

Lösungen zu Aufgabe 1, Zettel 8

Jendrik Stelzner

8. Juli 2016

1 Vorbereitung: Basiswechsellmatrizen

Lemma 1. Es sei $A \in M_n(\mathbb{K})$. Dann sind die folgenden Bedingungen äquivalent:

1. Die Matrix A ist invertierbar mit $A^{-1} = A^*$.
2. Es gilt $AA^* = I$.
3. Es gilt $A^*A = I$.
4. Die Spalten von A sind eine Orthonormalbasis von \mathbb{K}^n (als Spaltenvektoren gesehen).
5. Die Zeilen von A sind eine Orthonormalbasis von \mathbb{K}^n (als Zeilenvektoren gesehen).

Beweis. Die Äquivalenz der ersten drei Aussagen folgt, wie aus Lineare Algebra I bekannt, mithilfe der Dimensionsformel.

Dass $A^*A = I$ ist äquivalent dazu, dass $\sum_{l=1}^n \overline{a_{lj}} a_{lk} = \delta_{jk}$ für alle $j, k = 1, \dots, n$. Dies ist durch Konjugation äquivalent dazu, dass $\sum_{l=1}^n a_{lj} \overline{a_{lk}} = \delta_{j,k}$ für alle $j, k = 1, \dots, n$. Da der Ausdruck $\sum_{l=1}^n a_{lj} \overline{a_{lk}}$ das Standardskalarprodukt der j -ten und k -ten Spalten von A ist, bedeutet dies gerade, dass die Spalten von A eine Orthonormalbasis von \mathbb{K}^n bilden.

Analog ergibt sich, dass $AA^* = I$ äquivalent dazu ist, dass die Zeilen von A eine Orthonormalbasis von \mathbb{K}^n bilden. \square

Im Folgenden sei

$$D_\varphi := \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

die Drehmatrix um den Winkel $\varphi \in \mathbb{R}$. Für Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ sei

$$\text{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \in M_n(\mathbb{C})$$

die entsprechende Diagonalmatrix. Für Matrizen $A_1 \in M_{n_1}(\mathbb{C}), \dots, A_r \in M_{n_r}(\mathbb{C})$ sei

$$\text{block}(A_1, \dots, A_r) = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix} \in M_{n_1 + \dots + n_r}(\mathbb{C})$$

die entsprechende Blockdiagonalmatrix.

Theorem 2. 1. Ist $A \in M_n(\mathbb{C})$ normal, so gibt es eine unitäre Matrix $U \in U(n)$, so dass UAU^{-1} in Diagonalgestalt ist. Die Diagonaleinträge sind dabei bis auf Permutation eindeutig bestimmt.

2. Ist $A \in M_n(\mathbb{C})$ selbstadjungiert, so gibt es eine unitäre Matrix $U \in U(n)$, so dass UAU^{-1} in Diagonalgestalt mit reellen Diagonaleinträgen ist. Die Diagonaleinträge sind dabei bis auf Permutation eindeutig bestimmt.

3. Ist $A \in M_n(\mathbb{C})$ antiselbstadjungiert, so gibt es eine unitäre Matrix $U \in U(n)$, so dass UAU^{-1} in Diagonalgestalt mit rein imaginären Diagonaleinträgen ist. Die Diagonaleinträge sind dabei bis auf Permutation eindeutig bestimmt.

4. Ist $A \in M_n(\mathbb{C})$ unitär, so gibt es eine unitäre Matrix $U \in U(n)$, so dass UAU^{-1} in Diagonalgestalt ist, und alle Diagonaleinträge haben Betrag 1. Die Diagonaleinträge sind dabei bis auf Permutation eindeutig bestimmt.

5. Ist $A \in M_n(\mathbb{R})$ normal, so gibt es eine orthogonale Matrix $O \in O(n)$, so dass

$$OAO^{-1} = \text{block}(\lambda_1, \dots, \lambda_p, r_1 D_{\varphi_1}, \dots, r_q D_{\varphi_q}).$$

mit $\lambda_1, \dots, \lambda_p \in \mathbb{R}, r_1, \dots, r_q > 0$ und $\varphi_1, \dots, \varphi_q \in (0, \pi)$. Die Zahlen r und s sind dabei eindeutig bestimmt, und die Skalare $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ und Paare $(r_1, \varphi_1), \dots, (r_q, \varphi_q)$ sind dabei jeweils bis auf Permutation eindeutig bestimmt.

6. Ist $A \in M_n(\mathbb{R})$ selbstadjungiert, so gibt es eine orthogonale Matrix $O \in O(n)$, so dass OAO^{-1} in Diagonalgestalt ist. Die Diagonaleinträge sind dabei

7. Ist $A \in M_n(\mathbb{R})$ orthogonal, so gibt es eine orthogonale Matrix $O \in O(n)$, so dass

$$OAO^{-1} = \text{block}(1, \dots, 1, -1, \dots, -1, D_{\varphi_1}, \dots, D_{\varphi_r}).$$

mit Winkeln $\varphi_1, \dots, \varphi_r \in (0, \pi)$. Dabei eindeutig bestimmt, wie häufig die Einträge 1 und -1 vorkommen, und die Winkel $\varphi_1, \dots, \varphi_r$ sind bis auf Permutation eindeutig bestimmt.

Beweis. Wir betrachten den Fall, dass $A \in M_n(\mathbb{C})$ normal ist. Es sei $\mathcal{B} = (e_1, \dots, e_n)$ die Standardbasis von \mathbb{C}^n und $f: V \rightarrow V$ der eindeutige Endomorphismus mit $M_{\mathcal{B}}(f) = A$. Da \mathcal{B} eine Orthonormalbasis ist, folgt aus der Normalität von A , dass der Endomorphismus f normal ist. Da \mathbb{C}^n endlichdimensional ist, gibt es eine Orthonormalbasis $\mathcal{C} = (c_1, \dots, c_n)$ von \mathbb{C}^n aus Eigenvektoren von f . Für die Basiswechselmatrix $U := T_{\mathcal{C}}^{\mathcal{B}}$ gilt nun, dass

$$UAU^{-1} = T_{\mathcal{C}}^{\mathcal{B}} M_{\mathcal{B}}(f) T_{\mathcal{B}}^{\mathcal{C}} = M_{\mathcal{C}}(f)$$

eine Diagonalmatrix ist. Die Spalten der Matrix $U^{-1} = T_B^C$ sind genau die Spaltenvektoren $c_1, \dots, c_n \in \mathbb{C}^n$. Also sind die Spalten von U^{-1} eine Orthonormalbasis von \mathbb{C}^n , und U^{-1} somit unitär. Deshalb ist auch U unitär.

Das zeigt die erste Aussage. Die anderen Aussagen ergeben sich analog über die Normalformen der entsprechenden Endomorphismen. \square

Im Folgenden seien

$$I := \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \quad \text{und} \quad J := \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}.$$

Lemma 3. Für alle $r, \theta \in \mathbb{R}$ ist $\exp(\log(r)I + \theta J) = rD_\theta$.

Beweis. Da I und J kommutieren (denn I ist die Einheitsmatrix), kommutieren auch $\log(r)I$ und θJ . Daher ist

$$\exp(\log(r)I + \theta J) = \exp(\log(r)I) \exp(\theta J) = \exp(\log(r))I \exp(\theta J) = r \exp(\theta J).$$

Da $J^2 = -I$ gilt für alle $n \in \mathbb{N}$, dass

$$J^n = \begin{cases} I & \text{falls } n \equiv 0 \pmod{4} \\ J & \text{falls } n \equiv 1 \pmod{4} \\ -I & \text{falls } n \equiv 2 \pmod{4} \\ -J & \text{falls } n \equiv 3 \pmod{4}. \end{cases}$$

Damit ergibt sich, dass

$$\begin{aligned} \exp(\theta J) &= \sum_{n=0}^{\infty} \frac{(\theta J)^n}{n!} = \sum_{k=0}^{\infty} \frac{(\theta J)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(\theta J)^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{2k} I}{(2k)!} + \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{2k+1} J}{(2k+1)!} = \cos(\theta)I + \sin(\theta)J = D_\theta. \end{aligned}$$

Zusammengefasst ist also $\exp(\log(r)I + \theta J) = r \exp(\theta J) = rD_\theta$. \square

Bemerkung 4. Lemma 3 lässt sich auch konzeptioneller begründen: Es sei

$$C := \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\} \subseteq M_2(\mathbb{R}).$$

Die Abbildung $\Phi: \mathbb{C} \rightarrow C$ mit

$$\Phi(a + ib) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = aI + bJ \quad \text{für alle } a, b \in \mathbb{R}$$

ist bijektiv, und durch direktes Nachrechnen ergibt sich, dass Φ bezüglich der üblichen Matrixaddition und -multiplikation ein Ringhomomorphismus ist. (D.h. es ist $\Phi(z_1 + z_2) = \Phi(z_1) + \Phi(z_2)$, $\Phi(z_1 \cdot z_2) = \Phi(z_1) \cdot \Phi(z_2)$ und $\Phi(1) = I$ für alle $z_1, z_2 \in \mathbb{C}$.) Also ist Φ ein

Ringisomorphismus. Da \mathbb{C} ein Körper ist, folgt damit, dass C mit der üblichen Matrixaddition und -multiplikation ebenfalls ein Körper ist, und dass Φ ein Isomorphismus von Körpern ist. Dabei entspricht die Darstellung $a + ib$ einer komplexen Zahl der Darstellung $aI + bJ$ einer Matrix aus C , und die Polardarstellung $re^{i\varphi}$ einer komplexen Zahl der Darstellung rD_φ einer Matrix in C .

Neben diesen algebraischen Eigenschaft sind Φ und Φ^{-1} auch stetig. Somit ist Φ auch ein Homöomorphismus. (Insgesamt ist Φ also ein Isomorphismus von topologischen Körpern.)

Für alle $z_1, z_2 \in \mathbb{C}$ ist deshalb genau dann $z_2 = \exp(z_1)$, wenn $\Phi(z_2) = \exp(\Phi(z_1))$. Somit lassen sich Aussagen über das Matrixexponential auf C in Aussagen über die Exponentialabbildung auf \mathbb{C} übersetzen. Insbesondere übersetzt sich das Problem, einen Logarithmus einer Matrix $rD_\varphi \in C$ zu finden, darin, einen Logarithmus einer komplexen Zahl $re^{i\varphi}$ zu finden. Ein solcher ist durch $\log(r) + i\varphi$ gegeben. Diese komplexe Zahl entspricht zurückübersetzt der Matrix $\log(r)I + \varphi J \in C$.

Man bemerke, dass der obige Beweis von Lemma 3 eine Übersetzung des üblichen Beweises ist, dass $\exp(i\varphi) = \cos \varphi + i \sin \varphi$ für alle $\varphi \in \mathbb{R}$.

Lemma 5. Es sei $A \in M_n(\mathbb{C})$ und $U \in U(n)$. Dann ist $(UAU^{-1})^* = UA^*U^{-1}$.

Beweis. Es ist $(UAU^{-1})^* = (U^{-1})^*A^*U^* = (U^*)^{-1}A^*U^* = (U^{-1})^{-1}A^*U^{-1} = UAU^{-1}$. \square

Korollar 6. Es sei $A \in M_n(\mathbb{C})$ und $U \in U(n)$.

1. Die Matrix A ist genau dann normal, wenn UAU^{-1} normal ist.
2. Die Matrix A ist genau dann selbstadjungiert, wenn UAU^{-1} selbstadjungiert ist.
3. Die Matrix A ist genau dann antiselbstadjungiert, wenn UAU^{-1} antiselbstadjungiert ist.
4. Die Matrix A ist genau dann unitär, wenn UAU^{-1} unitär ist.

Bemerkung 7. Für nicht-unitäre Basiswechselmatrizen gilt zu zu Lemma 5 analoge Aussage nicht notwendigerweise. Für $S \in GL_n(\mathbb{C})$ gilt allgemeiner, dass genau dann $(SAS^{-1})^* = SA^*S^{-1}$ für alle $A \in M_n(\mathbb{C})$, wenn $S = \lambda U$ für ein eine unitäre Matrix $U \in U(n)$ und einen Skalar $\lambda \in \mathbb{C}^\times$; dabei lässt sich λ durch passende Wahl von U als positiven reellen Skalar $\lambda = \sqrt{\text{tr}(SS^*)/n}$ wählen.

2 Lösungen zu Aufgabe 1

Es sei $A \in GL_n(\mathbb{R})$.

a)

Angenommen, es ist $A = \exp(B)$ für selbstadjungiertes $B \in M_n(\mathbb{R})$. Da B selbstadjungiert ist, gibt es $O \in O(n)$, so dass OBO^{-1} eine Diagonalmatrix mit reellen Diagonaleinträgen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ ist. Da B selbstadjungiert ist, ist auch $A = \exp(B)$ selbstadjungiert, also symmetrisch. Da $OAO^{-1} = O \exp(B) O^{-1} = \exp(OBO^{-1})$ eine Diagonalmatrix mit

positiven reellen Diagonaleinträgen $\exp(\lambda_1), \dots, \exp(\lambda_n)$ ist, sind alle Eigenwerte von A positiv.

Angenommen, A ist symmetrisch, also selbstadjungiert, mit positiven Eigenwerten. Dann A selbstadjungiert ist, gibt es $O \in O(n)$, so dass OAO^{-1} eine Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Nach Annahme sind diese Diagonaleinträge positiv. Für die Diagonalmatrix $D \in M_n(\mathbb{R})$ mit Diagonaleinträgen $\log(\lambda_1), \dots, \log(\lambda_n)$ gilt nun, dass $OAO^{-1} = \exp(D)$, und somit $A = O^{-1} \exp(D) O = \exp(O^{-1} D O)$. Da A selbstadjungiert ist und O normal ist, ist auch $O^{-1} A \in M_n(\mathbb{R})$ selbstadjungiert.

b)

Angenommen, es ist $A = \exp(B)$ für antiselbstadjungiertes $B \in M_n(\mathbb{R})$. Da B antiselbstadjungiert ist, ist $\exp(A) \in M_n(\mathbb{R})$ unitär, und somit orthogonal. Da $B \in M_n(\mathbb{R})$ antiselbstadjungiert ist, sind die Diagonaleinträge von B alle Null, weshalb $\text{tr } B = 0$ und somit $\det A = \det \exp(B) = \exp(\text{tr } B) = \exp(0) = 1$.

Angenommen, A ist orthogonal mit $\det A = 1$. Da A orthogonal ist, gibt es eine orthogonale Matrix $O \in O(n)$, so dass

$$OAO^{-1} = \text{block}(1, \dots, 1, -1, \dots, -1, D_{\varphi_1}, \dots, D_{\varphi_r}),$$

wobei r die Vielfachheit des Diagonaleintrags -1 sei. Da

$$1 = \det A = \det(OAO^{-1}) = 1 \cdots 1 \cdot \underbrace{(-1) \cdots (-1)}_r \det(D_{\varphi_1}) \cdots \det(D_{\varphi_r}) = (-1)^r$$

muss r gerade sein. Für $s = r/2$ ist deshalb

$$OAO^{-1} = \text{block}(1, \dots, 1, \underbrace{-I, \dots, -I}_s, D_{\varphi_1}, \dots, D_{\varphi_r}).$$

Für die Matrix

$$B := \text{block}\left(0, \dots, 0, \underbrace{\frac{\pi}{2}J, \dots, \frac{\pi}{2}J}_s, \varphi_1 J, \dots, \varphi_r J\right)$$

gilt, dass $OAO^{-1} = \exp(B)$, und deshalb $A = O^{-1} \exp(B) O = \exp(O^{-1} B O)$. Da B antiselbstadjungiert ist (denn J ist antiselbstadjungiert) und O orthogonal ist, ist auch OBO^{-1} antiselbstadjungiert.

c)

Angenommen, es ist $A = \exp(B)$ für normales $B \in M_n(\mathbb{R})$. Als normale Matrix ist B über \mathbb{C} diagonalisierbar, d.h. es gibt $S \in GL_n(\mathbb{C})$ mit $SBS^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Da B eine reelle Matrix ist, ist für jeden nicht reellen Eigenwert λ von B auch $\bar{\lambda}$ ein Eigenwert von B , und λ und $\bar{\lambda}$ haben die gleichen algebraischen (und geometrischen) Vielfachheiten. Durch passende Wahl von S ist deshalb o.B.d.A.

$$SBS^{-1} = \text{diag}(\mu_1, \dots, \mu_r, \lambda_1, \bar{\lambda}_1, \dots, \lambda_s, \bar{\lambda}_s)$$

mit $\mu_1, \dots, \mu_r \in \mathbb{R}$ und $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ nicht reell. Daher ist

$$SAS^{-1} = S \exp(B) S^{-1} = \exp(SBS^{-1}) = \text{diag}(e^{\mu_1}, \dots, e^{\mu_r}, e^{\lambda_1}, \overline{e^{\lambda_1}}, \dots, e^{\lambda_s}, \overline{e^{\lambda_s}}).$$

Dabei wird genutzt, dass $\exp(\overline{z}) = \overline{\exp(z)}$ für alle $z \in \mathbb{C}$. Die Eigenwerte $e^{\mu_1}, \dots, e^{\mu_r}$ von A sind alle positiv. Ist λ ein negativer, reeller Eigenwert von A , so ist für alle $j = 1, \dots, s$ genau dann $\lambda = e^{\lambda_j}$, wenn $\overline{e^{\lambda_j}}$. Also tritt λ ein Eigenwert mit gerader Vielfachheit.

Angenommen, $A \in M_n(\mathbb{R})$ ist normal und invertierbar, so dass alle negativen reellen Eigenwerte von A gerade Vielfachheit haben. Da A normal ist, gibt es eine orthogonale Matrix $O \in O(n)$ mit

$$\begin{aligned} OAO^{-1} &= \text{block}(\lambda_1, \dots, \lambda_p, \mu_1, \mu_1, \dots, \mu_q, \mu_q, r_1 D_{\varphi_1}, \dots, r_s D_{\varphi_s}) \\ &= \text{block}(\lambda_1, \dots, \lambda_p, \mu_1 I, \dots, \mu_q I, r_1 D_{\varphi_1}, \dots, r_s D_{\varphi_s}), \end{aligned}$$

wobei $\lambda_1, \dots, \lambda_p > 0$ die positiven reellen Eigenwerte von A sind, $\mu_1, \dots, \mu_q < 0$ die negativen reellen Eigenwerte, $r_1, \dots, r_s > 0$ Radien und $\varphi_1, \dots, \varphi_s \in \mathbb{R}$ Drehwinkel. Für die Matrix

$$\begin{aligned} B := \text{block} &\left(\log(\lambda_1), \dots, \log(\lambda_p), \log(-\mu_1)I + \frac{\pi}{2}J, \dots, \log(-\mu_q)I + \frac{\pi}{2}J, \right. \\ &\left. \log(r_1)I + \varphi_1 J, \dots, \log(r_s)I + \varphi_s J \right) \end{aligned}$$

ist $OAO^{-1} = \exp(B)$ und somit $A = O^{-1} \exp(A) O = \exp(O^{-1}AO)$.

Die Matrix B ist normal: Es ist

$$\begin{aligned} B^* = B^T &= \text{block} \left(\log(\lambda_1), \dots, \log(\lambda_p), \log(-\mu_1)I + \frac{\pi}{2}J, \dots, \log(-\mu_q)I + \frac{\pi}{2}J, \right. \\ &\left. \log(r_1)I - \varphi_1 J, \dots, \log(r_s)I - \varphi_s J \right), \end{aligned}$$

und es genügt zu überprüfen, dass die einzelnen Blöcke von B und B^* kommutieren. Für die einzelnen Skalare $\log(\lambda_j)$ mit $j = 1, \dots, p$ ist dies klar. Für die Vielfachen der Einheitsmatrix $\log(-\mu_j)$ mit $j = 1, \dots, q$ ist dies ebenfalls klar. Für die (2×2) -Blöcke $\log(r_j)I + \varphi_j J$ und $\log(r_j)I - \varphi_j J$ mit $j = 1, \dots, s$ folgt dies daraus, dass I und J kommutieren. Also kommutieren B und B^* , weshalb B normal ist. Da O orthogonal ist, ist damit auch $O^{-1}BO$ normal.