GRUNDLAGEN DER SPEKTRALTHEORIE

Schursche- und Jordansche-Normalform

von
Sarah Nowak

Fakultät für Mathematik
Universität Bielefeld

Datum: 22. Juni 2017.

Inhaltsverzeichnis

Einleitung

Diese Ausarbeitung entstand im Rahmen des Proseminars 'Angewandte Lineare Algebra' im Sommersemester 2017 unter der Leitung von Herrn Prof. Dr. Wolf-Jürgen Beyn und Herrn Dr. Denny Otten. Die Ausarbeitung bezieht sich auf den ersten Vortrag mit dem Titel 'Grundlagen der Spektraltheorie'. Thematisiert werden hier besonders die Schursche-Normalform und die Jordan-Normalform. Falls nicht anders angegeben bezieht sich der Inhalt auf Allaire, Gregoire und Kaber - 'Numerical linear algebra'.

Bielefeld, April 2017 — S. Nowak

1. Grundlagen

Wir betrachten im Folgenden quadratische Matrizen mit komplexwertigen Einträgen, was den reellen Fall einschließt.

Definition 1.1. Sei $A \in \mathcal{M}_n(\mathbb{C})$. Das *charakteristische Polynom* von A ist das auf den komplexen Zahlen definierte Polynom $P_A(\lambda)$ mit

$$P_A(\lambda) = \det(A - \lambda I),$$

wobei det die *Determinante* von A und I die Einheitsmatrix bezeichnet.

Satz 1.2. *Das charakteristische Polynom und damit die Eigenwerte einer Matrix bleiben bei einem Basiswechsel unverändert, da*

$$\det(Q^{-1}AQ - \lambda I) = \det(A - \lambda I)$$

für eine invertierbare Matrix Q.

Bemerkung 1.3. Ist $A \in \mathcal{M}_n(\mathbb{C})$, so hat das zugehörige charakteristische Polynom $P_A(\lambda)$ den Grad n. Damit hat es n Nullstellen in $\mathbb{C}$.

Definition 1.4. Die n Nullstellen eines charakteristischen Polynoms von A werden als *Eigenwerte* bezeichnet. Im Folgenden bezeichnet $\lambda(A)$ einen Eigenwert von A.

Definition 1.5. Ein Vektor $x \in \mathbb{C}^n$ $(x \neq 0)$ mit $Ax = \lambda x$ heißt *Eigenvektor* von A zu dem Eigenwert λ.

Satz 1.6. *Ist λ ein Eigenwert von A, dann existiert immer ein zugehöriger Eigenvektor $x \in \mathbb{C}^n$ $(x \neq 0)$ mit $Ax = \lambda x$. Dieser ist nicht eindeutig. Existiert hingegen ein $x \neq 0$, so dass $Ax = \lambda x$, dann ist λ ein Eigenwert von A.*

Definition 1.7. Die Menge der Eigenwerte von A wird als *Spektrum* bezeichnet und erhält die Notation $\sigma(A)$.

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid \exists x \neq 0 : Ax = \lambda x\}.$$

Definition 1.8. Der größte Betrag aller Eigenwerte einer Matrix $A \in \mathcal{M}_n(\mathbb{C})$ heißt *Spektralradius* von A und wird mit $\varrho(A)$ bezeichnet.

$$\varrho(A) = \max\{|\lambda|\}\,, \lambda \in \sigma(A).$$

Definition 1.9. Ist $A \in \mathcal{M}_n(\mathbb{C})$, dann ist die zugehörige *adjungierte Matrix* $A^H \in \mathcal{M}_n(\mathbb{C})$ definiert durch

$$A^H = \overline{A}^T = \overline{A^T}.$$

Hier bezeichnet A^T die transponierte Matrix und $\overline{A}$ die konjugierte Matrix von A. Matrizen, die gleich ihrer Adjungierten sind, werden *hermitesche Matrizen* genannt.

Definition 1.10. Eine *unitäre* Matrix ist eine komplexe quadratische Matrix, deren Zeilen- und Spaltenvektoren orthonormal bezüglich des Standardskalarprodukts sind. Damit ist die *Inverse* einer unitären Matrix gleichzeitig ihre *Adjungierte*. Durch Multiplikation mit einer unitären Matrix bleibt sowohl die euklidische Norm als auch das Standardskalarprodukt zweier Vektoren erhalten. $A \in \mathcal{M}_n(\mathbb{C})$ ist unitär, wenn gilt

$$A^H A = I,$$

also das Produkt der Matrix mit ihrer Adjungierten die Einheitsmatrix ergibt.

Definition 1.11. Eine *hermitesche Matrix* ist eine komplexe quadratische Matrix, die gleich ihrer adjungierten Matrix ist.

$$A = A^H$$

Definition 1.12. i) Sei λ ein Eigenwert von A. Wir nennen den durch

$$E_\lambda = \ker(A - \lambda I)$$

gegebenen Untervektorraum den Eigenraum zum Eigenwert λ.
ii) Den Untervektorraum, der durch

$$F_\lambda = \bigcup_{k \geqslant 1} \ker(A - \lambda I)^k$$

gegeben ist, wird generalisierter Eigenraum oder auch Hauptraum genannt.

Bemerkung 1.13. Die in Definition 1.11 ii) genutzte Vereinigung der Kerne von $(A - \lambda I)^k$ ist endlich. Das bedeutet, dass eine ganze Zahl k_0 exisitert, so dass

$$F_\lambda = \bigcup_{1 \leqslant k \leqslant k_0} \ker(A - \lambda I)^k.$$

Die Folge der Untervektorräume $\ker(A - \lambda I)^k$ ist eine steigende, verschachtelte Folge in einem endlich dimensionalem Raum. Für eine ganze Zahl $k \geqslant k_0$ bleibt die Dimension konstant. Wäre dies nicht der Fall, so würde dies der endlichen Dimension des Raumes $\mathbb{C}^n$ widersprechen. Dementsprechend sind für $k \geqslant k_0$ alle Räume $\ker(A - \lambda I)^k$ äquivalent zu $\ker(A - \lambda I)^{k_0}$.

Definition 1.14. Sei $P(X) = \sum\limits_{i=1}^{d} a_i X^i$ ein Polynom mit komplexen Koeffizienten und A eine Matrix in $\mathcal{M}_n(\mathbb{C})$. Das Matrixpolynom $P(A)$ ist definiert durch $P(A) = \sum\limits_{i=1}^{d} a_i A^i$.

Lemma 1.15. *Sei $Ax = \lambda x$ mit $x \neq 0$. Dann gilt $P(A)x = P(\lambda)x$ für alle Polynome $P(X)$.*
Mit anderen Worten: Ist λ ein Eigenwert von A, so ist $P(\lambda)$ ein Eigenwert von $P(A)$.

Theorem 1.16 (Cayley-Hamilton)**.** *Sei $P_A(\lambda)$ das charakteristische Polynom von A. Es gilt*

$$P_A(A) = 0.$$

Dies bedeutet, dass jede quadratische Matrix Nullstelle ihres charakteristischen Polynoms ist.
Der Satz zeigt, dass der kleinste mögliche Grad für ein Polynom, welches in A verschwindet, kleiner oder gleich n ist.

Definition 1.17. Das Polynom mit minimalen Grad, das in A verschwindet und bei dem der Term mit dem höchsten Koeffizienten 1 ist, heißt *Minimalpolynom* von A.

Theorem 1.18 (Spektralzerlegung)**.** *Sei $A \in \mathcal{M}_n(\mathbb{C})$ eine Matrix mit p verschiedenen Eigenwerten $(\lambda_1, \dots, \lambda_p)$, mit $1 \leqslant p \leqslant n$ und algebraischer Vielfachheit $n_1, \dots, n_p$ mit $1 \leqslant n_i \leqslant n$ und $\sum\limits_{i=1}^{p} n_i = n$. Dann genügt der zugehörige Eigenraum*

$$\mathbb{C}^n = \bigoplus_{i=1}^{p} F_{\lambda_i}, \qquad F_{\lambda_i} = \ker(A - \lambda_i I)^{n_i} \quad \textit{und} \quad \dim(F_{\lambda i}) = n_i.$$

Hier meint $\oplus$ die direkte Summe von Unterräumen.
Genauer, $\mathbb{C}^n = \bigoplus\limits_{i=1}^{p} F_{\lambda_i}$ bedeutet, dass jeder Vektor $x \in \mathbb{C}^n$ eindeutig durch $x = \sum\limits_{i=1}^{p} x^i$ mit $x^i \in F_{\lambda_i}$ zerlegt werden kann.

Bemerkung 1.19. Theorem 1.17 kann wie folgt interpretiert werden. Sei $\mathcal{B}_i$ eine Basis des generalisierten Eigenraumes $F_{\lambda i}$. Die Vereinigung aller $(\mathcal{B}_i)_{1 \leqslant i \leqslant p}$ bildet eine Basis $\mathcal{B}$ des $\mathbb{C}^n$.
Sei P die Basiswechselmatrix. Da jedes $F_{\lambda i}$ in A fest ist, erhalten wir eine neue Matrix, welche blockweise diagonal unter der Basis $\mathcal{B}$ sind.

$$P^{-1}AP = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix},$$

wobei die A_i quadratische Matrizen von der Größe n_i sind.
Mit einer passenden Basiswahl $\mathcal{B}_i$, kann jeder Block A_i als eine obere Dreiecksmatrix, mit den Eigenwerten λ_i auf der Diagonalen, geschrieben werden. Dies ist eine Vorstufe zur *Jordan-Normalform.*

2. Triangulierung von Matrizen

Definition 2.1. i) Sei $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ eine Matrix. Sind die Einträge $a_{i,j} = 0$ für $i \neq j$, so heißt A *Diagonalmatrix.*
ii) Sind die Einträge $a_{i,j} = 0$ für $i > j$, nennen wir A *obere Dreiecksmatrix.*

iii) Sind die Einträge $a_{i,j} = 0$ für $i < j$, nennen wir A *untere Dreiecksmatrix.*

Da diese Formen Matrizen vereinfachen, werden sie häufig durch einen geeigneten Basiswechsel in eine Matrix der oben genannten Formen umgewandelt.

Definition 2.2. i) Eine Matrix $A \in \mathcal{M}_n(\mathbb{C})$ kann in Dreiecksform gebracht werden, wenn eine Matrix P mit $\det(P) \neq 0$ und eine Dreiecksmatrix T existieren, so dass

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

ii) Eine Matrix $A \in \mathcal{M}_n(\mathbb{C})$ kann in Diagonalform gebracht werden, wenn eine Matrix P mit $\det(P) \neq 0$ und eine Diagonalmatrix D existieren, so dass

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

Bemerkung 2.3. Die Matrizen A und T (beziehungsweise D) sind ähnlich, denn sie stimmen bis auf Basiswahl überein. P ist die Basiswechselmatrix. Kann A diagonalisiert werden, so sind die Spaltenvektoren von P gerade die Eigenvektoren von A.
Kann A auf Diagonalform oder auf Dreiecksform gebracht werden, dann befinden sich die Eigenwerte von A in ihrer algebraischen Vielfachheit $(\lambda_1, \dots, \lambda_n)$ auf der Diagonalen von D (beziehungsweise T).

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{beziehungsweise} \quad T = \begin{pmatrix} \lambda_1 & \dots & x \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix}$$

In beiden Fällen ist das charakteristische Polynom von A gegeben durch

$$P(\lambda) = \det(A - \lambda I) = \prod_{i=1}^{n} (\lambda_i - \lambda).$$

Proposition 2.4. *Jede Matrix $A \in \mathcal{M}_n(\mathbb{C})$ kann auf Dreiecksform gebracht werden.*

Beweis. Der Beweis erfolgt durch Induktion nach der Dimension n. Für $n = 1$ ist die Proposition wahr. Wir zeigen, dass sie für $n - 1$ wahr bleibt. Für $A \in \mathcal{M}_n(\mathbb{C})$, hat das charakteristische Polynom $\det(A - \lambda I)$ mindestens eine Nullstelle $\lambda_1 \in \mathbb{C}$ mit einem zugehörigen Eigenvektor $v_1 \neq 0$, so dass $Av_1 = \lambda_1 v_1$. Wir ergänzen v_1 durch die Vektoren $(v_2, \dots, v_n)$, um eine Basis des $\mathbb{C}^n$ zu erhalten. Für $2 \leqslant j \leqslant n$, existieren Koeffizienten α_j und $b_{i,j}$, so dass

$$Ae_j = \alpha v_1 + \sum_{i=2}^{n} b_{i,j} v_i. \tag{1}$$

Sei B die Matrix mit den $n-1$ Einträgen $(b_{i,j})_{2\leqslant i,j\leqslant n}$. Sei P_1 die zugehörige Basiswechselmatrix, in deren Spalten die Basisvektoren stehen, dann ist (1) äquivalent zu

$$P_1^{-1}AP_1 = \begin{pmatrix} \lambda_1 & \alpha_2 & \dots & \alpha_n \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}.$$

Mit der Induktionsvoraussetzung folgt, dass eine Matrix P_2 mit $\det(P_2) \neq 0$ von der Größe $n-1$ existiert, so dass $P_2^{-1}BP_2 = T_2$. T_2 ist eine obere Dreiecksmatrix mit der Ordnung $n-1$. Mit P_2 erzeugen wir eine Matrix P_3 mit der Größe n

$$P_3 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & P_2 & \\ 0 & & & \end{pmatrix}.$$

Wir setzen $P = P_1P_3$ und erhalten

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & \beta_2 & \dots & \beta_n \\ 0 & & & \\ \vdots & & P_2^{-1}BP_2 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & \beta_2 & \dots & \beta_n \\ 0 & & & \\ \vdots & & T_2 & \\ 0 & & & \end{pmatrix} = T,$$

wobei T eine obere Dreiecksmatrix ist und $(\beta_2, \dots, \beta_n) = (\alpha_2, \dots, \alpha_n)P_2$. □

Ziel des nächsten Abschnittes ist es zu zeigen, dass die Umformung einer Matrix $A \in \mathcal{M}_n(\mathbb{C})$ in Dreiecksform, durch den Wechsel der Orthonormalbasis zu erreichen ist.

3. Schursche-Normalform

Als Grundlage für diesen Abschnitt diente das Vorlesungsskript Numerik I [3]. Bei der Schurschen-Normalform wird durch orthogonale Ähnlichkeitstransformationen eine Matrix $A \in \mathcal{M}_n$ in obere Dreiecksform gebracht.

Satz 3.1 (Schur). *i) Zu jeder Matrix $A \in \mathcal{M}_n(\mathbb{C})$ existiert eine unitäre Matrix $Q \in \mathcal{M}_n(\mathbb{C})$ mit*

$$Q^H AQ = \begin{bmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}, \tag{2}$$

wobei $\lambda_j \in \mathbb{C}, j = 1, \dots, n$ die Eigenwerte von A sind und so oft auf der Diagonalen vorkommen, wie ihre algebraische Vielfachheit angibt. ii)

Zu jeder Matrix $A \in \mathcal{M}_n(\mathbb{R})$ existiert eine unitäre Matrix $Q \in \mathcal{M}_n(\mathbb{R})$ mit

$$Q^T A Q = \begin{bmatrix} \Lambda_1 & * & \dots & * \\ 0 & \Lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \Lambda_n \end{bmatrix}, \tag{3}$$

wobei entweder

$$\Lambda_j = \lambda_j \text{ und } \lambda_j \text{ ein reeller Eigenwert von } A \text{ ist}$$

oder

$$\Lambda_i = \begin{pmatrix} Re(\lambda_j) & -\alpha_j Im(\lambda_j) \\ \dfrac{1}{\alpha_j} Im(\lambda_j) & Re(\lambda_j) \end{pmatrix}, \qquad \alpha_j \neq 0, \alpha_j \in \mathbb{R}$$

und $\lambda_j = Re(\lambda_j) + iIm(\lambda_j)$, wobei $Im(\lambda_j) \neq 0$ ein echt komplexer Eigenwert von A ist. Λ_j tritt so oft auf der Diagonalen auf, wie es der algebraische Vielfachheit von λ_j entspricht.

Bemerkungen 3.2. i) Die 2×2-Blöcke Λ_j haben die Eigenwerte λ_j und $\overline{\lambda}_j$. Jedes Paar $\lambda_j, \overline{\lambda}_j$ von Eigenwerten einer reellen Matrix führt zu einem solchen 2×2-Diagonalblock.
ii) Man nennt (2) die *komplexe Schursche-Normalform* und (3) die *reelle Schursche-Normalform* von A.
iii) Falls $A \in \mathcal{M}_n(\mathbb{C})$ hermitesch beziehungsweise $A \in \mathcal{M}_n(\mathbb{R})$ symmetrisch ist, so sind alle Eigenwerte reell und $Q^H AQ$ beziehungsweise $Q^T AQ$ ist ebenfalls hermitesch beziehungsweise symmetrisch. Die Schursche-Normalform muss dann notwendig diagonal sein. Der Satz impliziert insbesondere, dass sich hermitesche [symmetrische] Matrizen mit einer unitären [orthogonalen] Ähnlichkeitstransformation diagonalisieren lassen.

Beweis. Zunächst wird die reelle Schursche-Normalform (3) bewiesen. Der Beweis erfolgt durch Induktion nach n. Für $n = 1$ ist die Bedingung trivial erfüllt. Sei also bis auf $n - 1$ alles bewiesen und ein $A \in \mathcal{M}_n(\mathbb{R})$ gegeben.
Fall 1: A besitzt einen reellen Eigenwert $\lambda \in \mathbb{R}$ mit zugehörigen Eigenvektor v_1. Wir ergänzen diesen mit $v_2, \dots, v_n$, so dass $v = (v_1, v_2, \dots, v_n)$ eine Orthonormalbasis des $\mathbb{C}^n$ bildet. Wir erhalten

$$P = \left(v \cdot \frac{1}{\|v\|} \mid \tilde{P} \right)$$

Da die Spalten von P orthonormal sind, ist P eine orthonormale Matrix. Durch Multiplikation mit A erhalten wir

$$AP = \left((Av \cdot \frac{1}{\|v\|} \mid A\tilde{P} \right) = \left(\frac{\lambda v}{\|v\|} \mid A\tilde{P} \right)$$

und weiter

$$
\begin{aligned}
P^T A P &= \begin{pmatrix} v^T \frac{1}{\|v\|} \\ \tilde{P}^T \end{pmatrix} \begin{pmatrix} \frac{\lambda v}{\|v\|} & | & A\tilde{P} \end{pmatrix} \\
&= \left(\begin{array}{c|c} \lambda & \frac{1}{\|v\|} v^T \tilde{P} \\ \hline \tilde{P}^T \lambda \frac{1}{\|v\|} v & \tilde{P}^T A \tilde{P} \end{array}\right) \\
&= \begin{pmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}.
\end{aligned}
$$

Hier geht ein, dass $v^T v = \|v\|^2$ gilt. Die Spalten von $\tilde{P}$ stehen orthogonal auf v. Wir benennen die erhaltene Form

$$
P^T A P = \begin{bmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix}, \quad B \in \mathcal{M}_{n-1}(\mathbb{R}).
$$

Nach Induktionsvoraussetzung existiert ein $Q \in \mathcal{M}_{n-1}(\mathbb{R})$, so dass $Q^T B Q$ die (reelle) Schursche-Normalform (3) hat. Auch die Matrix

$$
R = P \begin{bmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{bmatrix}
$$

ist orthogonal. Man erhält die Schursche-Normalform für A

$$
R^T A R = \begin{bmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & Q^T & \\ 0 & & & \end{bmatrix} P^T A P \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda & * & \dots & * \\ 0 & & & \\ \vdots & & Q^T B Q & \\ 0 & & & \end{bmatrix}.
$$

Fall 2: A besitzt einen komplexen Eigenwert $\lambda = \mu + i\rho$, $\rho \neq 0$ mit Eigenvektor $z = x + iy \neq 0$; $x, y \in \mathbb{R}^n$. Es gilt

$$
Az = Ax + iAy = (\mu + i\rho)(x + iy) = \mu x - \rho y + i(\rho x + \mu y),
$$

und damit

$$
Ax = \mu x - \rho y, \qquad Ay = \rho x + \mu y. \tag{4}
$$

Aus $z \neq 0$ und $\rho \neq 0$ folgt sowohl $x \neq 0$ sowie $y \neq 0$, denn ist $x = 0$, so erhalten wir $Ax = 0$. Andererseits gilt jedoch $Ax = \mu x - \rho y$. Um dies zu erfüllen müsste auch $y = 0$ sein, was im Widerspruch zu $z \neq 0$ steht. Der andere Fall folgt analog. Wir können nun außerdem o.B.d.A.

$$x^T y = 0 \text{ und } x^T x = 1$$

annehmen.
Es gilt $x + iy = z$. Betrachten wir $z^t z = x^t x - y^t y + 2ix^t y$.
Dabei ist $2ix^t y = \mathrm{Im}(z^t z)$.
Sei $\lambda = a + ib \in \mathbb{C}$. Wir erhalten

$$\mathrm{Im}((\lambda z)^t \lambda z) = \mathrm{Im}(\lambda^2 z^t z) = \mathrm{Re}(\lambda^2)\mathrm{Im}(z^t z) + \mathrm{Im}(\lambda^2)\mathrm{Re}(z^t z).$$

Ist $z^t z = 2x^t y = 0$, sind wir fertig. Sei also $\mathrm{Im}(z^t z) \neq 0$.
Setze

$$\alpha = \frac{\mathrm{Re}(z^t z)}{\mathrm{Im}(z^t z)}$$

Dann existiert ein $r > 0$, $\phi \in (\pi, -\pi)$ mit $\alpha + i = re^{i\phi}$.
Definiere

$$\lambda = \sqrt{r} e^{i\phi/2} = \sqrt{r} \cos(\phi/2) + ir \sin(\phi/2).$$

Damit folgt

$$\lambda^2 = \alpha + i,$$

also

$$\alpha = \frac{\mathrm{Re}(\lambda^2)}{\mathrm{Im}(\lambda^2)} = \alpha = \frac{\mathrm{Re}(z^t z)}{\mathrm{Im}(z^t z)}$$

Für die Matrix $P_1 = (x|(-\alpha y)) \in \mathcal{M}_{n,2}(\mathbb{R})$, $\alpha = \dfrac{1}{\|y\|_2}$ erhalten wir aus (4) die Beziehung

$$P_1^T A P_1 = \begin{pmatrix} \mu & -\alpha\rho \\ \frac{\rho}{\alpha} & \mu \end{pmatrix}.$$

Betrachte dazu das Bild von x, also $Ax = \mu x - \rho y = \mu x - \dfrac{\rho}{\alpha}\alpha y$ und das Bild von $-\alpha y$. $A(-\alpha y) = -\alpha A y = -\alpha(\rho x + \mu y) = -\alpha\rho x - \alpha\mu y$. Da P_1 zwei orthogonale Spalten hat, können wir diese Spalten zu einer orthogonalen Matrix

$$P = \begin{pmatrix} P_1 & P_2 \end{pmatrix}, \qquad P_2 \in \mathcal{M}_{n,n-2}(\mathbb{R})$$

ergänzen.
Somit hat $P^T AP$ die Form

$$P^T AP = \begin{bmatrix} \mu & -\alpha\rho & * & \dots & * \\ \frac{\rho}{\alpha} & \mu & * & \dots & * \\ 0 & 0 & & & \\ \vdots & \vdots & & B & \\ 0 & 0 & & & \end{bmatrix}, \quad B \in \mathcal{M}_{n-2}(\mathbb{R}).$$

Da der von x und y erzeugte Unterraum A-invariant ist, und damit auch der von x und $-\alpha y$, erhalten wir den 'Null'-Block unten links.

Die Matrix $\begin{pmatrix} \mu & -\alpha\rho \\ \frac{\rho}{\alpha} & \mu \end{pmatrix}$ besitzt genau die gewünschte Gestalt und wir können wie in Fall 1 vorgehen. Nach Induktionsannahme gibt es eine orthogonale Matrix Q $\in \mathcal{M}_{n-2}(\mathbb{R})$, so dass $Q^T BQ$ die Schursche-Normalform hat, und die orthogonale Matrix

$$R = P \begin{bmatrix} I_2 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{bmatrix}$$

bringt dann A auf ebenfalls auf die Schursche-Normalform.
Nun erfolgt der Beweis der komplexen Schurschen-Normalform (2) nach [1]. Es ist also zu zeigen, dass für eine Matrix $A \in \mathcal{M}_n(\mathbb{C})$ eine unitäre Matrix U (das heißt $U^{-1} = U^*$) existiert , so dass $U^{-1}AU$ eine Dreiecksmatrix ist.
Sei $(e_i)_{i=1}^n$ die kanonische Basis und sei $(f_i)_{i=1}^n$ die Basis mit der A Dreiecksgestalt hat. Nach Proposition 2.4 wissen wir, dass $A = PTP^{-1}$. Der Basiswechsel soll unitär sein. Wir wenden das Gram-Schmidt Verfahren auf die Basis $(f_i)_{i=1}^n$, was uns eine orthonormal Basis $(g_i)_{i=1}^n$ liefert, so dass für jedes $1 \leqslant i \leqslant n$,

$$\operatorname{span}\{g_1, \dots, g_i\} = \operatorname{span}\{f_1, \dots, f_i\}.$$

Wir erhalten hier beispielsweise $g_1 = \dfrac{f_1}{\|f_1\|}$.
Nun ist $g_2' = f_2 - \dfrac{\langle g_1, f_2\rangle}{g_1, g_1} g_1 = f_2 - \langle g_1, f_2\rangle g_1$ und damit $g_2 = \dfrac{g_2'}{\|g_2'\|}$. Wir erhalten damit $g_3' = f_3 - \langle g_1, f_3\rangle - \langle g_2, f_3\rangle g_2$ und so weiter. Wichtig hierbei ist, dass die Dreickesgestalt nicht verändert wird.
Da $AP = PT$, T eine obere Dreiecksmatrix, liefert und die Betrachtung der ersten i Spalten der Gleichung

$$\operatorname{span}\{Af_1, \dots, Af_i\} \subset \operatorname{span}\{f_1, \dots, f_i\}, \qquad \text{für alle } 1 \leqslant i \leqslant n.$$

Daraus folgern wir

$$\operatorname{span}\{Ag_1, \dots, Ag_i\} \subset \operatorname{span}\{g_1, \dots, g_i\}. \tag{5}$$

Für eine invertierbare Matrix A erhalten wir $\operatorname{span}\{Ag_1, \dots, Ag_i\} = \operatorname{span}\{g_1, \dots, g_i\}$. Umgekehrt impliziert (5), dass eine obere Dreiecksmatrix R existiert, so dass $AU = UR$, deren Spalten die orthonormal Vektoren $(g_i)_{i=1}^n$ sind. U ist eine unitäre Matrix. ☐

Beispiel 3.3. Sei $\epsilon \in \mathbb{R}$. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 1 \\ \epsilon & 1 \end{pmatrix}. \tag{6}$$

Im Fall $\epsilon = 0$ hat A bereits obere Dreiecksgestalt.
Betrachten wir zunächst $\epsilon > 0$. Das charakteristische Polynom von A

ist gegeben durch $(1-\lambda)^2-\epsilon$. Damit hat A die Eigenwerte $\pm\sqrt{\epsilon}+1$. $v = \frac{1}{\sqrt{1+\epsilon}}\begin{pmatrix}1\\ \sqrt{\epsilon}\end{pmatrix}$ ist der zugehörige Eigenvektor zum Eigenwert $1+\sqrt{\epsilon}$, da

$$Av = \frac{1}{\sqrt{1+\epsilon}}A\begin{pmatrix}1\\ \sqrt{\epsilon}\end{pmatrix} = \frac{1}{\sqrt{1+\epsilon}}\begin{pmatrix}1+\sqrt{\epsilon}\\ \sqrt{\epsilon}+\epsilon\end{pmatrix} = (1+\sqrt{\epsilon})v.$$

Außerdem gilt $\|v\| = 1$. Der Vektor $\begin{pmatrix}-\sqrt{\epsilon}\\ 1\end{pmatrix}$ ist orthogonal zu $\begin{pmatrix}1\\ \sqrt{\epsilon}\end{pmatrix}$. Sei nun $Q = \left(v|\frac{1}{\sqrt{1+\epsilon}}\begin{pmatrix}-\sqrt{\epsilon}\\ 1\end{pmatrix}\right) = \frac{1}{\sqrt{1+\epsilon}}\begin{pmatrix}1 & -\sqrt{\epsilon}\\ \sqrt{\epsilon} & 1\end{pmatrix}$. Wir erhalten damit

$$Q^T AQ = \begin{pmatrix}1+\sqrt{\epsilon} & 1-\epsilon\\ 0 & 1-\sqrt{\epsilon}\end{pmatrix}$$

als Schursche-Normalform von A.
Betrachten wir nun den Fall $\epsilon < 0$.
In diesem Fall erhalten wir als Eigenwerte

$$\lambda = 1+i\sqrt{-\epsilon} \quad \text{und} \quad \overline{\lambda} = 1-i\sqrt{-\epsilon}.$$

Wir betrachten nun

$$\begin{aligned} A\begin{pmatrix}z_1\\ z_2\end{pmatrix} = \begin{pmatrix}z_1+z_2\\ \epsilon z_1+z_2\end{pmatrix} &= (1+\lambda)\begin{pmatrix}z_1\\ z_2\end{pmatrix} \\ &= \begin{pmatrix}(1+\sqrt{-\epsilon}\, i)\, z_1\\ (1+\sqrt{-\epsilon}\, i)\, z_2\end{pmatrix} \\ &= \begin{pmatrix}z_1+\sqrt{-\epsilon}\, i\, z_1\\ z_2+\sqrt{-\epsilon}\, i\, z_2\end{pmatrix}\end{aligned}$$

Daraus können wir folgern, dass

$$z_2 = \sqrt{-\epsilon}\, i\, z_1 \quad \text{und} \quad \epsilon z_1 = \sqrt{-\epsilon}\, i\, z_2.$$

Damit gilt

$$\epsilon z_1 = \sqrt{-\epsilon}\, i(\sqrt{-\epsilon}\, i)z_1 = \epsilon z_1.$$

Daher darf nicht gelten $z_1 = 0$. Wählen wir $z_1 = 1$. Damit ist nun

$$\epsilon = \sqrt{-\epsilon}\, i\, z_2,$$

was äquivalent ist zu

$$z_2 = \frac{\epsilon}{\sqrt{-\epsilon}\, i} = -i\,\frac{\epsilon}{\sqrt{-\epsilon}} = i\,\frac{-\epsilon}{\sqrt{-\epsilon}} = i\,\frac{(\sqrt{-\epsilon})^2}{\sqrt{-\epsilon}} = i\,\sqrt{-\epsilon}.$$

Damit erhalten wir den Eigenvektor v zum Eigenwert λ

$$v = \begin{pmatrix}1\\ i\,\sqrt{-\epsilon}\end{pmatrix} = \begin{pmatrix}1\\ 0\end{pmatrix} + i\begin{pmatrix}0\\ \sqrt{-\epsilon}\end{pmatrix}.$$

Damit ist $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $y = \begin{pmatrix} 0 \\ \sqrt{-\epsilon} \end{pmatrix}$. Da Q von der Form $(x|(-\alpha y))$ ist und $\alpha^{-1} = \sqrt{-\epsilon}$ erhalten wir

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Die Schursche-Normalform im Fall $\epsilon < 0$ ist gegeben durch

$$Q^T AQ = \begin{pmatrix} 1 & -1 \\ -\epsilon & 1 \end{pmatrix}.$$

4. Jordan-Normalform

Als Grundlage für dieses Kapitel diente 'Matrix analysis' von Horn und Johnson [2].

Definition 4.1. Ein *Jordanblock* $J_k(\lambda)$ ist eine $k \times k$ große obere Dreiecksmatrix von der Form

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix}.$$

Es gibt $k-1$ Terme '+1' auf der Nebendiagonalen. Der Skalar λ erscheint k-mal auf der Hauptdiagonalen. Alle anderen Einträge sind Null.

Sei $A \in \mathcal{M}_n(\mathbb{C})$. Es existiert eine Matrix $S \in \mathcal{M}_n$ mit $\det(S) \neq 0$, so dass

$$A = S \begin{bmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_k}(\lambda_k) \end{bmatrix} S^{-1} = SJS^{-1}$$

und $n_1 + n_2 + \ldots + n_k = n$. Die Jordan-Matrix J von A ist eine direkte Summe von Jordanblöcken und ist bis auf Vertauschung dieser auf der Diagonalen eindeutig. Die Eigenwerte λ_i mit $i = 1, \ldots, k$ sind nicht notwendiger Weise verschieden. Wenn A eine reelle Matrix mit ausschließlich reellen Eigenwerten ist, so ist auch S eine reelle Matrix.

Beispiel 4.2. Sei $\epsilon \in \mathbb{R}$. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 1 \\ \epsilon & 1 \end{pmatrix}.$$

Im Fall $\epsilon = 0$ erhalten wir die Matrix $A(\epsilon) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, welche bereits Jordan-Normalform hat.

Im Fall $\epsilon > 0$ erhalten wir die Eigenwerte von A durch $\lambda_1 = 1 + \sqrt{\epsilon}$ und $\lambda_2 = 1 - \sqrt{\epsilon}$. Da wir zwei verschiedene Eigenwerte erhalten ergibt sich die Jordan-Normalform

$$J = \begin{pmatrix} 1 - \sqrt{\epsilon} & 0 \\ 0 & 1 + \sqrt{\epsilon} \end{pmatrix}.$$

Ist nun $\epsilon < 0$, so erhalten wir die Eigenwerte von A durch $\lambda_1 = 1 + \sqrt{\epsilon}\, i$ und $\lambda_2 = 1 - \sqrt{\epsilon}\, i$. Da wir zwei verschiedene Eigenwerte erhalten ergibt sich die Jordan-Normalform

$$J = \begin{pmatrix} 1 - \sqrt{\epsilon}\, i & 0 \\ 0 & 1 + \sqrt{\epsilon}\, i \end{pmatrix}.$$

5. Diagonalform

Proposition 5.1. *Sei $A \in \mathcal{M}_n(\mathbb{C})$ eine Matrix mit den verschiedenen Eigenwerten $(\lambda_1, \ldots, \lambda_p)$, $1 \leqslant p \leqslant n$. Die Matrix A ist genau dann diagonalisierbar, wenn*

$$\mathbb{C}^n = \bigoplus_{i=1}^{p} E_{\lambda_i},$$

oder, äquivalent dazu, genau dann, wenn $F_{\lambda_i} = E_{\lambda_i}$ für ein $1 \leqslant_i \leqslant p$.

Beweis. Wenn $\mathbb{C}^n = \bigoplus_{i=1}^{p} E_{\lambda_i}$, dann ist A diagonal mit einer Basis, welche die Vereinigung von Basen des Unterraumes E_{λ_i} ist. Existiert umgekehrt eine Matrix P mit $\det(P) \neq 0$, so dass $P^{-1}AP$ diagonal ist, so folgt $\mathbb{C}^n = \bigoplus_{i=1}^{p} E_{\lambda_i}$. Wir erhalten $E_{\lambda i} \subset F_{\lambda_i}$ und $\mathbb{C}^n = \bigoplus_{i=1}^{p} F_{\lambda i}$ mit Theorem 1.15. Daher müssen die Identitäten $E_{\lambda_i} = F_{\lambda_i}$ für alle $1 \leqslant i \leqslant p$ äquivalent seien, damit A diagonalisierbar ist. $\square$

Im Allgemeinen sind nicht alle Matrizen diagonalisierbar. Beschränken wir uns auf diagonalisierbare Matrizen in einer Orthonormalbasis von Eigenvektoren, so können diese elementar charakterisiert werden, nämlich als Menge diagonalisierbarer Matrize in einer Orthonormalbasis, die mit der Menge der *normalen* Matrizen übereinstimmen. Eine solche Matrix kommutiert mit ihrer adjungierten Matrix ($AA^* = A^*A$).

Theorem 5.2. *Sei $A \in \mathcal{M}_n(\mathbb{C})$ eine Matrix und U eine komplexe quadratische Matrix, deren Zeilen- und Spaltenvektoren orthonormal bezüglich des Standardskalarprodukts sind. Eine solche Matrix nennen wir* unitär. *A ist genau dann* normal, *wenn eine Matrix U exisitiert, so dass*

$$A = U \;\, diag(\lambda_1, \ldots, \lambda_n) U^{-1},$$

wobei $(\lambda_1, \ldots, \lambda_n)$ die Eigenwerte von A sind.

Beweis. Offensichtlich ist eine Matrix $A = UDU^*$ normal, wobei D eine Diagonalmatrix und U unitär ist. Auf der anderen Seite wissen wir mit Definition 2.3, dass jede Matrix A auf Diagonalform gebracht werden kann. In anderen Worten, es existiert eine unitäre Matrix U und eine obere Dreiecksmatrix T, so dass $A = UTU^*$ gilt. $AA^* = A^*A$ impliziert $TT^* = T^*T$, also ist T normal.
Wir zeigen nun: Ist eine Matrix sowohl in Dreiecksform als auch normal, so ist sie eine Diagonalmatrix. Nach Definition haben wir $T = (t_{i,j})_{1\leqslant i,j\leqslant n}$ mit $t_{i,j} = 0$ für $i > j$. Wir identifizieren wir den Eintrag in der ersten Zeile und der ersten Spalte des Produkts $TT^* = T^*T$, können wir folgern, dass

$$|t_{1,1}|^2 = \sum_{k=1}^{n} |t_{1,k}|^2,$$

was uns $t_{1,k} = 0$ liefert, für alle $2 \leqslant k \leqslant n$. Dies bedeutet, dass in der ersten Reihe von T nur Nullen stehen. Der Beweis erfolgt durch Induktion. Wir nehmen an, dass in den ersten $(i-1)$ Reihen von T, bis auf die Diagonaleinträge, nur Nullen stehen. Der Eintrag in der i-ten Zeile und i-ten Spalte des Produkts $TT^* = T^*T$ liefert

$$|t_{i,i}|^2 = \sum_{k=i}^{n} |t_{i,k}|^2,$$

so dass $t_{i,k} = 0$ für alle $i + 1 \leqslant k \leqslant n$. Damit sehen wir, dass die i-te Zeile von T ebenso nur Nullen außerhalb der Diagonalen hat. Damit ist gezeigt, dass T diagonal ist. $\square$

Literatur

[1] Allaire, Gregoire and Sidi M. Kaber. Numerical linear algebra. New York: Springer, 2008.
[2] Horn, Roger A. and Charles R. Johnson. Matrix analysis. Cambridgeshire New York: Cambridge University Press, 1985.
[3] W.-J. Beyn, Th. Hüls. Vorlesungsskript Numerik I (SoSe15)

Ich versichere hiermit, die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Sarah Nowak

Bielefeld, Tag Monat Jahr