi_B^B(\mathbf{b}_i) = \mathbf{e}_i$. □

Satz 6.5

Sei V ein Vektorraum mit den Basen $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ und $B' = (\mathbf{b}'_1, \dots, \mathbf{b}'_n)$:

- i $\Phi_{B'}^B(\text{id})\Phi_B^{B'} \circ \Phi_B^{-1} =: \mathbf{T}_{B'}^B$, Basistransformationsmatrix $\mathbf{s}_i(\mathbf{T}_{B'}^B) = \Phi_{B'}^B(\mathbf{b}_i)$
- ii $\mathbf{T}_{B'}^B$ ist regulär mit $(\mathbf{T}_{B'}^B)^{-1} = \mathbf{T}_B^{B'}$
- iii Seien B, B' Basen von V und C, C' Basen von W und $f \in \text{Hom}(V, W)$:

$$\Phi_{C'}^{B'} = \mathbf{T}_{C'}^C \Phi_C^B(f) (\mathbf{T}_B^{B'})^{-1}$$

Beweis. Zu i:

Zu ii:

$$\begin{aligned} \mathbf{T}_{B'}^B &= \Phi_{B'}^B(\text{id}) \\ (\mathbf{T}_{B'}^B)^{-1} &= (\Phi_{B'}^B(\text{id}))^{-1} = \Phi_B^{B'}(\text{id}^{-1}) = \Phi_B^{B'}(\text{id}) = \mathbf{T}_B^{B'} \end{aligned}$$

Zu iii:

Abbildung 2: Kommutatives Diagramm zur Bestimmung einer Matrixdarstellung mit Basiswechsel

□

Korollar 6.1

- i Zwei Matrixdarstellungen einer linearen Abbildung $f \in \text{Hom}(V, W)$ sind äquivalent
- ii Wenn $V = W$, dann sind zwei die Matrixdarstellungen $\Phi_B^B(f)$ und $\Phi_{B'}^{B'}$ sogar ähnlich
- iii Umgekehrt, wenn $\Phi_C^B(f) \sim M$, dann gibt es Basen $B' \subseteq V$ und $C' \subseteq W$, sodass $\mathbf{M} = \Phi_{C'}^{B'}$
- iv $\exists B', C': \Phi_{C'}^{B'} = I_{mn}^{(r)}$

Beweis. Zu ii:

$$\Phi_{B'}^{B'} = \mathbf{T}_{B'}^B \Phi_B^B(f) \mathbf{T}_B^{B'}$$

□

Unser Ziel wird es sein, eine Basis B zu finden, sodass $\Phi_B^B(f)$ möglichst einfache Gestalt hat¹.

¹also eine Diagonalmatrix

Kapitel VII Determinanten

Wie bereits erwähnt, wollen wir für eine gegebene lineare Abbildung A eine Basis finden, sodass die Matrixdarstellung für die lineare Abbildung möglichst einfach ist (also eine Diagonalform ist). Wir suchen also eine Basistransformation V , sodass VAV^{-1} eine Diagonalform ist, mit λ_i auf der Diagonale. Wir können nun den Rang als $|\{\lambda_i | \lambda_i \neq 0\}|$ bestimmen. Die Idee ist es, $\lambda \in \mathbb{K}$ zu finden, sodass $A - \lambda I$ nicht mehr invertierbar ist. Dazu benötigen wir ein Kriterium, ob eine Matrix invertierbar ist.

Betrachten wir den Fall in Dimension 1: $A = [a_{11}]$ ist invertierbar, wenn $a_{11} \neq 0$. Für Dimension 2 müssen wir etwas mehr Arbeiten verrichten:

$$\begin{aligned} & \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right] \\ a \neq 0 : & \xrightarrow{II - \frac{c}{a}I, \frac{1}{a}I} \left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{array} \right] \Rightarrow d - \frac{bc}{a} \neq 0 \\ c \neq 0 : & \xrightarrow{I - \frac{c}{d}II} \left[\begin{array}{cc|cc} 0 & b - \frac{ad}{c} & 1 & -\frac{a}{c} \\ 1 & \frac{d}{c} & 0 & \frac{1}{c} \end{array} \right] \Rightarrow b - \frac{ad}{c} \neq 0 \end{aligned}$$

Eine Matrix $A \in \mathbb{K}^{2 \times 2}$ ist invertierbar, wenn $ad - bc \neq 0$.

Definition 7.1: Determinante

Die Zahl

$$\det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

heißt die Determinante der Matrix.

Determinanten wurden 1683 von Leibniz und Seki Takaka erfunden, dabei war Leibniz' Ziel lineare Gleichungen zu lösen, Seki hingegen wollte nicht-lineare Gleichungen lösen. Leibniz nannten sie auch noch die Resultante. 1812 hat Cauchy ein Buch über Determinanten veröffentlicht. Erst 1850 wurde der Begriff der Matrix von Sylvester eingeführt. Zuletzt hat Cauchy 1858 das erste Buch über Matrizenmultiplikation veröffentlicht.

i Die Determinante ist **bilinear** in Zeilen und Spalten:

$$\mathbf{v}, \mathbf{w} \in \mathbb{K}^2 : \det(\mathbf{v}, \mathbf{w}) = v_1 w_2 - v_2 w_1$$

$$\det(\lambda \mathbf{v}' + \mu \mathbf{v}'', \mathbf{w}) = \lambda \det(\mathbf{v}', \mathbf{w}) + \mu \det(\mathbf{v}'', \mathbf{w})$$

$$\det(\mathbf{v}, \lambda \mathbf{w}' + \mu \mathbf{w}'') = \lambda \det(\mathbf{v}, \mathbf{w}') + \mu \det(\mathbf{v}, \mathbf{w}'')$$

$$\det(\lambda \mathbf{v}' + \mu \mathbf{v}'', \mathbf{w}) = \begin{vmatrix} \lambda a' + \mu a'' & c \\ \lambda b' + \mu b'' & d \end{vmatrix} = (\lambda a' + \mu a'')d - \lambda(b' + \mu b'')c = \lambda(a'd - b'c) + \mu(a''d - b''c)$$

ii $\det(\mathbf{v}, \mathbf{v}) = 0$

iii $\det(I_2) = 1$

Satz 7.1: Charakterisierung der Determinante

Die Eigenschaften i-iii charakterisieren die Determinante: Sei $\varphi: \mathbb{K}^2 \times \mathbb{K}^2 \rightarrow \mathbb{K}$ eine Abbildung, sodass:

i φ ist bilinear

ii $\varphi(\mathbf{v}, \mathbf{v}) = 0$

iii $\varphi(\mathbf{e}_1, \mathbf{e}_2) = 1$

Beweis. Sei $\mathbf{v} = \begin{bmatrix} a \\ c \end{bmatrix} = ae_1 + ce_2$ und $\mathbf{w} = \begin{bmatrix} b \\ d \end{bmatrix} = be_1 + de_2$:

$$\begin{aligned} \varphi(\mathbf{v}, \mathbf{w}) &= \varphi(ae_1 + ce_2, be_1 + de_2) \\ &= a\varphi(e_1, be_1 + de_2) + c\varphi(e_2, be_1 + de_2) \\ &= ab\varphi(e_1, e_1) + ad\varphi(e_1, e_2) + cb\varphi(e_2, e_1) + cd\varphi(e_2, e_2) \\ &= ad\underbrace{\varphi(e_1, e_2)}_{=1} + bc\varphi(e_2, e_1) \end{aligned}$$

Wir brauchen noch $\varphi(e_2, e_1)$:

$$\begin{aligned} 0 &= \varphi(e_1 + e_2, e_1 + e_2)\varphi(e_1, e_1) + \varphi(e_1, e_2) + \varphi(e_2, e_1) + \varphi(e_2, e_2) = \varphi(e_1, e_2) + \varphi(e_2, e_1) \\ &\Rightarrow \varphi(e_2, e_1) = -\varphi(e_1, e_2) = -1 \\ &\Rightarrow \varphi(\mathbf{v}, \mathbf{w}) = ad - bc \end{aligned}$$

□

Eine Folgerung aus diesem Beweis ist $\varphi(\mathbf{v}, \mathbf{w}) = -\varphi(\mathbf{w}, \mathbf{v})$. Die Determinante kann auch geometrisch gedeutet werden. $\det(\mathbf{v}, \mathbf{w})$ ist die von \mathbf{v} und \mathbf{w} aufgespannten Parallelogramms. Dabei beschreibt das Vorzeichen die Orientierung der Fläche.

Determinante als Fläche des aufgespannten Parallelogramms. Wir zeigen die Eigenschaften i-iii für die Fläche des Parallelogramms. Mit $\mathbf{e}_1, \mathbf{e}_2$, spannen wir ein Quadrat mit Seitenlänge 1 auf, womit $A(\mathbf{e}_1, \mathbf{e}_2) = 1$. Da (\mathbf{v}, \mathbf{v}) nur eine Linie ist, gilt $A(\mathbf{v}, \mathbf{v}) = 0$. Es verbleibt also die Bilinearität zu zeigen. Dabei genügt es die Linearität in \mathbf{v} zu zeigen:

1. $\mathbf{v} = \lambda\mathbf{w} \Rightarrow A(\mathbf{v}, \mathbf{w}) = 0$
2. $A(n\mathbf{v}, \mathbf{w}) = nA(\mathbf{v}, \mathbf{w}) \Rightarrow A(\frac{1}{n}\mathbf{v}, \mathbf{w}) = \frac{1}{n}A(\mathbf{v}, \mathbf{w})$ mit $n \in \mathbb{Z}$, somit kann $A(\frac{m}{n}\mathbf{v}, \mathbf{w}) = \frac{m}{n}A(\mathbf{v}, \mathbf{w})$, durch Approximation folgt, dass auch $A(\lambda\mathbf{v}, \mathbf{w}) = \lambda A(\mathbf{v}, \mathbf{w})$ für $\lambda \in \mathbb{R}$

Wir müssen noch die Additivität zeigen: $A(\mathbf{v}' + \mathbf{v}'', \mathbf{w}) = A(\mathbf{v}', \mathbf{w}) + A(\mathbf{v}'', \mathbf{w})$. Betrachten wir den Spezialfall $A(\mathbf{v} + \mathbf{w}, \mathbf{w}) = A(\mathbf{v}, \mathbf{w})$.

Betrachten wir weiterhin $A(\lambda\mathbf{v} + \mu\mathbf{w}, \mathbf{w}) = \lambda A(\mathbf{v}, \mathbf{w})$. $A(\lambda\mathbf{w} + \mu\mathbf{w}, \mathbf{w}) = \frac{1}{\mu}A(\lambda\mathbf{v} + \mu\mathbf{w}, \mu\mathbf{w}) = \frac{1}{\mu}A(\lambda\mathbf{v}, \mu\mathbf{w}) = A(\lambda\mathbf{v}, \mathbf{w}) = \lambda A(\mathbf{v}, \mathbf{w})$. Sind \mathbf{v}, \mathbf{w} linear unabhängig, dann bilden sie eine Basis des \mathbb{R}^2 , somit $\mathbf{v}' = \lambda_1\mathbf{v} + \mu_1\mathbf{w}$ und $\mathbf{v}'' = \lambda_2\mathbf{v} + \mu_2\mathbf{w}$:

$$\begin{aligned} A(\mathbf{v}' + \mathbf{v}'', \mathbf{w}) &= A(\lambda_1\mathbf{v} + \mu_1\mathbf{w} + \lambda_2\mathbf{v} + \mu_2\mathbf{w}, \mathbf{w}) \\ &= A((\lambda_1 + \lambda_2)\mathbf{v} + (\mu_1 + \mu_2)\mathbf{w}, \mathbf{w}) = \lambda_1 A(\mathbf{v}, \mathbf{w}) + \lambda_2 A(\mathbf{v}, \mathbf{w}) \end{aligned}$$

□

Definition 7.2: Determinanten auf endlichdimensionalen Vektorräumen

Sei V ein n -dimensionaler Vektorraum, dann ist eine Determinantenform eine Abbildung $\Delta : V^n \rightarrow \mathbb{K}$ mit den Eigenschaften:

- i Δ ist linear in jedem Argument (multilinear)
- ii $\Delta(a_1, \dots, a_n) = 0 \Leftrightarrow \exists k \neq l: a_k = a_l$
- Δ heißt nicht-trivial, wenn $\Delta \not\equiv 0$

Satz 7.2: Weitere Charakterisierungen der Determinantenform

Sei V ein Vektorraum mit $\dim(V) = n$ und $\Delta : V^n \rightarrow \mathbb{K}$ eine Determinantenform, dann gilt:

- iii $\Delta(a_1, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_n) = \Delta(a_1, \dots, a_n)$
- iv $\Delta(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -\Delta(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$ für $i \neq j$

Beweis. iii folgt direkt aus der Multilinearität:

$$\begin{aligned} &\Delta(a_1, \dots, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_n) \\ &= \Delta(a_1, \dots, a_{i-1}, a_i, \dots, a_n) + \lambda \Delta(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n) = \Delta(a_1, \dots, a_n) \end{aligned}$$

Da in $\lambda \Delta(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n)$ für $j \neq i$ a_j zweifach vorkommt ist diese Determinante 0.

Zu iv. Wir führen gezielt ein doppeltes Argument $a_i + a_j$ ein und nutzen wieder die Multilinearität:

$$\begin{aligned} 0 &= \Delta(a_1, \dots, a_i + a_j, \dots, a_i + a_j, \dots, a_n) \\ &= \Delta(a_1, \dots, a_i, \dots, a_i, \dots, a_n) + \Delta(a_1, \dots, a_j, \dots, a_i, \dots, a_n) \\ &\quad + \Delta(a_1, \dots, a_j, \dots, a_j, \dots, a_n) + \Delta(a_1, \dots, a_i, \dots, a_j, \dots, a_n) \\ &= \Delta(a_1, \dots, a_j, \dots, a_i, \dots, a_n) + \Delta(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = 0 \\ &\Leftrightarrow \Delta(a_1, \dots, a_j, \dots, a_i, \dots, a_n) = -\Delta(a_1, \dots, a_i, \dots, a_j, \dots, a_n) \end{aligned}$$

□

Definition 7.3: Permutationen n -ter Ordnung

Eine Permutation der Ordnung n ist eine bijektive Abbildung $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Weiters nennen wir \mathfrak{S}_n die Menge aller Permutationen der Ordnung n .

Satz 7.3: Permutationen als Gruppe

(\mathfrak{S}_n, \circ) ist eine Gruppe.

Beweis. $\mathfrak{S}_n \subseteq [n]^{[n]} = \{\varphi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}\}$, dabei ist $[n]^{[n]}$ ein Monoid mit neutralem Element id . Da \mathfrak{S}_n nur bijektive Funktionen σ enthält, ist jedes Element in \mathfrak{S}_n invertierbar, somit ist (\mathfrak{S}_n, \circ) also eine Gruppe. □

Wir nennen \mathfrak{S}_n die symmetrische Gruppe der Ordnung n . Im allgemeinen ist \mathfrak{S}_n für $n \geq 3$ nicht kommutativ.

Satz 7.4: Mächtigkeit der symmetrischen Gruppe

$$|\mathfrak{S}_n| = n!$$

Definition 7.4: Transposition

Eine Transposition (oder Vertauschung) ist eine Permutation der folgenden Form für $i \neq j$:

$$\tau_{ij}: \begin{cases} i \mapsto j \\ j \mapsto i \\ k \mapsto k \quad k \notin \{i, j\} \end{cases}$$

Wir sehen, dass Transpositionen zu sich selbst invers sind, so wie Zeilenvertauschungen zu sich selbst invers sind, also $\tau_{ij} \circ \tau_{ij} = \text{id}$.

Satz 7.5: Faktorisierung von Permutationen

Jede Permutation $\pi \in \mathfrak{S}_n$ lässt sich als Produkt von k Transpositionen schreiben:

$$\pi = \tau_1 \circ \tau_2 \circ \dots \circ \tau_k$$

Diese "Faktorisierung" ist allerdings nicht eindeutig.

Beweis.

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix}$$

Für den trivialen Fall $\pi = \text{id}$ müssen wir nichts weiter beweisen. Sei $k_1 = \min\{i | \pi(i) \neq i\}$, dann ist $\tau_1 = \tau_{k_1, \pi(k_1)}$ und $\pi_1 = \tau_1 \circ \pi$. Wenn nun $\pi_1 = \text{id}$, dann gilt $\tau_1 = \pi$, da Transposition zu sich selbst invers sind. Wenn dem

nicht so ist, wiederholen wir diese Prozedur und bestimmen $k_2 = \min\{i | \pi_1(i) \neq i\}$ und $\tau_2 = \tau_{k_2, \pi_1(k_2)}$. Durch dieses Vorgehen erhalten wir $k_1 < k_2 < \dots < k_r$, bis $\tau_r \circ \tau_{r-1} \circ \dots \circ \tau_1 \pi = \text{id}$, womit $\pi = \tau_1 \circ \tau_2 \circ \dots \circ \tau_r$. \square

Da die Faktorisierung nicht eindeutig ist, bestimmen wir eine Art "Metrik" auf \mathfrak{S}_n :

$$l(\pi) = \min\{r | \pi = \tau_1 \circ \dots \circ \tau_r\}$$

$$l(\text{id}) = 0$$

$d(\pi, \sigma) = l(\sigma^{-1}\pi)$ ist eine Metrik auf \mathfrak{S}_n .

Definition 7.5: Fehlstand

Sei $\pi \in \mathfrak{S}_n$ eine Permutation. Ein Fehlstand von π ist ein Paar (i, j) , sodass $i < j$ und $\pi(i) > \pi(j)$. Sei $F_\pi = \{(i, j) : i < j \wedge \pi(i) > \pi(j)\}$, dann ist $f_\pi = |F_\pi|$ die Anzahl der Fehlstände und wir nennen $\text{sign}(\pi) = (-1)^{f_\pi} = (-1)^\pi$ die Signatur von π .

Satz 7.6: Berechnung der Signatur

Sei $\pi \in \mathfrak{S}_n$, keine Transposition, dann gilt:

$$\text{sign}(\pi) = \prod_{i < j} \frac{\pi(j) - \pi(i)}{j - i}$$

Für Transpositionen τ ist $\text{sign}(\tau) = -1$.

Beweis. Zu i:

$$\prod_{i < j} \frac{\pi(j) - \pi(i)}{j - i} = \frac{\prod_{i < j} \pi(j) - \pi(i)}{\prod_{i < j} j - i}$$

Das Produkt läuft über alle Paare (i, j) mit $i < j$. Da π bijektiv ist, gilt $\pi_{i < j} \pi(j) - \pi(i) = \pm \prod_{i < j} (j - i)$. Betrachten wir die Signatur der einzelnen Faktoren, so liefern Fehlstände einen negativen Beitrag im Zähler. Um die Fehlstände in $\prod_{i < j} (j - i)$ zu berücksichtigen, betrachten wir $\text{sign}(\pi(j) - \pi(i))$, wobei sign in diesem Fall das Signum² ist. Jeder Fehlstand im Zähler liefert einen Faktor -1 , sprich das Vorzeichen des Zählers alterniert f_π -mal, womit $\prod_{i < j} \pi(j) - \pi(i) = (-1)^{f_\pi} \prod_{i < j} (j - i)$. Damit ergibt sich aber:

$$\text{sign}(\pi) = \frac{\prod_{i < j} \pi(j) - \pi(i)}{\prod_{i < j} j - i} = \frac{(-1)^{f_\pi} \prod_{i < j} j - i}{\prod_{i < j} j - i} = (-1)^{f_\pi}$$

Zu ii: Sei $\tau = \tau_{kl}$ mit $\tau(k) = l > k$:

$$\tau = \begin{pmatrix} 1 & 2 & \dots & k-1 & k & \dots & \tau(k) & \tau(k)+1 & \dots & n \\ 1 & 2 & \dots & k-1 & \tau(k) & \dots & k & \tau(k)+1 & \dots & n \end{pmatrix}$$

Wir suchen nun die Anzahl der Fehlstände von τ :

$$F_\tau = \{(k, k+1), \dots, (k, \tau(k)-1), (k, \tau(k)), (k+1, \tau(k)), (k+2, \tau(k)), \dots, (\tau(k)-1, \tau(k))\}$$

Wir haben $\tau(k) - k$ Fehlstände der Form (k, j) und $\tau(k) - 1 - k$ Fehlstände der Form $(j, \tau(k))$, insgesamt also:

$$\tau(k) - k + \tau(k) - 1 - k = 2(\tau(k) - k) - 1 \equiv 1 \pmod{2}$$

Somit gilt $(-1)^{f_\tau} = -1$. \square

²lat. Signum = Vorzeichen mit $\text{sign}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$

Das Bestimmen der Signatur ist also mit erheblichen Aufwand verbunden, wenn wir lediglich die Definition verwenden. Um uns die Arbeit zu vereinfachen, wollen wir zeigen, dass wir die Signatur einer Verknüpfung als Produkt der einzelnen Signaturen darstellen können:

Satz 7.7: Signatur von Verknüpfungen

- i $\text{sign}(\text{id}) = 1$
- ii Seien $\pi, \sigma \in \mathfrak{S}_n$, dann $\text{sign}(\pi \circ \sigma) = \text{sign}(\pi) \text{sign}(\sigma)$
- iii $\text{sign}(\pi^{-1}) = \text{sign}(\pi)$

Beweis. Zu i:

$$\prod_{i < j} \frac{j-i}{j-i} = \prod_{i < j} 1 = 1$$

Zu ii:

$$\begin{aligned} \text{sign}(\pi \circ \sigma) &= \prod_{i < j} \frac{\pi \circ \sigma(j) - \pi \circ \sigma(i)}{j - i} = \prod_{i < j} \left(\frac{\pi \circ \sigma(j) - \pi \circ \sigma(i)}{\sigma(j) - \sigma(i)} \cdot \frac{\sigma(j) - \sigma(i)}{j - i} \right) \\ &= \left(\prod_{i < j} \frac{\pi \circ \sigma(j) - \pi \circ \sigma(i)}{\sigma(j) - \sigma(i)} \right) \cdot \left(\prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i} \right) \\ &= \left(\prod_{k < l} \frac{\pi(l) - \pi(k)}{l - k} \right) \cdot \text{sign}(\sigma) = \text{sign}(\pi) \cdot \text{sign}(\sigma) \end{aligned}$$

iii folgt aus ii:

$$\pi \circ \pi^{-1} = \text{id} \Rightarrow \text{sign}(\pi \circ \pi^{-1}) = 1 \Leftrightarrow \text{sign}(\pi) \cdot \text{sign}(\pi^{-1}) = 1 \Leftrightarrow \text{sign}(\pi) = \text{sign}(\pi^{-1})$$

□

Korollar 7.1: Signatur aus Faktorisierung

- i Sei $\pi \in \mathfrak{S}_n$ mit einer minimalen Faktorisierung $\pi = \tau_1 \circ \dots \circ \tau_k$, dann gilt:

$$\text{sign}(\pi) = \text{sign}(\bigcirc_{i=1}^k \tau_i) = \prod_{i=1}^k \text{sign}(\tau_i) = (-1)^k$$

- ii $\mathcal{A}_n = \ker(\text{sign}) = \{\pi \mid \text{sign}(\pi) = 1\}$ ist eine Untergruppe und heißt **alternierende Gruppe** der Ordnung n

Korollar 7.2: Permutationen in Determinantenformen

Sei $\Delta: V^n \rightarrow \mathbb{K}$ eine Determinantenform auf V mit $\dim(V) = n$, dann gilt:

$$\forall \pi \in \mathfrak{S}_n: \forall \mathbf{a}_1, \dots, \mathbf{a}_n \in V: \Delta(\mathbf{a}_{\pi(1)}, \dots, \mathbf{a}_{\pi(n)}) = \text{sign}(\pi) \Delta(\mathbf{a}_1, \dots, \mathbf{a}_n)$$

Beweis. Die Idee läuft erneut darauf hinaus, dass eine Permutation als Produkt von Transpositionen geschrieben werden kann. Sei $\pi = \tau_1 \circ \pi_1$ mit $\pi_2 = \tau_2 \circ \dots \circ \tau_r$, so folgt aus Satz 7.2:

$$\Delta(\mathbf{a}_{\tau_1 \circ \pi_1(1)}, \dots, \mathbf{a}_{\tau_1 \circ \pi_1(n)}) = -\Delta(\mathbf{a}_{\pi_1(1)}, \dots, \mathbf{a}_{\pi_1(n)})$$

Da π nur aus r verschiedenen Transpositionen besteht, können wir diese Aufteilung wiederholen und dabei immer einen Faktor -1 "herausheben", bis wir bei $(-1)^r \Delta(\mathbf{a}_1, \dots, \mathbf{a}_n)$ angelangt sind. □

Satz 7.8: Determinanten

Sei $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V und $\mathbf{a}_1, \dots, \mathbf{a}_n \in V$ Vektoren mit Koordinaten:

$$\Phi_B(\mathbf{a}_j) = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

Mit $\mathbf{A} = [a_{ij}] = [\Phi_B(\mathbf{a}_1) \ \Phi_B(\mathbf{a}_2) \ \cdots \ \Phi_B(\mathbf{a}_n)] \in \mathbb{K}^{n \times n}$, dann gilt $\Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) = \det(\mathbf{A}) \cdot \Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)$, wobei:

$$\det(A) = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k)k} = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{k\pi(k)}$$

die **Determinante** von \mathbf{A} ist.

In dieser Darstellung benötigt die Berechnung der Determinante $n!$ Summanden, wird also dementsprechend schnell groß.

Beweis. $\mathbf{a}_j = \sum_{i=1}^n a_{ij} \mathbf{b}_i$, womit:

$$\begin{aligned} \Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) &= \Delta\left(\sum_{i_1=1}^n a_{i_1 1} \mathbf{b}_{i_1}, \sum_{i_2=1}^n a_{i_2 2} \mathbf{b}_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n} \mathbf{b}_{i_n}\right) \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n \left(\prod_{k=1}^n a_{i_k k}\right) \cdot \Delta(\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \dots, \mathbf{b}_{i_n}) \end{aligned}$$

Die Summe läuft also über alle Tupel (i_1, \dots, i_n) aus verschiedenen Indices $i_j \in \{1, \dots, n\}$ weswegen jede Zahl genau einmal vorkommt, sprich (i_1, \dots, i_n) ist eine Permutation von $(1, \dots, n)$ mit $\pi(j) = i_j$:

$$= \sum_{\pi \in \mathfrak{S}_n} \prod_{k=1}^n a_{\pi(k)k} \Delta(\mathbf{b}_{\pi(1)}, \mathbf{b}_{\pi(2)}, \dots, \mathbf{b}_{\pi(n)}) = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k)k} \Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)$$

□

Wir sehen, dass eine Determinantenform durch den Wert $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)$ auf einer Basis eindeutig festgelegt wird. Wünschenswert ist nun, dass Δ nicht trivial ist, also für eine Basis gilt $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) \neq 0$. Damit gilt aber für jede Basis $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) \neq 0$, denn sobald $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) = 0$ für eine beliebige Basis gilt, folgt $\Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$ für alle n -Tupel $\mathbf{a}_1, \dots, \mathbf{a}_n$.

Satz 7.9: Eindeutige Determinantenform als Determinante

Sei $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V und $c \in \mathbb{K}$, dann wird durch:

$$\Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) = c \cdot \det[\Phi_B(\mathbf{a}_1) \ \cdots \ \Phi_B(\mathbf{a}_n)]$$

eine Determinantenform auf V^n definiert, und zwar die eindeutige Determinantenform mit:

$$\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) = c$$

Beweis. Wir müssen zeigen, dass dieses Δ multilinear und alternierend ist.

1. Zur Multilinearität: Wir müssen zeigen:

1a $\Delta(\mathbf{a}_1, \dots, \lambda \mathbf{a}_k, \dots, \mathbf{a}_n) = \lambda \Delta(\mathbf{a}_1, \dots, \mathbf{a}_k, \dots, \mathbf{a}_n)$

1b $\Delta(\mathbf{a}_1, \dots, \mathbf{a}'_k + \mathbf{a}''_k, \dots, \mathbf{a}_n) = \Delta(\mathbf{a}_1, \dots, \mathbf{a}'_k, \dots, \mathbf{a}_n) + \Delta(\mathbf{a}_1, \dots, \mathbf{a}''_k, \dots, \mathbf{a}_n)$

Zu 1a:

$$\begin{aligned} \Delta(\mathbf{a}_1, \dots, \lambda \mathbf{a}_k, \dots, \mathbf{a}_n) &= c \cdot \text{det} [\Phi_B(\mathbf{a}_1), \dots, \lambda \Phi_B(\mathbf{a}_k), \dots, \Phi_B(\mathbf{a}_n)] \\ &= c \cdot \text{det} \begin{bmatrix} a_{11} & \cdots & \lambda a_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & \lambda a_{2k} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \lambda a_{nk} & \cdots & a_{nn} \end{bmatrix} = \lambda c \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k)k} = \lambda \Delta(\mathbf{a}_1, \dots, \mathbf{a}_k, \dots, \mathbf{a}_n) \end{aligned}$$

Zu 1b:

$$\begin{aligned} \Delta(\mathbf{a}_1, \dots, \mathbf{a}'_k + \mathbf{a}''_k, \dots, \mathbf{a}_n) &= c \cdot \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) a_{\pi(1)1} \cdot a_{\pi(2)2} \cdot \dots \cdot (a'_{\pi(k)k} + a''_{\pi(k)k}) \cdot \dots \cdot a_{\pi(n)n} \\ &= c \cdot \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \left(a_{\pi(1)1} \cdot a_{\pi(2)2} \cdot \dots \cdot a'_{\pi(k)k} \cdot \dots \cdot a_{\pi(n)n} + a_{\pi(1)1} \cdot a_{\pi(2)2} \cdot \dots \cdot a''_{\pi(k)k} \cdot \dots \cdot a_{\pi(n)n} \right) \\ &= \Delta(\mathbf{a}_1, \dots, \mathbf{a}'_k, \dots, \mathbf{a}_n) + \Delta(\mathbf{a}_1, \dots, \mathbf{a}''_k, \dots, \mathbf{a}_n) \end{aligned}$$

Wir prüfen noch, ob Δ alternierend ist, es muss also gelten $a_k = a_l$ für $k \neq l$, dass $\Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$. Sie oBdA $k < l$ und $\tau = \tau_{kl}$ eine Transposition:

$$\mathfrak{S}_n = \mathcal{A}_n \dot{\cup} (\mathfrak{S}_n \setminus \mathcal{A}_n) = \mathcal{A}_n \dot{\cup} \mathcal{A}_n \cdot \tau$$

Wir behaupten $\mathfrak{S}_n \setminus \mathcal{A}_n = \mathcal{A}_n \cdot \tau = \{\pi \circ \tau \mid \pi \in \mathcal{A}_n\}$, denn $\text{sign}(\pi) = -1 \Rightarrow \text{sign}(\pi \circ \tau) \text{sign}(\pi) \cdot \text{sign}(\tau) = 1$, aber $\tau \circ \tau = \text{id}$, womit $\pi = \pi \circ \tau \circ \tau = (\pi \circ \tau) \circ \tau$ und $\pi \circ \tau \in \mathcal{A}_n$, sprich wenn $\text{sign}(\pi) = -1$, dann folgt $\pi = \pi' \circ \tau$ mit $\pi' = \pi \circ \tau \in \mathcal{A}_n$.

Sei $a_k = a_l$:

$$\Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) = c \cdot \sum_{\pi \in \mathcal{A}_n \dot{\cup} \mathcal{A}_n \cdot \tau} \text{sign}(\pi) \cdot \prod_{j=1}^n a_{\pi(j)j} = c \cdot \left(\sum_{\pi \in \mathcal{A}_n} \prod_{j=1}^n a_{\pi(j)j} - \sum_{\pi \in \mathcal{A}_n} \prod_{j=1}^n a_{\pi \circ \tau(j)j} \right) = 0$$

Da $\mathbf{a}_k = \mathbf{a}_l$ folgt $a_{\pi(l)k} = a_{\pi(l)l}$ bzw. $a_{\pi(k)l} = a_{\pi(k)k}$. Damit ist Δ eine Determinantenform.

Weiterhin soll $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) = c \cdot \text{det} \mathbf{I} = c$ gelten:

$$\text{det} \mathbf{I} = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n \delta_{\pi(k)k}$$

Damit $\delta_{\pi(k)k} \neq 0$ muss $\pi = \text{id}$ gelten, womit $\text{sign}(\pi) = 1$. □

Satz 7.10: Lineare Unabhängigkeit aus Determinante

Sei $\Delta: V^n \rightarrow \mathbb{K}$ eine nicht-triviale Determinantenform, dann sind $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ linear unabhängig, wenn:

$$\Delta(\mathbf{v}_1, \dots, \mathbf{v}_n) = 0$$

Beweis. \Rightarrow Wenn $\mathbf{v}_1, \dots, \mathbf{v}_n$ linear unabhängig sind, dann bilden sie eine Basis, womit nach Satz 7.9 folgt $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) \neq 0$.

\Leftarrow Angenommen $\mathbf{v}_1, \dots, \mathbf{v}_n$ sind linear abhängig, dann folgt, dass eine Linearkombination $\sum_{k=1}^n \lambda_k \mathbf{v}_k = \mathbf{0}$ existiert. Sei oBdA $\lambda_k \neq 0$ und $\lambda_n = 01$, dann folgt:

$$\mathbf{v}_n = - \sum_{i=1}^{n-1} \lambda_i \mathbf{v}_i$$

Somit:

$$\begin{aligned}\Delta(\mathbf{v}_1, \dots, \mathbf{v}_n) &= \Delta\left(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, -\sum_{i=1}^{n-1} \lambda_i \mathbf{v}_i\right) \\ &= -\sum_{i=1}^{n-1} \lambda_i \Delta(\mathbf{v}_1, \dots, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}, \mathbf{v}_i) = 0\end{aligned}$$

□

Fassen wir zusammen:

- die Menge aller Determinantenformen $\Delta: V^n \rightarrow \mathbb{K}$ bildet einen 1-dimensionalen Vektorraum
- zu jeder Basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ gibt es eine eindeutige nicht-triviale Determinantenform mit $\Delta_1(\mathbf{b}_1, \dots, \mathbf{b}_n) = 1$

Damit ist eine beliebige Determinantenform Δ eindeutig durch $\Delta = c \cdot \Delta_1$ festgelegt (somit ist Δ_1 unsere Basis).

Lemma 7.1: Determinantenformen mit linearen Abbildungen

Seien V, W Vektorräume mit $\dim V = \dim W = n$, $\Delta W^n \rightarrow \mathbb{K}$ eine Determinantenform und $f \in \text{Hom}(V, W)$, dann ist $\Delta^f: V^n \rightarrow \mathbb{K}$ mit $\Delta^f(\mathbf{v}_1, \dots, \mathbf{v}_n) = \Delta(f(\mathbf{v}_1), \dots, f(\mathbf{v}_n))$ eine Determinantenform auf V^n .

Beweis. Wir zeigen zuerst, dass Δ^f multilinear ist:

$$\begin{aligned}\Delta^f(\mathbf{v}_1, \dots, \lambda \mathbf{v}'_k + \mu \mathbf{v}''_k, \dots, \mathbf{v}_n) &= \Delta(f(\mathbf{v}_1), \dots, \lambda f(\mathbf{v}'_k) + \mu f(\mathbf{v}''_k), \dots, f(\mathbf{v}_n)) \\ &= \lambda \Delta(f(\mathbf{v}_1), \dots, f(\mathbf{v}'_k), \dots, f(\mathbf{v}_n)) + \mu \Delta(f(\mathbf{v}_1), \dots, f(\mathbf{v}''_k), \dots, f(\mathbf{v}_n)) \\ &= \lambda \Delta^f(\mathbf{v}_1, \dots, \mathbf{v}'_k, \dots, \mathbf{v}_n) + \mu \Delta^f(\mathbf{v}_1, \dots, \mathbf{v}''_k, \dots, \mathbf{v}_n)\end{aligned}$$

Für $\mathbf{v}_k = \mathbf{v}_l$ mit $l \neq k$ soll gelten $\Delta^f(\mathbf{v}_1, \dots, \mathbf{v}_n) = 0$:

$$\Delta^f(\mathbf{v}_1, \dots, \mathbf{v}_n) = \Delta(f(\mathbf{v}_1), \dots, f(\mathbf{f}_n)) = 0$$

□

Korollar 7.3: Determinante einer Linearen Abbildung

Sei $\Delta: V^n \rightarrow \mathbb{K}$ eine Determinantenform und $f \in \text{End}(V)$, dann $\exists c_f \in \mathbb{K}$, sodass $\Delta^f = c_f \cdot \Delta$. Wir nennen $\det(f) := c_f$ die Determinante von f .

Beweis. Sei $\Lambda^n(V)$ der Raum der Determinantenformen und $\Delta_1, \Delta_2 \in \Lambda_n$, dann ist $\lambda \Delta_1 + \mu \Delta_2$ ebenfalls eine Determinantenform. Sei Δ_1 diejenige Determinantenform mit $\Delta_1(\mathbf{b}_1, \dots, \mathbf{b}_n) = 1$, dann ist $\Delta = c \cdot \Delta_1$ wobei $c = \Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)$.

Sei $\Delta: V^n \rightarrow V$ eine nicht-triviale Determinantenform und $f \in \text{End}(V)$, dann folgt aus Lemma 7.1, dass $\Delta^f: V^n \rightarrow V$ eine Determinantenform ist, also $\exists c_f \in \mathbb{K}: \Delta^f = c_f \cdot \Delta$. Weil $\dim \Lambda^n(V) = 1$ ist Δ eine Basis. □

Korollar 7.4

- i Für jede Basis $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ ist $\Delta(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)) = \det(f) \cdot \Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)$. Vergleichen wir mit der Basistransformation $\Phi_B(\mathbf{b}_i) \rightsquigarrow \Delta(\mathbf{a}_1, \dots, \mathbf{a}_n) = \det(\mathbf{A}) \cdot \Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)$, dann ist $\mathbf{A} = [\Phi_B(\mathbf{a}_1) \ \dots \ \Phi_B(\mathbf{a}_n)]$ bzw. $f(\mathbf{b}_i) = \mathbf{a}_i$. Dabei gilt:

$$\det(f) = \frac{\Delta(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n))}{\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)}$$

Womit $\det(f)$ unabhängig von der gewählten Basis ist.

- ii Ist $\mathbf{A} = \Phi_B^B(f)$ die Matrixdarstellung von f bezüglich B , dann ist $\det(f) = \det(\mathbf{A})$

Satz 7.11: Bijektivität aus Determinante

Sei $f \in \text{End}(V)$ dann ist f ein Isomorphismus, wenn $\det(f) \neq 0$.

Beweis. Sei Δ eine nicht-triviale Determinantenform und $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V , dann folgt $\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n) \neq 0$. Wenn f ein Isomorphismus ist, dann ist $f(B)$ eine Basis, dann folgt aber $\Delta(f(\mathbf{b}_1), \dots, f(\mathbf{b}_n)) \neq 0$ bzw, da $\det(\mathbf{b}_1, \dots, \mathbf{b}_n) \neq 0$ folgt direkt $\det(f) \neq 0$. \square

Eine direkte Folgerung aus Satz 7.11 ist, dass eine Matrix $A \in \mathbb{K}^{n \times n}$ invertierbar ist, wenn $\det(A) \neq 0$.

Satz 7.12: Determinante von Verknüpfungen

Seien $f, g \in \text{Hom}(V, W)$, dann gilt:

$$\det(f \circ g) = \det(f) \cdot \det(g)$$

Beweis. Sei $f \circ g$ ein Isomorphismus, dann sind f und g Isomorphismen und somit invertierbar mit $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. Weiterhin ist $f \circ g$ bijektiv weswegen $\text{im } f \circ g \subseteq \text{im } f$ und $g(\mathbf{x}) = \mathbf{0} \Rightarrow f \circ g(\mathbf{x}) = \mathbf{0}$.

Wir machen nun eine Fallunterscheidung.

Fall 1: $\det(f \circ g) = 0$. Dann ist $f \circ g$ nicht bijektiv, weshalb f oder g nicht bijektiv ist, sprich $\det(f) = 0 \vee \det(g) = 0$ womit $\det(f) \det(g) = 0$

Fall 2: $\det(f \circ g) \neq 0$, dann ist $f \circ g$ bijektiv und somit auch g . Wir wählen eine Basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ von V . Sei $\Delta: V^n \rightarrow \mathbb{K}$ eine nicht-triviale Determinantenform. Da g bijektiv ist, ist $g(\mathbf{b}_1), \dots, g(\mathbf{b}_n)$ eine Basis:

$$\begin{aligned} \det(f \circ g) &= \frac{\Delta(f \circ g(\mathbf{b}_1), \dots, f \circ g(\mathbf{b}_n))}{\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)} \\ &= \frac{\Delta(f(g(\mathbf{b}_1)), \dots, f(g(\mathbf{b}_n)))}{\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)} \cdot \frac{\Delta(g(\mathbf{b}_1), \dots, g(\mathbf{b}_n))}{\Delta(g(\mathbf{b}_1), \dots, g(\mathbf{b}_n))} \\ &= \frac{\Delta(f(g(\mathbf{b}_1)), \dots, f(g(\mathbf{b}_n)))}{\Delta(g(\mathbf{b}_1), \dots, g(\mathbf{b}_n))} \cdot \frac{\Delta(g(\mathbf{b}_1), \dots, g(\mathbf{b}_n))}{\Delta(\mathbf{b}_1, \dots, \mathbf{b}_n)} \\ &= \det(f) \cdot \det(g) \end{aligned}$$

\square

Korollar 7.5

Für $n \times n$ Matrizen gilt:

- i $\det(A \cdot B) = \det(A) \cdot \det(B)$
- ii $\det(A^{-1}) = \det(A)^{-1}$
- iii $\det(A) = 0 \Leftrightarrow \text{rg } A < n$
- iv $\det(A^T) = \det A$

Beweis. Zu i: Klar, da $\Phi_B^B(f \circ g) = \Phi_B^B(f) \cdot \Phi_B^B(g)$

Zu ii: $AA^{-1} = I \Rightarrow \det(AA^{-1}) = \det(I)$ womit $\det(A)\det(A^{-1}) = 1$

Zu iii: $\det(A) = 0$, dann ist A singulär und somit nicht invertierbar, womit $\text{rg } A < n$

Zu iv:

$$\begin{aligned} \det(A^T) &= \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k), k}^T = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{k, \pi(k)} \\ &= \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi^{-1}) \prod_{k=1}^n a_{k, \pi^{-1}(k)} = \det(A) \end{aligned}$$

\square

Da $\pi \in \mathfrak{S}_n$ bijektiv ist, $\exists \pi^{-1} \in \mathfrak{S}_n$, wobei $\pi = \bigcirc_{k=1}^r \tau_k$ und $\pi^{-1} = \bigcirc_{k=0}^{r-1} \tau_{r-k}$, gilt $\text{sign}(\pi^{-1}) = \text{sign}(\pi)$.

7.1 Berechnung von Determinanten

Wir haben bereits eine Möglichkeit erörtert Determinanten zu bestimmen:

$$\det(\mathbf{A}) = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k),k}$$

Das ist praktikabel für $n \leq 3$, da wir für $n = 3$ "nur" 6 Summanden bestimmen müssen, anstelle von 24 bei $n = 4$. Wir kennen außerdem schon eine Formel für Determinanten für Matrizen $\mathbf{A} \in \mathbb{K}^{2 \times 2}$:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Für 3×3 Matrizen wollen wir uns etwas näher mit der Summendarstellung beschäftigen.

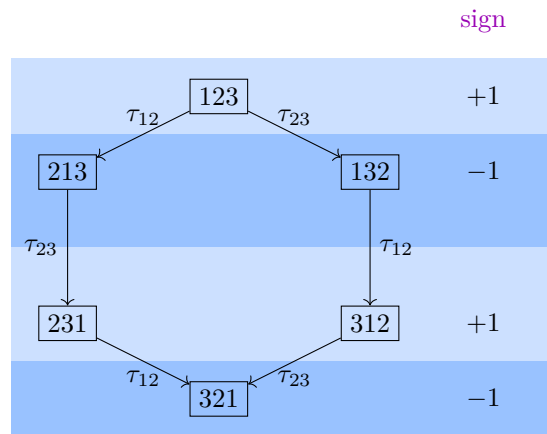


Abbildung 3: Bestimmung aller Permutationen $\pi \in \mathfrak{S}_3$ durch τ_{12} und τ_{23}

Sei $\mathbf{A} \in \mathbb{K}^{3 \times 3}$:

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13}$$

Diese Formel wird **Regel von Sarrus** genannt und gilt nur für 3×3 Matrizen. Sie kann nach dem folgenden Schema hergeleitet werden:

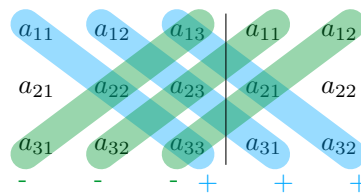


Abbildung 4: Regel von Sarrus

Für größerer Matrizen wird dieses Verfahren jedoch sehr unhandlich, da wir enorm viele Permutationen bestimmen müssen, was bei $n = 4$ machbar wäre, aber schon ab $n = 5$ mit 120 verschiedenen Permutationen in unnötig viel Rechenaufwand übergeht. Daher wollen wir größere Matrizen auf einfachere reduzieren.

Lemma 7.2: Determinanten von Dreiecksmatrizen

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$ eine Dreiecksmatrix, dann gilt:

$$\det(\mathbf{A}) = \prod_{k=1}^n a_{kk}$$

Beweis. Da bei einer (oberen) Dreiecksmatrix gilt $\forall i > j: a_{ij} = 0$, ist

$$\prod_{k=1}^n a_{\pi(k)k} = 0$$

für alle Permutationen $\exists k \in \{1, \dots, n\}: \pi(k) \geq k$, sprich nur **id** trägt zur Determinante bei. \square

Lemma 7.3: Elementare Zeilen- und Spaltenumformungen

Sei $\mathbf{A} = [a_{ij}]$ eine $n \times n$ Matrix, so gilt:

- i für $i \neq j$ gilt $\det(\begin{bmatrix} \mathbf{s}_1 & \dots & \mathbf{s}_i + \lambda \mathbf{s}_j & \dots & \mathbf{s}_n \end{bmatrix}) = \det(\mathbf{A})$
- ii für $i \neq j$ gilt $\det(\begin{bmatrix} \mathbf{z}_1 & \dots & \mathbf{z}_i + \lambda \mathbf{z}_j & \dots & \mathbf{z}_n \end{bmatrix}^T) = \det(\mathbf{A})$

Beweis. Wir haben i bereits über Determinantenformen bewiesen. Weiterhin gilt ii, da $\det(\mathbf{A}) = \det(\mathbf{A}^T)$ nach Korollar 7.5 \square

Um Determinanten zu berechnen, bringen wir Matrizen also in die Form einer oberen Dreiecksmatrix mittels Gaußscher Elimination.

Lemma 7.4: Nutzung von Nullspalten

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$, dann gilt:

$$\begin{vmatrix} a_{11} & \mathbf{a}_1 \\ \mathbf{0} & \mathbf{B} \end{vmatrix} = a_{11} \cdot \det(\mathbf{B})$$

$$\begin{vmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{a}_n & a_{nn} \end{vmatrix} = a_{nn} \cdot \det(\mathbf{B})$$

Beweis. Wir zeigen ii.

$$\det(\mathbf{A}) = \sum_{\pi \in \mathfrak{S}_n} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k)k} = \sum_{\substack{\pi \in \mathfrak{S}_n \\ \pi(n)=n}} \text{sign}(\pi) \prod_{k=1}^n a_{\pi(k)k}$$

Da wir $\pi(n) = n$ vorgeben, gilt $\pi \in \mathfrak{S}_{n-1}$, womit:

$$\det(\mathbf{A}) = \underbrace{\left(\sum_{\pi' \in \mathfrak{S}_{n-1}} \text{sign}(\pi) \prod_{k=1}^n a_{\pi'(k)k} \right)}_{=\det(\mathbf{B})} \cdot a_{nn} = \det(\mathbf{B}) \cdot a_{nn}$$

\square

Definition 7.6: Streichung von Zeilen und Spalten

Sie $\mathbf{A} \in \mathbb{K}^{n \times n}$, dann ist $\mathbf{A}_{k,l} \in \mathbb{K}^{(n-1) \times (n-1)}$ jene Matrix, die durch Streichen der k -ten Zeile und l -ten Spalte aus \mathbf{A} entsteht:

$$\mathbf{A}_{k,l} = \left[\begin{array}{ccc|ccc} a_{11} & \dots & a_{1,l-1} & a_{1,l+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k-1,1} & \dots & a_{k-1,l-1} & a_{k-1,l+1} & \dots & a_{k-1,n} \\ \hline a_{k+1,1} & \dots & a_{k+1,l-1} & a_{k+1,l+1} & \dots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,l-1} & a_{n,l+1} & \dots & a_{nn} \end{array} \right]$$

Satz 7.13: Entwicklungssatz von Laplace

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$ und $l \in \{1, \dots, n\}$ fest, dann gilt:

$$\det(\mathbf{A}) = \sum_{k=1}^n a_{kl} (-1)^{k+l} \det(\mathbf{A}_{k,l})$$

Wir nennen $(-1)^{k+l} \det(\mathbf{A}_{k,l})$ einen Kofaktor von \mathbf{A} und entwickeln die Determinante nach der l -ten Spalte.

Beweis. Die l -te Spalte von \mathbf{A} ist gegeben durch:

$$\mathbf{s}_l = \sum_{k=1}^n a_{kl} \mathbf{e}_k$$

Nach der Multilinearität der Determinante folgt daher:

$$\begin{aligned} \det(\mathbf{A}) &= \det \left(\begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_{l-1} & \sum_{k=1}^n a_{kl} \mathbf{e}_k & \mathbf{s}_{l+1} & \cdots & \mathbf{s}_n \end{bmatrix} \right) \\ &= \sum_{k=1}^n a_{kl} \det \left(\begin{bmatrix} \mathbf{s}_1 & \cdots & \mathbf{s}_{l-1} & \mathbf{e}_k & \mathbf{s}_{l+1} & \cdots & \mathbf{s}_n \end{bmatrix} \right) \\ &= \sum_{k=1}^n a_{kl} \begin{vmatrix} a_{11} & \cdots & a_{1,l-1} & 0 & a_{1,l+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{k1} & \cdots & a_{k,l-1} & 1 & a_{k,l+1} & \cdots & a_{kn} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,l-1} & 0 & a_{n,l+1} & \cdots & a_{nn} \end{vmatrix} \\ &= \sum_{k=1}^n a_{kl} (-1)^{l-1} \begin{vmatrix} 0 & a_{11} & \cdots & a_{1,l-1} & a_{1,l+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_{k1} & \cdots & a_{k,l-1} & a_{k,l+1} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \cdots & a_{n,l-1} & a_{n,l+1} & \cdots & a_{nn} \end{vmatrix} \\ &= \sum_{k=1}^n a_{kl} (-1)^{l-1+k-1} \begin{vmatrix} 1 & a_{k1} & \cdots & a_{k,l-1} & a_{k,l+1} & \cdots & a_{kn} \\ 0 & a_{11} & \cdots & a_{1,l-1} & a_{1,l+1} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{k-1,1} & \cdots & a_{k-1,l-1} & a_{k-1,l+1} & \cdots & a_{k-1,n} \\ 0 & a_{k+1,1} & \cdots & a_{k+1,l-1} & a_{k+1,l+1} & \cdots & a_{k+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n1} & \cdots & a_{n,l-1} & a_{n,l+1} & \cdots & a_{nn} \end{vmatrix} \\ &= \sum_{k=1}^n a_{kl} \begin{vmatrix} 1 & \mathbf{a}_k \\ \mathbf{0} & \mathbf{A}_{kl} \end{vmatrix} \end{aligned}$$

Nach Lemma 7.4 folgt daher:

$$\det(\mathbf{A}) = \sum_{k=1}^n a_{kl} (-1)^{k+l} \det(\mathbf{A}_{k,l})$$

□

7.2 Anwendungen

Satz 7.14: Bestimmung der Inversen

Sei $A \in \mathbb{K}^{n \times n}$ dann heißt $\hat{A} = [\hat{a}_{kl}]_{k,l=1}^n$, wobei $\hat{a}_{kl} = (-1)^{k+l} \det(A_{lk})$, die adjunkte oder Komplementärmatrix zu A . Die Zahlen \hat{a}_{kl} heißen Kofaktoren. Dann gilt:

$$\hat{A} = A^{-1} \cdot \det(A) \quad (1)$$

Mit anderen Worten:

$$(A^{-1})_{kl} = (-1)^{k+l} \det(A_{lk}) \quad (2)$$

Eine kleine Anmerkung:

$$\frac{\partial}{\partial a_{ik}} \det(A) = (-1)^{ik} \det(A_{ik})$$

Beweis. Wir zeigen $\hat{A}A = I \cdot \det(A)$. Sei $B = \hat{A}A$. Dann ist:

$$b_{kl} = \sum_{j=1}^n \hat{a}_{kj} \cdot a_{jl} = \sum_{j=1}^n (-1)^{k+j} \det(A_{jk}) \cdot a_{jl} \quad (3)$$

Fall 1: $k = l$

$$b_{kk} = \sum_{j=1}^n (-1)^{k+j} a_{jk} \det(A_{jk}) = \det(A)$$

Nach Satz 7.13

Fall 2: $k \neq l$

Sei oBdA $k < l$:

$$b_{kl} = \sum_{j=1}^n (-1)^{k+j} a_{jl} \det(A_{jk})$$

Betrachten wir nun folgende Matrix, bei der wir die k -te Spalte durch die l -te Spalte ersetzen:

$$\begin{vmatrix} a_{11} & \cdots & a_{1l} & \cdots & a_{1l} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nl} & \cdots & a_{nl} & \cdots & a_{nn} \end{vmatrix} = 0 = \sum_{j=1}^n a_{jl} (-1)^{j+l} \det(A_{jk})$$

Somit:

$$\sum_{j=1}^n (-1)^{k+j} a_{jl} \det(A_{jk}) = 0$$

Womit $\hat{A}A = I \cdot \det(A)$. □

Wir wollen nun die Inverse einer allgemeinen 2×2 Matrix bestimmen:

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

Eine alternative Schreibweise stammt von Cayley aus 1855

$$A^{-1} = \frac{1}{\nabla} \begin{bmatrix} \partial_a \nabla & \partial_c \nabla \\ \partial_b \nabla & \partial_d \nabla \end{bmatrix}$$

Wir können auch noch für 3×3 Matrizen eine explizite Formel für \mathbf{A}^{-1} bestimmen. Sei $\mathbf{A} \in \mathbb{K}^{3 \times 3}$:

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22}a_{33} - a_{32}a_{23} & a_{32}a_{13} - a_{12}a_{33} & a_{12}a_{23} - a_{22}a_{13} \\ a_{31}a_{23} - a_{21}a_{33} & a_{11}a_{33} - a_{31}a_{13} & a_{21}a_{13} - a_{11}a_{23} \\ a_{21}a_{32} - a_{31}a_{22} & a_{31}a_{12} - a_{11}a_{23} & a_{11}a_{22} - a_{21}a_{12} \end{bmatrix}$$

Sei nun ein Gleichungssystem $\mathbf{A}\mathbf{x} = \mathbf{b}$ gegeben, so können wir die Lösung explizit anschreiben, sofern \mathbf{A} regulär ist:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{\det(\mathbf{A})} \hat{\mathbf{A}}\mathbf{b}$$

Satz 7.15: Cramersche Regel

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$ eine reguläre Matrix mit Spaltenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ und $\mathbf{b} \in \mathbb{K}^n$, dann ist die eindeutige Lösung des linearen Gleichungssystems $\mathbf{A}\mathbf{x} = \mathbf{b}$ gegeben durch:

$$x_i = \frac{\Delta(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}{\det \mathbf{A}}$$

Beweis.

$$\begin{aligned} x_i &= (\mathbf{A}^{-1}\mathbf{b})_i = \left(\frac{1}{\det(\mathbf{A})} \hat{\mathbf{A}}\mathbf{b} \right)_i = \frac{1}{\det(\mathbf{A})} \sum_{j=1}^n \hat{a}_{ij} b_j \\ &= \frac{1}{\det(\mathbf{A})} \sum_{j=1}^n (-1)^{i+j} \det(\mathbf{A}_{ij}) b_j = \frac{1}{\det(\mathbf{A})} \sum_{j=1}^n \Delta(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{e}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) b_j \\ &= \frac{1}{\det(\mathbf{A})} \Delta \left(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \sum_{k=1}^n b_k \mathbf{e}_k, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n \right) \end{aligned}$$

□

Da die Determinante eine Matrix \mathbf{A} auf ein Polynom in $\mathbb{K}_{n^2}[x]$ abbildet, ist sie stetig und differenzierbar, was impliziert, dass die Menge der invertierbaren offen ist. Definieren wir etwa eine Metrik auf $\mathbb{R}^{n \times n}$:

$$\begin{aligned} d(\mathbf{A}, \mathbf{B}) &= \max_{i,j} |a_{ij} - b_{ij}| \\ d(\mathbf{A}, \mathbf{B}) &\geq 0 \quad d(\mathbf{A}, \mathbf{B}) = 0 \Leftrightarrow \mathbf{A} = \mathbf{B} \\ d(\mathbf{A}, \mathbf{B}) &\leq d(\mathbf{A}, \mathbf{C}) + d(\mathbf{C}, \mathbf{B}) \\ \mathbf{A} \neq 0 &\Rightarrow \exists \delta > 0: d(\mathbf{A}, \mathbf{B}) < \delta \Rightarrow \det(\mathbf{B}) \neq 0 \end{aligned}$$

Wenn also $\max |a_{ij} - b_{ij}| < \delta$, dann ist \mathbf{B} invertierbar, womit die Menge der invertierbaren Matrizen dicht in $\mathbb{K}^{n \times n}$ liegt. Da in dieser Menge alle Matrizen invertierbar sind, ist auch $\mathbf{A} \mapsto \frac{\hat{\mathbf{A}}}{\det(\mathbf{A})}$ stetig und differenzierbar auf $\mathbf{GL}(n, \mathbb{R})$.

Weiterhin ist die Funktion $F: \mathbf{GL}(n, \mathbb{R}) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $(\mathbf{A}, \mathbf{b}) \mapsto \mathbf{A}^{-1}\mathbf{b}$ ist stetig und differenzierbar, da

$$\mathbf{A}^{-1}\mathbf{b} = \frac{\Delta(\mathbf{a}_1, \dots, \mathbf{b}, \dots, \mathbf{a}_n)}{\det(\mathbf{A})}$$

Das hat insbesondere die Bedeutung, dass eventuell ungenaue Werte in \mathbf{b} trotzdem beschränkt bleiben.

Kapitel VIII Innere Produkte

Definition 8.1: Länge eines Vektors im \mathbb{R}^3 und \mathbb{R}^2

Sei $\mathbf{a} \in \mathbb{R}^2$, dann ist die Norm $\|\mathbf{a}\|$ gegeben durch:

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2}$$

Wenn a_1 und a_2 die Koordinaten von \mathbf{a} in der kanonischen Basis sind. Sei weiterhin $\mathbf{a} \in \mathbb{R}^3$, dann gilt:

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

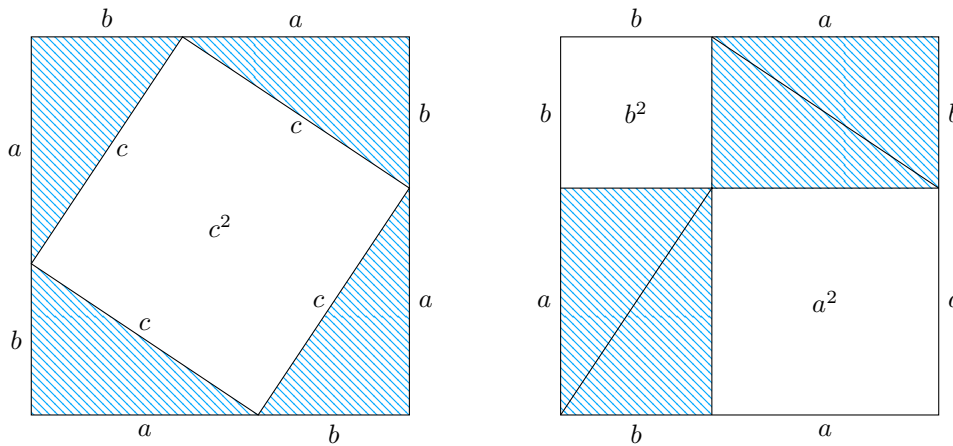


Abbildung 5: Beweisskizze für den Satz des Pythagoras

Wir sehen, dass beide Quadrate die gleiche Fläche haben und dass die blau schattierten Dreiecke ebenfalls die gleiche Fläche haben. Somit muss gelten $a^2 + b^2 = c^2$.

Ein Problem mit dieser Definition der Länge ist, dass sie nicht linear ist. Wir wollen daher \mathbf{a} auf $\|\mathbf{a}\|^2 = x_1^2 + x_2^2$ abbilden. Dazu verwenden wir das Skalarprodukt. Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ bzw. in \mathbb{R}^3 , dann ist das Skalarprodukt $\langle \mathbf{a}, \mathbf{b} \rangle$ gegeben durch:

$$\langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cdot \cos(\theta)$$

Dabei ist θ der von \mathbf{a} und \mathbf{b} aufgespannte Winkel.

Lemma 8.1: Eigenschaften des Skalarproduktes

- i $\langle \mathbf{a}, \mathbf{a} \rangle = \|\mathbf{a}\|^2$
- ii $\|\mathbf{a}\| = 0 \Leftrightarrow \mathbf{a} = \mathbf{0}$
- iii $\langle \mathbf{a}, \mathbf{b} \rangle = 0 \Leftrightarrow \mathbf{a} = \mathbf{0} \vee \mathbf{b} = \mathbf{0} \vee \theta = \pm \frac{\pi}{2}$ bzw. $\langle \mathbf{a}, \mathbf{b} \rangle > 0$ ist $|\theta| < \frac{\pi}{2}$ bzw. $\langle \mathbf{a}, \mathbf{b} \rangle < 0$ ist $\frac{\pi}{2} < |\theta| < \pi$

Wenn $\theta = \pm \frac{\pi}{2}$ schreiben wir $\mathbf{a} \perp \mathbf{b}$ und nennen \mathbf{a} orthogonal zu \mathbf{b} .

Satz 8.1: Weitere Eigenschaften des Skalarproduktes

- i $\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle$
 - ii $\langle \lambda \mathbf{a}, \mathbf{b} \rangle = \lambda \cdot \langle \mathbf{a}, \mathbf{b} \rangle$
 - iii $\langle \mathbf{a} + \mathbf{b}, \mathbf{c} \rangle = \langle \mathbf{a}, \mathbf{c} \rangle + \langle \mathbf{b}, \mathbf{c} \rangle$
- Also ist das Skalarprodukt bilinear.

Beweis. Zu ii müssen wir nur noch prüfen, ob auch negative λ das Skalarprodukt nicht verändern. Verwenden wir ein negatives λ , dann ändern wir die Orientierung von \mathbf{a} und erhalten einen neuen Winkel $\theta' = \pi - \theta$.

Aufgrund der Additionstheorem gilt allerdings $\cos(\pi - \theta) = -\cos(\theta)$, womit:

$$\langle \lambda \mathbf{a}, \mathbf{b} \rangle = \|\lambda \mathbf{a}\| \cdot \|\mathbf{b}\| \cdot \cos(\theta) = |\lambda| \cdot \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cdot \cos(\theta)$$

Wir zeigen noch iii. Sei $\|\mathbf{c}\| = 1$, womit $\langle \mathbf{a}, \mathbf{b} \rangle = \|\mathbf{a}\| \cdot \cos(\theta)$. Dabei ist $\langle \mathbf{a}, \mathbf{c} \rangle$ in diesem Fall die Länge der Projektion von \mathbf{a} auf \mathbf{c} , dann ist $\langle \mathbf{a} + \mathbf{b}, \mathbf{c} \rangle$ die Länge der Projektion von $\mathbf{a} + \mathbf{b}$ auf \mathbf{c} , die aber genau gegeben ist durch $\langle \mathbf{a}, \mathbf{c} \rangle + \langle \mathbf{b}, \mathbf{c} \rangle$. □

Satz 8.2: Skalarprodukt im \mathbb{R}^3

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, dann gilt:

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{k=1}^3 a_k b_k$$

Beweis.

$$\langle \mathbf{a}, \mathbf{b} \rangle = \langle a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3, \mathbf{b} \rangle$$

Nach der Linearität gilt:

$$\langle \mathbf{a}, \mathbf{b} \rangle = a_1 \langle \mathbf{e}_1, \mathbf{b} \rangle + a_2 \langle \mathbf{e}_2, \mathbf{b} \rangle + a_3 \langle \mathbf{e}_3, \mathbf{b} \rangle$$

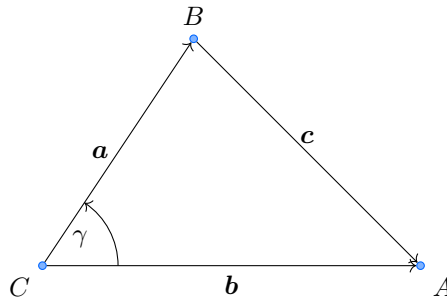
$$\langle \mathbf{e}_i, \mathbf{b} \rangle = \sum_{k=1}^3 b_k \langle \mathbf{e}_i, \mathbf{e}_k \rangle = b_i$$

$$\Rightarrow \langle \mathbf{a}, \mathbf{b} \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3$$

□

Wir sehen an Satz 8.2 schon, dass bei einer gegebenen Basis B mit Basisvektoren \mathbf{b}_i die i -te Koordinate eines Vektors $\mathbf{v} \in L(B)$ durch $\langle \mathbf{v}, \mathbf{b}_i \rangle$ bestimmbar ist.

Betrachten wir nun den Cosinussatz. Seien A, B, C Punkte im \mathbb{R}^2 , die ein Dreieck aufspannen:



Wenn $\gamma = \frac{\pi}{2}$, dann gilt der Satz von Pythagoras: $\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 = \|\mathbf{c}\|^2$. Im allgemeinen Fall gilt: $\mathbf{c} = \mathbf{b} - \mathbf{a}$, womit $\|\mathbf{c}\|^2 = \|\mathbf{b} - \mathbf{a}\|^2$ gilt:

$$\begin{aligned} \|\mathbf{c}\|^2 &= \langle \mathbf{b} - \mathbf{a}, \mathbf{b} - \mathbf{a} \rangle = \langle \mathbf{b}, \mathbf{b} \rangle - \langle \mathbf{a}, \mathbf{b} \rangle - \langle \mathbf{b}, \mathbf{a} \rangle + \langle \mathbf{a}, \mathbf{a} \rangle \\ &= \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\|\mathbf{b}\|\cos \gamma + \|\mathbf{a}\|^2 \end{aligned}$$

Wir nennen zwei Vektoren \mathbf{a} und \mathbf{b} orthogonal, wenn $\langle \mathbf{a}, \mathbf{b} \rangle = 0$, dann schreiben wir $\mathbf{a} \perp \mathbf{b}$. Gegeben ist nun \mathbf{a} . Wir suchen nun einen Vektor \mathbf{b} , sodass $\langle \mathbf{a}, \mathbf{b} \rangle = 0$. Wir suchen also $a_1 b_1 + a_2 b_2 = 0$ im \mathbb{R}^2 . Das ist ein lineares Gleichungssystem in zwei Variablen, da \mathbf{a} gegeben ist. Wir finden also eine ganze lineare Mannigfaltigkeit, deren Elemente orthogonal zu \mathbf{a} sind. Weiterhin erkennen wir, dass ein orthogonaler Vektor genau durch eine Drehung um $\frac{\pi}{2}$ aus \mathbf{a} entsteht:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -a_2 \\ a_1 \end{bmatrix}$$

Wie finden wir einen Normalvektor im \mathbb{R}^3 ? Hier ist das Problem, dass die Normalvektoren von $\mathbf{a} \in \mathbb{R}^3$ eine ganze Ebene aufspannen. Wir schränken uns also darauf ein, einen Normalvektor auf eine Ebene zu finden, die von zwei Vektoren \mathbf{a} und \mathbf{b} aufgespannt wird.

Definition 8.2: Vektorprodukt

Das Vektorprodukt (oder äußeres Produkt, Kreuzprodukt) ist nur im \mathbb{R}^3 definiert. Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, dann ist $\mathbf{a} \times \mathbf{b}$ derjenige Vektor, für den gilt:

1. $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cdot \sin(\theta)$
2. $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ und $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$
3. $(\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b})$ ist rechtsdrehend

Satz 8.3: Eigenschaften des Vektorproduktes

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, dann gilt:

- i $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
- ii $(\lambda \mathbf{a}) \times \mathbf{b} = \lambda \mathbf{a} \times \mathbf{b} = \mathbf{a} \times (\lambda \mathbf{b})$
- iii $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$

Beweis. Der Beweis für i und ii geht analog für Lemma 8.1. Hier ist zu beachten, dass sich bei ii das aufgespannte Parallelogramm invers orientiert ist. Zu iii müssen wir etwas mehr Arbeit leisten. Wenn $\mathbf{c} = \mathbf{0}$, dann ist das Vektorprodukt trivial. \square

Wir folgern:

$$\begin{aligned} \mathbf{e}_1 \times \mathbf{e}_2 &= \mathbf{e}_3 \\ \mathbf{e}_2 \times \mathbf{e}_3 &= \mathbf{e}_1 \\ \mathbf{e}_3 \times \mathbf{e}_1 &= \mathbf{e}_2 \end{aligned}$$

Die Idee hatte Hamilton in 1843 in seiner Untersuchung von Quaternionen, in der man die komplexen Zahlen um zwei weitere imaginäre Einheiten j und k erweitert mit der Multiplikationstabelle:

$$\begin{array}{lll} i^2 = -1 & ij = k & ji = -k \\ j^2 = -1 & jk = i & kj = -i \\ k^2 = -1 & ki = j & ik = -j \end{array}$$

Hamilton hatte somit die erste nicht kommutative Struktur gefunden.

Korollar 8.1: Äußeres Produkt in Koordinatenform

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ b_1 a_3 - a_1 b_3 \\ a_1 b_2 - b_1 a_2 \end{bmatrix} = \begin{vmatrix} a_1 & b_1 & \mathbf{e}_1 \\ a_2 & b_2 & \mathbf{e}_2 \\ a_3 & b_3 & \mathbf{e}_3 \end{vmatrix}$$

Bei der Determinantenform muss nach der letzten Spalte entwickelt werden.

Beweis.

$$\begin{aligned}
 & (a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3) \times (b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3) \\
 &= a_1 b_1 \mathbf{e}_1 \times \mathbf{e}_1 + a_1 b_2 \mathbf{e}_1 \times \mathbf{e}_2 + a_1 b_3 \mathbf{e}_1 \times \mathbf{e}_3 \\
 &+ a_2 b_1 \mathbf{e}_2 \times \mathbf{e}_1 + a_2 b_2 \mathbf{e}_2 \times \mathbf{e}_2 + a_2 b_3 \mathbf{e}_2 \times \mathbf{e}_3 \\
 &+ a_3 b_1 \mathbf{e}_3 \times \mathbf{e}_1 + a_3 b_2 \mathbf{e}_3 \times \mathbf{e}_2 + a_3 b_3 \mathbf{e}_3 \times \mathbf{e}_3 \\
 &= a_1 b_2 \mathbf{e}_3 - a_1 b_3 \mathbf{e}_2 - a_2 b_1 \mathbf{e}_3 + a_2 b_3 \mathbf{e}_1 + a_3 b_1 \mathbf{e}_2 - a_3 b_2 \mathbf{e}_1 \\
 &= (a_2 b_3 - a_3 b_2) \mathbf{e}_1 + (a_3 b_1 - a_1 b_3) \mathbf{e}_2 + (a_1 b_2 - a_2 b_1) \mathbf{e}_3
 \end{aligned}$$

□

Satz 8.4: Spatprodukt

Seien $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ so gilt:

$$\langle \mathbf{a} \times \mathbf{b}, \mathbf{c} \rangle = c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \det(\mathbf{a}, \mathbf{b}, \mathbf{c})$$

Der Name Spatprodukt kommt vom mineralogischen Begriffs des Spates, welches ein allgemeines Parallelepiped beschreibt.

Sei $E = \{\mathbf{x}_0 + \lambda \mathbf{a} + \mu \mathbf{b} | \lambda, \mu \in \mathbb{R}\}$ eine Ebene im \mathbb{R}^3 , dann können wir die Ebene auch durch den Normalvektor \mathbf{n} beschreiben. Die Ebene ist dann die Menge der Vektoren $\mathbf{x} \in \mathbb{R}^3$, die $\langle \mathbf{x} - \mathbf{x}_0, \mathbf{n} \rangle = 0$. Dabei handelt es sich um eine implizite Darstellung der Ebene.

Definition 8.3: Verallgemeinerungen

Eine inneres Produkt auf einem Vektorraum V über \mathbb{R} oder \mathbb{C} ist eine Abbildung $\langle \cdot, \cdot \rangle: V^2 \rightarrow \mathbb{K}$, welche die folgenden Eigenschaften erfüllt:

- i $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ für alle $\mathbf{x}, \mathbf{y} \in V$
- ii $\langle \lambda \mathbf{x}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$
- iii $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$

Insbesondere $\langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}$ (aus i).

Das innere Produkt heißt positiv semidefinit wenn $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ für $\mathbf{x} \in V$, und positiv definit wenn $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ für $\mathbf{x} \neq \mathbf{0}$. Wir sagen $\langle \cdot, \cdot \rangle$ ist negativ semidefinit wenn $\langle \mathbf{x}, \mathbf{x} \rangle \leq 0$, negativ definit, wenn $\langle \mathbf{x}, \mathbf{x} \rangle < 0$ für $\mathbf{x} \neq \mathbf{0}$ und indefinit wenn $\exists \mathbf{x} \in V: \langle \mathbf{x}, \mathbf{x} \rangle > 0$ und $\exists \mathbf{y} \in V: \langle \mathbf{y}, \mathbf{y} \rangle < 0$.

Für positiv definite Skalarprodukte gibt es auch alternative Bezeichnungen:

- Skalarprodukt $\mathbb{K} = \mathbb{R}$
- Hermitesches Produkt (unitäres Produkt) $\mathbb{K} = \mathbb{C}$
- Quadratische Form $\mathbb{K} = \mathbb{R}$
- Hermitesche Form $\mathbb{K} = \mathbb{C}$

Charles Hermit hat den Begriff hermitesch eingeführt³.

Lemma 8.2: Sesquilineares inneres Produkt

- i $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$
- ii $\langle \mathbf{x}, \lambda \mathbf{y} \rangle = \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle$
- iii $\langle \mathbf{x}, \mathbf{0} \rangle = 0$

Sesquilinear bedeutet ungefähr "eineinhalb-linear", da wir die Additivität erhalten, aber nicht ganz die Homogenität. Zwei einfache Beispiele sind etwa $V = \mathbb{R}^n$ mit $\mathbb{K} = \mathbb{R}$ und $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \langle \mathbf{x}, \mathbf{e}_i \rangle \langle \mathbf{y}, \mathbf{e}_i \rangle$. Hier gilt $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$, somit ist $\langle \cdot, \cdot \rangle$ positiv definit. Für $V = \mathbb{C}^n$ mit $\mathbb{K} = \mathbb{C}$ verwenden wir jedoch $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \langle \mathbf{x}, \mathbf{e}_i \rangle \overline{\langle \mathbf{y}, \mathbf{e}_i \rangle}$, dann gilt ebenfalls $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$.

³er hat unter anderem auch gezeigt, dass $e \notin \overline{\mathbb{Q}}$, sprich $\nexists p \in \mathbb{Q}[x]: p(e) = 0$, also dass die eulersche Zahl e nicht algebraisch ist

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$, wobei $a_{ij} = a_{ji}$ (also \mathbf{A} ist symmetrisch), dann können wir auch ein Skalarprodukt $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}}$ (bzw. eine quadratische Form) definieren:

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} = Q_{\mathbf{A}}(\mathbf{x}, \mathbf{y}) = \sum_{i,j} a_{ij} x_i y_j$$

Das Standardskalarprodukt entspricht dann $Q_{\mathbf{I}}$. Über den komplexen Zahlen gehen wir ähnlich vor. Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$:

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{A}} = \sum_{i,j} a_{ij} x_i \overline{y_j}$$

Aufgrund der Symmetrie des Skalarproduktes muss gelten: $a_{ij} = a_{ji}$, also erfüllt $\mathbf{A} = \overline{\mathbf{A}}^T$. In der Funktionalanalysis werden wir sehen, dass wir über $\mathcal{C}([a, b])$, der Menge der stetigen Funktionen auf $[a, b]$, ein Skalarprodukt z.B. über das Riemann-Integral definieren können:

$$\begin{aligned} \langle f, g \rangle &= \int_a^b f(x) \overline{g(x)} \, dx \\ \langle f, f \rangle &= \int_a^b |f(x)|^2 \, dx \geq 0 \end{aligned}$$

Definition 8.4: Norm

Sei V ein Vektorraum über \mathbb{R} oder \mathbb{C} . Eine Norm ist eine Abbildung $\|\cdot\|: V \rightarrow \mathbb{R}$, sodass:

- N1 $\forall \mathbf{x} \in V: \|\mathbf{x}\| \geq 0$ und $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
- N2 $\|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$
- N3 $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Eine Norm induziert eine Metrik $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, da $d(\mathbf{x}, \mathbf{y}) \geq 0$, $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$ und $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$. Diese Metrik ist **translationsinvariant**, da $d(\mathbf{x} + \mathbf{x}_0, \mathbf{y} + \mathbf{x}_0) = \|\mathbf{x} - \mathbf{y}\|$. Im \mathbb{R}^2 bzw. \mathbb{R}^3 nennen wir d die euklidische Metrik.

Betrachten wir ein Beispiel. Sei $V = \mathbb{K}^n$, dann definieren wir $\|\mathbf{x}\|_{\infty} = \max_{i=1, \dots, n} |x_i|$. Auch die Manhattan-Norm $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$.

Auch für Funktion in $\mathcal{C}([a, b], \mathbb{K})$ können wir eine Norm definieren $\|f\|_{\infty} = \max_{t \in [a, b]} |f(t)|$. Und:

$$\int_a^b |f(t)| \, dt$$

Allgemeiner definieren wir auf \mathbb{K}^n die p -Norm:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

Satz 8.5: Induzierte Norm

Sei $\langle \cdot, \cdot \rangle$ ein positiv-definites Produkt, dann ist $\|\mathbf{x}\| = (\langle \mathbf{x}, \mathbf{x} \rangle)^{\frac{1}{2}}$ eine Norm auf V .

Beweis. Da $\langle \cdot, \cdot \rangle$ positiv definit ist, gilt bereits $\|\mathbf{x}\| \geq 0$ und $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$. Auch die Homogenität folgt direkt:

$$\|\lambda \mathbf{x}\| = \sqrt{\langle \lambda \mathbf{x}, \lambda \mathbf{x} \rangle} = \sqrt{\lambda \overline{\lambda} \langle \mathbf{x}, \mathbf{x} \rangle} = |\lambda| \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

□

Lemma 8.3: Cauchy-Schwarz-Bunjakovski Ungleichung

Für ein positiv definites inneres Produkt gilt die Ungleichung:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

Sind \mathbf{x} und \mathbf{y} linear abhängig, dann gilt:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

Der Name dieser Ungleichung ist abhängig von der Literatur, die man liest. Cauchy hat bereits 1821 gezeigt:

$$\left| \sum x_i y_i \right| \leq \sqrt{\sum |x_i|^2} \cdot \sqrt{\sum |y_i|^2}$$

Bunjakovski, ein Student von Cauchy, hat die Ungleichung 1859 erweitert:

$$\left| \int f(t)g(t) dt \right| \leq \sqrt{\int |f(t)|^2 dt} \cdot \sqrt{\int |g(t)|^2 dt}$$

Zuletzt hat Schwarz 1883 einen Beweis für allgemeine Vektorräume gebracht:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

Den Beweis von Cauchy können wir über die Lagrange-Identität führen:

$$\begin{aligned} \sum_{i,j=1}^n (x_i y_j - x_j y_i)^2 &= \sum_{i=1}^n \sum_{j=1}^n x_i^2 y_j^2 - 2x_i x_j y_i y_j + x_j^2 y_i^2 \\ &= 2 \sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j^2 - 2 \sum_{i=1}^n x_i y_i \sum_{j=1}^n x_j y_j = 2\|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2 - 2(\langle \mathbf{x}, \mathbf{y} \rangle)^2 \geq 0 \\ &\Leftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle^2 + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (x_i y_j - x_j y_i)^2 = \|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2 \end{aligned}$$

Beweis zu Lemma 8.3. Falls $\mathbf{y} = \mathbf{0}$ gilt $\langle \mathbf{x}, \mathbf{0} \rangle \leq \|\mathbf{x}\| \cdot \|\mathbf{0}\|$

Sei $\lambda \in \mathbb{K}$ und $\mathbf{y} \neq \mathbf{0}$, so gilt $0 \leq \langle \mathbf{x} - \lambda \mathbf{x}, \mathbf{x} - \lambda \mathbf{y} \rangle$:

$$\begin{aligned} \langle \mathbf{x} - \lambda \mathbf{y}, \mathbf{x} - \lambda \mathbf{y} \rangle &= \langle \mathbf{x}, \mathbf{x} \rangle - \langle \lambda \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{x}, \lambda \mathbf{y} \rangle + \langle \lambda \mathbf{y}, \lambda \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle - \lambda \langle \mathbf{y}, \mathbf{x} \rangle - \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle + |\lambda|^2 \langle \mathbf{y}, \mathbf{y} \rangle \end{aligned}$$

Wir wählen $\lambda = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}$:

$$\begin{aligned} \langle \mathbf{x}, \mathbf{x} \rangle - \lambda \langle \mathbf{y}, \mathbf{x} \rangle - \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle + |\lambda|^2 \langle \mathbf{y}, \mathbf{y} \rangle &= \langle \mathbf{x}, \mathbf{x} \rangle - \frac{\langle \mathbf{x}, \mathbf{y} \rangle \langle \mathbf{y}, \mathbf{x} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} - \frac{\overline{\langle \mathbf{x}, \mathbf{y} \rangle} \langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} + \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\langle \mathbf{y}, \mathbf{y} \rangle} \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle - \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\langle \mathbf{y}, \mathbf{y} \rangle} \geq 0 \Rightarrow \|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2 = \|\mathbf{x} - \lambda \mathbf{y}\|^2 \cdot \|\mathbf{y}\|^2 \geq 0 \Rightarrow \|\mathbf{x} - \lambda \mathbf{y}\|^2 \cdot \|\mathbf{y}\|^2 = 0 \Leftrightarrow \mathbf{x} - \lambda \mathbf{y} = \mathbf{0} \end{aligned}$$

□

Fortsetzung von Satz 8.5.

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &\leq \langle \mathbf{x}, \mathbf{x} \rangle + |\langle \mathbf{x}, \mathbf{y} \rangle| + |\langle \mathbf{y}, \mathbf{x} \rangle| + \langle \mathbf{y}, \mathbf{y} \rangle \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \end{aligned}$$

□

Kleiner Einschub: Das Skalarprodukt ist im allgemeinen antilinear in der zweiten Komponente:

$$\begin{aligned}\langle \lambda \mathbf{x}, \mathbf{y} \rangle &= \lambda \langle \mathbf{x}, \mathbf{y} \rangle \\ \langle \mathbf{x}, \lambda \mathbf{y} \rangle &= \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle\end{aligned}$$

In der Physik arbeitet man genau anders herum:

$$\begin{aligned}\langle \mathbf{x} | \lambda \mathbf{y} \rangle &= \lambda \langle \mathbf{x} | \mathbf{y} \rangle \\ \langle \lambda \mathbf{x} | \mathbf{y} \rangle &= \bar{\lambda} \langle \mathbf{x} | \mathbf{y} \rangle\end{aligned}$$

Einschub Ende

Satz 8.6: Darstellung durch Matrizen

Sei V ein Vektorraum (mit $\mathbb{K} = \mathbb{R}$ oder \mathbb{C}), $\langle \cdot, \cdot \rangle$ ein inneres Produkt und $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V , dann $\exists! \mathbf{A} \in \mathbb{K}^{n \times n}$, wobei $\mathbf{A} = \mathbf{A}^H$ (\mathbf{A} ist hermitesch), sodass $\forall \mathbf{x}, \mathbf{y} \in V$: $\langle \mathbf{x}, \mathbf{y} \rangle = \Phi_B(\mathbf{x})^T \mathbf{A} \overline{\Phi_B(\mathbf{y})}$. Wenn $\langle \cdot, \cdot \rangle$ positiv definit ist, dann ist \mathbf{A} regulär.

Beweis. Sei $\mathbf{x} = \sum_{i=1}^n \xi_i \mathbf{b}_i$ und $\mathbf{y} = \sum_{i=1}^n \eta_i \mathbf{b}_i$:

$$\begin{aligned}\Phi_B(\mathbf{x}) &= \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} & \Phi_B(\mathbf{y}) &= \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \\ \langle \mathbf{x}, \mathbf{y} \rangle &= \left\langle \sum_{i=1}^n \xi_i \mathbf{b}_i, \sum_{i=1}^n \eta_i \mathbf{b}_i \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \xi_i \bar{\eta}_j \underbrace{\langle \mathbf{b}_i, \mathbf{b}_j \rangle}_{=a_{ij}} = \sum_{i=1}^n \sum_{j=1}^n \xi_i a_{ij} \bar{\eta}_j = \Phi_B(\mathbf{x})^T \mathbf{A} \overline{\Phi_B(\mathbf{y})} \\ a_{ij} &= \langle \mathbf{b}_i, \mathbf{b}_j \rangle = \overline{\langle \mathbf{b}_j, \mathbf{b}_i \rangle} = \overline{a_{ji}}\end{aligned}$$

Angenommen \mathbf{A} sei nicht regulär, dann ist $\ker \mathbf{A} \neq \{\mathbf{0}\}$. Sei $\xi \neq \mathbf{0}$ mit $\mathbf{A}\xi = \mathbf{0}$:

$$\begin{aligned}\mathbf{x} &= \sum_{i=1}^n \bar{\xi}_i \mathbf{b}_i & \overline{\Phi_B(\mathbf{x})} &= \xi \\ \mathbf{A}\Phi_B(\mathbf{x}) &= \mathbf{0} \Rightarrow \Phi_B(\mathbf{x})^T \mathbf{A} \overline{\Phi_B(\mathbf{x})} = \mathbf{0} = \langle \mathbf{x}, \mathbf{x} \rangle\end{aligned}$$

Wenn $\langle \cdot, \cdot \rangle$ positiv definit ist, dann muss aber gelten $\mathbf{x} = \mathbf{0}$ bzw. $\xi = \mathbf{0}$, womit aber nur $\ker \mathbf{A} = \{\mathbf{0}\}$. Das ist ein Widerspruch zu unserer Annahme, somit muss \mathbf{A} regulär sein. \square

Definition 8.5: Definitheit von Matrizen

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$, dann heißt $\mathbf{A}^* = \overline{\mathbf{A}^T}$ die zu \mathbf{A} adjungierte Matrix (nicht adjunkt). \mathbf{A} heißt selbst-adjungiert, wenn $\mathbf{A} = \mathbf{A}^*$. Reelle, selbst-adjungierte Matrizen sind symmetrisch, komplexe sind hermitesch.

\mathbf{A} heißt positiv semidefinit, $\langle \xi, \eta \rangle_{\mathbf{A}} = \xi^T \mathbf{A} \bar{\eta}$ positiv semidefinit ist, wir schreiben $\mathbf{A} \geq 0$. Analog ist \mathbf{A} negativ-semidefinit, wenn $\langle \xi, \eta \rangle_{\mathbf{A}}$ negativ-semidefinit ist, wir schreiben $\mathbf{A} \leq 0$.

Gleichfalls ist \mathbf{A} positiv definit, wenn $\langle \xi, \eta \rangle_{\mathbf{A}}$ positiv ist, wir schreiben $\mathbf{A} > 0$. Analog ist \mathbf{A} negativ definit, wenn $\langle \xi, \eta \rangle_{\mathbf{A}}$ negativ definit ist, wir schreiben $\mathbf{A} < 0$.

Wie erkennen wir positiv definite Matrizen?

Definition 8.6: Kongruenz von Matrizen

Zwei Matrizen $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$ heißen kongruent $\mathbf{A} \hat{=} \mathbf{B}$ wenn $\exists \mathbf{C} \in \text{GL}(n, \mathbb{K})$: $\mathbf{C}^* \mathbf{A} \mathbf{C} = \mathbf{B}$.

Dabei $\mathbf{A} \hat{=} \mathbf{B} \Rightarrow \mathbf{A} \simeq \mathbf{B}$

Satz 8.7: Kongruenz selbstadjungierter Matrizen

Jede selbstadjungierte Matrix A ist kongruent zu einer Matrix der Form

$$\text{diag}(+1, \dots, +1, -1, \dots, -1, 0, \dots, 0)$$

Beweis zu Satz 8.7. Konstruktiv: Wir finden eine Matrix $C^*AC = D$, wobei D eine Diagonalmatrix der beschriebenen Form aus Satz 8.7 ist. Für $n = 1$ gilt also $A = [a_{11}]$, für $a_{11} = 0$ sind wir bereits fertig, sonst wählen wir

$$C = \begin{bmatrix} \frac{1}{\sqrt{|a_{11}|}} \end{bmatrix} = C^*$$

$$A^*AC = \begin{bmatrix} \frac{a_{11}}{|a_{11}|} \end{bmatrix} = [\pm 1]$$

$n - 1 \mapsto n$

Wir machen eine Fallunterscheidung. Falls $A = 0$ ist $C = I_n$. Falls $a_{11} = 0$:

a $\exists j: a_{jj} \neq 0$. Sei $C = T_{1,j} = C^*$, dann hat C^*AC als ersten Eintrag $a_{jj} \neq 0$

b $a_{ii} = 0$ für $i = 1, \dots, n$, da $A \neq 0$ können wir ein a_{ij} mit $i \neq j$ finden, sodass $a_{ij} \neq 0$. Wir setzen $C = I_n + E_{ij}e^{i\theta}$, wobei θ der polare Winkel von a_{ij} ist:

$$C^*AC = A + e^{-i\theta}E_{ji}A + e^{i\theta}AE_{ij} + E_{ij}AE_{ij}$$

$$(C^*AC)_{jj} = a_{jj} + e^{-i\theta}a_{ij} + e^{i\theta}a_{ji} + a_{ii} = 2|a_{ij}| \neq 0$$

Somit können wir auf Fall 2a zurückgreifen

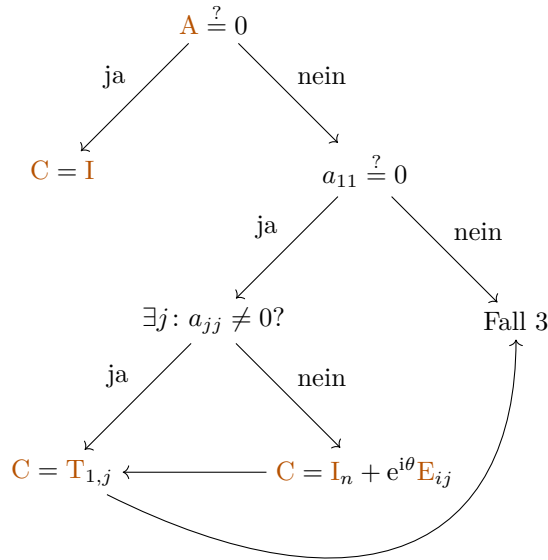


Abbildung 6: Entscheidungsbaum

Im dritten Fall gilt also $a_{11} \neq 0$, dann können wir in der ersten Spalte und Zeile Nullen erzeugen:

$$C = \begin{bmatrix} 1 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}$$

Womit

$$\mathbf{C}^* \mathbf{A} \mathbf{C} = \begin{bmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{A}} \end{bmatrix}$$

Nach unserer Induktionsvoraussetzung gibt es eine Matrix $\tilde{\mathbf{C}}$, die $\tilde{\mathbf{A}}$ auf die gewünschte Form bringt. Wir legen also fest:

$$\begin{aligned} \mathbf{C}_1 &= \begin{bmatrix} \frac{1}{\sqrt{|a_{11}|}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}} \end{bmatrix} \\ \mathbf{C}_1^* \mathbf{A} \mathbf{C}_1 &= \text{diag}(+1, \dots, +1, -1, \dots, -1, 0, \dots, 0) \end{aligned}$$

□

Betrachten wir das Ganze an einem Beispiel:

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

In diesem Fall sind alle Einträge der Diagonal 0, wir wählen also $\mathbf{C}_1 = \mathbf{I}_3 + \mathbf{E}_{12}$:

$$\mathbf{A}_1 = \mathbf{C}_1^* \mathbf{A} \mathbf{C}_1 = \begin{bmatrix} 0 & 1 & i \\ 1 & 2 & 1+i \\ -i & 1-i & 0 \end{bmatrix}$$

Jetzt vertauschen wir die zweite Spalte und Zeile mit der ersten, also $\mathbf{C}_2 = \mathbf{T}_{1,2}$:

$$\mathbf{A}_2 = \mathbf{C}_2^* \mathbf{A}_1 \mathbf{C}_2 = \begin{bmatrix} 2 & 1 & 1+i \\ 1 & 0 & i \\ 1-i & -i & 0 \end{bmatrix}$$

Jetzt erzeugen wir in der ersten Spalte und Zeile 0:

$$\begin{aligned} \mathbf{C}_3 &= \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1+i}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \Rightarrow \mathbf{A}_3 &= \mathbf{C}_3^* \mathbf{A}_2 \mathbf{C}_3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{i-1}{2} \\ 0 & -\frac{1+i}{2} & -1 \end{bmatrix} \\ \mathbf{A}_4 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{i-1}{2} \\ 0 & -\frac{1+i}{2} & -1 \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{A}} \end{bmatrix} \end{aligned}$$

Unser $\tilde{\mathbf{A}}$ erfüllt nun $\tilde{a}_{11} \neq 0$. Wir wählen daher

$$\begin{aligned} \tilde{\mathbf{C}}_5 &= \begin{bmatrix} 1 & i-1 \\ 0 & 1 \end{bmatrix} \\ \tilde{\mathbf{A}}_5 &= \tilde{\mathbf{C}}_5^* \tilde{\mathbf{A}} \tilde{\mathbf{C}}_5 = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \\ \Rightarrow \tilde{\mathbf{C}}_6 &= \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \tilde{\mathbf{A}}_6 = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \\ \mathbf{C}_5 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & i-1 \\ 0 & 0 & 1 \end{bmatrix} & \mathbf{C}_6 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{C} &= \mathbf{C}_1 \mathbf{C}_2 \mathbf{C}_3 \mathbf{C}_4 \mathbf{C}_5 \mathbf{C}_6 \\ \mathbf{C}^* \mathbf{A} \mathbf{C} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Hier ist anzumerken, dass $A^* = A \Rightarrow (C^*AC)^* = C^*AC$, wenn $A \geq 0$, dann $C^*AC \geq 0$ und wenn $A > 0$ und $C \in GL(n, \mathbb{K})$, dann ist $C^*AC > 0$.

Sylvester und Kayley habe die sogenannte Invariantentheorie entwickelt.

Satz 8.8: Trägheitssatz von Sylvester

Sei $A \in \mathbb{K}^{n \times n}$ eine hermitesche Matrix und $C \in GL(n, \mathbb{K})$, sodass

$$C^*AC = \text{diag}(+1, \dots, +1, -1, \dots, -1, 0, \dots, 0)$$

dann ist die Anzahl der $+1$ und -1 in C^*AC eindeutig bestimmt.

Beweis zu Satz 8.8. Seien $C, \tilde{C} \in GL(n, \mathbb{C})$, sodass

$$C^*AC = \text{diag}(\overbrace{1, \dots, 1}^r, \overbrace{-1, \dots, -1}^s, 0, \dots, 0) = D$$

$$\tilde{C}^*A\tilde{C} = \text{diag}(\overbrace{1, \dots, 1}^{\tilde{r}}, \overbrace{-1, \dots, -1}^{\tilde{s}}, 0, \dots, 0) = \tilde{D}$$

Dann ist $\text{rg}(A) = r + s = \tilde{r} + \tilde{s}$. Wir müssen also nur zeigen, dass $r = \tilde{r}$, bzw. $r \leq \tilde{r}$ aufgrund der Symmetrie. $C \in GL(n, \mathbb{C})$ entspricht einer Basistransformation in eine Basis $B = (b_1, \dots, b_n)$:

$$x^T A \bar{x} = \Phi_B(x)^T D \overline{\Phi_B(x)}$$

Somit ist auch \tilde{C} eine Basistransformation mit $\tilde{B} = (\tilde{b}_1, \dots, \tilde{b}_n)$, dann gilt für $x \in L(b_1, \dots, b_r) \setminus \{0\}$:

$$\begin{aligned} x^T A \bar{x} &= \Phi_B(x)^T D \overline{\Phi_B(x)} \\ &= \sum_{i=1}^r |\xi_i|^2 > 0 \end{aligned}$$

wobei $x = \sum_{i=1}^r \xi_i b_i$, während für $x \in \mathcal{L}(\tilde{b}_{\tilde{r}+1}, \dots, \tilde{b}_n)$ gilt $x^T A \bar{x} \leq 0$, dabei ist $\mathcal{L}(b_1, \dots, b_r) \cap \mathcal{L}(\tilde{b}_{\tilde{r}+1}, \dots, \tilde{b}_n) = \emptyset$. Womit $\dim(\mathcal{L}(b_1, \dots, b_r)) + \dim(\mathcal{L}(\tilde{b}_{\tilde{r}+1}, \dots, \tilde{b}_n)) \leq n$ bzw. $r + n - \tilde{r} \leq 0 \Rightarrow r \leq \tilde{r}$. \square

Definition 8.7: Index und Signatur von Matrizen

Sei $A = A^*$ kongruent zu $\text{diag}(\overbrace{1, \dots, 1}^r, \overbrace{-1, \dots, -1}^s, 0, \dots, 0)$ dann heißt $\text{ind}A = r$ der Index von A und $\text{sign}A = s$ die Signatur, mit $r + s = \text{rg}A$.

Eine Folgerung: Seien A, B hermitesche Matrizen. Sind A und B kongruent, dann gilt $\text{rank}A = \text{rank}B$, und $\text{ind}A = \text{ind}B$. A ist weiterhin positiv definit wenn $\text{ind}A = n$, und $A \geq 0 \Leftrightarrow \text{ind}A = \text{rank}A$.

Lemma 8.4: Determinanten hermitescher Matrizen

- i $\det A^* = \overline{\det A}$
- ii $A = A^* \Rightarrow \det A \in \mathbb{R}$
- iii $A = B^*$ und $B = B^*$ wobei $A \hat{=} B \Rightarrow \text{signdet}A = \text{signdet}B$
- iv $A = A^*$ und $A > 0$, dann $\det A > 0$

Beweis für Lemma 8.4. Zu i:

Sei A eine quadratische Matrix, dann ist $\det(A^*) = \det A^T = \det \bar{A} = \overline{\det A}$.

Zu ii:

Wenn $A = A^*$, dann gilt $\det A = \det A^* = \overline{\det A}$, Sei $z \in \mathbb{C}$: $z = \bar{z} \Rightarrow z \in \mathbb{R}$, womit $A = A^* \Rightarrow \det A \in \mathbb{R}$.

Zu iii: Sei $C^*AC = B$, dann gilt:

$$\det B = \det C^* \det A \det C = \det A \cdot |\det C|^2$$

Zu iv:

$A = A^*$ und $A > 0 \Rightarrow A \stackrel{\wedge}{=} I$, nach iii also $\det A > 0$. □

Definition 8.8: Minoren

Sei A eine $n \times n$ Matrix. Seien $I, J \subseteq \{1, \dots, n\}$ mit $|I| = |J| = r$. Wir bezeichnen $A_{I,J}$ als jene Matrix, in der nur Zeilenindizes aus I und Spaltenindizes aus J liegen. Es folgt direkt, dass $A_{I,J}$ eine $r \times r$ Matrix ist. Weiters bezeichnen wir $[A]_{I,J} = \det(A_{I,J})$ als **Minor** von $\det(A)$.

Wir nennen $\det(A_r)$ die **Hauptminoren**, wobei $A_r = A_{I,I}$ mit $I = \{1, \dots, r\}$.

Satz 8.9: Cholesky-Zerlegung

Sei A eine hermitesche Matrix, dann sind die folgenden Aussagen äquivalent:

- a $A > 0$
- b $\forall r \in \{1, \dots, n\}: \det A_r > 0$
- c es gibt eine untere Dreiecksmatrix C , sodass $A = CC^*$, bzw. $C^{-1}A(C^{-1})^* = I$, wobei $c_{ii} > 0$ für $i = 1, \dots, n$

Beweis für Satz 8.9. a) \Rightarrow b):

$$\begin{aligned} A > 0 &\Rightarrow A_r > 0 \\ A > 0 &\Leftrightarrow x^T A x > 0 \forall x \in \mathbb{K}^n: x \neq 0 \\ y \in \mathbb{K}^r \setminus \{0\} &\rightsquigarrow x = \begin{bmatrix} y^T & 0^T \end{bmatrix}^T \in \mathbb{K}^n \\ &\Rightarrow x^T A x = y^T A_r y > 0 \end{aligned}$$

Wir bestimmen C_r rekursiv, wobei $C_r C_r^* = A_r$. Wir erweitern C_r in C_{r+1} . Trivialerweise gilt $C_1 = \sqrt{a_1}$:

$$\begin{aligned} \begin{bmatrix} C_r & 0 \\ \bar{c}^T & \alpha \end{bmatrix} \begin{bmatrix} C_r^* & c \\ 0 & \alpha \end{bmatrix} &= \begin{bmatrix} A_r & b \\ \bar{b}^T & a_{r+1,r+1} \end{bmatrix} \\ b &= [a_{1,r+1} \quad \dots \quad a_{r,r+1}]^T \\ A_{r+1} &= \begin{bmatrix} A_r & C_r c \\ \bar{c}^T C_r^* & \bar{c}^T c + \alpha^2 \end{bmatrix} \\ C_r c &= b \Leftrightarrow c = C_r^{-1} b \\ \alpha^2 &= a_{r+1,r+1} - \bar{c}^T c = \beta > 0 \end{aligned}$$

Es gilt $\beta > 0$, da $\det(A_{r+1}) = \det(C_r) \alpha \det(C_r^*) \alpha = \alpha^2 \det(C_r C_r^*) = \alpha^2 \det(A_r) = \beta \det(A_r)$. Nach b) muss $\beta > 0$.

Woher wissen wir $\beta \in \mathbb{R}$? Wir wissen:

$$\begin{aligned} a_{ij} &= \sum_{k=1}^i c_{ik} c_{kj}^* = \sum_{k=1}^i c_{i,k} \overline{c_{j,k}} \\ a_{jj} &= \sum_{k=1}^{j-1} c_{j,k} \overline{c_{k,j}} + |c_{jj}|^2 \Rightarrow c_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} c_{j,k} \overline{c_{k,j}}} \\ c_{ij} &= \frac{1}{c_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} c_{ik} \overline{c_{jk}} \right) \end{aligned}$$

□

Definition 8.9: Bezeichnungen und Winkel

- i Sei V ein Vektorraum mit positiv definiten inneren Produkt $\langle \cdot, \cdot \rangle$ heißt euklidischer Raum wenn $\mathbb{K} = \mathbb{R}$ und unitärer Raum wenn $\mathbb{K} = \mathbb{C}$, und Prä-Hilbert-Raum, wenn $\dim V = \infty$
- ii $v \in V$ heißt normiert, wenn $\|v\| = 1$, wobei $\|\cdot\|$ die durch $\langle \cdot, \cdot \rangle$ induzierte Norm ist jenes $\varphi \in [0, \pi]$, für das $\cos(\varphi) = \frac{\Re\{\langle v, w \rangle\}}{\|v\| \cdot \|w\|}$
- iii $v, w \in V \setminus \{0\}$, wir Winkel φ zwischen v und w
- iv $v, w \in V$ heißen orthogonal bezüglich $\langle \cdot, \cdot \rangle$, wenn $\langle v, w \rangle = 0$, wir schreiben $v \perp w$

Satz 8.10: Cosinussatz und Parallelogrammgleichung

- i $\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\|v\| \cdot \|w\| \cdot \cos \varphi$
- ii $v \perp w \Rightarrow \|v + w\|^2 = \|v\|^2 + \|w\|^2$
- iii $\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$ (Parallelogrammgleichung)

Satz 8.11: Induzierte Normen

Eine Norm $\|\cdot\|$ auf V wird von einem Skalarprodukt induziert, genau dann wenn sie die Parallelogrammgleichung erfüllt.

8.1 Orthogonalität

Wir haben bereits kennengelernt, wie man den von zwei aufgespannten Vektoren x, y auffindet. Sind zwei Vektoren orthogonal, dann gilt der Satz von Pythagoras $\|x + y\| = \|x\| + \|y\|$.

Definition 8.10: Orthonormalbasis

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt. Eine Familie $(v_i)_{i \in I} \subseteq V$ heißt:

- i orthogonal wenn $\forall i \neq j: \langle v_i, v_j \rangle = 0$
- ii orthonormal wenn $\forall i, j \in I: \langle v_i, v_j \rangle = \delta_{ij}$
- iii Orthonormalbasis wenn sie zusätzlich eine Basis bilden

Ein klassisches Beispiel für eine Orthonormalbasis sind die kanonischen Einheitsvektoren e_i des \mathbb{R}^n mit dem Standardskalarprodukt $\langle x, y \rangle = \sum x_i y_i$, dann gilt sofort $\langle e_i, e_j \rangle = \delta_{ij}$.

Betrachten wir den Raum der stetigen Funktionen auf $[0, 1]$, also $\mathcal{C}[0, 1]$ mit Skalarprodukt:

$$\langle f, g \rangle = \int_0^1 f(x)g(x) \, dx$$

Dann sind $(\sqrt{2} \sin(2\pi n x))_{n \in \mathbb{N}}$ orthonormal, sowie auch $(\sqrt{2} \cos(2\pi n x))_{n \in \mathbb{N}}$. Zusammen mit $\{1\}$ bilden ein Orthonormalsystem. **Fourier** 1809 im Zuge seiner Untersuchung der Wärmeleitungsgleichung gefunden, dass man f folgendermaßen approximieren kann:

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi n x) + \sum_{n=1}^{\infty} b_n \sin(2\pi n x)$$

Das eröffnet das Feld der Fourier-Analyse, die hier jedoch nicht weiter behandelt wird.

Satz 8.12: Eigenschaften orthogonaler Familien

Sei $(v_i)_{i \in I} \subseteq V \setminus \{0\}$

- i $(v_i)_{i \in I}$ ist orthogonal \Leftrightarrow dann ist $(v_i \frac{1}{\|v_i\|})_{i \in I}$ orthonormal
- ii $(v_i)_{i \in I}$ ist orthogonal, dann ist $(v_i)_{i \in I}$ linear unabhängig

Beweis für Satz 8.12. i ist trivial

Zu ii:

Angenommen $\sum_{l=1}^n \lambda_l \mathbf{v}_{i_l} = \mathbf{0}$, dann folgt:

$$\begin{aligned} \left\langle \sum_{l=1}^n \lambda_l \mathbf{v}_{i_l}, \mathbf{v}_{i_k} \right\rangle &= \langle \mathbf{0}, \mathbf{v}_{i_k} \rangle = 0 \\ &= \sum_{l=1}^n \lambda_l \langle \mathbf{v}_{i_l}, \mathbf{v}_{i_k} \rangle = \lambda_k \langle \mathbf{v}_{i_k}, \mathbf{v}_{i_k} \rangle = \lambda_k \|\mathbf{v}_{i_k}\|^2 \end{aligned}$$

Da $\mathbf{v}_i \neq \mathbf{0}$ muss folgen $\lambda_k = 0$, somit müssen allen $\lambda_k = 0$ sein, womit $(\mathbf{v}_i)_{i \in I}$ linear unabhängig ist. \square

Satz 8.13: Koordinaten in Orthonormalbasen

Sei $B = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ eine Orthonormalbasis eines n -dimensionalen Vektorraumes V , dann gilt für $\mathbf{v}, \mathbf{w} \in V$ mit den Koordinatenvektoren $\Phi_B(\mathbf{v}) = [\lambda_1 \ \dots \ \lambda_n]^T$ und $\Phi_B(\mathbf{w}) = [\mu_1 \ \dots \ \mu_n]^T$:

- i $\lambda_i = \langle \mathbf{v}, \mathbf{u}_i \rangle \forall i$
- ii $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \lambda_i \overline{\mu_i}$

Beweis für Satz 8.13. Zu i:

$$\langle \mathbf{v}, \mathbf{u}_i \rangle = \left\langle \sum_{j=1}^n \lambda_j \mathbf{u}_j, \mathbf{u}_i \right\rangle = \sum_{j=1}^n \lambda_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = \lambda_i$$

Zu ii:

$$\begin{aligned} \langle \mathbf{v}, \mathbf{w} \rangle &= \left\langle \sum_{j=1}^n \lambda_j \mathbf{u}_j, \sum_{i=1}^n \mu_i \mathbf{u}_i \right\rangle = \sum_{j=1}^n \lambda_j \left\langle \mathbf{u}_j, \sum_{i=1}^n \mu_i \mathbf{u}_i \right\rangle \\ &= \sum_{j=1}^n \lambda_j \sum_{i=1}^n \overline{\mu_i} \langle \mathbf{u}_j, \mathbf{u}_i \rangle = \sum_{j=1}^n \lambda_j \sum_{i=1}^n \overline{\mu_i} \delta_{ji} = \sum_{j=1}^n \lambda_j \overline{\mu_j} \end{aligned}$$

\square

Definition 8.11: Orthogonales Komplement

Sei V ein Vektorraum auf dem $\langle \cdot, \cdot \rangle$ definiert ist. Sei $M \subseteq V$ eine Teilmenge. Dann heißt

$$M^\perp = \{\mathbf{v} \in V : \langle \mathbf{u}, \mathbf{v} \rangle = 0 \forall \mathbf{u} \in M\}$$

das orthogonale Komplement von M . Wenn $M = \{\mathbf{v}\}$, schreiben wir $\mathbf{v}^\perp = \{\mathbf{v}\}^\perp$.

Betrachten wir ein Beispiel im \mathbb{R}^2 :

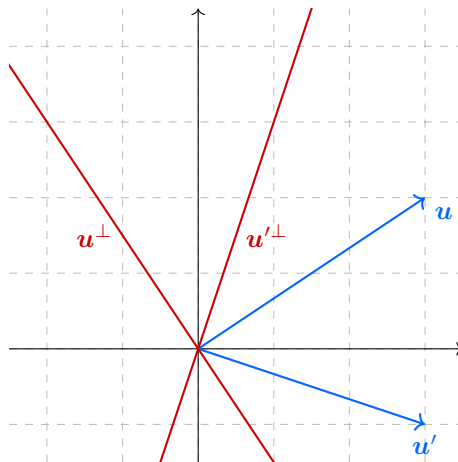


Abbildung 7: Beispiele für orthogonale Komplemente

Bestimmen wir $\{\mathbf{u}, \mathbf{u}'\}^\perp$ sehen wir schnell anhand Abbildung 7, dass nur $\mathbf{0} \in \{\mathbf{u}, \mathbf{u}'\}^\perp$.

Satz 8.14: Eigenschaften des orthogonalen Komplements

Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $M, N \subseteq V$:

- i M^\perp ist ein Unterraum
- ii $M \subseteq N \Rightarrow N^\perp \subseteq M^\perp$, insbesondere $(M_1 \cup M_2)^\perp = M_1^\perp \cap M_2^\perp$
- iii $\mathbf{0}^\perp = V$
- iv $V^\perp = \{\mathbf{0}\}$
- v $M \cap M^\perp \subseteq \{\mathbf{0}\}$
- vi $M^\perp = L(M)^\perp$
- vii $M \subseteq (M^\perp)^\perp$

Beweis. Zu i:

Wir verwenden das Unterraumkriterium. Da $\forall \mathbf{v} \in M: \langle \mathbf{v}, \mathbf{0} \rangle = 0$ gilt zumindest $M^\perp \supseteq \{\mathbf{0}\} \neq \emptyset$. Es verbleibt noch zu zeigen, dass $\forall \mathbf{v}, \mathbf{w} \in M^\perp: \forall \lambda, \mu \in \mathbb{K}: \lambda \mathbf{v} + \mu \mathbf{w} \in M^\perp$. Seien $\mathbf{v}, \mathbf{w} \in M^\perp$ und $\mathbf{u} \in M$ beliebig:

$$\langle \lambda \mathbf{v} + \mu \mathbf{w}, \mathbf{u} \rangle = \lambda \langle \mathbf{v}, \mathbf{u} \rangle + \mu \langle \mathbf{w}, \mathbf{u} \rangle = \lambda \cdot 0 + \mu \cdot 0 = 0$$

Alternativ definieren wir $T_{\mathbf{v}}: \mathbf{u} \mapsto \langle \mathbf{u}, \mathbf{v} \rangle$, dann ist $\mathbf{v}^\perp = \ker T_{\mathbf{v}}$, was natürlich ein Unterraum ist. Für beliebige M gilt:

$$M^\perp = \bigcap_{\mathbf{v} \in M} \mathbf{v}^\perp = \bigcap_{\mathbf{v} \in M} \ker T_{\mathbf{v}}$$

Der Durchschnitt von Unterräumen ist ein Unterraum, somit ist M^\perp ein Vektorraum.

Zu ii:

$$\begin{aligned} N^\perp &= \bigcap_{\mathbf{v} \in N} \mathbf{v}^\perp \subseteq \bigcap_{\mathbf{v} \in M} \mathbf{v}^\perp \\ (M_1 \cup M_2)^\perp &= \bigcap_{\mathbf{v} \in M_1 \cup M_2} \mathbf{v}^\perp = \bigcap_{\mathbf{v} \in M_1} \mathbf{v}^\perp \cap \bigcap_{\mathbf{v} \in M_2} \mathbf{v}^\perp = M_1^\perp \cap M_2^\perp \end{aligned}$$

Zu iii: $\forall \mathbf{v} \in V: \langle \mathbf{v}, \mathbf{0} \rangle = 0$.

Zu v:

Sei $\mathbf{v} \in M \cap M^\perp$, dann folgt:

$$\langle \mathbf{v}, \mathbf{u} \rangle \forall \mathbf{u} \in M$$

Insbesondere für $\mathbf{u} = \mathbf{v}$ folgt $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

Zu iv: $V^\perp = V \cap V^\perp = \{\mathbf{0}\}$

Zu vi:

Wir wissen bereits, dass $M \subseteq L(M)$. Nach ii folgt $L(M)^\perp \subseteq M^\perp$. Wir müssen noch zeigen, dass $M^\perp \subseteq L(M)^\perp$. Sei $\mathbf{v} \in M^\perp$ und $\mathbf{u} \in L(M)$, z.z. $\mathbf{v} \perp \mathbf{u}$. Sei $\mathbf{u} = \sum_{i=1}^n \lambda_i \mathbf{u}_i$ mit $\mathbf{u}_i \in M$ und $\lambda_i \in \mathbb{K}$, es folgt:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \sum_{i=1}^n \lambda_i \mathbf{u}_i, \mathbf{v} \right\rangle = \sum_{i=1}^n \lambda_i \langle \mathbf{u}_i, \mathbf{v} \rangle = 0$$

Somit $M^\perp = L(M)^\perp$

Zu vii:

Sei $\mathbf{u} \in M^\perp$ und $\mathbf{v} \in M$, dann $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Somit $(M^\perp)^\perp = \{\mathbf{v} \in V: \langle \mathbf{u}, \mathbf{v} \rangle = 0 \forall \mathbf{u} \in M^\perp\} \supseteq M$ □

An dieser Stelle sei angemerkt, dass $M \neq (M^\perp)^\perp$ möglich ist, wenn z.B. $\mathbf{0} \notin M$, denn $\mathbf{0}$ ist in jedem orthogonalen Komplement enthalten.

Sei $U \subseteq V$ ein Unterraum, dann ist $U \cap U^\perp = \{\mathbf{0}\}$, dann ist die Summe $U + U^\perp$ direkt und $(U + U^\perp)^\perp = \{\mathbf{0}\}$. Das gilt weil $U \cup U^\perp \subseteq (U + U^\perp) \subseteq (U \cup U^\perp)^\perp = U^\perp \cap (U^\perp)^\perp \subseteq \{\mathbf{0}\}$. Es muss jedoch nicht gelten $U + U^\perp = V$. Ein Beispiel hierfür ist $V = \ell^2 = \{(x_n)_{n \in \mathbb{N}} : \sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ mit $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \bar{y}_i$. Wir betrachten nun $U = L((e_i)_{i=1}^{\infty}) = \{(x_n)_{n \in \mathbb{N}} : \exists n_0 \in \mathbb{N} : n \geq n_0 \Rightarrow x_n = 0\}$, alternativ $U = L(M)$ mit $M = \{e_i | i \in \mathbb{N}\}$, dann gilt:

$$U^\perp = M^\perp = \{x : \langle x, e_i \rangle = 0 \forall i \in \mathbb{N}\} \Rightarrow \langle x, e_i \rangle = 0 \forall i \in \mathbb{N}$$

Sprich $x = 0$, womit $U^\perp = \{0\}$, aber $U \neq \ell^2$.

Wenn $V = U \dot{+} W$, dann $\forall v \in V : \exists! u \in U, w \in W : v = u + w$, mit $\pi_U : V \rightarrow U$ und $v \mapsto u$.

Definition 8.12: Konvexe Mengen

Eine Teilmenge $K \subseteq V$, wobei V ein Vektorraum ist, heißt konvex, wenn $\forall a, b \in K : \{a + \lambda(b - a) | \lambda \in [0, 1]\} \subseteq K$. Wir nennen $(1 - \lambda)a + \lambda b$ eine Konvexkombination.

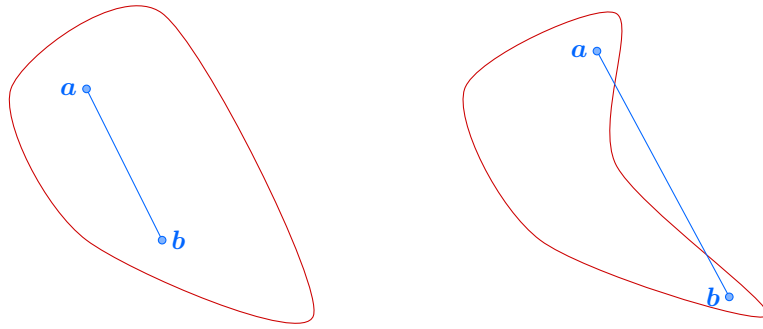


Abbildung 8: Beispiel für eine konvexe Menge (links) und eine nicht-konvexe Menge (rechts)

Eine direkte Folgerung ist, dass Unterräume konvex sind, da jede Linearkombination im Unterraum enthalten sein muss. Ein weiteres Beispiel ist die sogenannte Einheitskugel. Sei $(V, \|\cdot\|)$ ein normierter Raum, dann nennen wir $\mathcal{B}_1(0) = \{x \in V : \|x\| < 1\}$ die (offene) Einheitskugel, die konvex ist. Seien $x, y \in \mathcal{B}_1(0)$:

$$\|(1 - \lambda)x + \lambda y\| \leq (1 - \lambda)\|x\| + \lambda\|y\| < 1 - \lambda + \lambda = 1$$

Eine Translation einer konvexen Menge ist konvex. Sei K konvex, dann ist auch $x_0 + K$ konvex:

$$(1 - \lambda)(x_0 + a) + \lambda(x_0 + b) = (1 - \lambda)x_0 + (1 - \lambda)a + \lambda x_0 + \lambda b = x_0 + \underbrace{(1 - \lambda)a + \lambda b}_{\in K}$$

Sei $K \subseteq V$ konvex und $f \in \text{Hom}(V, W)$, dann ist $f(K)$ konvex.

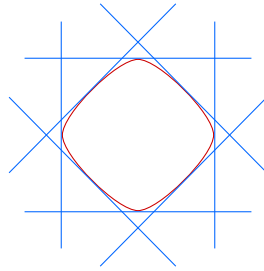


Abbildung 9: Linearisierung einer konvexen Menge

Besonders in der Optimierung stellen konvexe Mengen einen wichtigen Untersuchungsbereich dar, da sich viele Optimierungsaufgaben auf eine einfache Problemform reduzieren lassen: Man sucht den Punkt in der komplexen

Menge K , der einem beliebigen Punkt \mathbf{x} am nächsten ist. Das liefert bei konvexen Mengen immer eine eindeutige Lösung.

Satz 8.15: Hauptsatz der Optimierung

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt und $K \subseteq V$ eine konvexe Menge. Sei $\mathbf{x} \in V$ und $\mathbf{y}_0 \in K$. Dann sind äquivalent:

- i $\forall \mathbf{y} \in K: \|\mathbf{x} - \mathbf{y}_0\| \leq \|\mathbf{x} - \mathbf{y}\|$
- ii $\forall \mathbf{y} \in K: \Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle \leq 0$
- iii $\forall \mathbf{y} \in K \setminus \{\mathbf{y}_0\}: \|\mathbf{x} - \mathbf{y}_0\| < \|\mathbf{x} - \mathbf{y}\|$

Für den Fall, dass $K = U$ ein Unterraum ist, dann genügt ii':

$$\forall \mathbf{y} \in U: \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} \rangle = 0$$

Beweis für Satz 8.15. i \Rightarrow ii:

Sei $\mathbf{y}_\varepsilon = (1 - \varepsilon)\mathbf{y}_0 + \varepsilon\mathbf{y}$. Sei $0 < \varepsilon < 1$, dann ist $\mathbf{y}_\varepsilon \in K$, da K konvex ist.

$$\begin{aligned} 0 &\leq \|\mathbf{x} - \mathbf{y}_\varepsilon\|^2 - \|\mathbf{x} - \mathbf{y}_0\|^2 \\ \mathbf{x} - \mathbf{y}_\varepsilon &= \mathbf{x} - (1 - \varepsilon)\mathbf{y}_0 - \varepsilon\mathbf{y} = \mathbf{x} - \mathbf{y}_0 - \varepsilon(\mathbf{y} - \mathbf{y}_0) \\ \|\mathbf{x} - \mathbf{y}_\varepsilon\|^2 - \|\mathbf{x} - \mathbf{y}_0\|^2 &= \|\mathbf{x} - \mathbf{y}_0 - \varepsilon(\mathbf{y} - \mathbf{y}_0)\|^2 - \|\mathbf{x} - \mathbf{y}_0\|^2 \\ &= \|\mathbf{x} - \mathbf{y}_0\|^2 + \varepsilon^2\|\mathbf{y} - \mathbf{y}_0\|^2 - 2\varepsilon\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle - \|\mathbf{x} - \mathbf{y}_0\|^2 \\ &= \varepsilon(\varepsilon\|\mathbf{y} - \mathbf{y}_0\|^2 - 2\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle) \\ \varepsilon > 0 &\Rightarrow -2\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle + \varepsilon\|\mathbf{y} - \mathbf{y}_0\|^2 \geq 0 \\ &\Rightarrow \lim_{\varepsilon \rightarrow 0} -2\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle + \varepsilon\|\mathbf{y} - \mathbf{y}_0\|^2 \geq 0 \end{aligned}$$

ii \Rightarrow iii:

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= \|\mathbf{x} - \mathbf{y}_0 + \mathbf{y}_0 - \mathbf{y}\|^2 = \|(\mathbf{x} - \mathbf{y}_0) - (\mathbf{y} - \mathbf{y}_0)\|^2 \\ &= \|\mathbf{x} - \mathbf{y}_0\|^2 + \|\mathbf{y} - \mathbf{y}_0\|^2 - 2\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle \\ &\geq \|\mathbf{x} - \mathbf{y}_0\|^2 + \|\mathbf{y} - \mathbf{y}_0\|^2 > \|\mathbf{x} - \mathbf{y}_0\|^2 \end{aligned}$$

iii \Rightarrow i ist trivial, da für $\mathbf{y} = \mathbf{y}_0$ Gleichheit herrscht.

ii \Rightarrow ii'

Sei $K = U$ ein Unterraum, und es gelte $\forall \mathbf{y} \in U: \Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle \leq 0$ und $\mathbf{y}_0 \in U$. Sei $\mathbf{y}' = \mathbf{y} - \mathbf{y}_0 \in U$, dann $\forall \mathbf{y}' \in U: \Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y}' \rangle \leq 0$. Sei weiters $\mathbf{y}'' = -\mathbf{y}'$, dann $\forall \mathbf{y}'' \in U: \Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y}'' \rangle \geq 0$, sprich $\forall \mathbf{y} \in U: \Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} \rangle = 0$.

Für $\mathbb{K} = \mathbb{C}$ gilt $\mathbf{y} \in U \Rightarrow i\mathbf{y} \in U$, womit $\forall \mathbf{y} \in U: \Re \langle \mathbf{x} - \mathbf{y}_0, i\mathbf{y} \rangle = -\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} \rangle = \Im \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} \rangle$, womit folgt $\forall \mathbf{y} \in U: \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} \rangle = 0$.

ii' \Rightarrow ii:

$\langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} \rangle = 0 \forall \mathbf{y} \in U$, es gilt $\mathbf{y} \in U \Rightarrow \mathbf{y} - \mathbf{y}_0 \in U$ woraus folgt $\langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle = 0 \forall \mathbf{y} \in U$ und $\Re \langle \mathbf{x} - \mathbf{y}_0, \mathbf{y} - \mathbf{y}_0 \rangle = 0 \forall \mathbf{y} \in U$. \square

Wir folgern:

- i Wenn K konvex ist, dann hat das Optimierungsproblem $\min \|\mathbf{x} - \mathbf{y}\|$ für $\mathbf{y} \in K$ höchstens eine Lösung
- ii Wenn U ein Unterraum ist, dann gibt es höchstens ein $\mathbf{y}_0 \in U: \mathbf{x} - \mathbf{y}_0 \perp U$

Definition 8.13: Orthogonalprojektion

Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und U ein Unterraum, sodass $V = U \dot{+} U^\perp$, dann heißen die Abbildungen $\pi_U: V \rightarrow U$ und $\pi_{U^\perp}: V \rightarrow U^\perp$ Orthogonalprojektionen.

Wir wissen aus der allgemeinen Theorie der direkten Summen:

- i $\mathbf{x} \in U \Leftrightarrow \pi_U(\mathbf{x}) = \mathbf{x} \Leftrightarrow \pi_{U^\perp}(\mathbf{x}) = \mathbf{0}$
- ii $\pi_{U^\perp} = \mathbf{I} - \pi_U$
- iii $\pi_U(\pi_U(\mathbf{x})) = \pi_U(\mathbf{x})$ bzw. $\pi_U \circ \pi_U = \pi_U$, wir nennen π_U idempotent
- iv π_U ist linear

Satz 8.16: Skalarprodukt mit Orthogonalprojektionen

Sei $V = U \dot{+} U^\perp$, dann gilt:

- i $\forall \mathbf{x}, \mathbf{y} \in V: \langle \mathbf{x}, \pi_U(\mathbf{y}) \rangle = \langle \pi_U(\mathbf{x}), \mathbf{y} \rangle = \langle \pi_U(\mathbf{x}), \pi_U(\mathbf{y}) \rangle$
- ii $\forall \mathbf{x} \in V: \|\pi_U(\mathbf{x})\| \leq \|\mathbf{x}\|$, bzw. für $\mathbf{x} \in U$ gilt natürlich $\|\pi_U(\mathbf{x})\| = \|\mathbf{x}\|$

Beweis für Satz 8.16. Zu i:

$$\begin{aligned} \mathbf{x} &= \pi_U(\mathbf{x}) + \pi_{U^\perp}(\mathbf{x}) & \mathbf{y} &= \pi_U(\mathbf{y}) + \pi_{U^\perp}(\mathbf{y}) \\ \langle \mathbf{x}, \pi_U(\mathbf{y}) \rangle &= \langle \pi_U(\mathbf{x}) + \pi_{U^\perp}(\mathbf{x}), \pi_U(\mathbf{y}) \rangle = \langle \pi_U(\mathbf{x}), \pi_U(\mathbf{y}) \rangle + \langle \pi_{U^\perp}(\mathbf{x}), \pi_U(\mathbf{y}) \rangle = \langle \pi_U(\mathbf{x}), \pi_U(\mathbf{y}) \rangle \end{aligned}$$

Zu ii:

$$\begin{aligned} \|\mathbf{x}\|^2 &= \|\pi_U(\mathbf{x}) + \pi_{U^\perp}(\mathbf{x})\|^2 = \|\pi_U(\mathbf{x})\|^2 + \|\pi_{U^\perp}(\mathbf{x})\|^2 \geq \|\pi_U(\mathbf{x})\|^2 \\ \mathbf{x} \in U &\Rightarrow \pi_{U^\perp}(\mathbf{x}) = \mathbf{0} \Rightarrow \|\pi_U(\mathbf{x})\|^2 + \|\pi_{U^\perp}(\mathbf{x})\|^2 = \|\pi_U(\mathbf{x})\|^2 \end{aligned}$$

□

Im folgenden verwenden wir $[\mathbf{x}, \mathbf{y}] = \overline{\langle \mathbf{x}, \mathbf{y} \rangle} = \langle \mathbf{y}, \mathbf{x} \rangle$, sprich in \mathbb{C}^n gilt $[\mathbf{x}, \mathbf{y}] = \sum_{i=1}^n \bar{x}_i y_i = \mathbf{x}^* \mathbf{y}$

Definition 8.14: Gram-Matrix

Seien $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$, dann heißt die Matrix:

$$\text{Gram}(\mathbf{v}_1, \dots, \mathbf{v}_m) = \begin{bmatrix} [\mathbf{v}_1, \mathbf{v}_1] & [\mathbf{v}_1, \mathbf{v}_2] & \cdots & [\mathbf{v}_1, \mathbf{v}_m] \\ [\mathbf{v}_2, \mathbf{v}_1] & [\mathbf{v}_2, \mathbf{v}_2] & \cdots & [\mathbf{v}_2, \mathbf{v}_m] \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{v}_m, \mathbf{v}_1] & [\mathbf{v}_m, \mathbf{v}_2] & \cdots & [\mathbf{v}_m, \mathbf{v}_m] \end{bmatrix}$$

die Gram-Matrix, benannt nach Jørgen Pedersen [Gram](#), 1850-1916

Wenn $V = \mathbb{C}^n$ und $[\mathbf{x}, \mathbf{y}] = \mathbf{x}^* \mathbf{y}$:

$$\mathbf{M} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_m] \Rightarrow \mathbf{G} = \text{Gram}(\mathbf{v}_1, \dots, \mathbf{v}_m) = \mathbf{M}^* \mathbf{M}$$

Satz 8.17: Eigenschaften der Gram-Matrix

Sei $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ und $\mathbf{G} = \text{Gram}(\mathbf{v}_1, \dots, \mathbf{v}_m)$, dann gilt:

- i \mathbf{G} ist hermitesch und positiv semidefinit, für $\boldsymbol{\xi} \in \mathbb{K}^n$, gilt:

$$\boldsymbol{\xi}^* \mathbf{G} \boldsymbol{\xi} = \left\| \sum_{i=1}^m \xi_i \mathbf{v}_i \right\|^2$$

- ii $\boldsymbol{\xi} \in \ker \mathbf{G}$ wenn

$$\sum_{i=1}^m \xi_i \mathbf{v}_i = \mathbf{0}$$

- iii \mathbf{G} ist positiv definit, wenn $\mathbf{v}_1, \dots, \mathbf{v}_m$ linear unabhängig sind

Beweis für Satz 8.17. Zu i:

$$G_{ij} = [\mathbf{v}_i, \mathbf{v}_j] = \overline{[\mathbf{v}_j, \mathbf{v}_i]} = \overline{G_{ji}}$$

$$\begin{aligned} \boldsymbol{\xi}^* \mathbf{G} \boldsymbol{\xi} &= \sum_{i,j=1}^m \bar{\xi}_i G_{ij} \xi_j = \sum_{i,j=1}^m \bar{\xi}_i [\mathbf{v}_i, \mathbf{v}_j] \xi_j = \sum_{i,j=1}^m [\xi_i \mathbf{v}_i, \xi_j \mathbf{v}_j] \\ &= \left[\sum_{i=1}^m \xi_i \mathbf{v}_i, \sum_{j=1}^m \xi_j \mathbf{v}_j \right] = \left\| \sum_{i=1}^m \xi_i \mathbf{v}_i \right\|^2 \geq 0 \end{aligned}$$

Zu ii:

Sei $\boldsymbol{\xi} \in \ker \mathbf{G}$, sprich $\mathbf{G} \boldsymbol{\xi} = \mathbf{0}$, womit $\boldsymbol{\xi}^* \mathbf{G} \boldsymbol{\xi} = 0$, daher auch:

$$\sum_{i=1}^m \xi_i \mathbf{v}_i = \mathbf{0}$$

Angenommen $\sum_{i=1}^m \xi_i \mathbf{v}_i = \mathbf{0}$, dann:

$$(\mathbf{G} \boldsymbol{\xi})_i = \sum_{j=1}^m G_{ij} \xi_j = \sum_{j=1}^m [\mathbf{v}_i, \mathbf{v}_j] \xi_j = \left[\mathbf{v}_i, \sum_{j=1}^m \xi_j \mathbf{v}_j \right] = 0$$

Somit $\mathbf{G} \boldsymbol{\xi} = \mathbf{0} \Leftrightarrow \boldsymbol{\xi} \in \ker \mathbf{G}$

Zu iii:

Angenommen $\mathbf{G} > 0$, sprich $\boldsymbol{\xi}^* \mathbf{G} \boldsymbol{\xi} > 0 \forall \boldsymbol{\xi} \neq \mathbf{0}$, das ist äquivalent zu

$$\begin{aligned} \left\| \sum_{i=1}^m \xi_i \mathbf{v}_i \right\|^2 &> 0 \forall \boldsymbol{\xi} \neq \mathbf{0} \\ \Leftrightarrow \sum_{i=1}^m \xi_i \mathbf{v}_i &\neq \mathbf{0} \forall \boldsymbol{\xi} \neq \mathbf{0} \end{aligned}$$

Somit müssen $\mathbf{v}_1, \dots, \mathbf{v}_m$ linear unabhängig sein. □

Satz 8.18: Orthogonalprojektion aus der Gram-Matrix

Sei V ein Vektorraum mit Skalarprodukt $[\cdot, \cdot]$ und U ein endlich-dimensionaler Unterraum mit $\dim U = m < \infty$. Sei $\mathbf{u}_1, \dots, \mathbf{u}_m$ eine Basis von U . Wir bilden $\mathbf{G} = ([\mathbf{u}_i, \mathbf{u}_j])$ die Gramsche Matrix, dann ist für $\mathbf{x} \in V$ die Projektion auf U gegeben durch:

$$\pi_U(\mathbf{x}) = \sum_{j=1}^m \xi_j \mathbf{u}_j$$

wobei

$$\boldsymbol{\xi} = \mathbf{G}^{-1} \begin{bmatrix} [\mathbf{u}_1, \mathbf{x}] \\ \vdots \\ [\mathbf{u}_m, \mathbf{x}] \end{bmatrix}$$

Hier sei angemerkt, dass für $[\mathbf{u}_i, \mathbf{u}_j] = \delta_{ij}$, dann $\mathbf{G} = \mathbf{I}$ und $\pi_U(\mathbf{x}) = \sum_{i=1}^m [\mathbf{u}_i, \mathbf{x}] \mathbf{u}_i$.

Beweis für Satz 8.18. Sei $\mathbf{y}_0 = \sum_{i=1}^m \xi_i \mathbf{u}_i$, wobei

$$\boldsymbol{\xi} = \mathbf{G}^{-1} \begin{bmatrix} [\mathbf{u}_1, \mathbf{x}] \\ \vdots \\ [\mathbf{u}_m, \mathbf{x}] \end{bmatrix}$$

Zu zeigen ist:

$$\begin{aligned} \mathbf{x} - \mathbf{y}_0 &\in U^\perp = L(\mathbf{u}_1, \dots, \mathbf{u}_m)^\perp = \{\mathbf{u}_i, \dots, \mathbf{u}_m\}^\perp \\ &\Leftrightarrow [\mathbf{x} - \mathbf{y}_0, \mathbf{u}_i] = 0 \forall i \in \{1, \dots, m\} \\ &\Leftrightarrow [\mathbf{x}, \mathbf{u}_i] = [\mathbf{y}_0, \mathbf{u}_i] \forall i = 1, \dots, m \end{aligned}$$

$$\begin{aligned} [\mathbf{u}_i, \mathbf{y}_0] &= \left[\mathbf{u}_i, \sum_{j=1}^m \xi_j \mathbf{u}_j \right] = \sum_{j=1}^m [\mathbf{u}_i, \mathbf{u}_j] \xi_j = \sum_{j=1}^m G_{ij} \xi_j = [\mathbf{u}_i, \mathbf{x}] \\ &\Rightarrow [\mathbf{u}_i, \mathbf{x} - \mathbf{y}_0] = 0 \forall i = 1, \dots, m \Rightarrow \mathbf{x} - \mathbf{y}_0 \in U^\perp \end{aligned}$$

□

Betrachten wir ein Beispiel aus der angewandten Mathematik. Gesucht ist ein Polynom $p(t)$ vom Grad 2, sodass $\int_0^1 |t^3 - p(t)|^2 dt$ minimal ist. Wir setzen $V = \mathcal{C}[0, 1]$, mit

$$\langle f, g \rangle = \int_0^1 f(t)g(t) dt$$

$U = L(1, t, t^2)$, Gegeben ist $x = t^3 \notin U$. Gesucht $y_0 = p(t)$, sodass

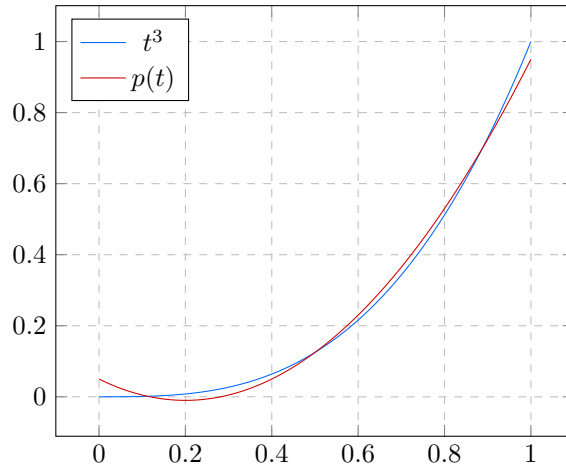
$$\|x - y_0\|^2 = \int_0^1 (t^3 - p(t))^2 dt$$

minimal ist. Wir bestimmen die Gram-Matrix:

$$\begin{aligned} G_{ij} &= \langle u_i, u_j \rangle = \int_0^1 t^{i-1} t^{j-1} dt = \int_0^1 t^{i+j-2} dt = \frac{1}{i+j-1} \\ &\Rightarrow \mathbf{G} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix} \end{aligned}$$

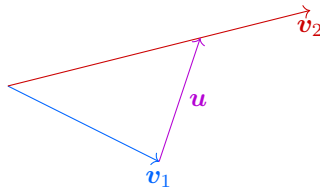
Wir nennen \mathbf{G} die **Hilbert-Matrix**. Gesucht ist jetzt $p(t) = \xi_1 + \xi_2 t + \xi_3 t^2$:

$$\begin{aligned} \boldsymbol{\xi} &= \mathbf{G}^{-1} \begin{bmatrix} \langle u_1, t^3 \rangle \\ \langle u_2, t^3 \rangle \\ \langle u_3, t^3 \rangle \end{bmatrix} \\ \langle u_i, t^3 \rangle &= \int_0^1 t^{i-1} t^3 dt = \int_0^1 t^{i+2} dt = \frac{1}{i+3} \\ &\Rightarrow \boldsymbol{\xi} = \mathbf{G}^{-1} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{5} \\ \frac{1}{6} \end{bmatrix} = \begin{bmatrix} \frac{1}{20} \\ -\frac{3}{5} \\ \frac{3}{2} \end{bmatrix} \\ &\Rightarrow p(t) = \frac{1}{20} - \frac{3}{5}t + \frac{3}{2}t^2 \end{aligned}$$

Abbildung 10: Vergleich von t^3 und $p(t)$ auf $[0, 1]$

8.2 Orthonormalisierungsverfahren

Wir wollen uns nun damit beschäftigen wie wir aus einer gegebenen Basis eine Orthonormalbasis konstruieren können. Im \mathbb{R}^2 können wir recht schnell ein geeignetes System finden:

Abbildung 11: Darstellung für die Bestimmung einer orthogonalen Basis des \mathbb{R}^2

Wir wissen $\mathbf{u} = \mathbf{v}_2 - \pi_1(\mathbf{v}_2)$, wobei $\pi_1(\mathbf{v}_2) \in \mathcal{L}(\mathbf{v}_1)$, womit $\mathbf{u} \in \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$. Nach dem Basis-Ergänzungssatz ist $(\mathbf{v}_1, \mathbf{u})$ immer noch eine Basis.

Korollar 8.2

Sei $U \subseteq V$ ein Unterraum des Vektorraumes V mit Skalarprodukt, mit der Orthonormalbasis $\mathbf{u}_1, \dots, \mathbf{u}_m$, dann gilt:

i $\forall \mathbf{v} \in V: \pi_U(\mathbf{v}) = \sum_{i=1}^m \langle \mathbf{v}, \mathbf{u}_i \rangle \mathbf{u}_i$

ii $\forall \mathbf{v} \in V: \sum_{i=1}^m |\langle \mathbf{v}, \mathbf{u}_i \rangle|^2 \leq \|\mathbf{v}\|^2$

Wir nennen ii die Besselsche Ungleichung, nach Friedrich [Bessel](#), 1784-1846

Für $\mathbf{v} \in U$ gilt $\sum_{i=1}^m |\langle \mathbf{v}, \mathbf{u}_i \rangle|^2 = \|\mathbf{v}\|^2$

Beweis für Korollar 8.2. $\mathbf{u} = \pi_U(\mathbf{v}) \Leftrightarrow \mathbf{u} \in U$ und $\mathbf{v} - \mathbf{u} \perp U$, also $\langle \mathbf{v} - \mathbf{u}, \mathbf{u}_i \rangle = 0 \forall i = 1, \dots, m$:

$$\begin{aligned} \langle \mathbf{v} - \mathbf{u}, \mathbf{u}_i \rangle &= \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle \mathbf{u}, \mathbf{u}_i \rangle = \langle \mathbf{v}, \mathbf{u}_i \rangle - \left\langle \sum_{j=1}^m \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j, \mathbf{u}_i \right\rangle \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - \sum_{j=1}^m \langle \mathbf{v}, \mathbf{u}_j \rangle \langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0 \end{aligned}$$

□

Satz 8.19: Gram-Schmidt Orthogonalisierung

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt und $(v_1, \dots, v_m) \subseteq V$ linear unabhängig, dann gibt es ein Orthonormalsystem (u_1, \dots, u_m) , sodass:

$$\mathfrak{L}(u_1, \dots, u_m) = \mathfrak{L}(v_1, \dots, v_m)$$

spricht, wenn (v_1, \dots, v_m) eine Basis von $U \subseteq V$ ist, dann ist (u_1, \dots, u_m) eine Orthonormalbasis von U .

Benannt nach **Gram** und **Schmidt** (1876-1959), allerdings wusste auch schon **Laplace** wie man eine Basis orthogonalisiert.

Beweis für Satz 8.19. Wir konstruieren u_i induktiv, beginnend mit u_1 :

$$u_1 = v_1 \frac{1}{\|v_1\|}$$

Für $k = 2, \dots, m$ gehen wir vor wie in Abbildung 11. Wir setzen:

$$\begin{aligned} \tilde{u}_2 &= v_2 - \langle v_2, u_1 \rangle u_1 = v_2 - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1 \\ u_2 &= \tilde{u}_2 \frac{1}{\|\tilde{u}_2\|} \end{aligned}$$

Allgemein:

$$\begin{aligned} \tilde{u}_k &= v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j \\ u_k &= \tilde{u}_k \frac{1}{\|\tilde{u}_k\|} \end{aligned}$$

Für $m = 1$ ist das Verfahren trivial, da $u_1 = v_1 \frac{1}{\|v_1\|}$ die selbe Linear Hülle hat. Wir führen nun einen Induktionsbeweis. Für den Schritt nehmen wir an, wir haben bereits $m-1$ Vektoren u_1, \dots, u_{m-1} bereits konstruiert, wobei $\mathfrak{L}(v_1, \dots, v_{m-1}) = \mathfrak{L}(u_1, \dots, u_{m-1})$ und $\langle u_i, u_j \rangle = \delta_{ij}$, dann gilt $v_m \notin \mathfrak{L}(u_1, \dots, u_{m-1})$.

$$\begin{aligned} \tilde{u}_m &= v_m - \underbrace{\sum_{k=1}^{m-1} \langle v_m, u_k \rangle u_k}_{=\pi_{U_{m-1}}(v_m)} \neq 0 \\ &\Rightarrow \tilde{u}_m \perp U_{m-1} \end{aligned}$$

Wir setzen nun $u_m = \tilde{u}_m \frac{1}{\|\tilde{u}_m\|}$, womit (u_1, \dots, u_m) ein Orthonormalsystem. Nach dem Basisaustauschsatz folgt:

$$v_m = \tilde{u}_m + \sum_{k=1}^{m-1} \langle v_m, u_k \rangle u_k \in \mathfrak{L}(u_1, \dots, u_m)$$

Es gilt somit auch $\mathfrak{L}(v_1, \dots, v_m) = \mathfrak{L}(u_1, \dots, u_m)$. □

Betrachten wir ein Beispiel. Sei $V = \mathbb{R}^3$ und $\langle x, y \rangle = x^T A y$ mit

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

Wir wollen die folgende Basis orthogonalisieren:

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

\mathbf{u}_1 ist bereits durch \mathbf{v}_1 gegeben:

$$\|\mathbf{v}_1\|^2 = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = a_{11} = 1 \Rightarrow \mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|} \mathbf{v}_1 = \mathbf{v}_1$$

Weiter mit \mathbf{u}_2 :

$$\begin{aligned} \tilde{\mathbf{u}}_2 &= \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{u}_1 \rangle \mathbf{u}_1 = \mathbf{v}_2 - \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ \|\tilde{\mathbf{u}}_2\|^2 &= \tilde{\mathbf{u}}_2 \mathbf{A} \tilde{\mathbf{u}}_2 = 2 \\ \Rightarrow \mathbf{u}_2 &= \frac{1}{\|\tilde{\mathbf{u}}_2\|} \tilde{\mathbf{u}}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

Zuletzt \mathbf{u}_3 :

$$\begin{aligned} \tilde{\mathbf{u}}_3 &= \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{v}_3, \mathbf{u}_2 \rangle \mathbf{u}_2 \\ &= \mathbf{v}_3 - \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{u}_1 - \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \tilde{\mathbf{u}}_2 \frac{1}{2} \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \mathbf{0} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ \|\tilde{\mathbf{u}}_3\|^2 &= \langle \tilde{\mathbf{u}}_3, \tilde{\mathbf{u}}_3 \rangle = 1 \Rightarrow \mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

Somit ist $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ eine Orthonormalbasis. Wäre die zu orthonormalisierende Familie nicht linear unabhängig, so gibt es zumindest ein $\mathbf{u}_k = \mathbf{0}$, diesen Vektor kann man aus dem Verfahren entfernen und mit den verbliebenen Vektoren fortfahren.

Mit der Orthogonalisierung erhalten wir eine Alternative zur Konstruktion der Orthogonalprojektion $\pi_U: V \rightarrow U$. Im Fall $V = \mathbb{C}^n$ oder \mathbb{R}^n hat die Orthogonalprojektion eine Matrixdarstellung $\pi_U(\mathbf{v}) = \mathbf{P}\mathbf{v}$. Im Fall einer Orthonormalbasis $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ von \mathbb{C}^n mit Standardskalarprodukt folgt:

$$\begin{aligned} \pi_U(\mathbf{v}) &= \sum_{k=1}^m \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k = \sum_{k=1}^m \mathbf{u}_k \mathbf{u}_k^* \mathbf{v} = \left(\sum_{k=1}^m \mathbf{u}_k \mathbf{u}_k^* \right) \mathbf{v} \\ \mathbf{P}_k &= \mathbf{u}_k \mathbf{u}_k^* \Rightarrow \mathbf{P}_k \mathbf{v} = \langle \mathbf{v}, \mathbf{u}_k \rangle \mathbf{u}_k \end{aligned}$$

Betrachten wir erneut das Polynombeispiel⁴ und orthonormalisieren die Basis:

$$\|v_1\|^2 = \int_0^1 dx = 1 \Rightarrow u_1 = 1$$

⁴Annäherung von x^3 durch $p(x) = ax^2 + bx + c$

$$\begin{aligned}
\tilde{u}_2 &= v_2 - \langle v_2, u_1 \rangle u_1 = x - \int_0^x x \, dx = x - \frac{1}{2} \\
\|\tilde{u}_2\| &= \int_0^2 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12} \\
\Rightarrow u_2 &= \sqrt{12} \left(x - \frac{1}{2}\right) \\
\tilde{u}_3 &= x^2 - x + \frac{1}{6} \Rightarrow u_3 = \sqrt{180} \left(x^2 - x + \frac{1}{6}\right) \\
\Rightarrow \pi_U(x^3) &= \langle x^3, u_1 \rangle u_1 + \langle x^3, u_2 \rangle u_2 + \langle x^3, u_3 \rangle u_3 \\
&= \int_0^1 x^3 dx + 12 \int_0^2 x^3 \left(x - \frac{1}{2}\right) dx \left(x - \frac{1}{2}\right) + \langle x^3, u_3 \rangle u_3 \\
&= \frac{1}{20} - \frac{3}{5}x + \frac{3}{2}x^2
\end{aligned}$$

Satz 8.20: Darstellungssatz von Riesz

Sei V ein endlich-dimensionaler Vektorraum mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$, dann ist $V \rightarrow V^* = \text{Hom}(V, \mathbb{K})$ mit $\mathbf{y} \mapsto T_{\mathbf{y}}: V \rightarrow \mathbb{K}$ mit $T_{\mathbf{y}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$ ein antilinearer Isomorphismus.

Dieser Satz stammt von [F. Riesz](#).

Das Skalarprodukt $\langle \mathbf{x}, \mathbf{y} \rangle$ ist sesquilinear, sprich linear in \mathbf{x} und antilinear in \mathbf{y} . Für ein fixes \mathbf{y} ist die Abbildung $T_{\mathbf{y}}(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$, sprich $\forall \mathbf{y} \in V: T_{\mathbf{y}} \in \text{Hom}(V, \mathbb{K})$. Satz 8.20 besagt, für jedes lineare Funktional $f: V \rightarrow \mathbb{K}$ gibt es ein $\mathbf{y} \in V$ sodass $f(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$.

Sei etwa $V = \mathcal{C}[0, 1]$ und $\Phi(f) = f(0)$, gibt es also eine Funktion y sodass

$$\int_0^1 f(x)y(x) \, dx = f(0)$$

Es gibt tatsächlich ein mathematisches Objekt, die sogenannte **Dirac-Delta-Distribution**⁵ $\delta(x)$, die folgendermaßen „definiert“ wird:

$$\begin{aligned}
\int_{\mathbb{R}} \delta(x) \, dx &= 1 \\
\delta(x) &= \begin{cases} 0 & x \neq 0 \\ x & x = 0 \end{cases}
\end{aligned}$$

Beweis für Satz 8.20.

$$\begin{aligned}
T_{\lambda \mathbf{y}_1 + \mu \mathbf{y}_2}(\mathbf{x}) &= \langle \mathbf{x}, \lambda \mathbf{y}_1 + \mu \mathbf{y}_2 \rangle = \langle \mathbf{x}, \mathbf{y}_1 \rangle \bar{\lambda} + \langle \mathbf{x}, \mathbf{y}_2 \rangle \bar{\mu} \\
&= \bar{\lambda} T_{\mathbf{y}_1}(\mathbf{x}) + \bar{\mu} T_{\mathbf{y}_2}(\mathbf{x}) \\
&= (\bar{\lambda} T_{\mathbf{y}_1} + \bar{\mu} T_{\mathbf{y}_2})(\mathbf{x})
\end{aligned}$$

Wir sehen, dass $T_{\mathbf{y}}$ antilinear ist. Es verbleibt zu zeigen, dass $T_{\mathbf{y}}$ ein Isomorphismus ist. Wir beginnen mit der Injektivität. Wenn $T_{\mathbf{y}_1} = T_{\mathbf{y}_2} \Rightarrow \mathbf{y}_1 = \mathbf{y}_2$, bzw:

$$T_{\mathbf{y}_1} - T_{\mathbf{y}_2} = 0 \Rightarrow T_{\mathbf{y}_1 - \mathbf{y}_2} = 0$$

Es folgt daher $\forall \mathbf{x} \in V: T_{\mathbf{y}_1 - \mathbf{y}_2}(\mathbf{x}) = 0$. Insbesondere $\mathbf{x} = \mathbf{y}_1 - \mathbf{y}_2$:

$$T_{\mathbf{y}_1 - \mathbf{y}_2}(\mathbf{y}_1 - \mathbf{y}_2) = \langle \mathbf{y}_1 - \mathbf{y}_2, \mathbf{y}_1 - \mathbf{y}_2 \rangle = \|\mathbf{y}_1 - \mathbf{y}_2\|^2 = 0 \Rightarrow \mathbf{y}_1 - \mathbf{y}_2 = \mathbf{0} \Leftrightarrow \mathbf{y}_1 = \mathbf{y}_2$$

Zur Surjektivität. Sei $f \in V^* = \text{Hom}(V, \mathbb{K})$, und $\mathbf{u}_1, \dots, \mathbf{u}_n$ eine Orthonormalbasis von V , dann ist

$$\mathbf{y} = \sum_{i=1}^n \overline{f(\mathbf{u}_i)} \mathbf{u}_i$$

⁵der Begriff der Distribution stammt von [Schwarz](#)

Es gilt somit für $x \in V$:

$$\begin{aligned} T_y(x) &= \langle x, y \rangle = \left\langle x, \sum_{i=1}^n \overline{f(u_i)} u_i \right\rangle = \sum_{i=1}^n f(u_i) \langle x, u_i \rangle \\ &= \sum_{i=1}^n f(\langle x, u_i \rangle u_i) = f\left(\sum_{i=1}^n \langle x, u_i \rangle u_i\right) = f(x) \end{aligned}$$

□

An dieser Stelle sei angemerkt, dass dieser Satz im Allgemeinen für $\dim V = \infty$ nicht, außer in sogenannten Hilbert-Räumen, benannt nach David Hilbert.

Einige weitere wichtige Bemerkungen zu Darstellungssatz von Riesz:

- i $v = 0 \Leftrightarrow \forall w \in V: \langle v, w \rangle = 0 \Leftrightarrow \forall f \in V^*: f(v) = 0$
- ii $\|v\| = \sup\{|\langle v, w \rangle| : \|w\| \leq 1\}$

Satz 8.21: Anwendung für den Satz von F. Riesz

Seien $(V, \langle \cdot, \cdot \rangle_V)$ und $(W, \langle \cdot, \cdot \rangle_W)$ endlich-dimensionale Vektorräume jeweils mit Skalarprodukt. Sei $g \in \text{Hom}(V, W)$:

- i für ein fixes $w \in W$ ist die Abbildung $V \rightarrow \mathbb{K}$ mit $v \mapsto g(v) \mapsto \langle g(v), w \rangle$ linear, bzw. $T_w \circ g \in V^*$
- ii $\forall w \in W: \exists! u \in V: \forall v \in V: \langle g(v), w \rangle = \langle v, u \rangle$, wir schreiben $g^*(w) = u$
- iii die Abbildung $g^*: W \rightarrow V$ ist linear und heißt **adjungierte Abbildung**
- iv die Abbildung $\text{Hom}(V, W) \rightarrow \text{Hom}(W, V)$ mit $g \mapsto g^*$ ist antilinear und es gilt $g^{**} = g$.

Beweis für Satz 8.21. Zu ii: Wir wenden den Darstellungssatz von Riesz an.

Zu iii: Zu zeigen ist, dass $g^*(\lambda w + \mu w') = \lambda g^*(w) + \mu g^*(w')$. $g^*(\lambda w + \mu w')$ ist jenes $u \in V$ für das gilt $\langle v, u \rangle = \langle g(v), \lambda w + \mu w' \rangle$.

$$\begin{aligned} \langle g(v), \lambda w + \mu w' \rangle &= \bar{\lambda} \langle g(v), w \rangle + \bar{\mu} \langle g(v), w' \rangle = \bar{\lambda} \langle v, g^*(w) \rangle + \bar{\mu} \langle v, g^*(w') \rangle \\ &= \langle v, \lambda g^*(w) \rangle + \langle v, \mu g^*(w') \rangle = \langle v, \lambda g^*(w) + \mu g^*(w') \rangle \Rightarrow g^*(\lambda w + \mu w') = \lambda g^*(w) + \mu g^*(w') \end{aligned}$$

Zu iv: Seien $g, h \in \text{Hom}(V, W)$, dann ist $\lambda g + \mu h \in \text{Hom}(V, W)$. Wir zeigen nun, dass $(\lambda g + \mu h)^* = \bar{\lambda} g^* + \bar{\mu} h^*$:

$$\begin{aligned} \langle v, (\lambda g + \mu h)^*(w) \rangle &= \langle (\lambda g + \mu h)(v), w \rangle = \lambda \langle g(v), w \rangle + \mu \langle h(v), w \rangle \\ &= \lambda \langle v, g^*(w) \rangle + \mu \langle v, h^*(w) \rangle = \langle v, \bar{\lambda} g^*(w) + \bar{\mu} h^*(w) \rangle \end{aligned}$$

Es verbleibt $g^{**} = g$, sprich es muss gelten $\forall v \in V: g(v) = g^{**}(v)$, bzw. $\forall v \in V: \forall w \in W: \langle g(v), w \rangle = \langle g^{**}(v), w \rangle$:

$$\langle g(v), w \rangle = \langle v, g^*(w) \rangle = \overline{\langle g^*(w), v \rangle} = \overline{\langle w, g^{**}(v) \rangle} = \langle g^{**}(v), w \rangle$$

□

Wen $\dim V = \infty$, dann hat nicht jede Abbildung eine adjungierte. Sei $V = \mathbb{R}[x]$ mit $\langle f, g \rangle = \int_0^1 f g \, dx$. Wir betrachten $D: V \rightarrow V$ mit $p \mapsto p'$. Angenommen es gibt $D^*: \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ existiert, also

$$\forall p, q \in \mathbb{R}[x]: \langle p', q \rangle = \langle p, D^*(q) \rangle$$

Was ist $D^*(x)$:

$$\langle x^n, D^*(x) \rangle = \langle D(x^n), x \rangle = \langle nx^{n-1}, x \rangle = \int_0^1 nx^{n-1}x \, dx = n \int_0^1 x^n \, dx = \frac{n}{n+1} \xrightarrow{n \rightarrow \infty} 0$$

Sei $D^*(x) = q(x) \in \mathbb{R}[x]$. Sei $M = \max_{x \in [0,1]} |q(x)| < \infty$, es muss gelten:

$$|\langle x^n, D^*(x) \rangle| = \left| \int_0^1 x^n q(x) \, dx \right| \leq \int_0^1 |x^n q(x)| \, dx \leq \int_0^1 x^n M \, dx = \frac{M}{n+1} \xrightarrow{n \rightarrow \infty} 0$$

Das ist ein Widerspruch.

Wie können wir $\langle g(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, g^*(\mathbf{y}) \rangle$ für Matrizen auslegen?

$$\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{A} \mathbf{x} = \overline{\mathbf{y}^T \mathbf{A} \mathbf{x}} = \overline{\mathbf{y}^T \mathbf{A} \mathbf{x}} = \overline{(\mathbf{A}^T \mathbf{y})^T \mathbf{x}} = (\overline{\mathbf{A}^T \mathbf{y}})^* \mathbf{x} = \langle \mathbf{x}, \overline{\mathbf{A}^T \mathbf{y}} \rangle$$

Satz 8.22: Adjungiertheit in Bezug auf Orthonormalbasen

Seien $B \subseteq V$ und $B' \subseteq W$ Orthonormalbasen, und $f \in \text{Hom}(V, W)$, dann gilt für die Matrixdarstellung

$$\Phi_B^{B'}(f^*) = \overline{\Phi_{B'}^B(f)}^T = \Phi_{B'}^B(f)^*$$

Beweis für Satz 8.22. Sei $\mathbf{A} = \Phi_B^{B'}(f^*) = [\Phi_B(f^*(\mathbf{b}_1)) \cdots \Phi_B(f^*(\mathbf{b}_n))]$, also ist a_{ij} ist die i -te Koordinate von $f^*(\mathbf{b}'_j)$, also $\langle f^*(\mathbf{b}'_j), \mathbf{b}_i \rangle$

$$\langle f^*(\mathbf{b}'_j), \mathbf{b}_i \rangle = \overline{\langle \mathbf{b}_i, f^*(\mathbf{b}'_j) \rangle} = \overline{\langle f(\mathbf{b}_i), \mathbf{b}'_j \rangle}$$

□

Satz 8.23: Eigenschaften der Adjungierten Abbildung

Seien U, V, W endlich-dimensional und $f \in \text{Hom}(U, V)$ und $g \in \text{Hom}(V, W)$:

- i $(g \circ f)^* = f^* \circ g^*$
- ii $f^{**} = f$
- iii $\ker f = (\text{im } f^*)^\perp$
- iv $\text{im } f = (\ker f^*)^\perp$
- v f ist injektiv wenn f^* surjektiv ist
- vi f ist surjektiv wenn f^* injektiv ist

Beweis für Satz 8.23. Zu i:

$$\begin{aligned} (g \circ f)(\mathbf{w}) &= T_{\mathbf{w}} \circ g \circ f \\ \Rightarrow \langle \mathbf{u}, (g \circ f)^*(\mathbf{w}) \rangle &= \langle (g \circ f)(\mathbf{u}), \mathbf{w} \rangle = \langle f(\mathbf{u}), g^*(\mathbf{w}) \rangle = \langle \mathbf{u}, f^*(g^*(\mathbf{w})) \rangle = \langle \mathbf{u}, (f^* \circ g^*)(\mathbf{w}) \rangle \end{aligned}$$

ii haben wir bereits bewiesen.

Zu iii: Sei $\mathbf{u} \in \ker f$, sprich $f(\mathbf{u}) = \mathbf{0} \Leftrightarrow \forall \mathbf{v} \in V: \langle f(\mathbf{u}), \mathbf{v} \rangle = 0$:

$$\Rightarrow \forall \mathbf{v} \in V: \langle \mathbf{u}, f^*(\mathbf{v}) \rangle = 0 \Leftrightarrow \mathbf{u} \perp f^*(\mathbf{v}) \forall \mathbf{v} \in V \Leftrightarrow \mathbf{u} \perp f^*(V) = \text{im } f^*$$

Zu iv:

$$\ker f^* = (\text{im } f^{**})^\perp = (\text{im } f)^\perp \Rightarrow (\ker f^*)^\perp = \text{im } (f)^\perp{}^\perp = \text{im } f$$

Zu v: f ist injektiv, wenn $\ker f = \{\mathbf{0}\} \Leftrightarrow (\text{im } f^*)^\perp = \{\mathbf{0}\} \Leftrightarrow \text{im } f^* = U$, also f^* ist surjektiv.

vi geht analog zu v.

□

Definition 8.15: Wichtige Klassen linearer Abbildungen

Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$

- i eine lineare Abbildung $f \in \text{End}(V)$ heißt selbstadjungiert, wenn $f = f^*$ (bei Matrizen also $\mathbf{A} = \mathbf{A}^*$, = hermitesch)
- ii $f \in \text{Hom}(V, W)$ heißt lineare Isometrie, wenn gilt: $\forall \mathbf{x}, \mathbf{y} \in V: \langle f(\mathbf{x}), f(\mathbf{y}) \rangle_W = \langle f(\mathbf{x}), \mathbf{y} \rangle_V$, insbesondere $\|f(\mathbf{x})\| = \|\mathbf{x}\|$

Einige Anmerkungen zu Definition 8.15:

i Isometrien sind injektiv

ii wenn $\dim V < \infty$ und $f \in \text{End}(V)$ ist eine Isometrie, dann ist f regulär und $f^{-1} = f^*$

iii wenn $\dim V = \infty$, dann muss das nicht gelten

Beweis dazu. Zu i: wenn $f(\mathbf{x}) = \mathbf{0}$, dann $\|\mathbf{x}\| = \|f(\mathbf{x})\| = 0 \Rightarrow \mathbf{x} = \mathbf{0}$. Zu ii:

$$\forall \mathbf{x}, \mathbf{y} \in V: \langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$$

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle \mathbf{x}, f^*(f(\mathbf{y})) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \Rightarrow f^*(f(\mathbf{y})) = \mathbf{y} \Leftrightarrow f^* \circ f = \text{id} \Rightarrow f^* = f^{-1}$$

Zu iii erbringen wir ein Gegenbeispiel. Sei $V = \ell^2$. Sei $S: V \rightarrow V$ mit $S((\xi_1, \xi_2, \dots)) = (0, \xi_1, \xi_2, \dots)$

$$\langle Sx, y \rangle = \langle (0, \xi_1, \xi_2, \dots), (\eta_1, \eta_2, \dots) \rangle = \sum_{i=1}^{\infty} \xi_i \eta_{i+1} = \langle (\xi_1, \xi_2, \dots), \eta_2, \eta_3, \dots \rangle = \langle x, S^*(y) \rangle$$

$$S^*S((\xi_1, \xi_2, \dots)) = S^*(0, \xi_1, \xi_2, \dots) = (\xi_1, \xi_2, \dots) \Rightarrow S^*S = \text{id}$$

$$SS^*((\xi_1, \xi_2, \dots)) = S((\xi_2, \xi_3, \dots)) = (0, \xi_2, \xi_3, \dots) \Rightarrow SS^* \neq \text{id}$$

Mit $P_1: V \rightarrow V$ und $(\xi_1, \xi_2, \dots) \mapsto (\xi_1, \dots)$ gilt $P_1 + SS^* = \text{id}$.

Definition 8.16: Unitäre Matrizen

i Eine Matrix $\mathbf{U} \in \mathbb{C}^{n \times n}$ heißt unitär, wenn $\mathbf{U}^* = \mathbf{U}^{-1}$

ii $\mathbf{U} \in \mathbb{R}^{n \times n}$ heißt orthogonal wenn $\mathbf{U}^* = \mathbf{U}^{-1}$

Satz 8.24: Eigenschaften unitärer Matrizen

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, dann sind folgende Aussagen äquivalent:

i \mathbf{A} ist unitär

ii $\forall \mathbf{x} \in \mathbb{C}^n: \|\mathbf{Ax}\| = \|\mathbf{x}\|$ (Isometrie)

iii $\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n: \Re\{\langle \mathbf{Ax}, \mathbf{Ay} \rangle\} = \Re\{\langle \mathbf{x}, \mathbf{y} \rangle\}$

iv $\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n: \langle \mathbf{Ax}, \mathbf{Ay} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$

v Die Spalten von \mathbf{A} bilden eine Orthonormalbasis

Wobei $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x}$

Beweis. i \Rightarrow ii:

$$\|\mathbf{Ax}\|^2 = \langle \mathbf{Ax}, \mathbf{Ax} \rangle = \langle \mathbf{x}, \mathbf{A}^* \mathbf{Ax} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$$

ii \Rightarrow iii:

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\Re\{\langle \mathbf{x}, \mathbf{y} \rangle\} + \|\mathbf{y}\|^2$$

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\Re\{\langle \mathbf{x}, \mathbf{y} \rangle\} + \|\mathbf{y}\|^2$$

$$\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 = 4\Re\{\langle \mathbf{x}, \mathbf{y} \rangle\}$$

$$\|\mathbf{A}(\mathbf{x} + \mathbf{y})\|^2 - \|\mathbf{A}(\mathbf{x} - \mathbf{y})\|^2 = \|\mathbf{Ax} + \mathbf{Ay}\|^2 - \|\mathbf{Ax} - \mathbf{Ay}\|^2 = 4\Re\{\langle \mathbf{Ax}, \mathbf{Ax} \rangle\}$$

iii \Rightarrow iv:

$$\Re\{\langle \mathbf{Ax}, \mathbf{Ax} \rangle\} = \Re\{\langle \mathbf{x}, \mathbf{y} \rangle\}$$

$$\Rightarrow \Re\{\langle \mathbf{Ax}, \mathbf{Ay} \rangle\} = \Re\{\langle \mathbf{x}, \mathbf{y} \rangle\}$$

$$\Rightarrow \Re\{-i \langle \mathbf{Ax}, \mathbf{Ax} \rangle\} = \Re\{-i \langle \mathbf{x}, \mathbf{y} \rangle\} = \Im\{\langle \mathbf{x}, \mathbf{y} \rangle\}$$

iv \Rightarrow v:

Wir wissen, dass $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$:

$$\langle \mathbf{Ae}_i, \mathbf{Ae}_j \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$$

$$\mathbf{Ae}_i = \mathbf{s}_i(\mathbf{A}) \rightsquigarrow \langle \mathbf{s}_i(\mathbf{A}), \mathbf{s}_j(\mathbf{A}) \rangle = \delta_{ij}$$

$v \Rightarrow i$:

$$(A^*A)_{ij} = s_i(A)^* s_j(A) = \langle s_j(A), s_i(A) \rangle = \delta_{ij}$$

□

Satz 8.25: Wichtige Gruppen von Matrizen

Die folgenden Mengen sind Gruppen:

- $\mathcal{O}(n) = \{U \in \mathbb{R}^{n \times n} : U^T U = I\}$ orthogonale Gruppe
- $\mathcal{U}(n) = \{U \in \mathbb{C}^{n \times n} : |U^* U = I\}$ unitäre Gruppe
- $SO(n) = \{U \in \mathcal{O}(n) : \det U = 1\}$ spezielle orthogonale Gruppe
- $SU(n) = \{U \in \mathcal{U}(n) : \det U = 1\}$ spezielle unitäre Gruppe

Beweis. Seien $U_1, U_2 \in \mathcal{U}(n)$:

$$\begin{aligned} (U_1 U_2)^* (U_1 U_2) &= U_2^* U_1^* U_1 U_2 \\ (U^{-1})^* (U_1^{-1}) &= U_1 U_1^{-1} = I \Rightarrow U_1^{-1} \in \mathcal{U}(n) \end{aligned}$$

Seien $U_1, U_2 \in SU(n)$:

$$\det(U_1 U_2) = \det(U_1) \det(U_2) = 1$$

□

8.3 Nachtrag zu Kapitel 5

$V^* = \text{Hom}(V, \mathbb{K})$ heißt der Dualraum von V , bzw. ist der Raum der linearen Funktionalen $V \rightarrow \mathbb{K}$. Wir schreiben: $u^* \in V^*$ mit $u^*(v) = \langle v^*, v \rangle$, also ist u eine Bilinearform $V^* \times V \rightarrow \mathbb{K}$ mit $(\langle 0|u|0 \rangle^*, v) \mapsto u^*(v) = \langle v^*, v \rangle$

Hier ist jedoch zu beachten, dass u^* linear in beiden Argumenten ist (auch über \mathbb{C}).

Satz

Sei V ein Vektorraum über \mathbb{K} :

- wenn $\dim V = n < \infty$ dann ist $\dim V^* = n$
 - wenn (b_1, \dots, b_n) eine Basis von V ist, dann bilden die Funktionalen $b_k^*: V \rightarrow \mathbb{K}$ mit $b_i \rightarrow \delta_{ik}$ eine Basis von V^* , die sogenannte duale Basis
 - wenn $b_k^*(b_i) = \delta_{ij}$ ein duales Paar von Basen bilden, dann gilt $\forall v \in V: v = \sum_{i=1}^n b_i^*(v) b_i = \sum_{i=1}^n \langle b_i^*, v \rangle b_i$ bzw. $\forall v^* \in V^*: v^* = \sum_{i=1}^n \langle v^*, b_i \rangle b_i^*$
- wenn $\dim V = \infty$, dann stimmt das nicht

Beweis. Sei (v_i) eine Basis von W und (w_j) eine Basis von W , wir betrachten $f_{ij}: V \rightarrow W$ mit $v_i \mapsto w_j$ und $v_k \mapsto 0$ für $k \neq i$

Wenn $v = \sum_{i=1}^n \lambda_i b_i$, dann $b_k^*(v) = \sum_{i=1}^n \lambda_i b_k^*(b_i) = \lambda_k$

Zu ii:

Sei $(b_i)_{i \in I}$ eine Basis von V , wir betrachten $f: V \rightarrow \mathbb{K}$ mit $f(b_i) = 1$ für $i \in I$:

$$f\left(\sum_{k=1}^n \lambda_k b_{i_k}\right) = \sum_{k=1}^n \lambda_k$$

Sei $b_i^*: V \rightarrow \mathbb{K}$ mit $b_i^*(b_j) = \delta_{ij}$, angenommen $f \in \mathcal{L}(b_i^* | i \in I)$, dann $\exists i_1, \dots, i_n: \exists \mu_1, \dots, \mu_n$ sodass:

$$f = \sum_{k=1}^n \mu_k b_{i_k}^*$$

Sei $j \notin \{i_1, \dots, i_n\}$, dann :

$$f(\mathbf{b}_i) = \sum_{k=1}^n \mu_k \mathbf{b}_{i_k}^*(\mathbf{b}_j) = 0$$

Wir müssten also $f = \sum_{i \in I} \mathbf{b}_i^*$ setzen. Für jede $J \subseteq I$ ist das Funktional

$$f_J: \mathbf{b}_i \mapsto \begin{cases} 1 & i \in J \\ 0 & i \notin J \end{cases}$$

□

Betrachten wir ein konkretes Beispiel. Sei $V = \mathbb{K}[x]$, also:

$$V = \left\{ \sum_{k=0}^n a_k x^k \mid n \in \mathbb{N}, a_k \in \mathbb{K} \right\}$$

Als Basis verwenden wir $(x^k)_{k=0}^\infty$ womit $\dim V = \aleph_0$. Sei $f \in V^*$, eindeutig bestimmt durch $\langle f, x^k \rangle = a_k$, d.h:

$$f = \sum_{k=0}^{\infty} a_k t^k$$

Wir nennen f eine formale Potenzreihe, wobei alle Rechenoperationen wohldefiniert sind. $\langle t^k, x^l \rangle = \delta_{kl}$

$$\mathbb{K}[[x]] = \left\{ \sum_{k=0}^{\infty} a_k t^k \mid a_k \in \mathbb{K} \right\}$$

Übung: $\dim K[[x]] \geq |\mathbb{R}|$

Sei $f \in \text{Hom}(V, W)$ und $w^*: W \rightarrow \mathbb{K}$, dann ist $f^T(w^*) = w^* \circ f \in V^*$: $\langle f^T(w^*), \mathbf{v} \rangle = \langle w^*, f(\mathbf{v}) \rangle$, dabei gilt $f^T \in \text{Hom}(W^*, V^*)$ und $\text{Hom}(V, W) \rightarrow \text{Hom}(W^*, V^*)$ mit $f \mapsto f^T$ ist ebenfalls linear.

Wenn $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine Basis von V ist, $B^* = (\mathbf{b}_1^*, \dots, \mathbf{b}_n^*)$ eine Basis von V^* , $C = (\mathbf{c}_1, \dots, \mathbf{c}_m)$ eine Basis von W und $C^* = (\mathbf{c}_1^*, \dots, \mathbf{c}_m^*)$ die Basis von W^* . Wir betrachten $\Phi_{B^*}^{C^*}(f^T) = \Phi_C^B(f)^T$

Satz

Sei V ein Vektorraum mit Dualraum V^* und $(V^*)^* = V^{**}$ der Bidualraum.

Für $\mathbf{v} \in V$ ist die Abbildung $\iota: V^* \rightarrow \mathbb{K}$ mit $\mathbf{v}^* \mapsto \langle \mathbf{v}^*, \mathbf{v} \rangle$ linear, dann ist $\iota: V \rightarrow V^{**}$ linear und injektiv.

Wenn $\dim V < \infty$, dann ist ι ein Isomorphismus.

Beweis. $V^* \times V \rightarrow \mathbb{K}$ mit $(\mathbf{v}^*, \mathbf{v}) \mapsto \mathbf{v}^*(\mathbf{v}) = \langle \mathbf{v}^*, \mathbf{v} \rangle$ ist bilinear. Zur Injektivität. Angenommen $\iota(\mathbf{v}_1) = \iota(\mathbf{v}_2)$, d.h. $\langle \iota(\mathbf{v}_1), \mathbf{v}^* \rangle = \langle \iota(\mathbf{v}_2), \mathbf{v}^* \rangle \forall \mathbf{v}^* \in V^*$: Wir zeigen $\langle \iota(\mathbf{v}_1 - \mathbf{v}_2), \mathbf{v}^* \rangle = 0 \forall \mathbf{v}^* \in V^*$.

Nach dem Lemma unten folgt:

$$\langle \mathbf{v}^*, \mathbf{v}_1 - \mathbf{v}_2 \rangle = 0 \Rightarrow \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$$

□

Lemma

$\forall \mathbf{v} \in V: \mathbf{v} = \mathbf{0} \Leftrightarrow \forall \mathbf{v}^* \in V^*: \mathbf{v}^*(\mathbf{v}) = 0$

Unter der Annahme des Auswahlaxioms

Beweis. zu zeigen ist: $\forall \mathbf{v} \neq \mathbf{0}: \exists \mathbf{v}^* \in V^*: \mathbf{v}^*(\mathbf{v}) \neq 0$

Wir ergänzen \mathbf{v} zu einer Basis von V und definieren $\mathbf{v}^*: V \rightarrow \mathbb{K}$ mit $\mathbf{v} \mapsto 1$ und $\mathbf{b}_i \mapsto 0$ für $i \in I$, womit $\mathbf{v}^* \in V^*$ und $\mathbf{v}^*(\mathbf{v}) \neq 0$ □

Kapitel IX Polynome und Algebra

Wir wissen bereits $\mathbb{K}[x]$ ist ein Vektorraum und ein Ring zugleich.

Definition 9.1: Algebra

Sei \mathbb{K} ein Körper. Eine assoziative \mathbb{K} -Algebra ist ein Vektorraum über \mathbb{K} mit einer Multiplikation $*$: $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$, sodass $(\mathcal{A}, +, \cdot)$ ein Ring ist, und zusätzlich $\lambda(\mathbf{a} \cdot \mathbf{b}) = (\lambda\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\lambda\mathbf{b})$.

Betrachten wir ein Beispiel. $(\mathbb{K}, +, \cdot, \cdot)$ ist eine assoziative Algebra. Weiters sind bilden die $(\mathbb{K}^{n \times n}, +, \cdot, *)$, wobei $*$ die Matrizenmultiplikation ist, eine Algebra. Auch $\text{Hom}(V, V)$ mit $f * g = f \circ g$. Weiters $\mathcal{C}[0, 1]$ mit $f * g = f(t)g(t)$ ist eine kommutative Algebra. Ein Beispiel einer nicht assoziativen Algebra ist etwa $(\mathbb{R}^3, +, \cdot, \times)$, wobei \times das Kreuzprodukt ist. Anstatt der Assoziativität verwendet man die Jacobi-Identität:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{x}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$$

Ein weiteres Beispiel mit dem **Kommutator**, mit $\mathcal{A} = \mathbb{K}^{n \times n}$, wobei $\mathbf{A} \times \mathbf{B} = [\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$, ist nicht assoziativ mit Jacobi-Identität. Eine solche Algebra nennt man eine **Lie-Algebra**, benannt nach Sophus Lie.

Analog gibt es eine Jordan-Algebra, in der man stattdessen die Bedingung $\mathbf{A} * \mathbf{B} = \mathbf{AB} + \mathbf{BA}$ verwendet, wobei $\mathcal{A} = \{\mathbf{A} \in \mathbb{K}^{n \times n} | \mathbf{A} = \mathbf{A}^T\}$.

Ein besonders wichtiges Beispiel für eine assoziative, kommutative Algebra bildet $\mathbb{K}[x]$.

Definition 9.2: Raum der Polynome

$P_{\mathbb{K}} \subseteq \mathbb{K}^{\infty} = \{(a_0, a_1, \dots), | a_i \in \mathbb{K}\}$, also $P_{\mathbb{K}} = \text{span}\{\mathbf{e}_i, i \geq 0\}$, mit der Basis $\mathbf{e}_i = x^i$. Wir definieren die Multiplikation folgendermaßen:

$$(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} = (c_n)_{n \in \mathbb{N}}$$

$$c_n = \sum_{k=0}^n a_k b_{n-k}$$

Woraus folgt $x^k x^l = x^{k+l}$

Satz 9.1: Polynome formen eine Algebra

- i $(P_{\mathbb{K}}, *)$ ist eine assoziative, kommutative \mathbb{K} -Algebra mit 1-Elemente $x^0 = (1, 0, \dots)$ und heißt Polynomalgebra $\mathbb{K}[x]$
- ii $(\mathbb{K}^{\infty}, *)$ ist eine assoziative kommutative \mathbb{K} -Algebra mit

$$\mathbb{K}[[x]] = \left\{ \sum_{k=0}^{\infty} a_k x^k | a_k \in \mathbb{K} \right\}$$

Wir nennen $\mathbb{K}[[x]]$ die Algebra der formalen Potenzreihen
Wir können eine Metrik $d(f, g) = d(f - g, 0)$ definieren. Sei $h = \sum_{k=0}^{\infty} a_k x^k$, dann gilt

$$d(h, 0) = 2^{-\min\{k | a_k \neq 0\}}$$

Dann gilt:

$$d\left(\sum_{k=0}^n a_k x^k, \sum_{k=0}^{\infty} a_k x^k\right) \xrightarrow{n \rightarrow \infty} 0$$

Definition 9.3: Grad eines Polynoms

Sei $p \in \mathbb{K}[x]$, dann ist $\deg p = \max\{k | a_k \neq 0\}$, und $\deg 0 = -\infty$.

Interessante Beobachtungen zu den Polynomen:

- i $\deg(pq) = \deg p + \deg q$
- ii $\mathbb{K}[x]$ ist nullteilerfrei

Definition 9.4: Induzierte Funktionen

Jedes Polynom $p \in \mathbb{K}[x]$ induziert eine Polynomfunktion $p: \mathbb{K} \rightarrow \mathbb{K}$ mit $\alpha \mapsto p(\alpha)$. Umgekehrt definiert jedes Element $\alpha \in \mathbb{K}$ eine lineare Abbildung $\mathbb{K}[x] \rightarrow \mathbb{K}$ mit $p \mapsto p(\alpha)$:

$$\begin{aligned}(\lambda p + \mu q)(\alpha) &= \lambda p(\alpha) + \mu q(\alpha) \\ (pq)(\alpha) &= p(\alpha)q(\alpha)\end{aligned}$$

Es sei folgendes angemerkt:

- i in $\mathbb{R}[x]$ oder $\mathbb{C}[x]$ ist ein Polynom durch die entsprechende Polynomfunktion eindeutig festgelegt
- ii wenn $|\mathbb{K}| < \infty$, folgt $|\mathbb{K}^{\mathbb{K}}| < \infty$, dann ist jede Funktion $f: \mathbb{K} \rightarrow \mathbb{K}$ als Polynomfunktion darstellbar, also nicht eindeutig
- ii sehen wir bereits in $\mathbb{Z}_2[x]$, da $1+x$ und $1+x^2$ die gleiche Funktion definieren.

Definition 9.5: Algebra-Homomorphismen

Seien \mathcal{A}, \mathcal{B} \mathbb{K} -Algebra. Ein \mathbb{K} -Algebra-Homomorphismus ist eine lineare Abbildung $\psi: \mathcal{A} \rightarrow \mathcal{B}$, sodass:

$$\begin{aligned}\psi(a+b) &= \psi(a) + \psi(b) \\ \psi(\lambda a) &= \lambda \psi(a) \\ \psi(a * b) &= \psi(a) * \psi(b)\end{aligned}$$

Ein Beispiel. Sei $\alpha \in \mathbb{K}$ und $\delta_\alpha: \mathbb{K}[x] \rightarrow \mathbb{K}$ mit $p \mapsto p(\alpha)$, dann ist δ_α ein Homomorphismus. Auch $\mathbb{K}^{\mathbb{K}}$ mit $(f * g)(\alpha) = f(\alpha)g(\alpha)$ mit $\delta_\alpha = f(\alpha)$ ist ein Homomorphismus.

Ebenfalls die Zuordnung $F: \mathbb{K}[x] \rightarrow \mathbb{K}^{\mathbb{K}}$ ist ein Homomorphismus, dann ist $\delta_\alpha = \delta_\alpha \circ F$.

$$\begin{array}{ccc} \mathbb{K}[x] & \xrightarrow{F} & \mathbb{K}^{\mathbb{K}} \\ & \searrow \delta_\alpha & \swarrow \delta_\alpha \\ & \mathbb{K} & \end{array}$$

Satz 9.2: Einsetzungssatz

Sei \mathcal{A} eine assoziative Algebra mit $1_{\mathcal{A}}$, dann ist für jedes $a \in \mathcal{A}$ die Abbildung $\psi_a: \mathbb{K}[x] \rightarrow \mathcal{A}$ gegeben mit:

$$\sum_{k=0}^n c_k x^k \mapsto \sum_{k=0}^n c_k a^k \in \mathcal{A}$$

ein eindeutiger Algebra-Homomorphismus mit der Eigenschaft $\psi_a(x) = a$.

In diesem Zusammenhang, hat $\mathbb{K}[x]$ die universelle Eigenschaft, dass für jede Algebra \mathcal{A} und jedem $a \in \mathcal{A}$ ein eindeutiger Homomorphismus $\psi_a: \mathbb{K}[x] \rightarrow \mathcal{A}$ existiert, sodass $\psi_a(x) = a$. Im Vergleich steht der Fortsetzungssatz, wenn $\mathbf{b}_1, \dots, \mathbf{b}_n$ eine Basis von V ist, dann gibt es zu jedem Vektorraum W und $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$ ein eindeutige Monom $f: V \rightarrow W$, sodass $f(\mathbf{v}_i) = \mathbf{w}_i$.

Einige Beispiele:

- $\mathcal{A} = \mathbb{K}$, dann $p(\alpha) = \psi_\alpha(p(x))$
- $\mathcal{A} = \mathbb{K}^{n \times n}$, dann $p(\mathbf{A}) = \psi_{\mathbf{A}}(p(x)) = \sum_{k=0}^n c_k \mathbf{A}^k$

Definition 9.6: Nullstellen von Polynomen

Sei $p \in \mathbb{K}[x]$. Eine Nullstelle ist ein $\xi \in \mathbb{K}$, sodass $p(\xi) = \psi_\xi(p(x)) = 0$, sprich $p(x) \in \ker \psi_\xi$.

Sei etwa $p(x) = a_0$ hat keine Nullstelle außer $a_0 = 0$. Für $p(x) = a_0 + a_1x$ hat eine Nullstelle $x_0 = -\frac{a_0}{a_1}$. Auch $p(x) = a_0 + a_1x + a_2x^2$ hat 2 Nullstellen, die man mit der bekannten Formel ausrechnen kann. Auch Polynome vom Grad 3 und 4 haben Formeln zur Bestimmung der jeweiligen Nullstellen. [Cardano](#) hat bereits im 16. Jahrhundert die Formel für die Lösung der Polynomgleichung dritten Grades gefunden, in seinem Buch *Ars Magna*. Allerdings hatten auch andere Mathematiker ähnliche Gleichungen gelöst, wie etwa Scipione [DEL Ferro](#), der Gleichungen wie $n = ax + bx^3$, $n = ax^2 + bx^3$ und $ax + n = bx^3$ gelöst hatte⁶. Der Schüler von [DEL Ferro](#) hatte in einer Art mathematischer „Millionenshow“ den Mathematiker [Tartaglia](#) herausgefordert $n = ax + bx^3$ zu lösen, was ihm auch gelungen ist.

Die Formel von [Cardano](#) enthält notwendigerweise die imaginäre Einheit i . Betrachten wir ein Beispiel. $x^3 + 6x = 20$. Der Ansatz ist nun der Folgende. Wir substituieren $x = u + v$:

$$\begin{aligned} u^3 + 3u^2v + 3uv^2 + v^3 + 6(u + v) &= 20 \\ \Leftrightarrow u^3 + (3uv + 6)(u + v) + v^3 &= 20 \end{aligned}$$

Wir wählen nun u und v so, dass $3uv + 6 = 0$, womit wir erhalten $u^3 + v^3 = 20$ und $uv = -2$, dann folgt $u^3v^3 = -8$. Setzen wir ein, erhalten wir $v^3 = 20 - u^3$ und weiter $u^3(20 - u^3) = -8$, was sich in eine quadratische Gleichung in u^3 vereinfacht:

$$\begin{aligned} (u^3)^2 - 20u^3 + 8 &= 0 \\ u^3 &= \frac{20 \pm \sqrt{400 + 32}}{2} = 10 \pm \sqrt{108} \\ v^3 &= 20 - u^3 = 10 \pm \sqrt{108} \\ x &= \sqrt[3]{10 + \sqrt{108}} + \sqrt[3]{10 - \sqrt{108}} \end{aligned}$$

Für Grad 5 gibt es keine allgemeine Lösungsformel. Das wurde erst im 19. Jahrhundert von [Abel](#) und [Ruffini](#) gelöst, hat also noch sehr lange gedauert.

Noch ein Nachtrag zum Lösen Polynomgleichungen zweiten Grades. François [Viète](#) hat herausgefunden, dass für zwei reelle Zahlen $\xi_1, \xi_2 \in \mathbb{R}$ gilt:

$$(x - \xi_1)(x - \xi_2) = x^2 - (\xi_1 + \xi_2)x + \xi_1\xi_2$$

Dann sind ξ_1 und ξ_2 Nullstellen der obigen Gleichung. Kennt man also die Nullstellen eines Polynoms kann man direkt die Koeffizienten bestimmen, nämlich $a_0 = \xi_1\xi_2$ und $-a_1 = \xi_1 + \xi_2$. Anhand dieser Darstellung können wir auch die bekannte Lösungsformel herleiten:

$$\begin{aligned} \xi_1 &= \frac{1}{2}(\xi_1 + \xi_2 + (\xi_1 - \xi_2)) \\ \xi_2 &= \frac{1}{2}(\xi_1 + \xi_2 - (\xi_1 - \xi_2)) \\ \xi_{1,2} &= \frac{1}{2} \left(\xi_1 + \xi_2 \pm \sqrt{(\xi_1 - \xi_2)^2} \right) \\ (\xi_1 - \xi_2)^2 &= \xi_1^2 + \xi_2^2 - 2\xi_1\xi_2 = (\xi_1 + \xi_2)^2 - 4\xi_1\xi_2 = a_1^2 - 4a_0 \\ \xi_{1,2} &= \frac{1}{2} \left(-a_1 \pm \sqrt{a_1^2 - 4a_0} \right) \end{aligned}$$

Hier sei noch schnell angemerkt, dass wir ein beliebige nicht-monisches Polynom $a_2x^2 + a_1x + a_0 = 0$ immer auf ein monisches umformen können:

$$a_2x^2 + a_1x + a_0 = 0 \Leftrightarrow x^2 + \frac{a_1}{a_2}x + \frac{a_0}{a_2} = 0$$

Sofern natürlich $a_2 \neq 0$.

⁶zu dieser Zeit hatte man das Konzept der negativen Zahlen noch nicht allgemein akzeptiert, daher gibt es für ähnliche Gleichungen oftmals unterschiedliche Darstellungen hat, um negative Zahlen zu vermeiden.

9.1 Division mit Rest

Wir wollen nun das Konzept der Division mit Rest auf Polynome anwenden. Seien $p, q \in \mathbb{K}[x]$, wobei $q(x) \neq 0$, dann $\exists! s, r \in \mathbb{K}[x]$ sodass gilt $p = sq + r$, wobei $\deg r < \deg q$. Betrachten wir dazu zuerst ein Beispiel. Sei $p(x) = 3x^5 - x^4 + 2x^3 + x^2 + 1$ und $q(x) = x^2 - 3x + 1$:

$$\begin{array}{r}
 3x^5 - x^4 + 2x^3 + x^2 + 1 = (x^2 - 3x + 1)(3x^3 + 8x^2 + 23x + 62) + 163x - 61 \\
 \underline{- 3x^5 + 9x^4 - 3x^3} \\
 8x^4 - x^3 + x^2 \\
 \underline{- 8x^4 + 24x^3 - 8x^2} \\
 23x^3 - 7x^2 \\
 \underline{- 23x^3 + 69x^2 - 23x} \\
 62x^2 - 23x + 1 \\
 \underline{- 62x^2 + 186x - 62} \\
 163x - 61
 \end{array}$$

Wir sehen, dass $s(x) = 3x^3 + 8x^2 + 23x + 62$ und $r(x) = 163x - 61$.

Definition 9.7: Polynomteiler

Seien $p, q \in \mathbb{K}[x]$. Wir sagen q **teilt** p , wenn $\exists s \in \mathbb{K}[x]: p = sq$.

Satz 9.3: Bestimmung von Nullstellen durch Algebraische Methoden

- i wenn $p(x) = s(x)(x - \xi) + r$, dann ist $p(\xi) = r$
- ii ξ ist eine Nullstelle von $p(x)$, wenn $(x - \xi)$ p teilt

Eine praktische Anwendung von Satz 9.3 ist im sogenannten **Horner-Schema**, benannt nach William George **Horner** (1786-1837), obwohl dieser das Schema bereits von Joseph-Louis **Lagrange** kennengelernt hat. Allerdings war das Schema bereits 600 Jahre zuvor dem Chinesen Qin **Jiushao** (ca 1202-1261) bekannt.

Sei $p \in \mathbb{K}[x]$ und $\lambda \in \mathbb{K}$, gesucht ist nun $p(\lambda)$. Der naive Ansatz wäre nun $\sum_{k=0}^n a_k \lambda^k$ auszurechnen, was zu $\frac{n(n+1)}{2} \in \mathcal{O}(n^2)$ Multiplikation führt, also dementsprechend viel Rechenaufwand. Die Idee beim Horner-Schema ist es nun, den gemeinsamen Faktor λ aus allen entsprechenden Summanden herauszuheben:

$$\sum_{k=0}^n a_k \lambda^k = \lambda \sum_{k=1}^n a_k \lambda^{k-1} + a_0 = \lambda \left(\lambda \sum_{k=2}^n a_k \lambda^{k-2} + a_1 \right) + a_0$$

Wiederholt man diese Idee solange, bis die höchste Potenz von λ genau λ ist, hat man nur noch n Multiplikation⁷. Für ein Polynom dritten Grades erhalten wir somit:

$$a_3 \lambda^3 + a_2 \lambda^2 + a_1 \lambda + a_0 = ((a_3 \lambda + a_2) \lambda + a_1) \lambda + a_0$$

Wir können $p(\lambda)$ nun über eine Folge ausdrücken. Wir setzen $\xi_n = a_n$ und bestimmen für $k = n-1, n-2, \dots, 0$ jeweils $\xi_k = \lambda \xi_{k+1} + a_k$, und erhalten somit für $p(\lambda) = \xi_0$. Wenden wir diese rekursive Methode auf $p(x) = 3x^5 - x^4 + 2x^3 + x^2 + 1$, um $p(5)$ zu bestimmen, an erhalten wir:

$$\begin{aligned}
 \xi_5 &= 3 \\
 \xi_4 &= 3 \cdot 5 - 1 = 14 \\
 \xi_3 &= 14 \cdot 5 + 2 = 72 \\
 \xi_2 &= 72 \cdot 5 + 1 = 361 \\
 \xi_1 &= 361 \cdot 5 + 0 = 1805 \\
 \xi_0 &= 1805 \cdot 5 + 1 = 9026
 \end{aligned}$$

Vergleichen wir mit der Polynomdivision $p(x)/(x - 5)$:

⁷und entsprechend viele vergleichsweise einfache Additionen

$$\begin{array}{r}
3x^5 - x^4 + 2x^3 + x^2 \\
- 3x^5 + 15x^4 \\
\hline
14x^4 + 2x^3 \\
- 14x^4 + 70x^3 \\
\hline
72x^3 + x^2 \\
- 72x^3 + 360x^2 \\
\hline
361x^2 \\
- 361x^2 + 1805x \\
\hline
1805x + 1 \\
- 1805x + 9025 \\
\hline
9026
\end{array}
\quad + 1 = (x - 5) (3x^4 + 14x^3 + 72x^2 + 361x + 1805) + 9026$$

Wir sehen, dass die Koeffizienten von s genau den ξ_k für $k = n, \dots, 1$ entsprechen und dann $r = \xi_0$.

Definition 9.8: Reduzierbarkeit

Sei $p \in \mathbb{K}[x]$. Wir nennen p **reduzibel** über \mathbb{K} , wenn $\exists p_1, p_2 \in \mathbb{K}[x]$ mit $\deg(p_1) < \deg(p)$ und $\deg(p_2) < \deg(p)$, sodass $p = p_1 p_2$, als p_1 und p_2 echte Teiler von p sind.

Falls p nicht in so ein Produkt zerlegbar ist, nennen wir p **irreduzibel**.

Im Vergleich nennen wir $n \in \mathbb{N}$ irreduzibel, wenn $n \in \mathbb{P}$. Hier muss man jedoch aufpassen, dass das im Allgemeinen nicht gilt. Weiters sei angemerkt, dass irreduzible Polynome keine Nullstellen haben. Man muss bei diesen Betrachtungen sehr auf den zugrundeliegenden Körper achten. Das Polynom $x^2 - 2$ ist irreduzibel über \mathbb{Q} , da $\sqrt{2} \notin \mathbb{Q}$, denn $(x + \sqrt{2})(x - \sqrt{2}) = x^2 - 2$. Allerdings ist $x^2 - 2$ reduzibel über \mathbb{R} . Im Gegenzug dazu ist $(x^2 - 2)(x^2 + 2) = x^4 - 4$ reduzibel über \mathbb{Q} hat aber keine rationalen Nullstellen. Auch $x^2 + 1$ ist über \mathbb{R} irreduzibel, aber reduzibel über \mathbb{C} .

Klar ist natürlich, dass Polynome vom Grad 1 irreduzibel sind, unabhängig des zugrunde liegenden Körpers. Über \mathbb{R} haben alle irreduziblen Polynome höchstens Grad 2. Über $\mathbb{Z}/2\mathbb{Z}$ gibt es zu jedem Grad n ein irreduzibles Polynom vom Grad n . Gehen wir wie [Cardano](#) vor und erweitern $\mathbb{Z}/2\mathbb{Z}$ um α mit $\alpha^2 = \alpha + 1$, dann ist α eine Nullstelle für $x^2 + x + 1$. Verallgemeinern wir dieses Konzept, können wir einen endlichen Körper $\mathbf{GF}(p^k)$ bestimmen, der genau p^k Elemente enthält, wobei $p \in \mathbb{P}$ und $k \in \mathbb{N}$. Hier ist anzumerken, dass wir alle endliche Körper über $\mathbf{GF}(p^k)$ auffinden können. Hierbei steht **GF** für Galois-Feld⁸, nach Évariste [Galois](#) (1811-1832).

Satz 9.4: Hauptsatz der Algebra

\mathbb{C} ist algebraisch abgeschlossen, d.h. jedes Polynom $p \in \mathbb{C}[x]$ hat eine Nullstell über \mathbb{C} .

Das Polynom $x^2 + 1$ hat keine Nullstell über \mathbb{R} . Führen wir die imaginäre Einheit i ein und erweitern \mathbb{R} um i auf $\mathbb{R}(i) = \{a + ib | a, b \in \mathbb{R}\} = \mathbb{C}$. Der Beweis für den Hauptsatz der Algebra basiert nicht auf algebraischen Methoden, sondern die sogenannte **Funktionentheorie**⁹. Eine direkte Folgerung aus Satz 9.4 ist, dass ein Polynom $p \in \mathbb{C}[x]$ irreduzibel ist, wenn $\deg p \leq 1$. Weitaus bedeutender ist hierbei jedoch, dass jedes Polynom $p \in \mathbb{C}[x]$ eine Faktorisierung in Linearfaktoren hat:

$$p(x) = a \prod_{k=1}^n (x - \xi_k)$$

Dieser Umstand lässt sich relativ leicht zeigen. Da \mathbb{C} algebraisch abgeschlossen ist, hat p eine Nullstelle ξ_1 und lässt sich mittels Polynomdivision in $p(x) = s_1(x)(x - \xi_1)$ zerlegen. s hat nun wieder eine Nullstelle ξ_2 , für die natürlich auch gilt $p(\xi_2) = 0$, womit $p(x) = s_2(x)(x - \xi_2)(x - \xi_1)$. Da p endlichen Grad haben muss, erhalten wir ein irreduzibles s_n nach n Iterationen.

Definition 9.9: Vielfachheit

Eine Nullstelle ξ eines Polynoms $p \in \mathbb{C}[x]$ hat Vielfachheit m wenn gilt $(x - \xi)^m | p(x)$, sprich $p(x) = s(x)(x - \xi)^m$, wobei $s(\xi) \neq 0$.

⁸field ist die englische Bezeichnung für Körper

⁹diese Theorie behandelt Funktionen mit komplexen Argumenten und Werten

Eine Nullstelle ξ hat mindestens Vielfachheit 2, genau dann wenn, ξ eine Nullstelle von $\text{ggT}(p, p')$ ist. Dabei bezeichnet p' die Ableitungsfunktion von p . Der ggT für Polynome kann analog zum ggT natürlicher Zahlen mit dem Euklidischen Algorithmus bestimmt werden.

Kapitel X Eigenwerte und Eigenvektoren

Unser Ziel ist es nun eine invertierbare Matrix V zu finden, sodass:

$$V^{-1}AV = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

Sprich es gilt $Ab_i = \lambda_i b_i$, wobei b_i die Spalten von V sind, welche eine Basis sind. Es folgt also, dass $(\lambda_i I - A)b_i = 0$ keine triviale Lösung hat, womit $\ker(\lambda_i I - A) \neq \{0\}$, womit auch $\det(\lambda I - A) = 0$ sein muss.

Definition 10.1: Eigenwerte- und vektoren

Sei V ein Vektorraum über \mathbb{K} und $f \in \text{End}(V)$. Wir nennen $\lambda \in \mathbb{K}$ einen Eigenwert¹⁰ von A , wenn $\exists v \in V \setminus \{0\}: f(v) = \lambda v$. v heißt der zugehörige Eigenvektor¹¹. Wir bezeichnen als das Spektrum von f , $\text{spec}(f)$, die Menge der Eigenwerte von f . Weiters nennen wir $\mathcal{N}_\lambda = \{v \in V: f(v) = \lambda v\}$, den Null-Space¹².

Lemma 10.1: Unterraum der Eigenvektoren

Sei $f \in \text{End}(V)$ mit einem Eigenwert $\lambda \in V$, dann bildet \mathcal{N}_λ einen Unterraum, den sogenannten **Eigenraum** zum Eigenwert λ .

Beweis für Lemma 10.1. Sei $v \in \mathcal{N}_\lambda$, dann folgt:

$$\begin{aligned} f(v) &= \lambda v \\ \Leftrightarrow \lambda v - f(v) &= 0 \\ \Leftrightarrow (\lambda \text{id} - f)(v) &= 0 \\ \Leftrightarrow v &\in \ker(\lambda \text{id} - f) \end{aligned}$$

Da der Kern einen Unterraum bildet, ist \mathcal{N}_λ ein Unterraum. □

Ein triviales Beispiel ist etwa $f = \alpha v$, dann ist $f(v) = \alpha v$, dann ist $\text{spec}(\alpha \text{id}) = \{\alpha\}$ und es gilt $\mathcal{N}_\lambda = V$. Ein etwas gehaltvolleres Beispiel ist der Fall $A = \text{diag}(\lambda_1, \dots, \lambda_n)$, dann gilt $Ae_i = \lambda_i e_i$, wobei e_i der i -te kanonische Basisvektor ist. Hier gilt natürlich $\text{spec}(A) = \{\lambda_1, \dots, \lambda_n\}$. Hier muss auf die Vielfachheit der λ_i geachtet werden, denn $\mathcal{N}_\lambda = \text{span}\{e_i | \lambda_i = \lambda\}$.

Wir werden im Allgemeinen endlich-dimensionalen Vektorräume betrachten, allerdings kann das Konzept auf unendlich-dimensionalen Räumen ebenfalls nützlich sein. Betrachten wir etwa $V = C^\infty(\mathbb{R})$, also den Raum der glatten¹³ Funktionen. Die Ableitung $\frac{d}{dx} \in \text{End}(C^\infty(\mathbb{R}))$ ist linear und wir suchen $f \in C^\infty(\mathbb{R})$, sodass $f' = \lambda f$. Wir sehen, dass $f = e^{\lambda x}$ diese Bedingung erfüllt. Kennt man diese Eigenschaft der Exponentialfunktion nicht, muss man eine Differentialgleichung lösen:

$$\frac{dy}{dx} = \lambda y \Leftrightarrow \frac{dy}{y} = \lambda dx$$

Integriert man auf beiden Seiten erhält man:

$$\ln(y) = \lambda x + C \Leftrightarrow y = e^{\lambda x + C} = c_0 e^{\lambda x}$$

¹⁰engl. eigenvalue

¹¹engl. eigenvector

¹²aus dem Englischen, da $(\lambda \text{id} - f)(v) = 0$, also der Kern von $\lambda \text{id} - f$

¹³unendlich-oft differenzierbaren Funktionen

Findet man eine Basis aus Eigenvektoren für \mathcal{C}^∞ , kann man über die Fourier-Transformation die meisten Funktionen folgendermaßen ausdrücken:

$$f(x) = \int \hat{f}(t) e^{xt} dt$$

Das ist hier etwas vereinfacht und nicht ganz korrekt ausgedrückt, vermittelt aber die Idee sehr gut.

Betrachten wir $V = \mathcal{C}^\infty[0, L]$ mit $f \mapsto f''$, können wir iterieren und erhalten:

$$\begin{aligned} \frac{d^2}{dx^2} e^{\lambda x} &= \lambda^2 e^{\lambda x} \\ \frac{d^2}{dx^2} e^{i\omega x} &= (i\omega)^2 e^{i\omega x} = -\omega^2 e^{i\omega x} \\ \frac{d^2}{dx^2} \sin(\omega x) &= -\omega^2 \sin(\omega x) & \frac{d^2}{dx^2} \cos(\omega x) &= -\omega^2 \cos(\omega x) \end{aligned}$$

In der Physik erhält man Sinus und Cosinus durch den **harmonischen Oszillator**¹⁴. Wir sehen, dass die $\omega \in \mathbb{R}$ die Eigenwerte der zweiten Ableitung bilden. Schränken wir uns auf eine feste Länge L ein, so muss zusätzlich gelten $\sin(\omega 0) = 0$ und $\sin(\omega L) = 0$, womit nur noch $\omega = \frac{k\pi}{L}$ Eigenwerte sind.

Definition 10.2: Links- und Rechtseigenwerte

Sei $A \in \mathbb{K}^{n \times n}$, gibt es $x \neq 0$, sodass $Ax = \lambda x$, dann nennen wir λ einen Rechtseigenwert und x einen Rechtseigenvektor.

Analog nennen wir $x \neq 0$, für den gilt $x^T A = \lambda x^T$, einen Linkseigenvektor und den zugehörigen Eigenvektor λ einen Linkseigenwert.

Lemma 10.2: Gleichheit von Links- und Rechtseigenwerten

Linkseigenwerte und Rechtseigenwerte sind ident.

Beweis für Lemma 10.2. Sei λ ein Rechtseigenwert von A , dann gibt es $x \neq 0$ sodass $Ax - \lambda x = 0$ bzw. $(A - \lambda I)x = 0$ oder aber $\ker(A - \lambda I) \neq \{0\}$, was natürlich ident ist zu $\text{rank}(A - \lambda I) < n$ bzw. $\text{rank}(A^T - \lambda I) < n$. Daraus folgt natürlich, dass λ ein Linkseigenwert sein muss. \square

Es sei noch angemerkt, dass x ein Rechtseigenvektor von A ist, wenn x^T ein Linkseigenvektor von A^T ist. Weiters gilt Lemma 10.2 für $\dim = \infty$ nicht. Betrachten wir den Shift-Operator $S: (\xi_1, \xi_2, \dots) \mapsto (0, \xi_1, \xi_2, \dots)$, da S injektiv ist, folgt $\ker(S) = \{0\}$, somit ist 0 kein Eigenwert. Allerdings ist $S^T: (\xi_1, \xi_2, \dots) \mapsto (\xi_2, \xi_3, \dots)$ nicht injektiv, da $(1, 0, \dots) \in \ker(S^T)$, womit ein Rechtseigenwert von S^T ist. Beide Operatoren S und S^T sind nicht invertierbar, daher definiert man für unendlich-dimensionale Räume das Spektrum $\text{spec}(T) = \{\lambda \mid \lambda \text{Id} - T \text{ nicht invertierbar}\}$. Verwendet man diese Definition folgt dann $\text{spec}(T) = \text{spec}(T^T)$.

Definition 10.3: Spektrum

Für $A \in \mathbb{K}^{n \times n}$ ist $\text{spec}(A) = \{\lambda \in \mathbb{K} \mid Ax = \lambda x\} = \{\lambda \in \mathbb{K} \mid x^T A = \lambda x^T\}$ das Spektrum von A .

Lemma 10.3: Spektrum ist unabhängig von der Basis

Sei $f \in \text{End}(V)$ mit $\dim V = n < \infty$, dann gilt $\text{spec}(f) = \text{spec}(\Phi_B^B(f))$ für jede Basis B .

Beweis für Lemma 10.3. $\lambda \notin \text{spec}(f)$ wenn $\lambda \text{Id} - f$ invertierbar ist. Das ist der Fall, wenn $\lambda I - \Phi_B^B(f)$ regulär ist, also $\lambda \notin \text{spec}(\Phi_B^B(f))$. \square

¹⁴alternativ lassen sich Schwingungen auch als Impulsantworten des Systems $\dot{x} = \begin{bmatrix} \delta & -\omega \\ \omega & \delta \end{bmatrix} x$ beschreiben

Wir folgern daraus, dass sich das Spektrum einer Matrix A nach einem Basiswechsel nicht ändert. Bilden die Spalten von V eine Basis, dann gilt $\text{spec}(V^{-1}AV) = \text{spec}(A)$. Ein direkter Beweis für diesen Umstand folgt: Sei $Ax = \lambda x$, dann gilt $AVV^{-1}x = \lambda x$, bzw:

$$\begin{aligned} V^{-1}AV \underbrace{V^{-1}x}_y &= \lambda V^{-1}x \\ \Leftrightarrow V^{-1}AVy &= \lambda y \end{aligned}$$

Satz 10.1: Bestimmung der Eigenwerte

Sei $A \in \mathbb{K}^{n \times n}$:

- i $\chi_A(\lambda) = \det(\lambda I - A)$ ist ein Polynom von λ und heißt das **charakteristische Polynom** von A
- ii λ ist ein Eigenwert von A , genau dann wenn $\chi_A(\lambda) = 0$

Beweis für Satz 10.1. Zu i:

$$\det(\lambda I - A) = \sum_{\pi \in \mathfrak{S}_n} (-1)^\pi \prod_{i=1}^n (\lambda \delta_{i, \pi(i)} - a_{i, \pi(i)})$$

Die höchste auftretende Potenz von λ ist λ^n . Für $\pi = \text{id}$ erhalten wir $\prod_{k=1}^n (\lambda - a_{k,k})$, was zu einem Polynom mit normierten λ^n wird, womit $\chi_A(\lambda) = \lambda^n + \sum_{k=0}^{n-1} c_k \lambda^k$. \square

Betrachten wir ein Beispiel:

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

Das charakteristische Polynom ist nun:

$$\begin{aligned} \chi_A(\lambda) &= \begin{vmatrix} \lambda + 1 & -1 & -2 \\ 1 & \lambda + 5 & -2 \\ -2 & 2 & \lambda + 4 \end{vmatrix} = \begin{vmatrix} \lambda & -1 & -2 \\ \lambda + 6 & \lambda + 5 & -2 \\ 0 & 2 & \lambda + 4 \end{vmatrix} \\ &= \lambda \begin{vmatrix} \lambda + 5 & -2 \\ 2 & \lambda + 4 \end{vmatrix} - (\lambda + 6) \begin{vmatrix} -1 & -2 \\ 2 & \lambda + 4 \end{vmatrix} \\ &= \lambda((\lambda + 5)(\lambda + 4) + 5) - (\lambda + 6)(-(\lambda + 4) + 4) = \lambda^3 + 10\lambda^2 + 30\lambda = \lambda(\lambda^2 + 10\lambda + 30) \\ &\Rightarrow \lambda_1 = 0 \\ \lambda_{2,3} &= -5 \pm \sqrt{25 - 30} = -5 \pm i\sqrt{5} \end{aligned}$$

Exkurs Anfang Das sich die Eigenwerte numerisch alternativ bestimmen lassen, ohne Nullstellen eines Polynoms zu finden, kann man auch umgekehrt arbeiten und eine Matrix suchen mit dem gegebenen charakteristischen Polynom und deren Eigenwerte numerisch bestimmen, um alle Nullstellen des gegebenen Polynoms zu finden.

Exkurs Ende

Betrachten wir eine 2×2 Matrix finden wir eine gewisse Struktur im charakteristischen Polynom:

$$\begin{vmatrix} \lambda a_{11} & -a_{12} \\ -a_{12} & \lambda - a_{22} \end{vmatrix} = \lambda^2 - \lambda \underbrace{(a_{22} + a_{11})}_{=\text{Tr}(A)} + \det(A)$$

Für $n = 3$ ergibt sich:

$$\begin{aligned} \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix} &= \lambda^3 + \lambda \left(\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right) \\ &\quad - \lambda^2 \underbrace{(a_{33} + a_{22} + a_{11})}_{=\text{Tr}(A)} - \det(A) \end{aligned}$$

Satz 10.2: Allgemeine Form des charakteristischen Polynoms

Sei $A \in \mathbb{K}^{n \times n}$ gilt:

$$\chi_A(\lambda) = \sum_{k=0}^n (-1)^k c_k(A) x^k$$

$$c_k(A) = \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I| = n-k}} \det(A_{I,I})$$

Dabei ist $A_{I,I}$ jene Teilmatrix mit Indizes aus I , also beschreibt c_k die Summe der Minoren k -ter Ordnung. Insbesondere $c_n = 1$, $c_{n-1} = \text{Tr}(A)$ und $c_0 = \det(A)$.

Beweis für Satz 10.2. Vergleichen wir zuerst (als Motivation):

$$\prod_{k=1}^n (x - y_1) = \sum_{I \subseteq \{1, \dots, n\}} x^{|I|} \prod_{j \in I^C} (-y_j) = \sum_{I \subseteq \{1, \dots, n\}} x^{|I|} (-1)^{n-|I|} \prod_{j \in I^C} y_j$$

Wir arbeiten nun mit der Determinantenform $\Delta(e_1, \dots, e_n) = 1$, dann ist $\det(A) = \Delta(s_1, \dots, s_n)$, dabei sind s_i die Spalten von A . Damit folgt:

$$\begin{aligned} \det(\lambda I - A) &= \Delta(\lambda e_1 - s_1, \lambda e_2 - s_2, \dots, \lambda e_n - s_n) \\ &= \sum_{I \subseteq \{1, \dots, n\}} \Delta(y_1, \dots, y_n) \\ y_i &= \begin{cases} \lambda e_i & i \in I \\ -s_i & i \in I^C \end{cases} \\ \Rightarrow \det(\lambda I - A) &= \sum_{I \subseteq \{1, \dots, n\}} \Delta(y_1, \dots, y_{k-1}, \lambda e_k, y_{k+1}, \dots, y_n) \\ \Delta(y_1, \dots, y_{k-1}, \lambda e_k, y_{k+1}, \dots, y_n) &= \det \begin{bmatrix} y_1 & y_2 & \cdots & y_{k-1} & \lambda e_1 & y_{k+1} & \cdots & y_n \end{bmatrix} \\ &= \det \begin{bmatrix} \tilde{y}_1 & \cdots & \tilde{y}_{k-1} & \tilde{y}_{k+1} & \cdots & \tilde{y}_n \end{bmatrix} (-1)^{2k} \end{aligned}$$

Dabei entsteht \tilde{y}_i durch Streichen der i -ten Zeile aus y_i . Zu jeder Teilmenge $I \subseteq \{1, \dots, n\}$ streichen wir die entsprechenden Zeilen und Spalten, womit folgt:

$$\Delta(y_1, \dots, y_n) = \lambda^{|I|} \det(A_{I^C, I^C})$$

Insgesamt erhalten wir:

$$\sum_{I \subseteq \{1, \dots, n\}} \lambda^{|I|} (-1)^{n-|I|} \det(A_{I^C, I^C})$$

□

Ähnliche Matrizen haben das gleiche charakteristische Polynom:

$$\begin{aligned} \chi_{VAV^{-1}}(\lambda) &= \det(\lambda I - VAV^{-1}) = \det(\lambda VV^{-1} - VAV^{-1}) \\ &= \det(V(\lambda I - A)V^{-1}) = \det(V) \cdot \det(\lambda I - A) \cdot \det(V^{-1}) = \det(\lambda I - A) = \chi_A(\lambda) \end{aligned}$$

Definition 10.4: Diagonalisierbarkeit

Sei $A \in \mathbb{K}^{n \times n}$. A heißt diagonalisierbar, wenn

$$\exists V \in \text{GL}(n, \mathbb{K}): V^{-1}AV = \text{diag}(\lambda_1, \dots, \lambda_n)$$

also A ähnlich zu einer Diagonalmatrix ist.

Es gilt natürlich $\chi_A(\lambda) = \chi_{\text{diag}(\lambda_1, \dots, \lambda_n)}(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i)$.

Satz 10.3: Diagonalisierbarkeit anhand der Eigenvektoren

Sei $A \in \mathbb{K}^{n \times n}$. A ist diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.

Beweis für Satz 10.3. A ist diagonalisierbar, wenn $\exists V \in \text{GL}(n, \mathbb{K}): V^{-1}AV = \text{diag}(\lambda_1, \dots, \lambda_n)$, es folgt:

$$\begin{aligned} AV &= V \text{diag}(\lambda_1, \dots, \lambda_n) \\ V &= [b_1 \quad b_2 \quad \dots \quad b_n] \\ AV &= V \text{diag}(\lambda_1, \dots, \lambda_n) \\ \Leftrightarrow [Ab_1 \quad Ab_2 \quad \dots \quad Ab_n] &= [\lambda_1 b_1 \quad \lambda_2 b_2 \quad \dots \quad \lambda_n b_n] \\ \Leftrightarrow \forall i = 1, \dots, n: Ab_i &= \lambda_i b_i \end{aligned}$$

Da V regulär ist, bilden die b_i eine Basis □

Betrachten wir ein kurzes Beispiel.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow \chi_A(\lambda) = \lambda^2 \Rightarrow \text{spec}(A) = \{0\}$$

Somit folgt $\ker A = \text{span} \left(\begin{bmatrix} 1 & 0 \end{bmatrix}^T \right)$. Da $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$ keine Basis ist, ist A nicht diagonalisierbar. Betrachten wir noch ein Beispiel einer diagonalisierbaren Matrix.

$$A = \begin{bmatrix} -1 & 2 & 4 \\ 4 & -3 & -8 \\ -2 & 2 & 5 \end{bmatrix} \Rightarrow \chi_A(\lambda) = (\lambda - 1)(\lambda - 1)(\lambda + 1) = (\lambda - 1)^2(\lambda + 1)$$

Somit $\text{spec}(A) = \{1, -1\}$, wobei 1 ein doppelter Eigenwert ist. Wir suchen nun $\ker(I - A)$:

$$\begin{bmatrix} 2 & -2 & -4 \\ -4 & 4 & 8 \\ 2 & -2 & -5 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & -1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Wir sehen, dass $\dim(\ker(I - A)) = 2$, mit $x_1 = x_2 + 2x_3$, somit die Basis:

$$\ker(I - A) = \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Für $\lambda = -1$:

$$\begin{bmatrix} 0 & -2 & -4 \\ -4 & 2 & 8 \\ 2 & -2 & 6 \end{bmatrix} \xrightarrow{II+2III} \begin{bmatrix} 0 & -2 & -4 \\ 0 & -2 & -4 \\ 2 & -2 & -6 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & -2 & -4 \\ 0 & 0 & 0 \\ 2 & -2 & -4 \end{bmatrix}$$

Wir sehen $\dim(\ker(-I - A)) = 1$ mit $x_1 = x_2 + 3x_3 = x_3$, da $x_2 = -2x_3$, und erhalten

$$\ker(-I - A) = \text{span} \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

Wir setzen nun V :

$$V = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow V^{-1}AV = \text{diag}(1, 1, -1)$$

Eine wichtige Anwendung der Diagonalmatrix $\mathbf{A} = \mathbf{V}^{-1} \mathbf{A} \mathbf{V} = \text{diag}(\lambda_1, \dots, \lambda_n)$ ist die Bestimmung der Potenzen \mathbf{A}^k . Wollen wir etwa \mathbf{A}^{10} bestimmen, können wir die bestehenden Potenzen verwenden und 10 als Summe von Potenzen von 2 darstellen:

$$\begin{aligned}\mathbf{A}^2 &= \mathbf{A} \mathbf{A} \\ \mathbf{A}^4 &= \mathbf{A}^2 \mathbf{A}^2 \\ \mathbf{A}^8 &= \mathbf{A}^4 \mathbf{A}^4 \\ \mathbf{A}^{10} &= \mathbf{A}^8 \mathbf{A}^2\end{aligned}$$

Effizienter können wir jedoch mit der Diagonalmatrix arbeiten. Sei $\mathbf{A} = \mathbf{V} \mathbf{A} \mathbf{V}^{-1}$, dann folgt:

$$\mathbf{A}^k = (\mathbf{V} \mathbf{A} \mathbf{V}^{-1})^k = \prod_{i=1}^k \mathbf{V} \mathbf{A} \mathbf{V}^{-1} = \mathbf{V} \mathbf{A} \mathbf{V}^{-1} \mathbf{V} \mathbf{A} \mathbf{V}^{-1} \dots \mathbf{V}^{-1} \mathbf{V} \mathbf{A} \mathbf{V}^{-1} = \mathbf{V} \mathbf{A}^k \mathbf{V}^{-1}$$

Da $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n)$ folgt $\mathbf{A}^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$.

Betrachten wir die Fibonacci-Zahlen mit den Startwerten $F_0 = F_1 = 1$ und der rekursiven Definition $F_{n+1} = F_n + F_{n-1}$ für $n \geq 2$. Diese Rekursion ist linear. Wir schreiben sie also als Matrix-Vektor-Produkt:

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} F_n + F_{n-1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \dots = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Bestimmen wir nun die Eigenwerte für \mathbf{A} :

$$\chi_{\mathbf{A}}(\lambda) = \lambda(\lambda - 1) - 1 = \lambda^2 - \lambda - 1 \Rightarrow \lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

Für die Eigenvektoren erhalten wir:

$$\mathbf{V} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$$

Somit folgt:

$$\begin{aligned}F_n &= \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right) \sim \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^{n+1} \\ \frac{F_{n+1}}{F_n} &\xrightarrow{n \rightarrow \infty} \frac{1+\sqrt{5}}{2} = \varphi\end{aligned}$$

Wir nennen φ den **goldenen Schnitt**, die $\varphi = \frac{1+\sqrt{5}}{2}$ erfüllt.

Lemma 10.4: Lineare Unabhängigkeit der Eigenvektoren

Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Beweis für Lemma 10.4. Seien $\lambda_1, \dots, \lambda_r$ verschiedene Eigenwerte von \mathbf{A} und $\mathbf{v}_1, \dots, \mathbf{v}_r$ die zugehörigen Eigenvektoren, mit $\sum_{k=1}^r \alpha_k \mathbf{v}_k = \mathbf{0}$ für alle $\alpha_k = 0$. Wir zeigen das mittels Induktion. Für $r = 1$ ist die Aufgabe trivial. Sei $\sum_{k=1}^r \alpha_k \mathbf{v}_k = \mathbf{0}$:

$$\begin{aligned}(i) : \mathbf{A} \sum_{k=1}^r \alpha_k \mathbf{v}_k = \mathbf{0} &\Leftrightarrow \sum_{k=1}^r \alpha_k \mathbf{A} \mathbf{v}_k = \mathbf{0} \Leftrightarrow \sum_{k=1}^r \alpha_k \lambda_k \mathbf{v}_k = \mathbf{0} \\ (ii) : \lambda_1 \sum_{k=1}^r \alpha_k \mathbf{v}_k &= \mathbf{0} \\ (i) - (ii) : \sum_{k=1}^r \alpha_k (\lambda_k - \lambda_1) \mathbf{v}_k &= \mathbf{0}\end{aligned}$$

Nach der Induktionsvoraussetzung gilt $\alpha_i(\lambda_i - \lambda_r) = 0$, da $\lambda_r \neq \lambda_i$, womit $\alpha_i = 0$, woraus folgt $\alpha_r \mathbf{v}_r = \mathbf{0} \Rightarrow \alpha_r = 0$. \square

Daraus folgern wir, dass eine Matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ mit n verschiedenen Eigenwerten diagonalisierbar ist, da n linear unabhängige Eigenvektoren existieren. Im Allgemeinen Fall haben wir r verschiedene Eigenwert $\lambda_1, \dots, \lambda_r$ mit den Eigenräumen E_1, \dots, E_r mit den Basen B_1, \dots, B_r , wobei $B = \bigcup_{i=1}^r B_i$ linear-unabhängig ist. Die Frage ist nun, ob B eine Basis für V ist.

Definition 10.5: Vielfachheiten von Eigenwerten

Sei $\lambda_i \in \text{spec}(\mathbf{A})$, dann ist $k(\lambda_i)$ die Vielfachheit von λ_i als Nullstelle von $\chi_{\mathbf{A}}(\lambda)$, und wir nennen $k(\lambda_i)$ die **algebraische Vielfachheit** von λ_i . Weiters nennen wir $d(\lambda_i) = \dim(\ker(\lambda_i \mathbf{I} - \mathbf{A}))$ die **geometrische Vielfachheit** von λ_i .

Satz 10.4: Zusammenhang geometrischer und algebraischer Vielfachheit

Für alle Eigenwerte $\lambda_i \in \text{spec}(\mathbf{A})$ gilt:

$$d(\lambda_i) \leq k(\lambda_i)$$

Beweis für Satz 10.4. Sei $\lambda_i \in \text{spec}(\mathbf{A})$ und $d = d(\lambda) = \dim(\ker(\lambda_i \mathbf{I} - \mathbf{A}))$, sei $(\mathbf{b}_1, \dots, \mathbf{b}_d)$ eine Basis von $\ker(\lambda_i \mathbf{I} - \mathbf{A})$. Wir ergänzen diese Basis nun zu einer Basis $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ von \mathbb{K}^n und schreiben $\mathbf{V} = [\mathbf{b}_1 \dots \mathbf{b}_n]$. Bilden wir $\mathbf{A}\mathbf{V}$ erhalten wir:

$$\begin{aligned} \mathbf{A}\mathbf{V} &= [\mathbf{A}\mathbf{b}_1 \dots \mathbf{A}\mathbf{b}_d \quad \mathbf{A}\mathbf{b}_{d+1} \dots \mathbf{A}\mathbf{b}_n] \\ &= [\lambda_i \mathbf{b}_1 \dots \lambda_i \mathbf{b}_d \quad \mathbf{A}\mathbf{b}_{d+1} \dots \mathbf{A}\mathbf{b}_n] \\ &= [\mathbf{b}_1 \dots \mathbf{b}_d \quad \mathbf{b}_{d+1} \dots \mathbf{b}_n] \begin{bmatrix} \text{diag}_d \lambda & \mathbf{X} \\ \mathbf{0} & \tilde{\mathbf{A}} \end{bmatrix} \end{aligned}$$

Wobei $\text{diag}_d \in \mathbb{K}^{d \times d}$ ist, also eine Diagonalmatrix mit d Diagonalelementen. Es folgt somit:

$$\begin{aligned} \mathbf{A} &= \mathbf{V} \begin{bmatrix} \text{diag}_d(\lambda_i) & \mathbf{X} \\ \mathbf{0} & \tilde{\mathbf{A}} \end{bmatrix} \mathbf{V}^{-1} \\ \Rightarrow \chi_{\mathbf{A}}(\lambda) &= \det(\lambda \mathbf{I} - \mathbf{A}) = \det \begin{bmatrix} \lambda \mathbf{I}_d - \text{diag}_d(\lambda_i) & \mathbf{X} \\ \mathbf{0} & \lambda \tilde{\mathbf{I}}_{n-d} - \tilde{\mathbf{A}} \end{bmatrix} \\ &= (\lambda - \lambda_i)^d \cdot \det(\lambda \tilde{\mathbf{I}}_{n-d} - \tilde{\mathbf{A}}) = (\lambda - \lambda_i)^d \chi_{\tilde{\mathbf{A}}}(\lambda) \end{aligned}$$

Somit teilt $(\lambda - \lambda_i)^d$ das charakteristische Polynom $\chi_{\mathbf{A}}(\lambda)$, woraus folgt $d(\lambda_i) \leq k(\lambda_i)$. \square

Eine interessante Beobachtung hierbei ist, dass für algebraisch abgeschlossene Körper \mathbb{K} folgendes gilt:

$$\begin{aligned} \sum_{\lambda \in \text{spec}(\mathbf{A})} k(\lambda) &= n \\ \sum_{\lambda \in \text{spec}(\mathbf{A})} d(\lambda) &\leq n \end{aligned}$$

Die Eigenvektoren einer Matrix \mathbf{A} bilden nur dann eine Basis von \mathbb{K}^n , wenn $\forall i = 1, \dots, n: k(\lambda_i) = d(\lambda_i)$. Wir erhalten damit das folgende Lemma:

Lemma 10.5: Diagonalisierbarkeit von Matrizen

Eine Matrix $\mathbf{A} \in \text{End}(V)$ ist diagonalisierbar genau dann wenn:

$$\sum_{\lambda \in \text{spec}(V)} d(\lambda) = \dim V$$

Beweis für Lemma 10.5. Sei B eine Basis aus Eigenvektoren von V , es gilt also:

$$V = \bigoplus_{\lambda \in \text{spec}(\mathbf{A})} \mathcal{N}_\lambda = \bigoplus_{\lambda \in \text{spec} \mathbf{A}} \ker(\lambda \mathbf{I} - \mathbf{A})$$

Dabei bezeichnet \oplus die direkte Summe von Unterräumen. Diese Summe ist direkt, da nach Lemma 10.4 die Eigenvektoren verschiedener Eigenwerte linear unabhängig sind. Da die Summe direkt ist folgt hernach ebenfalls:

$$\dim V = \sum_{\lambda \in \text{spec} \mathbf{A}} d(\lambda)$$

Sei B_k eine Basis mit $B_k = \{\mathbf{b}_k^{(1)}, \dots, \mathbf{b}_k^{(d(\lambda_k))}\}$ eine Basis von \mathcal{N}_{λ_k} , mit $1 \leq k \leq r$, dabei ist r die Anzahl der verschiedene Eigenwerte. Somit ist $\bigcup_{k=1}^r B_k$ linear unabhängig:

$$\begin{aligned} \sum_{k=1}^r \left(\sum_{j=1}^{d(\lambda_k)} \alpha_k^{(j)} \mathbf{b}_k^{(j)} \right) &= \mathbf{0} \\ \sum_{j=1}^{d(\lambda_k)} \alpha_k^{(j)} \mathbf{b}_k^{(j)} &= \mathbf{y} \in \mathcal{N}_{\lambda_k} \Rightarrow \mathbf{A} \mathbf{y} = \lambda_k \mathbf{y} \end{aligned}$$

Hier greift erneut wieder Lemma 10.4, denn die obige Summer ergibt $\mathbf{0}$ wenn $\forall k = 1, \dots, r: \mathbf{y}_k = \mathbf{0}$, was wiederum bedeutet, dass $\forall k = 1, \dots, r \wedge \forall j = 1, \dots, d(\lambda_k): \alpha_k^{(j)} = 0$, womit $\mathbf{b}_k^{(1)}, \dots, \mathbf{b}_k^{(d(\lambda_k))}$ linear unabhängig sind. Somit ist $\bigcup_{k=1}^r B_k$ linear unabhängig, womit:

$$\left| \bigcup_{k=1}^r B_k \right| = \sum_{k=1}^r d(\lambda_k) = n$$

Womit $\bigcup_{k=1}^r B_k$ eine Basis von V ist. □

Wir folgern damit, dass eine Matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ diagonalisierbar ist, genau dann wenn $\chi_{\mathbf{A}}(\lambda)$ in Linearfaktoren zerfällt

$$\chi_{\mathbf{A}}(\lambda) = \prod_{k=1}^r (\lambda - \lambda_k)^{d(\lambda_k)}$$

und

$$\forall i = 1, \dots, r: d(\lambda_i) = k(\lambda_i)$$

Hierbei muss man darauf achten, dass unsere erste Bedingung nur über algebraisch abgeschlossenen Körpern automatisch erfüllt ist.

Kapitel XI Die Jordansche Normalform

Dieser Typ von Matrizen geht auf [C. Jordan](#)¹⁵ zurück. Ein dritter Mathematiker mit Nachnamen Jordan war Pascual [P. Jordan](#), nach dem das Konzept der Jordan-Algebra benannt ist. [C. Jordan](#) hat unter anderem auch den **Jordanschen Kurvensatz**¹⁶ bewiesen.

Definition 11.1: Invarianter Unterraum

1. Sei $f \in \text{End}(V)$. Ein Unterraum $U \subseteq V$ heißt **invariant** unter f , wenn $f(U) \subseteq U$

Triviale Beispiele sind natürlich der triviale Unterraum $\{0\}$ und V selbst. Der Kern von f ist invariant, da $x \in \ker f \Rightarrow f(x) = 0$, womit $f(\ker f) = f(0) = 0$. Für das Bild argumentieren wir analog, da $x \in \text{im } f \Rightarrow f(x) \in \text{im } f$. Sei $\lambda \in \text{spec}(f)$, dann gilt für $x \in \mathcal{N}_\lambda$, dass $f(x) = \lambda x$, woraus folgt $f(f(x)) = f(\lambda x) = \lambda f(x) = \lambda^2 x$. Ein invarianter Unterraum der Dimension 1 ist ein Eigenraum. Dafür wählen wir $x \in U \setminus \{0\}$, womit $\{x\}$ eine Basis von U ist, insbesondere gilt $f(x) \in U$, sprich $\exists \lambda \in \mathbb{K}: f(x) = \lambda x$. Ist A eine obere Dreiecksmatrix, dann ist $U_k = \text{span}(e_1, \dots, e_k)$ invariant.

Satz 11.1: Eigenschaften invarianter Unterräume

Sei $A \in \mathbb{K}^{n \times n}$

- i Sei $U \subseteq V$ ein unter A invarianter Unterraum, und $p \in \mathbb{K}[x]$, werten wir $p(A) = \Phi_A(p)$, dann ist U auch invariant unter $p(A)$
- ii Seien U_1, \dots, U_m invariant, dann sind auch $\sum_{k=1}^m U_k$ und $\bigcup_{k=1}^m U_k$ invariant

Beweis für Satz 11.1. Sei $p \in \mathbb{K}[X]$:

$$p(A) = \sum_{k=0}^n c_k A^k$$

Sei $u \in U$, dann gilt $Au \in U$, dann gilt auch $A^k u \in U$. Da U ein Unterraum ist, ist jede Linearkombination in U enthalten:

$$\sum_{k=0}^n c_k \underbrace{A^k u}_{=y_k \in U} = \sum_{k=0}^n c_k y_k \in U$$

Zu ii:

Es gilt:

$$u = \sum_{k=1}^m u_k \in \sum_{k=1}^m U_k \Rightarrow Au = \sum_{k=1}^m Au_k \in \sum_{k=1}^m U_k$$

□

Lemma 11.1: Einschränkung auf invariante Unterräume

Sei $f \in \text{End}(V)$ und $U \subseteq V$ invariant unter f , dann ist $f|_U \in \text{End}(U)$.

Satz 11.2: Matrixdarstellung in Bezug auf invariante Basen

Sei $f \in \text{End}(V)$ und $U, W \subseteq V$ invariant, sodass $U \oplus W = V$. Sei B eine Basis von U und C eine Basis von W (also ist $B \cup C$ eine Basis von V), dann hat die Matrixdarstellung die folgende Form:

$$\Phi_{B \cup C}^{B \cup C}(f) = \begin{bmatrix} \Phi_B^B(f|_U) & 0 \\ 0 & \Phi_C^C(f|_W) \end{bmatrix}$$

¹⁵nicht zu verwechseln mit dem Wilhelm [W. Jordan](#) der den Gauss-Jordan-Algorithmus in seinem *Handbuch für Vermessungskunde* (ein Buch über Geodäsie) erweitert hat

¹⁶ein geschlossener Weg $\gamma \subseteq \mathbb{R}^2$ teilt die Ebene \mathbb{R}^2 in genau zwei Teile, nämlich „innen“ und „außen“, sofern γ sich nicht selbst schneidet

Korollar 11.1: zu Satz 11.2

Sei $f \in \text{End}(V)$ und $V = U_1 \dot{+} U_2 \dot{+} \dots \dot{+} U_m$, wobei die U_i invariant unter f sind, und $B = B_1 \cup \dots \cup B_m$, dann gilt:

$$\Phi_B^B(f) = \text{diag}_{k=1}^m (\Phi_{B_k}^{B_k}(f|_{U_k}))$$

Ist (b_1, \dots, b_k) eine Basis des Eigenraumes \mathcal{N}_λ , dann gilt $f(b_i) = \lambda b_i$, dann ist natürlich auch $\text{span}(b_i)$ invariant unter f . Somit folgt:

$$\Phi_B^B(f|_{\mathcal{N}_\lambda}) = \text{diag}_{d(\lambda)}(\lambda)$$

Korollar 11.2: aus Korollar 11.1

Sei $f \in \text{End}(V)$, mit $f(x) = Ax$, und U_i invariant unter f , wobei $V = U_1 \dot{+} \dots \dot{+} U_m$, dann folgt:

$$\chi_A(\lambda) = \prod_{k=1}^m \chi_{A|_{U_k}}(\lambda)$$

Lemma 11.2: Lemma von Fitting

Sei V ein Vektorraum mit $\dim V = n$ und $f \in \text{End}(V)$:

- a $\{0\} \subseteq \ker(f) \subseteq \ker(f^2) \subseteq \dots$
- b $\text{im}(f) \supseteq \text{im}(f^2) \supseteq \text{im}(f^3) \supseteq \dots$
- c $\exists m \leq n: \ker(f^m) = \ker(f^{m+1})$
- d $\exists m' \leq n: \text{im}(f^{m'}) = \text{im}(f^{m'+1})$
- e die folgenden Aussagen sind äquivalent:
 - i $\ker(f^m) = \ker(f^{m+1})$
 - ii $\text{im}(f^m) = \text{im}(f^{m+1})$
 - iii $\forall k \geq 1: \ker(f^m) = \ker(f^{m+k})$
 - iv $\forall k \geq 1: \text{im}(f^m) = \text{im}(f^{m+k})$
 - v $\ker(f^m) \cap \text{im}(f^m) = \{0\}$
 - vi $V = \ker(f^m) \dot{+} \text{im}(f^m)$

Lemma 11.2 ist nach Hans **Fitting** benannt.

Betrachten wir vor dem Beweis erst ein Beispiel, um ein Gefühl dafür zu bekommen, wie sich die wiederholte Anwendung von f auf den Kern und das Bild auswirkt.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A^3 = 0$$

Wir sehen direkt:

$$\ker A = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \quad \text{im} A = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Für A^2 folgt:

$$\ker(A^2) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \quad \text{im}(A^2) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Wir sehen, dass $\ker(A) \subseteq \ker(A^2)$ und $\text{im}(A) \supseteq \text{im}(A^2)$:

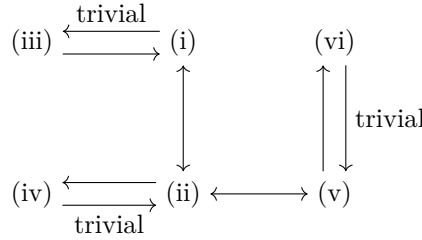
$$\ker(A^3) = \mathbb{R}^3 \quad \text{im}(A^3) = \{0\}$$

Eine weitere Anwendung erkennen wir bei Betrachtung der folgenden Matrix:

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

Wir bestimmen nun $\ker(\lambda I - A) \subset \ker((\lambda I - A)^2) \subset \ker((\lambda I - A)^3)$. Sei $x \in \ker((\lambda I - A)^2)$, dann folgt $(\lambda I - A)x \in \ker(\lambda I - A)$. Es gilt daher $(\lambda I - A)x = v$ ist ein Eigenvektor und es folgt $Ax = \lambda x - v$.

Beweis für Lemma 11.2. Wir arbeiten die folgenden Äquivalenzen ab:



Wir beginnen mit (i) \rightarrow (iii). Sei $\ker(f^m) = \ker(f^{m+1})$, dann müssen wir zeigen, dass folgt $\ker(f^m) = \ker(f^{m+k}) \forall k \geq 1$. Wir wissen $\ker(f^m) \subseteq \ker(f^{m+1})$. Sei $m_0 = \min\{m \in \mathbb{N} : \ker(f^m) = \ker(f^{m+1})\}$. Wir behaupten nun, dass $\forall k \geq 1: \ker(f^{m_0+k}) = \ker(f^{m_0})$. Sei $x \in \ker(f^{m_0+k+1})$, womit $f^{m_0+k+1}(x) = 0$, was äquivalent ist zu $f^{m_0+1}(f^k(x)) = 0$, womit $f^k(x) \in \ker(f^{m_0+1}) = \ker(f^{m_0})$, daher auch $x \in \ker(f^{m_0+k})$.

Für (ii) \rightarrow (iv) gehen wir ähnlich vor. Angenommen $\text{im}(f^m) = \text{im}(f^{m+1})$ gilt, wir wollen zeigen, dass daraus folgt $\text{im}(f^{m+k+1}) = \text{im}(f^{m+k}) \forall k \geq 0$. Sei $y \in \text{im}(f^{m+k})$, wir suchen nun z sodass $y = f^{m+k+1}(z)$. Da $y \in \text{im}(f^{m+k})$ folgt $\exists x: f^{m+k}(x) = y$, bzw. $f^k(f^m(x))$. Es folgt $f^m(x) \in \text{im}(f^m) = \text{im}(f^{m+1})$ womit weiter gilt $\exists z: f^m(x) = f^{m+1}(z)$.

Weiter mit (i) \leftrightarrow (ii). Wir wissen $\ker(f^m) \subseteq \ker(f^{m+1})$ und $\text{im}(f^m) \supseteq \text{im}(f^{m+1})$. Es gilt $\ker(f^m) = \ker(f^{m+1})$ wenn $\dim(\ker(f^m)) = \dim(\ker(f^{m+1}))$. Nach dem Dimensionssatz gilt $\dim(\ker(f^m)) + \dim(\text{im}(f^m)) = n$. Wenn $\dim(\ker(f^{m+1})) + \dim(\text{im}(f^{m+1})) = n$, dann gilt $\ker(f^m) = \ker(f^{m+1})$ und $\text{im}(f^m) = \text{im}(f^{m+1})$.

Da (i) \Leftrightarrow (ii) \Leftrightarrow (iii) \Leftrightarrow (iv), können wir alle 4 Behauptungen verwenden. Sei $W = \text{im}(f^m)$ und $g = f^m|_W$, also $g: W \rightarrow W$. Es gilt $\ker g = \ker(f^m) \cap W = \ker(f^m) \cap \text{im}(f^m)$. Das ist der Fall wenn $\ker(g) = \{0\}$, also ist g bijektiv, womit $\text{img} = W$, da $\text{img} = g(W) = f^m(f^m(V)) = f^{2m}(V) = \text{im}(f^{2m})$, nach (iv) folgt somit $\text{im}(f^{m+1}) = \text{im}(f^m)$.

Wir schließen mit (v) \rightarrow (vi) ab. Sei $\ker(f^m) \cap \text{im}(f^m) = \{0\}$, sprich $\ker(f^m) + \text{im}(f^m)$ ist direkt. Nach dem Dimensionssatz folgt $\dim(\ker(f^m)) + \dim(\text{im}(f^m)) = n$, daher $\ker(f^m) + \text{im}(f^m) = V$. \square

Definition 11.2: Hauptraum

Sei $A \in \mathbb{K}^{n \times n}$ und $\lambda \in \text{spec}(A)$, dann heißt $\ker((\lambda I - A)^n)$ der verallgemeinerte Eigenraum, bzw. der **Hauptraum** von A zum Eigenwert λ .

Da $\exists m \leq n: \ker((\lambda I - A)^m) = \ker((\lambda I - A)^{m+1})$, folgt $\ker((\lambda I - A)^m) = \ker((\lambda I - A)^n)$.

Lemma 11.3

Seien $\lambda_1, \dots, \lambda_k$ verschiedene Eigenwerte einer Matrix A und $\ker((\lambda_i I - A)^{r_i})$ die zugehörigen Haupträume. Dann ist

$$\left(\bigcap_{i=1}^k \text{im}((\lambda_i I - A)^{r_i}) \right) \cap \ker \left(\prod_{i=1}^k (\lambda_i I - A)^{r_i} \right) = \{0\}$$

Betrachten wir dazu wieder erst ein Beispiel.

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{spec}(A) = \{1, 2, \pm i\}$$

Über den reellen Zahlen gilt natürlich $\text{spec}(A) = \{1, 2\}$. Für $\lambda = 1$ erhalten wir die folgende Matrix:

$$I - A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \Rightarrow \ker(I - A) = \text{span}(e_1, e_3)$$

$$(I - A)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \end{bmatrix} \Rightarrow \ker((I - A)^2) = \text{span}(e_1, e_2, e_3)$$

Hier gilt bereits $\ker((I - A)^2) = \ker((I - A)^3)$, womit $\ker((I - A)^2)$ der Hauptraum zum Eigenwert 1 ist. Für $\lambda = 2$ erhalten wir:

$$2I - A = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix} \Rightarrow \ker(2I - A) = \text{span}(e_4)$$

$$(2I - A)^2 = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 3 \end{bmatrix} \Rightarrow \ker((2I - A)^2) = \text{span}(e_4, e_5)$$

$$\text{im}((I - A)^2) = \text{span}(e_4, e_5, e_6, e_7, e_8)$$

Zuletzt erhalten wir:

$$(2\mathbf{I} - \mathbf{A})^3 = \begin{bmatrix} 1 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 11 \\ 0 & 0 & 0 & 0 & 0 & 0 & -11 & 2 \end{bmatrix} \Rightarrow \ker((2\mathbf{I} - \mathbf{A})^3) = \text{span}(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6)$$

$$\text{im}((2\mathbf{I} - \mathbf{A})^3) = \text{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_7, \mathbf{e}_8)$$

Der Hauptraum zum Eigenwert $\lambda = 2$ ist also $\ker((2\mathbf{I} - \mathbf{A})) = \text{span}(\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6)$.

$$\bigcap_{i=1}^2 \text{im}((\lambda_i \mathbf{I} - \mathbf{A})^{r_i}) = \text{im}((\mathbf{I} - \mathbf{A})^2) \cap \text{im}((2\mathbf{I} - \mathbf{A})^3) = \text{span}(\mathbf{e}_7, \mathbf{e}_8)$$

$$\ker((\mathbf{I} - \mathbf{A})^2(2\mathbf{I} - \mathbf{A})^3) \supseteq \ker((2\mathbf{I} - \mathbf{A})^3) + \ker((\mathbf{I} - \mathbf{A})^2) = \text{span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6)$$

Wonach der gegebene Durchschnitt aus Lemma 11.3 tatsächlich trivial ist.

Beweis für Lemma 11.3. Es ist zu zeigen. Wenn $\mathbf{x} \in \bigcap_{i=1}^k \text{im}((\lambda_i \mathbf{I} - \mathbf{A})^{r_i})$ und $(\prod_{i=1}^k (\lambda_i \mathbf{I} - \mathbf{A})^{r_i})\mathbf{x} = \mathbf{0}$, dann ist $\mathbf{x} = \mathbf{0}$. Dazu führen wir eine Induktion nach k durch. Für $k = 1$ folgt die Aussage direkt aus dem Lemma von Fitting.

Sei nun $\mathbf{x} \in \bigcap_{i=1}^{k+1} \text{im}((\lambda_i \mathbf{I} - \mathbf{A})^{r_i})$ und $(\prod_{i=1}^{k+1} (\lambda_i \mathbf{I} - \mathbf{A})^{r_i})\mathbf{x} = \mathbf{0}$.

$$\mathbf{y} = (\lambda_{k+1} \mathbf{I} - \mathbf{A})^{r_{k+1}} \mathbf{x} \in \ker \left(\prod_{i=1}^k (\lambda_i \mathbf{I} - \mathbf{A})^{r_i} \right)$$

Wir behaupten nun:

$$\mathbf{y} \in \bigcap_{i=1}^k \text{im}((\lambda_i \mathbf{I} - \mathbf{A})^{r_i})$$

Da $\mathbf{x} \in \bigcap_{i=1}^k \text{im}((\lambda_i \mathbf{I} - \mathbf{A})^{r_i})$, gilt:

$$\begin{aligned} \exists i: \exists \mathbf{u}_i: \mathbf{x} &= (\lambda_i \mathbf{I} - \mathbf{A})^{r_i} \mathbf{u}_i \\ \Rightarrow \mathbf{y} &= (\lambda_{k+1} \mathbf{I} - \mathbf{A})^{r_{k+1}} (\lambda_i \mathbf{I} - \mathbf{A})^{r_i} \mathbf{u}_i = (\lambda_i \mathbf{I} - \mathbf{A})^{r_i} (\lambda_{k+1} \mathbf{I} - \mathbf{A})^{r_{k+1}} \mathbf{u}_i \in \text{im}((\lambda_i \mathbf{I} - \mathbf{A})^{r_i}) \end{aligned}$$

Nach der Induktionsvoraussetzung folgt damit $\mathbf{y} = \mathbf{0}$, wonach

$$\mathbf{x} \in \ker((\lambda_k \mathbf{I} - \mathbf{A})^{r_{k+1}}) \cap \text{im}((\lambda_{k+1} \mathbf{I} - \mathbf{A})^{r_{k+1}})$$

Nach Lemma 11.2 folgt somit $\mathbf{x} = \mathbf{0}$: □

Lemma 11.4

- i $\forall \lambda \neq \mu, \forall k, l \geq 1: \ker((\lambda \mathbf{I} - \mathbf{A})^k) \cap \ker((\mu \mathbf{I} - \mathbf{A})^l) = \{\mathbf{0}\}$
 ii die Summe $\sum_{i=1}^k \ker((\lambda_i \mathbf{I} - \mathbf{A})^{r_i})$ ist direkt

Beweis für Lemma 11.4. Für $k = l = 1$ sind die Kerne genau die Eigenräume, womit auch der Durchschnitt leer sein muss.

Zu ii:

Wir führen eine Induktion nach k . Für $k = 1$ besteht die Summe nur aus einem Summanden. Im Induktionsschritt zeigen wir nun, wenn $\mathbf{v}_i \in \ker((\lambda_i \mathbf{I} - \mathbf{A})^{r_i})$ für $i = 1, \dots, k+1$ und die Summe $\mathbf{v}_1 + \dots + \mathbf{v}_{k+1} = \mathbf{0}$, dann sind

alle $\mathbf{v}_i = \mathbf{0}$. Sei $\mathbf{w}_i = (\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}\mathbf{v}_i$ für $i = 1, \dots, k+1$, dann ist $\mathbf{w}_{k+1} = \mathbf{0}$, da $\mathbf{v}_{k+1} \in \ker((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}})$, womit:

$$\sum_{i=1}^k \mathbf{w}_i = \sum_{i=1}^{k+1} \mathbf{w}_i = \sum_{i=1}^{k+1} (\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}\mathbf{v}_i = (\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}} \underbrace{\sum_{i=1}^{k+1} \mathbf{v}_i}_{=\mathbf{0}} = \mathbf{0}$$

Wir behaupten nun $\mathbf{w}_i \in \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$:

$$\begin{aligned} (\lambda_i\mathbf{I} - \mathbf{A})^{r_i}\mathbf{w}_i &= (\lambda_i\mathbf{I} - \mathbf{A})^{r_i}(\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}\mathbf{v}_i \\ &= (\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}} \underbrace{(\lambda_i\mathbf{I} - \mathbf{A})^{r_i}\mathbf{v}_i}_{=\mathbf{0}} = \mathbf{0} \end{aligned}$$

Somit sind $\mathbf{w}_i \in \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$ und $\sum_{i=1}^k \mathbf{w}_k = \mathbf{0}$. Nach der Induktionsvoraussetzung folgt damit, dass alle $\mathbf{w}_i = \mathbf{0}$, womit $\mathbf{v}_i \in \ker((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}) \cap \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i}) = \{\mathbf{0}\}$. Somit sind $\mathbf{v}_i = \mathbf{0}$, wonach auch $\mathbf{v}_{k+1} = \mathbf{0}$. \square

Satz 11.3

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$ und $\lambda_1, \dots, \lambda_k \in \text{spec}(\mathbf{A})$ verschieden:

$$\text{i } \left(\bigoplus_{i=1}^k \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i}) \right) \oplus \underbrace{\bigcap_{i=1}^k \text{im}((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})}_{=W} = V$$

ii W ist invariant unter \mathbf{A}

Beweis für Satz 11.3. Für i führen wir einen Induktionsbeweis. Für $k = 1$ gilt nach dem Lemma von Fitting wieder $V = \ker((\lambda_1\mathbf{I} - \mathbf{A})^{r_1}) \oplus \text{im}((\lambda_1\mathbf{I} - \mathbf{A})^{r_1})$. Im Induktionsschritt nehmen wir an, dass i für $\lambda_1, \dots, \lambda_k$ gilt. Sei $W_k = \bigcap_{i=1}^k \text{im}((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$, dann ist W_k invariant unter \mathbf{A} . Sei $\mathbf{y} \in \text{im}((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$, dann $\exists \mathbf{x}_i: \mathbf{y} = (\lambda_i\mathbf{I} - \mathbf{A})^{r_i}\mathbf{x}_i$ und es folgt $\mathbf{A}\mathbf{y} = \mathbf{A}(\lambda_i\mathbf{I} - \mathbf{A})^{r_i}\mathbf{x}_i = (\lambda_i\mathbf{I} - \mathbf{A})^{r_i}\mathbf{A}\mathbf{x}_i \in \text{im}((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$. Es folgt daher $\forall i: \mathbf{A}\mathbf{y} \in \text{im}((\lambda_i\mathbf{I} - \mathbf{A})^{r_i}) \Rightarrow \mathbf{A}\mathbf{y} \in W_k$. Wir betrachten nun $\mathbf{A}|_{W_k}$. Nach dem Lemma von Fitting gilt $W_k = \ker((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}) \oplus \text{im}((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}) \subseteq V$, insbesondere:

$$W_k \subseteq \ker((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}) + \underbrace{W_k \cap \text{im}((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}})}_{=W_{k+1}}$$

Es folgt daher $V \subseteq \ker((\lambda_1\mathbf{I} - \mathbf{A})^{r_1}) + \dots + \ker((\lambda_k\mathbf{I} - \mathbf{A})^{r_k}) + \ker((\lambda_{k+1}\mathbf{I} - \mathbf{A})^{r_{k+1}}) + W_{k+1}$. Es verbleibt zu zeigen, dass diese Summe direkt ist. Seien $\mathbf{x}_i \in \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$ und $\mathbf{y} \in W_{k+1}$. Wenn $\mathbf{x}_1 + \dots + \mathbf{x}_{k+1} + \mathbf{y} = \mathbf{0}$, dann $\forall i: \mathbf{x}_i = \mathbf{0}$ und $\mathbf{y} = \mathbf{0}$:

$$\begin{aligned} \mathbf{0} &= \prod_{i=1}^{k+1} (\lambda_i\mathbf{I} - \mathbf{A})^{r_i} \left(\sum_{i=1}^{k+1} \mathbf{x}_i + \mathbf{y} \right) = \sum_{i=1}^{k+1} \mathbf{0} + \prod_{i=1}^{k+1} (\lambda_i\mathbf{I} - \mathbf{A})^{r_i} \mathbf{y} \\ &\Rightarrow \mathbf{y} \in \ker \left(\prod_{i=1}^{k+1} (\lambda_i\mathbf{I} - \mathbf{A})^{r_i} \right) \cap W_{k+1} = \{\mathbf{0}\} \end{aligned}$$

Somit ist die Summe direkt und V entsteht genau durch i.

Zu ii: $\lambda_i \notin \text{spec}(\mathbf{A}|_{W_k})$, da $\ker(\lambda_i\mathbf{I} - \mathbf{A}) \cap W_k \subseteq \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i}) \cap W_k = \{\mathbf{0}\}$. \square

Korollar 11.3: zu Satz 11.3

Wenn \mathbb{K} algebraisch abgeschlossen ist und $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ alle Eigenwerte der Matrix \mathbf{A} sind, dann gilt:

$$\mathbb{K}^n = \bigoplus_{i=1}^n \ker((\lambda_i\mathbf{I} - \mathbf{A})^{r_i})$$

Beweis für Korollar 11.3. $\mathbf{A}|_{W_k}$ hat keine Eigenwerte, womit $W_k = \{\mathbf{0}\}$. \square

Sei $A_i = A|_{\ker((\lambda_i I - A)^n)}$, dann gilt $\text{spec}(A_i) = \{\lambda_i\}$ und $(\lambda_i I - A)^n = 0$.

Definition 11.3: Nilpotente Matrizen

Eine Matrix (bzw. lineare Abbildung) $A \in \mathbb{K}^{n \times n}$ heißt nilpotent, falls $\exists r \in \mathbb{N}: A^r = 0$.

Nilpotente Matrizen sind nicht diagonalisierbar, denn sie besitzen nur den Eigenwert 0, womit $\chi_A(\lambda) = x^n$, es gilt also $\dim(\ker A) < n$. Unser Ziel ist nun, nilpotente Matrizen möglichst einfach darzustellen. Das werden wir durch eine geschickte Wahl der Basis erreichen.

Lemma 11.5

Sei $\ker(f^k) \subseteq \ker(f^{k+1}) \subseteq \ker(f^{k+2})$. Sei (u_1, \dots, u_p) eine Basis von $\ker(f^k)$ und $(u_1, \dots, u_p, v_1, \dots, v_q)$ eine Basis von $\ker(f^{k+1})$, bzw. $(u_1, \dots, u_p, v_1, \dots, v_q, w_1, \dots, w_r)$ eine Basis von $\ker(f^{k+2})$, dann ist $(u_1, \dots, u_p, f(w_1), \dots, f(w_r))$ linear unabhängig.

Betrachten wir ein Beispiel zu Lemma 11.5:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Wir sehen direkt an A , dass $\ker A = \text{span}(e_1)$, $\ker(A^2) = \text{span}(e_1, e_2)$ und $\ker(A^3) = \text{span}(e_1, e_2, e_3)$. Sei $v \in \ker(A^3) \setminus \ker(A^2)$ etwa $v = e_3$, dann $Ae_3 = e_2 \in \ker(A^2) \setminus \ker(A^3)$.

Beweis für Lemma 11.5. Wir wissen $f(\ker(f^{k+2})) \subseteq \ker(f^{k+1})$. Angenommen

$$\sum_{i=1}^p \lambda_i u_i + \sum_{j=1}^r \mu_j f(w_j) = 0$$

dann folgt

$$\begin{aligned} 0 &= f^k(0) = \sum_{i=1}^p \underbrace{\lambda_i f^k(u_i)}_{=0} + \sum_{j=1}^r \mu_j f^{k+1}(w_j) = f^{k+1} \left(\sum_{j=1}^r \mu_j w_j \right) \\ &\Rightarrow \sum_{j=1}^r \mu_j w_j \in \ker(f^{k+1}) = \text{span}(u_1, \dots, u_p, v_1, \dots, v_q) \\ \ker(f^{k+2}) &= \ker(f^{k+1}) \oplus \text{span}(w_1, \dots, w_r) \\ &\Rightarrow \ker(f^{k+1}) \cap \text{span}(w_1, \dots, w_r) = \{0\} \end{aligned}$$

Somit folgt:

$$\sum_{j=1}^r \mu_j w_j = 0 \Rightarrow \mu_j = 0 \Rightarrow \sum_{i=1}^p \lambda_i u_i = 0 \Rightarrow \lambda_i = 0$$

□

Satz 11.4: Jordansche Normalform

Im einfachsten Fall hat **A** nur einen Eigenwert, wobei $\dim(V) = n$. Sei $f \in \text{End}(V)$ nilpotent vom Index p , es gilt also $f^p = 0$, und $d = \dim(\ker(f))$, dann gibt es eine Basis B von V , sodass

$$\Phi_B^B(f) = \text{diag}_1^d(\mathbf{N}_i) \quad \mathbf{N}_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix} \in \mathbb{K}^{n_i \times n_i}$$

Sei $n_1 = p \geq n_2 \geq \dots \geq n_d$, wobei $\sum_{i=1}^d n_i = n$.

Im zweiten Fall betrachten wir $\text{spec}(\mathbf{A}) = \{\lambda\}$ und \mathbb{K} algebraisch abgeschlossen, dann ist $(\mathbf{A} - \lambda \mathbf{I})$ nilpotent, womit $\Phi_B^B(\mathbf{A} - \lambda \mathbf{I}) = \text{diag}_i^d(\mathbf{N}_i)$, wonach

$$\Phi_B^B(\mathbf{A}) = \text{diag}_1^d(\mathbf{N}_i(\lambda)) \quad \mathbf{N}_i(\lambda) = \mathbf{N}_i + \lambda \mathbf{I}_{n_i}$$

Beweis für Satz 11.4. Sei $U_k = \ker(f^k)$ für $k = 1, \dots, p$ und $\dim U_k = m_k$, wobei $m_1 = d$, dann gilt $d = m_1 \leq m_2 \leq \dots \leq m_p = n$. Wir wissen bereits $f(U_k) \subseteq U_{k-1}$. Wir wählen nun eine Basis, wobei jeweils die erweiternden Basisvektoren der U_k in $\mathbf{u}_{m_{k-1}+1}, \dots, \mathbf{u}_{m_k}$ gruppiert sind:

$$\underbrace{(\mathbf{u}_1, \dots, \mathbf{u}_{m_1})}_{U_1}, \underbrace{\mathbf{u}_{m_1+1}, \dots, \mathbf{u}_{m_2}}_{U_2}, \underbrace{\mathbf{u}_{m_2+1}, \dots, \mathbf{u}_{m_3}}_{U_3}, \dots, \underbrace{\mathbf{u}_{m_p}, \dots, \mathbf{u}_{m_p+n}}_{U_p=V}$$

Wir beginnen nun von hinten und setzen $\mathbf{v}_1^{(p)} = \mathbf{u}_{m_{p-1}+1}, \mathbf{v}_2^{(p)} = \mathbf{u}_{m_{p-1}+2}, \dots, \mathbf{v}_{m_p-m_{p-1}}^{(p)} = \mathbf{u}_{m_p}$, worauf wir wiederum f anwenden:

$$\mathbf{v}_1^{(p-1)} = f\left(\mathbf{v}_1^{(p)}\right) \quad \dots \quad \mathbf{v}_{m_p-m_{p-1}}^{(p-1)} = f\left(\mathbf{v}_{m_p-m_{p-1}}^{(p)}\right)$$

Dabei beachten wir $\mathbf{v}_k^{(p-1)} \in U_{p-1} \setminus U_{p-2}$ nach Lemma 11.5. Hernach sind $\mathbf{v}_k^{(p-1)}$ linear unabhängig von $\mathbf{u}_1, \dots, \mathbf{u}_{m_{p-2}}$ und liegen in U_{p-1} . Wir ergänzen nun $\mathbf{v}_1, \dots, \mathbf{u}_{m_{p-2}}, \mathbf{v}_1^{(p-1)}, \dots, \mathbf{v}_{m_p-m_{p-1}}^{(p-1)}$ zu einer Basis von U_{p-1} mit $\mathbf{v}_{m_p-m_{p-1}+1}^{(p-1)}, \dots, \mathbf{v}_{m_p-m_{p-2}}^{(p-1)}$ und wenden darauf wieder f an:

$$\mathbf{v}_1^{(p-2)} = f\left(\mathbf{v}_1^{(p-1)}\right) \quad \dots \quad \mathbf{v}_{m_{p-1}-m_{p-2}}^{(p-2)} = f\left(\mathbf{v}_{m_{p-1}-m_{p-2}}^{(p-1)}\right)$$

Aufgrund von Lemma 11.5 sind die $\mathbf{v}_i^{(p-2)} \in U_{p-2} \setminus U_{p-3}$, wir ergänzen

$$u_1, \dots, v_{m_{p-3}}, v_1^{(p-2)}, \dots, v_{m_{p-1}-m_{p-2}}^{(p-2)}$$

zu einer Basis von U_{p-2} . Das wiederholen wir, bis wir die folgende Situation erhalten:

Basis von:

$$\begin{array}{ccccccc}
U_p \ominus U_{p-1} & & \mathbf{v}_1^{(p)} & \mathbf{v}_2^{(p)} & \cdots & \mathbf{v}_{m_p-m_{p-1}}^{(p)} & \\
\downarrow f & & & & & & \\
U_{p-1} \ominus U_{p-2} & & \mathbf{v}_1^{(p-1)} & \mathbf{v}_2^{(p-1)} & \cdots & \mathbf{v}_{m_p-m_{p-1}}^{(p-1)} & \cdots \quad \mathbf{v}_{m_{p-1}-m_{p-2}}^{(p-1)} \\
\downarrow f & & & & & & \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\
U_1 & & \mathbf{v}_1^{(1)} & \mathbf{v}_2^{(1)} & \cdots & \cdots & \cdots \quad \mathbf{v}_{m_{p-1}-m_{p-2}}^{(1)} \quad \cdots \quad \mathbf{v}_{m_1}^{(1)}
\end{array}$$

Es gilt:

- i die Elemente der letzten k Zeilen bilden eine Basis von U_k
- ii f bildet die k -te Zeile auf die $k+1$ Zeile ab
- iii wir bilden die Basis B , indem wir die Elemente der Spalten von unten nach oben und von links nach rechts „auffädeln“

Wir erhalten somit die folgende Matrix:

$$\Phi_B^B(f) = \text{diag}_1^d(N_i)$$

Für $\lambda \neq 0$ ist $(A - \lambda I)$ nilpotent □

Definition 11.4: Allgemeiner Jordanblock

Eine Matrix der Form

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \lambda \end{bmatrix} \in \mathbb{K}^{k \times k}$$

heißt **Jordanblock** der Länge k zum Eigenwert λ .

Da $J_k(\lambda)$ eine obere Dreiecksmatrix ist, folgt natürlich direkt:

$$\chi_{J_k(\lambda_0)}(\lambda) = (\lambda - \lambda_0)^k$$

Weiterhin gilt natürlich:

$$(J_k(\lambda) - \lambda I_k)^k = 0$$

Satz 11.5: Jordansche Normalform

Sei \mathbb{K} algebraisch abgeschlossen, dann ist jede Matrix $A \in \mathbb{K}^{n \times n}$ ähnlich zu einer Matrix in Jordanscher Normalform:

$$\exists B \in \text{GL}(n, \mathbb{K}): B^{-1}AB = \text{diag}_1^q(J_i)$$

wobei die J_i die Jordanblöcke zu den Eigenwerten von A sind. Zum Eigenwert λ_i gibt es d_i Blöcke, wobei $d_i = \dim W_{\lambda_i}$.

Beweis für Satz 11.5. Seien $\lambda_1, \dots, \lambda_m$ die verschiedenen Eigenwerte von A und $U_i = \ker((\lambda_i - A)^n)$ die Haupträume, da \mathbb{K} algebraisch abgeschlossen ist, folgt aus Korollar 11.3, dass $\mathbb{K}^n = \bigoplus_{i=1}^m U_i$ und es genügt $A|_{U_i}$ zu betrachten. Wir wissen $(A - \lambda_i I)|_{U_i}$ ist nilpotent, womit es nach Satz 11.4 eine Basis B_i von U_i gibt, sodass:

$$\begin{aligned} \Phi_{B_i}^{B_i}((A - \lambda_i I)|_{U_i}) &= \text{diag}_1^{d_i}(N_{n_{ij}}) \\ \Rightarrow \Phi_{B_i}^{B_i}(A|_{U_i}) &= \text{diag}_1^{d_i}(J_{n_{ij}}(\lambda_i)) \end{aligned}$$

Hernach gilt für $B = \bigcap_{i=1}^m B_i$:

$$\Phi_B^B(A) = \text{diag}_1^m(\Phi_{B_i}^{B_i}(A|_{U_i}))$$

□

Korollar 11.4: zu Satz 11.5

Sei $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ in Jordanscher Normalform:

- i $\sum_{i=1}^m k_i = n$
- ii für $\lambda \in \text{spec}(\mathbf{A})$ gilt $d(\lambda)$ entspricht der Anzahl der Blöcke und weiter gilt $\sum_{\lambda_i=\lambda} k_i = k(\lambda)$
- iii die kleinste Potenz r_i für die $\ker((\lambda_i \mathbf{I} - \mathbf{A})^{r_i}) = \ker((\lambda_i \mathbf{I} - \mathbf{A})^n)$ ist genau $\max\{k_i | \lambda = \lambda_i\}$, also die Größe des größten Blocks zum Eigenwert λ
- iv $|\{i | \lambda_i = \lambda \wedge k_i \geq j + 1\}|$:

$$\text{rank}((\lambda \mathbf{I} - \mathbf{A})^k) - \text{rank}((\lambda \mathbf{I} - \mathbf{A})^{k+1}) = \dim(\ker(\lambda \mathbf{I} - \mathbf{A})^{k+1}) - \dim(\ker(\lambda \mathbf{I} - \mathbf{A})^k)$$

v die Jordanblöcke sind bis auf ihre Reihenfolge eindeutig bestimmt

11.1 Anwendungen

Betrachten wir erneut den Einsetzungshomomorphismus, $\psi_{\mathbf{A}}: \mathbb{K}[x] \rightarrow \mathbb{K}^{n \times n}$ mit $p \mapsto p(\mathbf{A})$. Ist $\mathbf{A} = \text{diag}(\lambda_i)$, dann gilt:

$$p(\mathbf{A}) = \text{diag}(p(\lambda_i))$$

Ist analog $\mathbf{A} = \text{diag}(\mathbf{A}_i)$, dann gilt

$$p(\mathbf{A}) = \text{diag}(p(\mathbf{A}_i))$$

Wenn $\mathbf{A} = \mathbf{T}^{-1}\mathbf{B}\mathbf{T}$, dann gilt $p(\mathbf{A}) = \mathbf{T}^{-1}p(\mathbf{B})\mathbf{T}$

Lemma 11.6: Auswertung von Jordanblöcken

Sei \mathbf{J} ein Jordanblock zum Eigenwert λ , dann gilt:

$$p(\mathbf{J})_{ij} = \begin{cases} p^{(j-i)}(\lambda) & j \geq i \\ 0 & j < i \end{cases}$$

wobei $p^{(k)} = \frac{d^k p}{dx^k}$.

Beweis für Lemma 11.6. Für kommutierende Matrizen gilt:

$$(\mathbf{A} + \mathbf{B})^n = \sum_{k=0}^n \binom{n}{k} \mathbf{A}^k \mathbf{B}^{n-k}$$

Da $\mathbf{J} = \lambda \mathbf{I} - \mathbf{N}$ und \mathbf{N} nilpotent ist, gilt daher:

$$(\mathbf{J}_\lambda)^m = (\lambda \mathbf{I} - \mathbf{N})^m = \sum_{k=0}^m \binom{m}{k} \lambda^{m-k} \mathbf{N}^k$$

□

Betrachten wir erneut das Beispiel mit der Fibonacci-Folge. Sei $\varphi_{\pm} = \frac{1 \pm \sqrt{5}}{2}$:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \mathbf{B}^{-1} \begin{bmatrix} \varphi_+ & 0 \\ 0 & \varphi_- \end{bmatrix} \mathbf{B} \\ \Rightarrow \mathbf{A}^n &= \mathbf{B}^{-1} \begin{bmatrix} \varphi_+^n & 0 \\ 0 & \varphi_-^n \end{bmatrix} \mathbf{B} \end{aligned}$$

In der Biologie betrachtet man sogenannte **Räuber-Beute-Modelle** der folgenden Form:

$$\begin{aligned} F_{n+1} &= qF_n + qH_n \\ H_{n+1} &= -tF_n + gH_n \end{aligned}$$

Sei F_n die Anzahl der Füchse und H_n die Anzahl der Hasen, jeweils zum Zeitpunkt n . Dieses System können wir in Matrix-Form bringen:

$$\begin{bmatrix} F_{n+1} \\ H_{n+1} \end{bmatrix} = \underbrace{\begin{bmatrix} p & q \\ -t & g \end{bmatrix}}_{=A} \begin{bmatrix} F_n \\ H_n \end{bmatrix} \Rightarrow \begin{bmatrix} F_n \\ H_n \end{bmatrix} = A^n \begin{bmatrix} F_0 \\ H_0 \end{bmatrix}$$

$$A = B^{-1} \text{diag}(\lambda_1 \quad \lambda_2) B \quad \lambda_2) B \Rightarrow A^n = [B] \text{diag}(\lambda_1^n, \lambda_2^n) B$$

Für $n \rightarrow \infty$ ist die Entwicklung des Systems abhängig von λ_i . Für $\lambda_i > 1$ gilt natürlich, dass das System divergiert, also $\| \begin{bmatrix} F_n & H_n \end{bmatrix} \| \rightarrow \infty$. Für $\lambda_i < 1$ gilt $\begin{bmatrix} F_n & H_n \end{bmatrix} \xrightarrow{n \rightarrow \infty} \mathbf{0}$. Sind etwa $\lambda_1 > 1$ und $\lambda_2 < 1$, dann ist die Entwicklung abhängig ob $\begin{bmatrix} F_0 & H_0 \end{bmatrix}$ ein Eigenvektor zu λ_1 oder λ_2 ist.

Man kann auch kontinuierliche Systeme betrachten, die auf lineare Differentialgleichungen führen:

$$\frac{dy}{dx} = Ay \Rightarrow y(x) = e^{Ax} y_0$$

Diese Idee können wir auf beliebige Funktionen f erweitern. Sei A diagonalisierbar mit Eigenwerten $\lambda_1, \dots, \lambda_n$, dann gilt $f(A) = T^{-1} \text{diag}(f(\lambda_i)) T$, womit wir ψ_A auf $\psi_A: \mathbb{K} \rightarrow \mathbb{K}^{n \times n}$ erweitern können. Für Jordanblöcke muss man die Definition auf analytische Funktionen einschließen. Eine Funktion ist analytisch, falls:

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

Sofern die Reihe einen Konvergenzradius $R > 0$ hat, können wir $f(A)$ bestimmen, sofern $\forall i: |\lambda_i| < R$. Somit erhalten wir die Matrix-Exponentialfunktion:

$$e^A = \sum_{k=0}^{\infty} A^k \frac{1}{k!}$$

Ist A in Jordanscher Normalform gegeben, so gilt:

$$e^A = B \text{diag}(e^{J_i}) B^{-1}$$

$$e^{J_i} = e^{\lambda_i I + N} = e^{\lambda_i I} e^N = e^{\lambda_i} \sum_{k=0}^{k_i} N^k \frac{1}{k!}$$

Sofern A und B kommutieren gilt $e^{AB} = e^A e^B$. Kommutieren A und B nicht, dann gibt es eine Matrix C sodass $e^A e^B = e^C$. Es gilt

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] + \dots$$

wobei $[A, B] = AB - BA$ der Kommutator ist. Die Reihe für C nennt man die **Campbell-Baker-Hausdorff-Reihe**.

Betrachten wir die Pendelgleichung:

$$m\ddot{x} = F$$

$$ml\ddot{\varphi} = -\omega^2 \sin(\varphi)$$

Wir führen nun eine Zusatzvariable $\psi = \frac{\dot{\varphi}}{\omega}$ ein und erhalten somit eine Gleichung erster Ordnung, dafür ein System zweiter Ordnung. Weiters linearisieren wir das Problem, da für kleine Winkel φ gilt $\sin(\varphi) \approx \varphi$:

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{\omega}{ml} \\ \omega & 0 \end{bmatrix} \begin{bmatrix} \varphi \\ \psi \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \varphi(t) \\ \psi(t) \end{bmatrix} = e^{At} \begin{bmatrix} \varphi_0 \\ \psi_0 \end{bmatrix}$$

Da die Eigenwerte dieser Matrix komplex sind, sehen wir, dass das Pendel schwingt, und zwar mit der Frequenz ω .

Satz 11.6: Minimalpolynom

Sei $A \in \mathbb{K}^{n \times n}$:

- a $\exists p \in \mathbb{K}[x]: p(A) = 0$, sprich jede Matrix ist algebraisch
- b $\exists! m_A \in \mathbb{K}[x]$ mit minimalen Grad und fhrenden Koeffizienten 1, sodass $m_A(A) = 0$, dieses Polynom nennen wir **Minimalpolynom** von A
- c $\text{Ann}(A) = \{p \in \mathbb{K}[x]: p(A) = 0\}$ heit **Annihilator** von A und es gilt:

$$\forall p \in \text{Ann}(A): m_A | p$$

- d $\forall \lambda \in \text{spec}(A): m_A(\lambda) = 0$

Beweis fr Satz 11.6. Zu a:

$I, A, A^2, \dots, A^{n^2}$ sind linear abhngig, sprich:

$$\exists c_0, \dots, c_{n^2}: \sum_{k=0}^{n^2} c_k A^k = 0$$

$$p(x) = \sum_{k=0}^{n^2} c_k x^k \Rightarrow p \in \text{Ann}(A)$$

Zu b und c:

Sei $d_0 = \min\{\deg p | p \in \text{Ann}(A)\}$. Sei p ein Polynom mit minimalen Grad d_0 und oBdA $p(x) = x^{d_0}$. Sei $q \in \text{Ann}(A)$, dann gilt $\deg p \leq \deg q$:

$$\begin{aligned} q &= ap + r \\ \Rightarrow 0 &= q(A) = a(A) \underbrace{p(A)}_{=0} + r(A) \end{aligned}$$

Somit folgt $r(A) = 0$, womit jedes $q \in \text{Ann}(A)$ von p geteilt wird, und p eindeutig ist. \square

Satz 11.7: Cayley-Hamilton

Sei $A \in \mathbb{K}^{n \times n}$:

$$\chi_A \in \text{Ann}(A)$$

Korollar 11.5: zu Satz 11.7

Sei $A \in \mathbb{K}^{n \times n}$:

$$m_A | \chi_A$$

Daher haben m_A und χ_A die gleichen Nullstellen, sind aber nicht notwendigerweise ident, da die Eigenwerte verschiedene Vielfachheiten haben knnen.

Beweis 1 fr Satz 11.7. Angenommen \mathbb{K} ist algebraisch abgeschlossen. Wir bilden die Jordansche Normalform, dann gilt:

$$\chi_A(A) = B^{-1} \text{diag}(\chi_A(J_i)) B$$

Es gengt zu zeigen, dass $\chi_A(J_i) = 0 \forall i$. Sei $J_s = J_{k_s}(\lambda_s)$ und $k_s \leq k(\lambda_s)$, womit:

$$(J_s - \lambda_s I)^{k(\lambda_s)} = 0$$

da $\chi_A(\lambda) = (\lambda - \lambda_s)^{k(\lambda_s)} p(\lambda)$ ist auch $\chi_A(J_s) = 0$. \square

Beweis 2 für Satz 11.7. Ein nicht ganz korrekter Beweis:

$$\chi_A(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) \Rightarrow \chi_A(\mathbf{A}) = \det(\mathbf{A} \mathbf{I} - \mathbf{A}) = 0$$

Dieser Beweis stimmt nicht, da es sich bei $\mathbf{A} \mathbf{I}$ nicht um das klassische Matrix-Produkt handelt. \square

Beweis 3 für Satz 11.7. Wir erinnern uns an die Komplementärmatrix $\hat{\mathbf{A}}$, wobei $\mathbf{A} \hat{\mathbf{A}} = \det(\mathbf{A}) \mathbf{I}$, und $\hat{\mathbf{A}}_{ij} = (-1)^{i+j} \det(\mathbf{A}_{ji})$. Wir betrachten nun:

$$\widehat{x\mathbf{I} - \mathbf{A}} \in (\mathbb{K}[x])^{n \times n} \quad \widehat{x\mathbf{I} - \mathbf{a}} = [b_{ij}(x)] \quad \deg(b_{ij}) \leq n-1$$

Alternativ gilt:

$$\widehat{x\mathbf{I} - \mathbf{A}} = \sum_{k=0}^{n-1} \mathbf{B}_k x^k \in \mathbb{K}^{n \times n}[x]$$

$$\begin{aligned} (x\mathbf{I} - \mathbf{A})(\widehat{x\mathbf{I} - \mathbf{A}}) &= \det(x\mathbf{I} - \mathbf{A}) \mathbf{I} = \chi_A(x) \mathbf{I} \\ &= (x\mathbf{I} - \mathbf{A}) \left(\sum_{k=0}^{n-1} \mathbf{B}_k x^k \right) = \mathbf{I} \sum_{k=0}^n c_k x^k \end{aligned}$$

Wir führen einen Koeffizientenvergleich durch:

$$-\mathbf{A} \mathbf{B}_0 + \sum_{k=1}^{n-1} (\mathbf{B}_k - \mathbf{A} \mathbf{B}_{k-1}) x^k + \mathbf{B}_{n-1} x^n = \sum_{k=0}^1 c_k x^k \mathbf{I}$$

Es folgt $c_0 \mathbf{I} = -\mathbf{A} \mathbf{B}_0$, $\mathbf{I} = \mathbf{B}_{n-1}$ und $c_k \mathbf{I} = \mathbf{B}_{k-1} - \mathbf{A} \mathbf{B}_k$ für $0 < k < n$. Multiplizieren wir mit \mathbf{A}^k von links:

$$\begin{aligned} c_0 \mathbf{I} &= -\mathbf{A} \mathbf{B}_0 \\ c_1 \mathbf{A} &= \mathbf{A} \mathbf{B}_0 - \mathbf{A}^2 \mathbf{B}_1 \\ c_2 \mathbf{A}^2 &= \mathbf{A}^2 \mathbf{B}_1 - \mathbf{A}^3 \mathbf{B}_2 \\ &\vdots \\ c_{n-1} \mathbf{A}^{n-1} &= \mathbf{A}^{n-1} \mathbf{B}_{n-2} - \mathbf{A}^n \mathbf{B}_{n-1} \\ \mathbf{A}^n &= \mathbf{A}^n \mathbf{B}_{n-1} \end{aligned}$$

Wir sehen, dass die Summe der rechten Seiten $\mathbf{0}$ ergibt und die der linken Seite $\chi_A(\mathbf{A})$, womit $\chi_A(\mathbf{A}) = 0$. Der Vorteil dieses Beweises ist es, dass er nicht voraussetzt, dass \mathbb{K} algebraisch abgeschlossen ist. \square

Das Minimalpolynom m_A wird folgendermaßen gebildet:

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

wobei m_i die Größe des größten Jordanblocks der zu λ_i zugehörigen Jordanblöcke ist. Ist \mathbf{A} diagonalisierbar, dann gilt:

$$m_A(x) = \prod_{i=1}^r (x - \lambda_i)$$

Betrachten wir noch eine Anwendung. Wir haben bereits die Exponentialfunktion $\exp(\mathbf{A})$ folgendermaßen definiert:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \mathbf{A}^k \frac{1}{k!}$$

Wir können eine Vereinfachung vornehmen:

$$e^{\alpha \mathbf{I} + \mathbf{A}} = e^{\alpha} e^{\mathbf{A}}$$

Sei oBdA $\text{Tr}(\mathbf{A}) = 0$, dann schreiben wir $\overset{\circ}{\mathbf{A}} = \mathbf{A} - \frac{1}{n} \text{Tr}(\mathbf{A}) \mathbf{I}$. Betrachten wir dazu kurz ein Beispiel:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \Rightarrow \chi_{\mathbf{A}}(\lambda) = (\lambda - a)(\lambda + a) - bc = \lambda^2 - \text{det}(\mathbf{A}) = \lambda^2 - \delta$$

Nach Cayley-Hamilton gilt in diesem Fall $\mathbf{A}^2 = \delta \mathbf{I}$. Allgemein gilt

$$\mathbf{A}^n = \sum_{i=0}^{n-1} c_i \mathbf{A}^i$$

Hernach können wir $e^{\mathbf{A}}$ (für das Beispiel) umschreiben:

$$\begin{aligned} e^{\mathbf{A}} &= \sum_{n=0}^{\infty} \mathbf{A}^n \frac{1}{n!} = \sum_{n=0}^{\infty} \mathbf{A}^{2n} \frac{1}{(2n)!} + \sum_{n=0}^{\infty} \mathbf{A}^{2n+1} \frac{1}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \mathbf{I} \frac{\delta^n}{(2n)!} + \sum_{n=0}^{\infty} \frac{\delta^n}{(2n+1)!} \mathbf{A} = \mathbf{I} \cosh(\sqrt{\delta}) + \frac{\sinh(\sqrt{\delta})}{\sqrt{\delta}} \mathbf{A} \end{aligned}$$

Im Allgemeinen Fall ist $e^{\mathbf{A}}$ ein Polynom vom Grad $n-1$ in \mathbf{A} .

Satz 11.8: Spektralabbildungssatz

Wenn \mathbb{K} algebraisch abgeschlossen ist und $\mathbf{A} \in \mathbb{K}^{n \times n}$, dann $\text{spec}(p(\mathbf{A})) = \{p(\lambda) | \lambda \in \text{spec}(\mathbf{A})\} = p(\text{spec}(\mathbf{A}))$ für $p \in \mathbb{K}[x]$.

Beweis für Satz 11.8. Sei $\mu \in \text{spec}(p(\mathbf{A}))$. Zu zeigen ist $\exists \lambda \in \text{spec}(\mathbf{A}): p(\lambda) = \mu$. Sei \mathbf{x} Eigenvektor von $p(\mathbf{A})$ zum Eigenwert μ , es gilt also $p(\mathbf{A})\mathbf{x} = \mu\mathbf{x}$. Wir legen $q(x) = p(x) - \mu$, womit $q(\mathbf{A})\mathbf{x} = \mathbf{0}$. Da \mathbb{K} algebraisch abgeschlossen ist, und $q \in \mathbb{K}[x]$, können wir q in Linearfaktoren zerlegen:

$$q(x) = \prod_{i=1}^k (x - \mu_i)$$

und daher:

$$\left(\prod_{i=1}^k (\mathbf{A} - \mu_i \mathbf{I}) \right) \mathbf{x} = \mathbf{0}$$

Somit $\mathbf{x} \in \ker(q(\mathbf{A}))$, hernach ist $q(\mathbf{A})$ singulär, sprich $\exists j: \mathbf{A} - j\mu_j \mathbf{I}$ singular, womit $\mu_j \in \text{spec}(\mathbf{A})$. Weiter gilt $q(\mu_j) = 0$ und damit $p(\mu_j) - \mu = 0$, sprich $\mu = p(\mu_j)$ mit $\mu_j \in \text{spec}(\mathbf{A})$. \square

Ist \mathbb{K} nicht algebraisch abgeschlossen, dann ist Satz 11.8 nicht erfüllt, z.B:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \chi_{\mathbf{A}}(\lambda) = \lambda^2 + 1$$

Über \mathbb{R} ist $\text{spec}(\mathbf{A}) = \emptyset$, aber $\mathbf{A}^2 = -\mathbf{I}$ hat Eigenwert -1 .

Kapitel XII Normale Matrizen

Wir kommen zurück zum Inneren Produkt, also $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V hat immer ein inneres Produkt $\langle \cdot, \cdot \rangle$. Sei $f \in \text{Hom}(V, W)$, dann gibt es $f^*: W \rightarrow V$, nach dem Darstellungssatz von Riesz gilt dann

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

Für die Matrixdarstellung bezüglich einer Orthonormalbasis gilt dann $(A^*)_{ij} = \overline{a_{ji}}$. Weiters nennen wir f selbstadjungiert, falls $V = W$ und $f^* = f$, bzw. bei Matrizen $A = A^*$.

Definition 12.1: Normalität von Matrizen

Eine Matrix A heißt **normal** wenn:

$$AA^* = A^*A \Leftrightarrow [A, A^*] = 0$$

Insbesondere für $A = A^*$.

Wir sehen direkt, dass selbstadjungierte und unitäre Matrizen¹⁷ normal sind. Sind A, B normal, dann ist $A + B$ normal, und AB , falls $AB = BA$.

Lemma 12.1: Kern normaler Matrizen

Sei $A \in \mathbb{C}^{n \times n}$ normal:

- i $\ker(A) = \ker(A^*)$
- ii $\ker(A) = \ker(A^2) = \ker(A^*A)$

Hier sehen wir bei ii, dass das Lemma von Fitting greift, woran zu erkennen ist, dass normale Matrizen immer diagonalisierbar sein werden.

Beweis für Lemma 12.1. Zu i:

$$\begin{aligned} x \in \ker(A) &\Leftrightarrow Ax = 0 \Leftrightarrow \|Ax\|^2 = 0 \\ &\Leftrightarrow \langle Ax, Ax \rangle = 0 \Leftrightarrow \langle x, A^*Ax \rangle = 0 \Leftrightarrow \langle x, AA^*x \rangle = 0 \\ &\Leftrightarrow \langle A^*x, A^*x \rangle = 0 \Leftrightarrow \|A^*x\|^2 = 0 \Leftrightarrow A^*x = 0 \Leftrightarrow x \in \ker(A^*) \end{aligned}$$

Zu ii:

Sei $x \in \ker(A^2)$, sprich $A^2x = 0 \Rightarrow Ax \in \ker(A) = \ker(A^*)$:

$$\Rightarrow A^*Ax = 0 \Rightarrow \langle A^*Ax, x \rangle = 0 \Rightarrow \langle Ax, Ax \rangle = 0 \Rightarrow \|Ax\|^2 = 0 \Rightarrow Ax = 0 \Rightarrow x \in \ker(A)$$

□

Korollar 12.1

Sie $A \in \mathbb{C}^{n \times n}$ normal:

- i $\ker(\lambda I - A) = \ker(\overline{\lambda}I - A^*)$
- ii A ist diagonalisierbar

Lemma 12.2

Sei $A \in \mathbb{C}^{n \times n}$ normal, und $\lambda, \mu \in \text{spec}(A)$ wobei $\lambda \neq \mu$:

$$\ker(\lambda I - A) \perp \ker(\mu I - A)$$

¹⁷ $U^{-1} = U^*$, womit $UU^* = UU^{-1} = I = U^{-1}U = U^*U$

Beweis für Lemma 12.2. Sei $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ und $\mathbf{A}\mathbf{y} = \mu\mathbf{y} \Rightarrow \mathbf{A}^*\mathbf{y} = \bar{\mu}\mathbf{y}$:

$$\begin{aligned}\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{x}, \mathbf{A}^*\mathbf{y} \rangle \Leftrightarrow \langle \lambda\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \bar{\mu}\mathbf{y} \rangle \Leftrightarrow \lambda \langle \mathbf{x}, \mathbf{y} \rangle = \bar{\mu} \langle \mathbf{x}, \mathbf{y} \rangle \\ &\Leftrightarrow (\lambda - \bar{\mu}) \langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad \lambda \neq \bar{\mu} \Rightarrow \langle \mathbf{x}, \mathbf{y} \rangle = 0\end{aligned}$$

□

Satz 12.1

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, dann sind die folgenden Aussagen äquivalent:

- i \mathbf{A} ist normal
- ii es gibt eine Orthonormalbasis aus Eigenvektoren
- iii \mathbf{A} ist unitär diagonalisierbar¹⁸

Beweis für Satz 12.1. i \Rightarrow ii haben wir bereits in Lemma 12.2 gesehen, sprich

$$\mathbb{C}^n = \bigoplus_{i=1}^k \mathcal{N}_{\lambda_i}$$

und $\forall i, j: i \neq j: \mathcal{N}_{\lambda_i} \perp \mathcal{N}_{\lambda_j}$. Bilden wir die Orthonormalbasen B_j der Eigenräume, dann ist

$$B = \bigcup_{i=1}^k B_i$$

eine Orthonormalbasis von \mathbb{C}^n .

ii \Rightarrow iii:

Sei $\mathbf{u}_1, \dots, \mathbf{u}_n$ ein Orthonormalbasis aus Eigenvektoren, dann:

$$\mathbf{A}\mathbf{U} = \mathbf{U}\text{diag}(\lambda_1, \dots, \lambda_n)$$

Wobei $\mathbf{U} = [\mathbf{u}_1 \ \dots \ \mathbf{u}_n]$ unitär ist, dann:

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

iii \Rightarrow i:

Angenommen $\mathbf{A} = \mathbf{U}\text{diag}(\lambda_1, \dots, \lambda_n)\mathbf{U}^*$:

$$\begin{aligned}\mathbf{A}^* &= (\mathbf{U}^*)^* \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) \mathbf{U}^* = \mathbf{U} \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) \mathbf{U}^* \\ \Rightarrow \mathbf{A}\mathbf{A}^* &= \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{U}^* \mathbf{U} \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) \mathbf{U}^* \\ &= \mathbf{U} \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{U}^* \\ &= \mathbf{U} \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n) \mathbf{U}^* \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{U}^* \\ &= \mathbf{A}^*\mathbf{A}\end{aligned}$$

□

Satz 12.2: Schursche Normalform

Benannt nach Issai *Schur*.

- i Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, dann $\exists \mathbf{U} \in \mathcal{U}(n): \mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{R}$ eine obere Dreiecksmatrix mit $\text{diag}(\mathbf{R}) = \text{spec}(\mathbf{A})$
- ii wenn $\mathbf{A} \in \mathbb{R}^{n \times n}$ und $\chi_{\mathbf{A}}(\lambda)$ über \mathbb{R} in Linearfaktoren zerfällt¹⁹, dann kann $\mathbf{U} \in \mathcal{O}(n)$ gewählt werden

¹⁸ $\exists \mathbf{U} \in \mathcal{U}(n): \mathbf{U}^*\mathbf{A}\mathbf{U} = \text{diag}(\lambda_1, \dots, \lambda_n)$

¹⁹also $\text{spec}(\mathbf{A}) \subset \mathbb{R}$

Beweis für Satz 12.2. Wir führen einen Induktionsbeweis. Sei $\lambda \in \text{spec}(\mathbf{A})$ und \mathbf{u} der zugehörige Eigenvektor mit $\|\mathbf{u}\| = 1$. Wir beginnen damit, \mathbf{u} zu einer Orthonormalbasis von \mathbb{C}^n zu erweitern, $\mathbf{u}, \mathbf{w}_1, \dots, \mathbf{w}_{n-1}$:

$$\begin{aligned} \mathbf{W} &= [\mathbf{w}_1 \ \cdots \ \mathbf{w}_{n-1}] \\ \mathbf{U}_1 &= [\mathbf{u} \ \mathbf{w}_1 \ \cdots \ \mathbf{w}_{n-1}] \\ \Rightarrow \mathbf{AU}_1 &= [\mathbf{Au} \ \mathbf{Aw}_1 \ \cdots \ \mathbf{Aw}_{n-1}] = [\lambda \mathbf{u} \ \mathbf{AW}] \\ &\Rightarrow \mathbf{U}_1^* \mathbf{AU}_1 = \begin{bmatrix} \mathbf{u}^* \\ \mathbf{W}^* \end{bmatrix} [\lambda \mathbf{u} \ \mathbf{AW}] \\ &= \begin{bmatrix} \lambda & \mathbf{u}^* \mathbf{AW} \\ \mathbf{0} & \mathbf{W}^* \mathbf{AW} \end{bmatrix} \end{aligned}$$

Es gilt: $\chi_{\mathbf{A}}(x) = \chi_{\mathbf{U}_1^* \mathbf{AU}_1}(x) = (\lambda - x) \chi_{\mathbf{W}^* \mathbf{AW}}(x)$. Für $\mathbb{K} = \mathbb{R}$ zerfällt $\chi_{\mathbf{W}^* \mathbf{AW}}(x)$ immer noch in Linearfaktoren. Nach der Induktionsvoraussetzung hat $\mathbf{W}^* \mathbf{AW} \in \mathbb{C}^{(n-1) \times (n-1)}$ eine Schursche Normalform, sprich

$$\begin{aligned} \exists \mathbf{U}_2 \in \mathcal{U}(n-1): \mathbf{U}_2^* (\mathbf{W}^* \mathbf{AW}) \mathbf{U}_2 &= \mathbf{R}_2 \\ \Rightarrow \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{bmatrix}^* \mathbf{U}_1^* \mathbf{AU}_1 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{bmatrix} &= \begin{bmatrix} \lambda & \mathbf{u}^* \mathbf{AW} \\ \mathbf{0} & \mathbf{R}_2 \end{bmatrix} \\ \Rightarrow \mathbf{U}_1 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} &= \mathbf{U} \end{aligned}$$

Wiederholen wir diese Prozedur, erhalten wir $\mathbf{U}^* \mathbf{AU} = \mathbf{R}$, und da $\chi_{\mathbf{A}}(\lambda) = \chi_{\mathbf{U}^* \mathbf{AU}}(\lambda) = \prod_{i=1}^n (\lambda - r_{ii})$, womit r_{ii} die Eigenwerte von \mathbf{A} sind. \square

Alternativer Beweis für Satz 12.1. Sei \mathbf{A} unitär diagonalisierbar impliziert, dass \mathbf{A} normal ist. Wir wollen nun prüfen, ob auch Normalität die unitäre Diagonalisierbarkeit impliziert. Wir bilden die Schursche Normalform von $\mathbf{A} = \mathbf{URU}^*$. Da \mathbf{A} normal ist, gilt $\mathbf{AA}^* = \mathbf{A}^* \mathbf{A}$, was äquivalent ist zu:

$$\mathbf{RR}^* = \mathbf{R}^* \mathbf{R}$$

\square

Satz 12.3: Eigenwerte selbstadjungierter Matrizen

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ mit $\mathbf{A} = \mathbf{A}^*$, dann $\text{spec}(\mathbf{A}) \subset \mathbb{R}$.

Beweis für Satz 12.3. $\mathbf{A} = \mathbf{A}^*$, womit \mathbf{A} normal ist, also auch $\lambda \mathbf{I} - \mathbf{A}$. Woraus folgt $\ker(\lambda \mathbf{I} - \mathbf{A}) = \ker((\lambda \mathbf{I} - \mathbf{A})^*) = \ker(\overline{\lambda} \mathbf{I} - \mathbf{A})$:

$$\forall \mathbf{x} \in \ker(\lambda \mathbf{I} - \mathbf{A}): \mathbf{Ax} = \lambda \mathbf{x} \wedge \mathbf{Ax} = \overline{\lambda} \mathbf{x} \Rightarrow \lambda = \overline{\lambda}$$

\square

Korollar 12.2

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch, womit $\text{spec}(\mathbf{A}) \subset \mathbb{R}$ und $\forall i: \mathcal{N}_i \subset \mathbb{R}^n$. Sprich:

$$\exists \mathbf{Q} \in \mathcal{O}(n): \mathbf{Q}^T \mathbf{AQ} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Satz 12.4: Definitheit selbstadjungierter Matrizen

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, wobei $\mathbf{A} = \mathbf{A}^*$:

- i $\mathbf{A} > 0 \Leftrightarrow \text{spec}(\mathbf{A}) \subset \mathbb{R}^+$
- ii $\mathbf{A} \geq 0 \Leftrightarrow \text{spec}(\mathbf{A}) \subset \mathbb{R}^+ \cup \{0\}$
- iii $\mathbf{A} < 0 \Leftrightarrow \text{spec}(\mathbf{A}) \subset \mathbb{R}^-$
- iv \mathbf{A} ist indefinit $\exists \lambda_1, \lambda_2 \in \text{spec}(\mathbf{A}): \lambda_1 > 0 \wedge \lambda_2 < 0$

Beweis für Satz 12.4. Wir betrachten nur i, da die übrigen Punkte analog zu beweisen sind. Sei $A > 0$ und $\lambda \in \text{spec}(A)$, sprich $\forall \mathbf{x} \neq \{\mathbf{0}\}: \langle A\mathbf{x}, \mathbf{x} \rangle > 0$. Sei \mathbf{x} ein Eigenvektor zum Eigenwert λ :

$$\langle A\mathbf{x}, \mathbf{x} \rangle = \langle \lambda\mathbf{x}, \mathbf{x} \rangle = \lambda \cdot \|\mathbf{x}\|^2 > 0 \Leftrightarrow \lambda > 0$$

Sei $\text{spec}(A) \subset \mathbb{R}^+$. Sei $\mathbf{u}_1, \dots, \mathbf{u}_n$ eine Orthonormalbasis aus Eigenvektoren. Sei $\mathbf{x} \in \mathbb{C}^n$ mit $\mathbf{x} \neq \mathbf{0}$, dann:

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{u}_i$$

Wir betrachten nun:

$$\begin{aligned} \langle A\mathbf{x}, \mathbf{x} \rangle &= \left\langle A \sum_{i=1}^n \alpha_i \mathbf{u}_i, \sum_{j=1}^n \alpha_j \mathbf{u}_j \right\rangle = \sum_{i,j=1}^n \alpha_i \overline{\alpha_j} \langle A\mathbf{u}_i, \mathbf{u}_j \rangle \\ &= \sum_{i,j=1}^n \alpha_i \overline{\alpha_j} \lambda_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i,j=1}^n \alpha_i \overline{\alpha_j} \lambda_i \delta_{ij} = \sum_{i=1}^n |\alpha_i|^2 \lambda_i > 0 \end{aligned}$$

□

12.1 Anwendungen

Betrachten wir eine Anwendung für die Taylor-Formel in mehreren Dimensionen. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \mathbf{H} (\mathbf{x} - \mathbf{x}_0) + \mathcal{O}(\|\mathbf{x} - \mathbf{x}_0\|^3)$$

Dabei ist \mathbf{H} die **Hesse-Matrix**:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Mit \mathbf{H} können wir Extrema von f klassifizieren. Wenn $\mathbf{H} > 0$, dann ist das Extremum \mathbf{x}_0 ein Minimum. Analog bedeutet $\mathbf{H} < 0$, dass \mathbf{x}_0 ein Maximum ist. Andernfalls handelt es sich bei \mathbf{x}_0 um einen sogenannten *Sattelpunkt*.

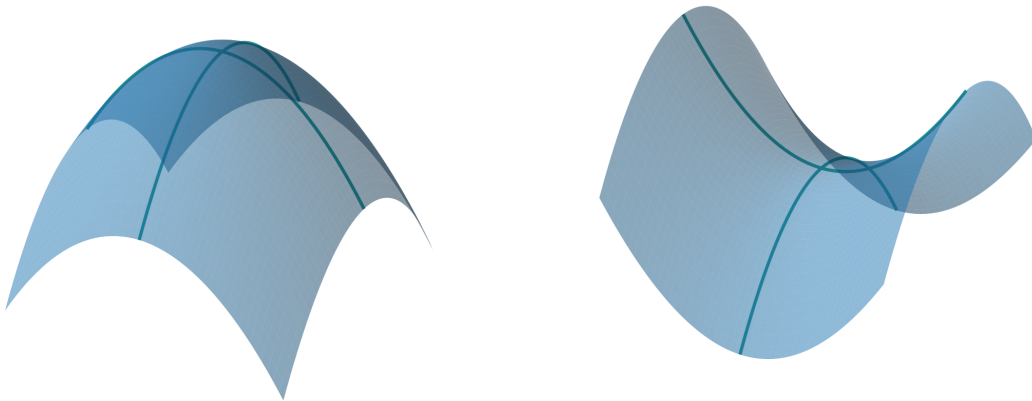


Abbildung 12: Ein Maximum (links) und ein Sattelpunkt (rechts)

Auch Kegelschnitte kann man untersuchen.

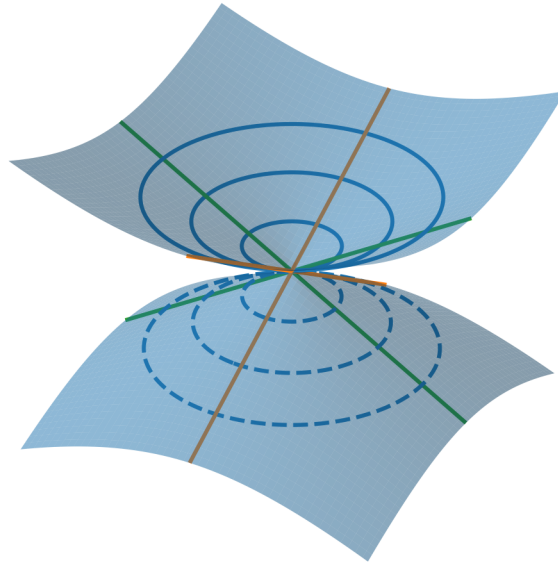


Abbildung 13: Ein einfacher Kegel mit einigen Konturlinien

Ein (Doppel)-Kegel entsteht durch Rotation einer Geraden $g = \{\lambda \mathbf{x} | \lambda \in \mathbb{R}\}$, um die z-Achse, er wird also durch die folgende Gleichung beschrieben:

$$x_1^2 + x_2^2 = x_3^2$$

Ein Kegelschnitt entsteht nun durch das Schneiden des Kegels mit einer Ebene E :

$$E = \{\mathbf{u} + \xi \mathbf{v} + \eta \mathbf{w} | \xi, \eta \in \mathbb{R}\}$$

Dabei sind \mathbf{v} und \mathbf{w} linear unabhängig. TODO: Graphik (/) Wir wollen anhand E herausfinden, welcher Fall von Kegelschnitt eintritt. Sei nun $\mathbf{x} \in E$, dann setzen wir \mathbf{x} in die Gleichung des Kegels ein:

$$(u_3 + \xi v_3 + \eta w_3)^2 = (u_1 + \xi v_1 + \eta w_1)^2 + (u_2 + \xi v_2 + \eta w_2)^2$$

Es handelt sich um eine quadratische Gleichung in ξ und η :

$$\begin{aligned} a\xi^2 + 2b\xi\eta + c\eta^2 + d\xi + e\eta &= f \\ \Leftrightarrow \begin{bmatrix} \xi & \eta \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} &= f \end{aligned}$$

Wir verlegen den Ursprung so, dass der lineare Term $f = 0$ wird:

$$\xi = x + x_0 \quad \eta = y + y_0$$

$$a(x + x_0)^2 + 2b(x + x_0)(y + y_0) + c(y + y_0)^2 + d(x + x_0) + e(y + y_0) - f = 0$$

$$\rightsquigarrow ax^2 + 2bxy + cy^2 + (2ax_0 + 2by_0 + d)x + (2by_0 + 2cy_0 + e)y + ax_0^2 + 2bx_0y_0 + cy_0^2 + dx_0 + ey_0 - f = 0$$

Wir wählen x_0 und y_0 , sodass die linearen Terme in x und y verschwinden:

$$2ax_0 + 2by_0 + d = 0$$

$$2by_0 + 2cy_0 + e = 0$$

In Matrix-Schreibweise:

$$\underbrace{\begin{bmatrix} a & b \\ b & c \end{bmatrix}}_{=A} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} d \\ e \end{bmatrix}$$

Wir machen nun eine Fallunterscheidung.

Fall 1 $\det(\mathbf{A}) = 0$, dann ist ein Eigenwert 0. Wir diagonalisieren mit $\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}(\lambda_1, 0)$:

$$\begin{bmatrix} \xi \\ y\eta \end{bmatrix} = \mathbf{Q} \begin{bmatrix} x \\ y \end{bmatrix}$$

Da \mathbf{Q} orthogonal ist, beschreibt \mathbf{Q} eine Drehung oder Spiegelung:

$$\begin{aligned} \begin{bmatrix} x & y \end{bmatrix} \mathbf{Q}^T \mathbf{A} \mathbf{Q} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \mathbf{Q} \begin{bmatrix} x \\ y \end{bmatrix} &= f \\ \rightsquigarrow \lambda_1 x^2 + \tilde{d}x + \tilde{e}y &= f \\ \begin{bmatrix} \tilde{d} \\ \tilde{e} \end{bmatrix} &= \mathbf{Q} \begin{bmatrix} d \\ e \end{bmatrix} \end{aligned}$$

a) $\tilde{e} \neq 0$, dann haben wir einen expliziten Ausdruck für y :

$$d = \frac{f - \lambda_1 x^2 - \tilde{d}x}{\tilde{e}}$$

b) $\tilde{e} = 0$ ergibt eine quadratische Gleichung in x :

$$\begin{aligned} \lambda_1 x^2 + \tilde{d}x &= f \\ \Rightarrow x &= \frac{-\tilde{d} \pm \sqrt{\tilde{d}^2 + 4\lambda_1 f}}{2} \end{aligned}$$

Wenn $\tilde{d}^2 + 4\lambda_1 f < 0$ gibt es keine Lösungen, was aber nicht eintritt. Ansonsten erhalten wir eine oder zwei²⁰ Geraden, die parallel zur y Achse sind

Fall 2 $\det(\mathbf{A}) \neq 0$, wir können also x_0 und y_0 direkt bestimmen. Wir erhalten damit die Gleichung:

$$\begin{aligned} ax^2 + 2bx + y^2 &= g = f - ax_0^2 - 2bx_0y_0 - cy_0^2 - 2dx_0 - 2ey_0 \\ \Leftrightarrow \begin{bmatrix} x & y \end{bmatrix} \mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix} &= g \end{aligned}$$

Sei $\mathbf{Q} \in \mathcal{O}(2)$, sodass $\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2)$:

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix} &= \mathbf{Q} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} \rightsquigarrow \begin{bmatrix} \tilde{x} & \tilde{y} \end{bmatrix} \mathbf{Q}^T \mathbf{A} \mathbf{Q} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = g \\ \Leftrightarrow \lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 &= g \end{aligned}$$

a) $g = 0$ liefert $\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 = 0$, womit $\tilde{y} = -\frac{\lambda_1}{\lambda_2} \tilde{x}^2$. Wenn $\lambda_1 \lambda_2 > 0$, dann $\tilde{y} = \tilde{x} = 0$. Wenn $\lambda_1 \lambda_2 < 0$, dann:

$$|\tilde{y}| = \sqrt{-\frac{\lambda_1}{\lambda_2}} |\tilde{x}|$$

Das sind wiederum zwei Geraden.

b) $g \neq 0$ liefert:

$$\frac{\lambda_1}{g} \tilde{x}^2 + \frac{\lambda_2}{g} \tilde{y}^2 = 1$$

Wenn $\frac{\lambda_1}{g} < 0$ und $\frac{\lambda_2}{g} < 0$, dann gibt es keine Lösung. Sind beide Faktoren positiv, dann $\alpha^2 = \frac{g}{\lambda_1}$ und $\beta^2 = \frac{g}{\lambda_2}$ ergibt eine Ellipse:

$$\frac{\tilde{x}^2}{\alpha^2} + \frac{\tilde{y}^2}{\beta^2} = 1$$

Wenn $\frac{\lambda_1}{g} > 0$ und $\frac{\lambda_2}{g} < 0$, dann $\alpha^2 = \frac{g}{\lambda_1}$ und $\beta^2 = -\frac{g}{\lambda_2}$, ergibt eine Hyperbel:

$$\frac{\tilde{x}^2}{\alpha^2} - \frac{\tilde{y}^2}{\beta^2} = 1$$

²⁰eine Gerade, falls $\tilde{d}^2 + 4\lambda_1 f = 0$

Definition 12.2: Quadriken

Eine **Quadrik** im \mathbb{R}^n ist die Lösungsmenge einer quadratischen Gleichung:

$$\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = 0$$

wobei $\mathbf{A} = \mathbf{A}^T$.

Satz 12.5: Hauptachsentransformation

Es gibt eine Transformation der Gestalt

$$\mathbf{x} = \mathbf{Q} \boldsymbol{\xi} + \mathbf{x}_0$$

mit $\mathbf{Q} \in \mathcal{O}(n)$ sodass die Gleichung die folgende Gestalt bekommt:

$$\sum_{i=1}^k \lambda_i \xi_i^2 + \sum_{i=k+1}^n \tilde{b}_i \xi_i + \tilde{c} = 0$$

Satz 12.6: Wurzeln von Matrizen

Sei $\mathbf{A} > 0$, dann $\exists! \mathbf{B} \geq 0$ sodass $\mathbf{B}^2 = \mathbf{A}$, wir schreiben dann $\mathbf{B} = \mathbf{A}^{\frac{1}{2}}$.

Beweis für Satz 12.6. Wir beginnen mit der Existenz von \mathbf{B} . Sei $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n) > 0$, dann folgt direkt $\mathbf{B} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Falls $\mathbf{A} > 0$ diagonalisieren wir \mathbf{A} und erhalten:

$$\mathbf{B} = \mathbf{Q} \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) \mathbf{Q}^T$$

Sie $\mathbf{B} > 0$ und $\mathbf{B}^2 = \mathbf{A}$. Wir können \mathbf{B} diagonalisieren. Sei dazu $\mathbf{u}_1, \dots, \mathbf{u}_n$ eine Orthonormalbasis von Eigenvektoren für \mathbf{A} , dann gilt:

$$\mathbf{B}^2 \mathbf{u}_i = \mathbf{A} \mathbf{u}_i = \lambda_i \mathbf{u}_i \Rightarrow \text{spec}(\mathbf{B}^2) = \{\mu_i^2 \mid \mu_i \in \text{spec}(\mathbf{B})\}$$

Zu jedem λ_i gibt es $\mu_i \in \text{spec}(\mathbf{B})$: $\lambda_i = \mu_i^2$. Da $\mathbf{B} \geq 0$ muss $\mu_i > 0$. Da $\mathbf{A}\mathbf{B} = \mathbf{B}^2\mathbf{B} = \mathbf{B}\mathbf{B}^2 = \mathbf{B}\mathbf{A}$ sind \mathbf{u}_i auch Eigenvektoren von \mathbf{B} , somit gilt $\mathbf{B}\mathbf{u}_i = \mu_i \mathbf{u}_i$, wonach \mathbf{B} eindeutig sein muss. \square

Definition 12.3: Singulärwerte

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, dann gilt:

$$\mathbf{A}^* \mathbf{A} = \mathbf{V} \text{diag}(s_1^2, \dots, s_n^2) \mathbf{V}^*$$

Da $\mathbf{A}^* \mathbf{A} > 0$ können wir $(\mathbf{A}^* \mathbf{A})^{\frac{1}{2}}$ bilden:

$$(\mathbf{A}^* \mathbf{A})^{\frac{1}{2}} = \mathbf{V} \text{diag}(s_1, \dots, s_n) \mathbf{V}^*$$

Wir nennen s_i die **Singulärwerte** von \mathbf{A} .

Satz 12.7: Polarzerlegung

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, dann legen wir fest:

$$|\mathbf{A}| = (\mathbf{A}^* \mathbf{A})^{\frac{1}{2}}$$

Dann gibt es eine unitäre Matrix $\mathbf{U} \in \mathcal{U}(n)$ sodass:

$$\mathbf{A} = \mathbf{U} \cdot |\mathbf{A}|$$

Beweis für Satz 12.7. Falls A regulär ist, dann gilt:

$$U = A((A^*A)^{\frac{1}{2}})^{-1} = A(A^*A)^{-\frac{1}{2}}$$

Wir prüfen noch ob U unitär ist:

$$\begin{aligned} U^*U &= ((A^*A)^{-\frac{1}{2}}A^*)(A(A^*A)^{-\frac{1}{2}}) \\ &= (A^*A)^{-\frac{1}{2}}A^*A(A^*A)^{-\frac{1}{2}} \\ &= (A^*A)^{-\frac{1}{2}}(A^*A)^{\frac{1}{2}}(A^*A)^{\frac{1}{2}}(A^*A)^{-\frac{1}{2}} \\ &= I \end{aligned}$$

Falls A singulär ist, ist $A^*A \geq 0$. Für eine hermitesche Matrix H gilt $\text{im } H \perp \ker H$ und $H|_{\text{im } H}$ ist regulär. Damit ist $\mathbb{C}^n = \text{im}(A^*A) \oplus \ker(A^*A)$ und entsprechend gilt:

$$A^*A = \text{diag}(P, 0)$$

Wobei $P = A^*A|_{\text{im}(A^*A)}$. Sei $x \in \mathbb{C}^n$, dann $\exists! u \in \text{im}(A^*A), v \in \ker(A^*A): x = u + v$, wobei $u \perp v$, und weiter $A^*Ax = Pu$. Wir definieren uns $V = AP^{-\frac{1}{2}} \in \text{End}(\text{im}(A^*A))$. Sei

$$U = V \oplus I|_{\ker(A^*A)}$$

Es gilt zu zeigen, dass $A = U(A^*A)^{\frac{1}{2}}$. Dazu sei $x \in \mathbb{C}^n$ mit $x = u + v$:

$$\begin{aligned} U(A^*A)^{\frac{1}{2}}x &= U(A^*A)^{\frac{1}{2}}u = UP^{\frac{1}{2}}u \\ &= AP^{-\frac{1}{2}}P^{\frac{1}{2}}u = Au = Ax \end{aligned}$$

□

Korollar 12.3: Singulärwertzerlegung

Seien s_i die Singulärwerte von A und U bildet eine Orthonormalbasis aus Eigenvektoren, dann folgt mit der Polarzerlegung:

$$A = U(A^*A)^{\frac{1}{2}} = \sum_{k=1}^n s_i \langle \cdot, u_i \rangle Uu_i$$

Wir nennen $Uu_i = v_i$, und da U unitär ist, bilden die v_i erneut eine Orthonormalbasis. Sprich es gibt Orthonormalbasen u_1, \dots, u_n und v_1, \dots, v_n sodass:

$$Ax = \sum_{i=1}^n s_i \langle x, u_i \rangle v_i$$

Eine Anwendung für die Singulärwertzerlegung liegt in der Computertomographie. Die Grundlegende Idee ist es, mittels Röntgenstrahlen eine Dichtefunktion $\varrho(r, \theta)$ durch Messungen zu beschreiben. Integriert man die Funktion entlang einer Geraden $L(s, \varphi)$

$$\int_{L(s, \varphi)} \varrho(r, \theta) dx$$

so erhält man die sogenannte **Radon**-Transformation $\varrho(r, \theta) \mapsto R\varrho(s, \varphi)$, benannt nach Johann **Radon**. Die Messung der Computertomographie ergibt genau $R\varrho(s, \varphi)$. Die Frage ist nun, um man die inverse Transformation finden und anwenden kann. Per se gibt es keine inverse Transformation R^{-1} , da die Singulärwerte $(s_i)_{i=1}^{\infty} \rightarrow 0$. Wir bilden daher die Orthonormalbasen u_i und v_i für die Singulärwertzerlegung und bildet eine endliche Approximation:

$$Rf = \sum_{i=1}^N s_i \langle f, u_i \rangle v_i$$

Die u_i und v_i wurden von **Godfrey Hounsfield** und **Allan Cormack** gefunden, was ihnen den Nobelpreis für Medizin eingebracht hat.

Kapitel XIII Eigenwertabschätzungen

Definition 13.1: Numerischer Wertebereich

Sei $A \in \mathbb{C}^{n \times n}$, dann heißt

$$W = \{\langle Ax, x \rangle \mid x \in \mathbb{C}^n: \|x\| = 1\}$$

der **numerische Wertebereich** der Matrix A . Weiters nennt man

$$w(A) = \sup\{|z|, z \in W(A)\}$$

den **numerischen Radius** von A .

Lemma 13.1: Eigenwerte im numerischen Wertebereich

Sei $A \in \mathbb{C}^{n \times n}$, dann

$$\text{spec}(A) \subset W(A)$$

Beweis für Lemma 13.1. Sei u_i ein Eigenvektor zum Eigenwert λ_i und $\|u_i\| = 1$, dann $\langle Au_i, u_i \rangle \in W(A)$, bzw:

$$\langle \lambda_i u_i, u_i \rangle = \lambda_i$$

□

Satz 13.1: Satz von Töplitz-Hausdorff

$W(A)$ ist konvex. Ist A normal, dann gilt sogar:

$$W(A) = \text{co}(\text{spec}(A))$$

Dabei beschreibt $\text{co}(A)$ die konvexe Hülle, welche die kleinste konvexe Menge ist, die A enthält. Für Satz 13.1 also:

$$\text{co}(\text{spec}(A)) = \left\{ \sum_{i=1}^n \alpha_i \lambda_i \mid 0 \leq \alpha_i \wedge \sum_{i=1}^n \alpha_i = 1 \right\}$$

Sind etwa $x, y, z \in \mathbb{R}^2$ paarweise verschieden, dann ist

$$\text{co}(\{x, y, z\}) = \{\alpha x + \beta y + \gamma z \mid \alpha, \beta, \gamma \geq 0 \wedge \alpha + \beta + \gamma = 1\}$$

also genau das von x, y, z aufgespannte Dreieck.

Beweis für Satz 13.1. Für allgemeine Matrizen $A \in \mathbb{C}^{n \times n}$ wird der Beweis nachgereicht.

Sei $A \in \mathbb{C}^{n \times n}$ normal, dann ist A diagonalisierbar mit $A = V \text{diag}(\lambda_1, \dots, \lambda_n) V^{-1}$. Wir bilden eine Orthonormalbasis aus Eigenvektoren u_1, \dots, u_n . Sei $x \in \mathbb{C}^n$ mit $\|x\| = 1$. Es gilt $W(A) = \text{co}(\text{spec}(A))$, wenn $\forall x \in \mathbb{C}^n: \|x\| = 1$ gilt, dass $\exists \alpha_i \geq 0: \sum_{i=1}^n \alpha_i \lambda_i = \langle Ax, x \rangle \in \text{co}(\text{spec}(A))$ wobei $\sum_{i=1}^n \alpha_i = 1$.

$$\begin{aligned} \langle Ax, x \rangle &= \left\langle A \sum_{i=1}^n \xi_i u_i, \sum_{j=1}^n \xi_j u_j \right\rangle = \sum_{i=1}^n \lambda_i \xi_i \left\langle u_i, \sum_{j=1}^n \xi_j u_j \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \xi_i \overline{\xi_j} \langle u_i, u_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \lambda_i |\xi_i|^2 \delta_{ij} = \sum_{i=1}^n \lambda_i |\xi_i|^2 \end{aligned}$$

Da $|\xi_i|^2 \geq 0$ können wir $\alpha_i = |\xi_i|^2$ wählen. Wenn $\sum_{i=1}^n \alpha_i = 1$, dann ist $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \in \text{co}(\text{spec}(\mathbf{A}))$:

$$\begin{aligned} 1 &= \|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \left\langle \sum_{i=1}^n \xi_i \mathbf{u}_i, \sum_{j=1}^n \xi_j \mathbf{u}_j \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n |\xi_i|^2 \delta_{ij} = \sum_{i=1}^n |\xi_i|^2 = \sum_{i=1}^n \alpha_i \end{aligned}$$

□

Satz 13.2: Rayleigh-Ritz

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ selbstadjungiert und $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ die Eigenwerte von \mathbf{A} , dann:

$$\begin{aligned} \lambda_1 &= \min W(\mathbf{A}) = \min \left\{ \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \mid \mathbf{x} \neq \mathbf{0} \right\} \\ \lambda_n &= \max W(\mathbf{A}) = \max \left\{ \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \mid \mathbf{x} \neq \mathbf{0} \right\} \end{aligned}$$

Wir nennen

$$\frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}$$

den **Rayleigh-Quotient**.

Beweis 1 für Satz 13.2. Es gilt

$$\max_{\mathbf{x} \neq \mathbf{0}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \max_{\|\mathbf{x}\|=1} \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$$

Sei etwa $\|\mathbf{x}\| = 1$, dann

$$\frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \Rightarrow \max_{\|\mathbf{x}\|=1} \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \leq \max_{\mathbf{x} \neq \mathbf{0}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

Andernfalls, wenn $\mathbf{x} \neq \mathbf{0}$, dann

$$\frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \left\langle \mathbf{A} \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\rangle$$

Wir wissen aus Satz 13.1, dass $\lambda_1, \lambda_n \in W(\mathbf{A})$, womit $\lambda_1 \geq \min W(\mathbf{A})$ und $\lambda_n \leq \max W(\mathbf{A})$.

Sei $\mathbf{u}_1, \dots, \mathbf{u}_n$ eine Orthonormalbasis aus Eigenvektoren von \mathbf{A} , wobei $\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i$. Sei $\mathbf{x} \in \mathbb{C}^n$ mit $\|\mathbf{x}\| = 1$, dann können wir $\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{u}_i$ darstellen:

$$\begin{aligned} \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle &= \left\langle \sum_{i=1}^n \lambda_i \alpha_i \mathbf{u}_i, \sum_{j=1}^n \alpha_j \mathbf{u}_j \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\alpha_j} \lambda_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i=1}^n |\alpha_i|^2 \lambda_i \\ \lambda_i \geq \lambda_1 &\Rightarrow \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq \sum_{i=1}^n \lambda_1 |\alpha_i|^2 = \lambda_1 \\ \lambda_i \leq \lambda_n &\Rightarrow \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \leq \sum_{i=1}^n \lambda_n |\alpha_i|^2 = \lambda_n \end{aligned}$$

Somit $\lambda_1 = \min W(\mathbf{A})$ und $\lambda_n = \max W(\mathbf{A})$.

□

Beweis 2 für Satz 13.2. Für $\mathbb{K} = \mathbb{R}$. Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ mit $\mathbf{A} = \mathbf{A}^T$. Wir suchen

$$\sup\{\mathbf{x}^T \mathbf{A} \mathbf{x} \mid \mathbf{x}^T \mathbf{x} = 1\}$$

Dabei handelt es sich um ein Maximum, da $S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$ kompakt ist. Weiters ist $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ stetig, somit nimmt f in einem Punkt $\mathbf{x}_0 \in S^{n-1}$ ein Maximum an. Sprich für jede Kurve $\gamma(t)$, mit $\gamma(0) = \mathbf{x}_0$ hat die Funktion $t \mapsto f(\gamma(t))$ ein Maximum bei 0, sprich

$$\frac{df \circ \gamma}{dt}(0) = 0$$

Sei

$$\gamma(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

mit $\gamma(0) = \mathbf{x}_0$ und $\|\gamma(t)\|^2 = \gamma(t)^T \gamma(t) = 1$:

$$\frac{df \circ \gamma}{dt} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \frac{dx_i}{dt} = \nabla f(\gamma(t))^T \dot{\gamma}(t)$$

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{i,j=1}^n a_{ij} x_i x_j \\ \frac{\partial f}{\partial x_k} &= \sum_{i,j=1}^n a_{ij} \frac{\partial x_i x_j}{\partial x_k} = \sum_{i,j=1}^n a_{ij} (\delta_{ik} x_j + x_i \delta_{jk}) \\ &= \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n a_{ik} x_i = 2 \sum_{j=1}^n a_{kj} x_j \Rightarrow \nabla f = 2 \mathbf{A} \mathbf{x} \end{aligned}$$

Somit erhalten wir:

$$(2 \mathbf{A} \gamma(0))^T \dot{\gamma}(0) = 0 \Leftrightarrow 2 \gamma(0)^T \mathbf{A} \dot{\gamma}(0) = 0$$

Sei $\gamma_0 = \gamma(0)$, dann $\gamma_0^\perp \subseteq \{\mathbf{A} \gamma_0\}^\perp$. Es können nun verschiedene Fälle auftreten. Fall $\gamma_0 \in \ker(\mathbf{A})$, dann ist $f(\gamma_0) = 0$ und $\lambda = 0$ ist der kleinste Eigenwert. Falls $\gamma_0 \notin \ker(\mathbf{A})$, dann gilt natürlich $\gamma_0^\perp = (\mathbf{A} \gamma_0)^\perp$, dann gilt aber auch $(\gamma_0^\perp)^\perp = \text{span}(\gamma_0) = \text{span}(\mathbf{A} \gamma_0)$, sprich $\exists \lambda: \mathbf{A} \gamma_0 = \lambda \gamma_0$ da die Unterräume eindimensional sind. \square

Wie kommt man nun an die anderen Eigenwerte? Wir gehen wie für λ_1 vor:

$$\begin{aligned} \lambda_2 &= \min_{\mathbf{x} \neq \mathbf{0}} \frac{\left\langle \sum_{i=2}^n \lambda_i \alpha_i \mathbf{u}_i, \sum_{j=2}^n \alpha_j \mathbf{u}_j \right\rangle}{\left\| \sum_{l=2}^n \alpha_l \mathbf{u}_l \right\|^2} \\ &= \min_{\mathbf{x} \in \mathbf{u}_1^\perp} \frac{\langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \end{aligned}$$

Wir folgern aus dem ersten Beweis für Satz 13.2, dass

$$\begin{aligned} \lambda_k &= \min_{\mathbf{x} \in \{\mathbf{u}_1, \dots, \mathbf{u}_{k-1}\}^\perp} \frac{\langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \min_{\mathbf{x} \in \text{span}(\mathbf{u}_k, \dots, \mathbf{u}_n)} \frac{\langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \\ &= \max_{\mathbf{x} \in \{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}^\perp} \frac{\langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \max_{\mathbf{x} \in \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_k)} \frac{\langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \end{aligned}$$

Satz 13.3: Min-Max Prinzip von Courant-Fischer-Weyl

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ selbstadjungiert und $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ die Eigenwerte von \mathbf{A} , dann gilt:

i

$$\lambda_k = \max_{\substack{W \subseteq V \\ \dim W = k-1}} \min_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

ii

$$\lambda_{n+1-k} = \min_{\substack{W \subseteq V \\ \dim W = k-1}} \max_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

Beweis für Satz 13.3. Wir wissen bereits:

$$\lambda_k = \min_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

für $W = \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})$, wobei \mathbf{u}_i Eigenvektoren von \mathbf{A} sind. Hierbei gilt dann

$$\lambda_k \leq \max_{\dim W = k} \min_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

Wenn W ein Unterraum ist mit $\dim W = k-1$, dann muss gelten:

$$\min_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \leq \lambda_k$$

Wenn $\dim W = k-1$, dann $\dim W^\perp = n - (k-1) = n - k + 1$, womit

$$W^\perp \cap \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_k) \neq \{\mathbf{0}\}$$

Sei $\mathbf{v} \in (W^\perp \cap \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_k)) \setminus \{\mathbf{0}\}$, dann

$$\begin{aligned} \mathbf{v} &= \sum_{i=1}^k \alpha_i \mathbf{u}_i \\ \Rightarrow \frac{\langle \mathbf{A}\mathbf{v}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} &= \frac{\langle \sum_{i=1}^k \alpha_i \lambda_i \mathbf{u}_i, \sum_{j=1}^k \alpha_j \mathbf{u}_j \rangle}{\|\sum_{i=1}^k \alpha_i \mathbf{u}_i\|^2} = \frac{\sum_{i=1}^k \lambda_i |\alpha_i|^2}{\sum_{i=1}^k |\alpha_i|^2} \leq \lambda_k \\ \Rightarrow \min_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} &\leq \lambda_k \end{aligned}$$

Zu ii:

Wir betrachten $-\mathbf{A}$ mit den Eigenwerten $-\lambda_n \leq -\lambda_{n-1} \leq \dots \leq -\lambda_1$:

$$\begin{aligned} -\lambda_{n+1-k} &= \max_{\substack{W \\ \dim W = k-1}} \min_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle -\mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \max_{\substack{W \\ \dim W = k-1}} \left(- \max_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \right) \\ &= - \min_{\substack{W \\ \dim W = k-1}} \max_{\mathbf{x} \in W^\perp \setminus \{\mathbf{0}\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \end{aligned}$$

□

Satz 13.4: Schachtelungssatz von Cauchy

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ selbstadjungiert und $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ die Eigenwerte von \mathbf{A} . Sei $\mathbf{B} = [a_{ij}]_{i,j=1,\dots,n-1}$ mit Eigenwerten $\mu_1 \leq \mu_2 \leq \dots \leq \mu_{n-1}$, dann gilt:

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \dots \leq \lambda_{n-1} \leq \mu_{n-1} \leq \lambda_n$$

Allgemeiner: Sei \mathbf{P} eine Orthogonalprojektion von Rang $n-1$, dann sind die Eigenwerte von \mathbf{PAP} geschachtelt zwischen den Eigenwerten von \mathbf{A} .

Beweis für Satz 13.4. Sei $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^* & \gamma \end{bmatrix}$. Sei $\mathbf{w}_1, \dots, \mathbf{w}_n$ eine Orthonormalbasis aus Eigenvektoren, sodass $\mathbf{B}\mathbf{w}_i = \mu_i \mathbf{w}_i$. Wir bestimmen nun

$$\mathbf{u}_i = \begin{bmatrix} \mathbf{w}_i \\ 0 \end{bmatrix} \quad \mathbf{u}_n = \mathbf{e}_n$$

Somit ist $\mathbf{u}_1, \dots, \mathbf{u}_n$ eine Orthonormalbasis. Sei $W_k = \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}, \mathbf{u}_n)$, dann gilt

$$\lambda_{k+1} = \max_{\dim W = k} \min_{\mathbf{x} \in W^\perp \setminus \{0\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \geq \min_{\mathbf{x} \in W_k^\perp \setminus \{0\}} \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

Sei $\mathbf{x} \in W^\perp$, dann $\mathbf{x} \perp \mathbf{u}_n$, womit $x_n = 0$, sprich $\mathbf{x}^T = [\mathbf{y}^T \ 0]$ und $\mathbf{y} \in \mathbb{C}^{n-1}$:

$$\frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \frac{\left\langle \begin{bmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^* & \gamma \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} \right\rangle}{\|\mathbf{y}\|^2} = \frac{\langle \mathbf{B}\mathbf{y}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2}$$

Hernach:

$$\lambda_{k+1} \geq \min_{\mathbf{y} \in \{\mathbf{w}_1, \dots, \mathbf{w}_{k-1}\}^\perp \setminus \{0\}} \frac{\langle \mathbf{B}\mathbf{y}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2} = \mu_k$$

Umgekehrt betrachten wir wieder $-\mathbf{A}$, dann gilt $\tilde{\lambda}_i = -\lambda_{n+1-i}$, analog für $\tilde{\mu}_i$. Wie zuvor zeigen wir $\lambda_{k+1} \geq \tilde{\mu}_k$, womit

$$-\lambda_{n+1-(k+1)} \geq -\mu_{n+1} \Leftrightarrow \mu_{n-k} \geq \lambda_{n-k} \Leftrightarrow \mu_k \geq \lambda_k$$

□

Satz 13.5: Weylsche Ungleichungen

Seien $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ selbstadjungiert, dann gilt:

$$\lambda_k(\mathbf{A}) + \lambda_1(\mathbf{B}) \leq \lambda_k(\mathbf{A} + \mathbf{B}) \leq \lambda_k(\mathbf{A}) + \lambda_n(\mathbf{B})$$

Sind die Eigenwerte λ_k von \mathbf{A} geordnet, dann ist $\lambda_k(\mathbf{A})$ der k -te Eigenwert.

Beweis für Satz 13.5.

$$\frac{\langle (\mathbf{A} + \mathbf{B})\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} = \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} + \frac{\langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} \in \left[\frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} + \lambda_1(\mathbf{B}), \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2} + \lambda_n(\mathbf{B}) \right]$$

□

Wenn $\mathbf{B} \geq 0$, dann $\lambda_k(\mathbf{B}) \geq 0$ und es folgt:

$$\lambda_k(\mathbf{A} + \mathbf{B}) \geq \lambda_k(\mathbf{A}) + \lambda_1(\mathbf{B}) \geq \lambda_k(\mathbf{A})$$

Seien $\lambda_1, \dots, \lambda_n$ und μ_1, \dots, μ_n gegeben, jeweils Eigenwerte von \mathbf{A} und \mathbf{B} . Was sind die Eigenwerte von $\mathbf{A} + \mathbf{B}$? Dieses Problem ist als *Horns-Problem* bekannt.

Satz 13.6: Satz von Gešgarin

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, und

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

dann

$$\text{spec}(\mathbf{A}) \subseteq \bigcup_{i=1}^n \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

Wir nennen $\{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$ die **Grešgarin-Kreise**.

Beweis für Satz 13.6. Sei $\lambda \in \text{spec}(\mathbf{A})$ und $\mathbf{x} \in \mathbb{C}^n$ ein Eigenvektor von \mathbf{A} mit $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Sei oBdA $\|\mathbf{x}\|_\infty = \max |x_i| = 1$, sprich $\forall j: |x_j| \leq 1$ und $\exists i: |x_i| = 1$. Wir betrachten nun:

$$\begin{aligned} \sum_{j=1}^n a_{ij}x_j &= (\mathbf{A}\mathbf{x})_i = \lambda x_i \\ \Leftrightarrow \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j + a_{ii}x_i &= \lambda x_i \Rightarrow \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j = (\lambda - a_{ii})x_i \\ \Rightarrow \left| \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \right| &= |\lambda - a_{ii}| \cdot |x_i| = |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \cdot |x_j| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| = r_i \end{aligned}$$

□

Wir betrachten nun noch zuletzt die numerische Berechnung von Eigenwerten und Eigenvektoren. In der Theorie verwenden wir das charakteristische Polynom $\chi_{\mathbf{A}}(\lambda)$, was die Berechnung einer Determinante erfordert und das Auffinden von Nullstellen eines Polynoms vom Grad n , und die Bestimmung der Eigenvektoren durch das Lösen von Gleichungssystemen. Alle drei Verfahren sind numerisch instabil. Daher ist es bedeutend sinnvoller Iterationsverfahren anzuwenden. Sei dazu \mathbf{A} gegeben und ein beliebiger Startvektor \mathbf{x}_0 . Angenommen \mathbf{A}^∞ existiert, wenn $\mathbf{x}_\infty = \mathbf{A}^\infty \mathbf{x}_0$, dann $\mathbf{A}\mathbf{x}_\infty = \mathbf{x}_\infty$. Dieser Ansatz ist aber relativ uninteressant. Wir konstruieren daher

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{A}\mathbf{x}_k \\ \mathbf{x}_{k+1} &= \mathbf{w}_{k+1} \frac{1}{\|\mathbf{w}_{k+1}\|} \\ \lambda_{k+1} &= \mathbf{x}_{k+1}^T \mathbf{A}\mathbf{x}_{k+1} \end{aligned}$$

Angenommen \mathbf{A} ist diagonalisierbar und $|\lambda_n| > |\lambda_{n-1}| \geq \dots \geq |\lambda_1|$ und $\mathbf{v}_1, \dots, \mathbf{v}_n$ die Eigenvektoren, dann $\mathbf{x}_0 = \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n$, dann gilt:

$$\begin{aligned} \mathbf{x}_k &= \frac{\mathbf{A}^k \mathbf{x}_0}{\|\mathbf{A}^k \mathbf{x}_0\|} = \frac{\sum_{i=1}^n \alpha_i \lambda_i^k \mathbf{v}_i}{\left\| \sum_{i=1}^n \alpha_i \lambda_i^k \mathbf{v}_i \right\|} = \frac{\alpha_n \lambda_n^k}{|\alpha_n \lambda_n^k|} \frac{\sum_{i=1}^n \frac{\alpha_i \lambda_i^k}{\alpha_n \lambda_n^k} \mathbf{v}_i}{\left\| \sum_{i=1}^n \frac{\alpha_i \lambda_i^k}{\alpha_n \lambda_n^k} \mathbf{v}_i \right\|} \\ &= e^{i\varphi_k} \left(\sum_{i=1}^n \frac{\alpha_i}{\alpha_n} \left(\frac{\lambda_i}{\lambda_n} \right)^k \mathbf{v}_i \right) \cdot \left\| \sum_{i=1}^n \frac{\alpha_i}{\alpha_n} \left(\frac{\lambda_i}{\lambda_n} \right)^n \mathbf{v}_i \right\|^{-1} \xrightarrow{k \rightarrow \infty} e^{i\varphi_\infty} \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|} \end{aligned}$$

Sofern λ_n der betragsmäßig größte Eigenwert ist. Je weiter $|\lambda_n|$ von $|\lambda_{n-1}|$ entfernt ist, desto schneller konvergiert die Iteration. Wir nennen $|\lambda_n| - |\lambda_{n-1}|$ die **spectral-gap**.

Verzeichnisse

Abbildungsverzeichnis

1	Kommutatives Diagramm zur Matrixdarstellung eines Basiswechsels	2
2	Kommutatives Diagramm zur Bestimmung einer Matrixdarstellung mit Basiswechsel	3
3	Bestimmung aller Permutationen $\pi \in \mathfrak{S}_3$ durch τ_{12} und τ_{23}	13
4	Regel von Sarrus	13
5	Beweisskizze für den Satz des Pythagoras	18
6	Entscheidungsbaum	25
7	Beispiele für orthogonale Komplemente	30
8	Beispiel für eine konvexe Menge (links) und eine nicht-konvexe Menge (rechts)	32
9	Linearisierung einer konvexen Menge	32
10	Vergleich von t^3 und $p(t)$ auf $[0, 1]$	37
11	Darstellung für die Bestimmung einer orthogonalen Basis des \mathbb{R}^2	37
12	Ein Maximum (links) und ein Sattelpunkt (rechts)	77
13	Ein einfacher Kegel mit einigen Konturlinien	78

Sätze

6.1	1
6.2	2
6.3	2
6.4	2
6.5	3
7.1	Charakterisierung der Determinante	4
7.2	Weitere Charakterisierungen der Determinantenform	5
7.3	Permutationen als Gruppe	6
7.4	Mächtigkeit der symmetrischen Gruppe	6
7.5	Faktorisierung von Permutationen	6
7.6	Berechnung der Signatur	7
7.7	Signatur von Verknüpfungen	8
7.8	Determinanten	9
7.9	Eindeutige Determinantenform als Determinante	9
7.10	Lineare Unabhängigkeit aus Determinante	10
7.11	Bijektivität aus Determinante	12
7.12	Determinante von Verknüpfungen	12
7.13	Entwicklungssatz von Laplace	15
7.14	Bestimmung der Inversen	16
7.15	Cramersche Regel	17
8.1	Weitere Eigenschaften des Skalarproduktes	18
8.2	Skalarprodukt im \mathbb{R}^3	19
8.3	Eigenschaften des Vektorproduktes	20
8.4	Spatprodukt	21
8.5	Induzierte Norm	22
8.6	Darstellung durch Matrizen	24
8.7	Kongruenz selbstadjungierter Matrizen	25
8.8	Trägheitssatz von Sylvester	27
8.9	Cholesky-Zerlegung	28
8.10	Cosinussatz und Parallelogrammgleichung	29
8.11	Induzierte Normen	29
8.12	Eigenschaften orthogonaler Familien	29
8.13	Koordinaten in Orthonormalbasen	30
8.14	Eigenschaften des orthogonalen Komplements	31
8.15	Hauptsatz der Optimierung	33
8.16	Skalarprodukt mit Orthogonalprojektionen	34
8.17	Eigenschaften der Gram-Matrix	34
8.18	Orthogonalprojektion aus der Gram-Matrix	35

8.19	Gram-Schmidt Orthogonalisierung	38
8.20	Darstellungssatz von Riesz	40
8.21	Anwendung für den Satz von <i>F. Riesz</i>	41
8.22	Adjungiertheit in Bezug auf Orthonormalbasen	42
8.23	Eigenschaften der Adjungierten Abbildung	42
8.24	Eigenschaften unitärer Matrizen	43
8.25	Wichtige Gruppen von Matrizen	44
9.1	Polynome formen eine Algebra	46
9.2	Einsetzungssatz	47
9.3	Bestimmung von Nullstellen durch Algebraische Methoden	49
9.4	Hauptsatz der Algebra	50
10.1	Bestimmung der Eigenwerte	54
10.2	Allgemeine Form des charakteristischen Polynoms	55
10.3	Diagonalisierbarkeit anhand der Eigenvektoren	56
10.4	Zusammenhang geometrischer und algebraischer Vielfachheit	58
11.1	Eigenschaften invarianter Unterräume	60
11.2	Matrixdarstellung in Bezug auf invariante Basen	60
11.3	65
11.4	Jordansche Normalform	67
11.5	Jordansche Normalform	68
11.6	Minimalpolynom	71
11.7	Cayley-Hamilton	71
11.8	Spektralabbildungssatz	73
12.1	75
12.2	Schursche Normalform	75
12.3	Eigenwerte selbstadjungierter Matrizen	76
12.4	Definitheit selbstadjungierter Matrizen	76
12.5	Hauptachsentransformation	80
12.6	Wurzeln von Matrizen	80
12.7	Polarzerlegung	80
13.1	Satz von Töplitz-Hausdorff	82
13.2	Rayleigh-Ritz	83
13.3	Min-Max Prinzip von Courant-Fischer-Weyl	85
13.4	Schachtelungssatz von Cauchy	86
13.5	Weylsche Ungleichungen	86
13.6	Satz von Gešgarin	87

Definitionen

7.1	Determinante	4
7.2	Determinanten auf endlichdimensionalen Vektorräumen	5
7.3	Permutationen n -ter Ordnung	6
7.4	Transposition	6
7.5	Fehlstand	7
7.6	Streichung von Zeilen und Spalten	14
8.1	Länge eines Vektors im \mathbb{R}^3 und \mathbb{R}^2	18
8.2	Vektorprodukt	20
8.3	Verallgemeinerungen	21
8.4	Norm	22
8.5	Definitheit von Matrizen	24
8.6	Kongruenz von Matrizen	24
8.7	Index und Signatur von Matrizen	27
8.8	Minoren	28
8.9	Bezeichnungen und Winkel	29
8.10	Orthonormalbasis	29
8.11	Orthogonales Komplement	30
8.12	Konvexe Mengen	32
8.13	Orthogonalprojektion	33

8.14	Gram-Matrix	34
8.15	Wichtige Klassen linearer Abbildungen	42
8.16	Unitäre Matrizen	43
9.1	Algebra	46
9.2	Raum der Polynome	46
9.3	Grad eines Polynoms	46
9.4	Induzierte Funktionen	47
9.5	Algebra-Homomorphismen	47
9.6	Nullstellen von Polynomen	48
9.7	Polynomteiler	49
9.8	Reduzierbarkeit	50
9.9	Vielfachheit	50
10.1	Eigenwerte- und vektoren	52
10.2	Links- und Rechtseigenwerte	53
10.3	Spektrum	53
10.4	Diagonalisierbarkeit	55
10.5	Vielfachheiten von Eigenwerten	58
11.1	Invarianter Unterraum	60
11.2	Hauptraum	62
11.3	Nilpotente Matrizen	66
11.4	Allgemeiner Jordanblock	68
12.1	Normalität von Matrizen	74
12.2	Quadriken	80
12.3	Singulärwerte	80
13.1	Numerischer Wertebereich	82

Lemmata

6.1	1
7.1	Determinantenformen mit linearen Abbildungen	11
7.2	Determinanten von Dreiecksmatrizen	13
7.3	Elementare Zeilen- und Spaltenumformungen	14
7.4	Nutzung von Nullspalten	14
8.1	Eigenschaften des Skalarproduktes	18
8.2	Sesquilineares inneres Produkt	21
8.3	Cauchy-Schwarz-Bunjakovski Ungleichung	23
8.4	Determinanten hermitescher Matrizen	27
10.1	Unterraum der Eigenvektoren	52
10.2	Gleichheit von Links- und Rechtseigenwerten	53
10.3	Spektrum ist unabhängig von der Basis	53
10.4	Lineare Unabhängigkeit der Eigenvektoren	57
10.5	Diagonalisierbarkeit von Matrizen	58
11.1	Einschränkung auf invariante Unterräume	60
11.2	Lemma von Fitting	61
11.3	62
11.4	64
11.5	66
11.6	Auswertung von Jordanblöcken	69
12.1	Kern normaler Matrizen	74
12.2	74
13.1	Eigenwerte im numerischen Wertebereich	82

Korollare

6.1	3
7.1	Signatur aus Faktorisierung	8
7.2	Permutationen in Determinantenformen	8
7.3	Determinante einer Linearen Abbildung	11

7.4	11
7.5	12
8.1	Äußeres Produkt in Koordinatenform	20
8.2	37
11.1	zu Satz 11.2	61
11.2	aus Korollar 11.1	61
11.3	zu Satz 11.3	65
11.4	zu Satz 11.5	69
11.5	zu Satz 11.7	71
12.1	74
12.2	76
12.3	Singulärwertzerlegung	81