CHRISTOPH NIEHOFF
VORLESUNG FREITAG

FERIENKURS LINEARE ALGEBRA FÜR PHYSIKER
WS 2008/09

In diesem Teil des Ferienkurses beschäftigen wir uns mit der Transformation linearer Abbildungen auf eine einfachere Gestalt. Die vorrangigen Themen werden *Ähnlichkeit*, *Diagonalisierung* und *Trigonalisierung* sein. Wir beschränken uns hier auf den Fall endlichdimensionaler $\mathbb{K}$-Vektorräume, wobei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

§1 Ähnlichkeit

[1.1] **Definition.** *Zwei Matrizen* $\underline{\mathbf{A}}, \underline{\mathbf{B}} \in \mathbb{K}^{n\times n}$ *heißen* ***ähnlich****, falls es eine invertierbare Matrix* $\underline{\mathbf{T}} \in \mathrm{GL}(n, \mathbb{K})$ *gibt, so dass*

$$\underline{\mathbf{A}} = \underline{\mathbf{T}}^{-1}\underline{\mathbf{B}}\ \underline{\mathbf{T}} \tag{1.1}$$

□

Ähnliche Matrizen bilden **Äquivalenzklassen** mit einigen Invarianten.

[1.2] **Satz.** *Seien* $\underline{\mathbf{A}}, \underline{\mathbf{B}} \in \mathbb{K}^{n\times n}$ *zwei ähnliche Matrizen, dann gilt*

(i) $\det(\underline{\mathbf{A}}) = \det(\underline{\mathbf{B}})$

(ii) $\chi_A(\lambda) = \chi_B(\lambda)$

Beweis. Sei $\underline{\mathbf{T}}$ die Transformationsmatrix zwischen $\underline{\mathbf{A}}$ und $\underline{\mathbf{B}}$.

(i) $\det(\underline{\mathbf{A}}) = \det(\underline{\mathbf{T}}^{-1}\underline{\mathbf{B}}\ \underline{\mathbf{T}}) = \det(\underline{\mathbf{T}}^{-1})\det(\underline{\mathbf{B}})\det(\underline{\mathbf{T}}) = \det(\underline{\mathbf{T}})^{-1}\det(\underline{\mathbf{B}})\det(\underline{\mathbf{T}}) = \det(\underline{\mathbf{B}})$

(ii)
$$\begin{aligned}\chi_B(\lambda) &= \det(\underline{\mathbf{B}} - \lambda\underline{\mathbf{E_n}}) = \det(\underline{\mathbf{T}}^{-1}\underline{\mathbf{A}}\ \underline{\mathbf{T}} - \lambda\underline{\mathbf{E_n}}) = \det(\underline{\mathbf{T}}^{-1}\underline{\mathbf{A}}\ \underline{\mathbf{T}} - \underline{\mathbf{T}}^{-1}\underline{\mathbf{E_n}}\ \underline{\mathbf{T}})\\ &= \det(\underline{\mathbf{T}}^{-1}(\underline{\mathbf{A}} - \lambda\underline{\mathbf{E_n}})\underline{\mathbf{T}}) = \det(\underline{\mathbf{A}} - \lambda\underline{\mathbf{E_n}}) = \chi_A(\lambda)\end{aligned}$$

□

Aus der Invarianz des charakteristischen Polynoms folgt, dass ebenso die Eigenwerte (inkl. Vielfachheit) invariant unter Ähnlichkeitstransformationen sind. Die Eigenvektoren sind aber **keine** Invarianten.

§2 Diagonalisierung

[2.1] **Definition.** *Eine Matrix* $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$ *heißt* ***diagonalisierbar****, wenn es eine Basis des* $\mathbb{K}^n$ *aus Eigenvektoren von* $\underline{\mathbf{A}}$ *gibt. Dies ist äquivalent dazu, dass sich der Gesamtraum als direkte Summe der Eigenräume von* $\underline{\mathbf{A}}$ *darstellen lässt:*

$$\mathbb{K}^n = \bigoplus_i \mathrm{Eig}_A(\lambda_i) \tag{2.1}$$

□

Um zu charakterisieren, wann dieses möglich ist, ist es notwendig die Begriffe der *algebraischen* und *geometrischen Vielfachheit* einzuführen.

[2.2] Definition. *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$. Das charakteristische Polynom χ_A zerfalle in Linearfaktoren*

$$\chi_A(\lambda) = \prod_i (\lambda - \lambda_i)^{\zeta_i}$$

mit paarweise unterschiedlichen λ_i. Dann definieren wir

(i) ***geometrische Vielfachheit:*** $\gamma_A(\lambda_i) := \dim \mathrm{Eig}_A(\lambda_i)$

(ii) ***algebraische Vielfachheit:*** $\alpha_A(\lambda_i) := \zeta_i$

□

Im Allgemeinen müssen geometrische und algebraische Vielfachheit **nicht** gleich sein. Es gilt jedoch stets:

$$1 \leq \gamma_A(\lambda) \leq \alpha_A(\lambda) \tag{2.2}$$

[2.3] Satz. *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$, dann ist $\underline{\mathbf{A}}$ genau dann diagonalisierbar, wenn χ_A in Linearfaktoren zerfällt* ***und*** *$\gamma_A(\lambda) = \alpha_A(\lambda)$ für alle Eigenwerte λ von $\underline{\mathbf{A}}$.*

□

Für den Fall $\mathbb{K} = \mathbb{C}$ zerfällt das charakteristische Polynom nach dem Fundamentalsatz der Algebra immer in Linearfaktoren; die zweite Bedingung muss aber nicht immer erfüllt sein. Hat die Matrix $\underline{\mathbf{A}}$ paarweise verschiedene Eigenwerte, so sind die Eigenvektoren linear unabhängig und $\underline{\mathbf{A}}$ ist also diagonalisierbar.

Nun zur eigentlichen Diagonalisierung.

[2.4] Satz. *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$ diagonalisierbar, dann ist $\underline{\mathbf{A}}$ ähnlich zu einer Diagonalmatrix $\underline{\mathbf{D}}$. Das heißt, es existiert ein $\underline{\mathbf{T}} \in \mathrm{GL}(n, \mathrm{K})$, sodass*

$$\underline{\mathbf{D}} = \underline{\mathbf{T}}^{-1} \underline{\mathbf{A}}\ \underline{\mathbf{T}} \tag{2.3}$$

mit $\underline{\mathbf{D}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Hierbei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von $\underline{\mathbf{A}}$ und die Spalten von $\underline{\mathbf{T}}$ sind die jeweiligen Eigenvektoren.

□

Diagonalisierung normaler Matrizen.
Eine große Rolle spielt die Diagonalisierung für *normale* Matrizen. Wichtige Teilmengen der normalen Matrizen sind die *komplex hermiteschen* und *reell symmetrischen* Matrizen. Für diese Matrizen ergibt sich eine einfache Form der Diagonalisierung. Hierfür werden zunächst einige Matrizengruppen eingeführt.

[2.5] Definition.

(i) ***Orthogonale Matrizen*** $\mathrm{O}(n) = \{\underline{\mathbf{A}} \in \mathrm{GL}(n, \mathbb{R}) : \underline{\mathbf{A}}^{-1} = \underline{\mathbf{A}}^T\}$

(ii) ***Unitäre Matrizen*** $\mathrm{U}(n) = \{\underline{\mathbf{A}} \in \mathrm{GL}(n, \mathbb{C}) : \underline{\mathbf{A}}^{-1} = \bar{\underline{\mathbf{A}}}^T\}$

(iii) ***Spezielle orthogonale Matrizen*** $\mathrm{SO}(n) = \{\underline{\mathbf{A}} \in \mathrm{O}(n) : \det(\underline{\mathbf{A}}) = 1\}$

(iv) ***Spezielle unitäre Matrizen*** $\mathrm{SU}(n) = \{\underline{\mathbf{A}} \in \mathrm{U}(n) : \det(\underline{\mathbf{A}}) = 1\}$

□

Im $\mathbb{R}^n$ stellen die orthogonalen Matrizen Drehungen und Spiegelungen dar. Insbesondere lassen sie Winkel und Längen invariant. Ihre Zeilen (bzw. Spalten) bilden Orthonormalbasen. Die unitären Matrizen sind in gewisser Weise ihre komplexe Verallgemeinerung.

[2.6] **Definition.** *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times m}$. Dann wird die **adjungierte** Matrix $\underline{\mathbf{A}}^* \in \mathbb{K}^{m\times n}$ folgendermaßen definiert:*

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{A}}\underline{\mathbf{y}} \rangle = \langle \underline{\mathbf{A}}^*\underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle \qquad \forall \underline{\mathbf{x}} \in \mathbb{K}^m, \underline{\mathbf{y}} \in \mathbb{K}^n \tag{2.4}$$

Für die adjungierte Matrix gilt:

$$\underline{\mathbf{A}}^* = \bar{\underline{\mathbf{A}}}^T \tag{2.5}$$

*Die Matrix $\underline{\mathbf{A}}$ heißt **normal**, falls*

$$\underline{\mathbf{A}}^*\underline{\mathbf{A}} = \underline{\mathbf{A}}\ \underline{\mathbf{A}}^* \tag{2.6}$$

□

[2.7] **Satz.** *Seien $\underline{\mathbf{B}} \in \mathbb{R}^{n\times n}$ und $\underline{\mathbf{C}} \in \mathbb{C}^{n\times n}$. Dann gilt:*

*(i) Ist $\underline{\mathbf{B}}$ **symmetrisch** (d.h. $\underline{\mathbf{B}} = \underline{\mathbf{B}}^T$), so ist $\underline{\mathbf{B}}$ auch normal.*

*(ii) Ist $\underline{\mathbf{C}}$ **hermitesch** (d.h. $\underline{\mathbf{C}} = \bar{\underline{\mathbf{C}}}^T$), so ist $\underline{\mathbf{C}}$ auch normal.*

Beweis. Die Behauptungen folgen direkt aus der Definition [2.6], da gilt:

$$\underline{\mathbf{B}}^* = \underline{\mathbf{B}} \qquad \text{und} \qquad \underline{\mathbf{C}}^* = \underline{\mathbf{C}} \tag{2.7}$$

□

[2.8] **Satz.** *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$ normal. Dann existiert eine Orthonormalbasis aus Eigenvektoren von $\underline{\mathbf{A}}$.*

□

Hieraus folgt nun unmittelbar, dass jede normale Matrix (also auch jede komplex hermitesche oder reell symmetrische Matrix) diagonalisierbar ist.

[2.9] **Satz.** *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$ normal. Dann ist $\underline{\mathbf{A}}$ diagonalisierbar und es existiert ein $\underline{\mathbf{T}} \in \mathrm{U}(n)$ derart, dass*

$$\underline{\mathbf{T}}^*\underline{\mathbf{A}}\ \underline{\mathbf{T}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \tag{2.8}$$

wobei $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von $\underline{\mathbf{A}}$ sind und die normierten Eigenvektoren von $\underline{\mathbf{A}}$ die Spalten von $\underline{\mathbf{T}}$ bilden. Es gilt auch die Umkehrung.

Ist $\underline{\mathbf{A}}$ reell symmetrisch, so ist $\underline{\mathbf{T}} \in \mathrm{SO}(n)$ wählbar.
Ist $\underline{\mathbf{A}}$ komplex hermitesch, so ist $\underline{\mathbf{T}} \in \mathrm{SU}(n)$ wählbar.

□

Geometrisch anschaulich kann man es sich im reellen Fall vorstellen, dass der Endomorphismus, der durch die Matrix $\underline{\mathbf{A}}$ dargestellt wird, so in ein Koordinatensystem gedreht wird, dass er in Richtung der Koordinatenachsen skaliert.

Abschließend zu diesem Thema soll nun noch ein "**Kochrezept**" zur Diagonalisierung einer Matrix $\underline{\mathbf{A}}$ angegeben und an einem Beispiel verdeutlicht werden.

[2.10] Zusammenfassung.

Schritt 1: Berechne das charakteristische Polynom $\chi_A(\lambda) = \det(\underline{\mathbf{A}} - \lambda\underline{\mathbf{E_n}})$.

Schritt 2: Zerlege $\chi_A(\lambda)$ in Linearfaktoren: $\chi_A(\lambda) = \prod_i (\lambda - \lambda_i)^{\zeta_i}$. Ist dieses nicht möglich, so ist $\underline{\mathbf{A}}$ nicht diagonalisierbar. Aus dieser Zerlegung kann die algebraische Vielfachheit $\alpha_A(\lambda_i)$ angelesen werden.

Schritt 3: Berechne zu jedem Eigenwert λ_i eine Basis des Eigenraums $\mathrm{Eig}_A(\lambda_i) = \mathrm{Kern}(\underline{\mathbf{A}} - \lambda_i\underline{\mathbf{E_n}})$. Hieraus können die geometrische Vielfachheit $\gamma_A(\lambda_i)$ und die Eigenvektoren $\underline{\mathbf{v_i}}$ abgelesen werden. Für normale Matrizen können die Eigenvektoren normiert werden. Dadurch kann am Ende das Invertieren der Transformationsmatrix durch Adjungieren ersetzt werden.

Schritt 4: Überprüfe, ob für für <u>jeden</u> Eigenwert λ_i die Beziehung $\alpha_A(\lambda_i) = \gamma_A(\lambda_i)$ gilt. Ist dies nicht der Fall, so ist $\underline{\mathbf{A}}$ nicht diagonalisierbar. Dieser Schritt ist für normale Matrizen nicht notwendig.

Schritt 5: Bestimme die Diagonalmatrix und die Transformationsmatrix als $\underline{\mathbf{D}} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ und $\underline{\mathbf{T}} = (\underline{\mathbf{v_1}}, \dots, \underline{\mathbf{v_n}})$. Hierbei ist allerdings wichtig, dass die Reihenfolge der Indizes beide Male gleich ist.

□

[2.11] Beispiel.
Es soll folgende Matrix diagonalisiert werden:

$$\underline{\mathbf{A}} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2+i & 1+2i \\ 0 & -1-2i & 2+i \end{pmatrix} \in \mathbb{C}^{3\times 3}$$

Hierzu sehen wir zuerst

$$\underline{\mathbf{A}}\ \underline{\mathbf{A}}^* = \begin{pmatrix} 16 & 0 & 0 \\ 0 & 10 & 6i \\ 0 & -6i & 10 \end{pmatrix} = \underline{\mathbf{A}}^*\underline{\mathbf{A}}$$

Die Matrix ist also normal. Wir erwarten also Diagonalisierbarkeit und eine unitäre Transformationsmatrix.
Als nächstes berechnen wir das charakteristische Polynom.

$$\chi_A(\lambda) = \det \begin{pmatrix} 4-\lambda & 0 & 0 \\ 0 & 2+i-\lambda & 1+2i \\ 0 & -1-2i & 2+i-\lambda \end{pmatrix} = \lambda^3 - (8+2i)\lambda^2 + (16+16i)\lambda - 32i = (\lambda-4)^2(\lambda-2i)$$

Das charakteristische Polynom zerfällt in Linearfaktoren. Es ergeben sich zwei Eigenwerte $\lambda_1 = 4$ und $\lambda_2 = 2i$ mit den algebraischen Vielfachheiten $\alpha_A(4) = 2$ und $\alpha_A(2i) = 1$. Nun werden die Eigenräume zu den verschiedenen Eigenwerten berechnet.

$$\mathrm{Eig}_A(4) = \mathrm{Kern}(\underline{\mathbf{A}} - 4\underline{\mathbf{E_3}}) = \mathrm{Kern} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2+i & 1+2i \\ 0 & -1-2i & -2+i \end{pmatrix}$$

Die Matrix wird durch Umformungen auf die spezielle Zeilenstufenform gebracht.

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -2+i & 1+2i \\ 0 & -1-2i & -2+i \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & -2+i & 1+2i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Mit $\varsigma, \tau \in \mathbb{C}$ als freie Variablen ergibt sich für den Eigenraum:

$$\mathrm{Eig}_A(4) = \left\{ \varsigma \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \tau \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix} : \varsigma, \tau \in \mathbb{C} \right\}$$

Damit erhalten wir zwei normierte Eigenvektoren zum Eigenwert $\lambda_1 = 4$:

$$\underline{\mathbf{v_1}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{und} \qquad \underline{\mathbf{v_2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix} \tag{2.9}$$

Für den Eigenwert $\lambda_2 = 2i$ ergibt sich analog:

$$\mathrm{Eig}_A(2i) = \mathrm{Kern}(\underline{\mathbf{A}} - 2i\underline{\mathbf{E_3}}) = \mathrm{Kern} \begin{pmatrix} 4-2i & 0 & 0 \\ 0 & 2-i & 1+2i \\ 0 & -1-2i & 2-i \end{pmatrix}$$

Die spezielle Zeilenstufenform ist hier:

$$\begin{pmatrix} 4-2i & 0 & 0 \\ 0 & 2-i & 1+2i \\ 0 & -1-2i & 2-i \end{pmatrix} \longrightarrow \begin{pmatrix} 4-2i & 0 & 0 \\ 0 & 2-i & 1+2i \\ 0 & 0 & 0 \end{pmatrix}$$

Hier ergibt sich ein freier Parameter $\varsigma \in \mathbb{C}$:

$$\mathrm{Eig}_A(2i) = \left\{ \varsigma \begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix} : \varsigma \in \mathbb{C} \right\}$$

Somit erhalten wir einen weiteren normierten Eigenvektor:

$$\underline{\mathbf{v_3}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -i \\ 1 \end{pmatrix} \tag{2.10}$$

Außerdem erhalten wir $\gamma_A(4) = 2 = \alpha_A(4)$ und $\gamma_A(2i) = 1 = \alpha_A(2i)$. Die Matrix ist also diagonalisierbar. Außerdem sind die Eigenvektoren paarweise orthogonal. Damit erhalten wir:

$$\underline{\mathbf{D}} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2i \end{pmatrix} \tag{2.11}$$

$$\underline{\mathbf{T}} = (\underline{\mathbf{v_1}}, \underline{\mathbf{v_2}}, \underline{\mathbf{v_3}}) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & i & -i \\ 0 & 1 & 1 \end{pmatrix} \in \mathrm{U}(3) \tag{2.12}$$

Zur Kontrollen berechnen wir:

$$\underline{\mathbf{T}}\,\underline{\mathbf{D}}\,\underline{\mathbf{T}}^* = \frac{1}{2} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & i & -i \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2i \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & -i & 1 \\ 0 & i & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2+i & 1+2i \\ 0 & -1-2i & 2+i \end{pmatrix} = \underline{\mathbf{A}}$$

Die Eigenvektoren hätten nicht notwendigerweise normiert werden müssen. Mit nicht-normierten Eigenvektoren hätte jedoch Satz [2.4] verwendet werden müssen, sodass die Transformationsmatrix hätte invertiert werden müssen. In diesem Fall einer normalen Matrix erspart die Verwendung einer unitären Transformationsmatrix einige Rechenarbeit.

□

§3 Trigonalisierung

Wie wir in §2 gesehen haben, ist für die Diagonalisierbarkeit einer Matrix notwendig, dass das charakteristische Polynom in Linearfaktoren zerfällt und die Eigenräume maximale Dimension haben. Daher stellt sich die Frage, was man machen kann, wenn diese Bedingung nicht erfüllt ist. Es ergibt sich, dass eine Matrix in diesem Fall oft *trigonalisierbar* ist.

[3.1] Satz (Satz von Schur). *Sei $\underline{\mathbf{A}} \in \mathbb{K}^{n\times n}$ eine Matrix, deren charakteristisches Polynom in Linearfaktoren zerfalle:*

$$\chi_A(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i).$$

Dann ist $\underline{\mathbf{A}}$ ähnlich zu einer oberen Dreiecksmatrix $\underline{\mathbf{R}} \in \mathbb{K}^{n\times n}$, d.h. es existiert ein $\underline{\mathbf{T}} \in \mathrm{GL}(n,\mathbb{K})$, sodass

$$\underline{\mathbf{T}}^{-1}\underline{\mathbf{A}}\ \underline{\mathbf{T}} = \underline{\mathbf{R}} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \tag{3.1}$$

□

Die Bedingung, dass nur das charakteristische Polynom in Linearfaktoren zerfallen muss, ist um einiges schwächer als die Bedingung für Diagonalisierbarkeit. Insbesondere ergibt sich mit dem Fundamentalsatz der Algebra, dass für $\mathbb{K} = \mathbb{C}$ **jede** Matrix trigonalisierbar ist.