Lösungen zu Aufgabe 1, Zettel 8

Jendrik Stelzner

17. Juli 2016

1 Vorbereitung: Basiswechselmatrizen

Lemma 1. *Es sei $A \in \mathrm{M}_n(\mathbb{K})$. 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.

Die Gleichheit $A^*A = I$ ist in den Einträgen äquivalent dazu, dass $\sum_{l=1}^n \overline{a_{lj}} a_{lk} = \delta_{jk}$ für alle $j, k = 1, \dots, n$. Durch komplexe Konjugation ist dies äquivalent zu $\sum_{l=1}^n a_{lj} \overline{a_{lk}} = \delta_{j,k}$ 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 Spaltenvektoren 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. □

Im Folgenen sei

$$D(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

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

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \in \mathrm{M}_n(\mathbb{C})$$

die entsprechende Diagonalmatrix. Für Matrizen $A_1 \in \mathrm{M}_{n_1}(\mathbb{C}), \ldots, A_r \in \mathrm{M}_{n_r}(\mathbb{C})$ sei

$$\operatorname{block}(A_1, \ldots, A_r) = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix} \in \mathrm{M}_{n_1+\cdots+n_r}(\mathbb{C})$$

die entsprechende Blockdiagonalmatrix.

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

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

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

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

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

$$OAO^{-1} = \operatorname{block}(\lambda_1, \ldots, \lambda_p, r_1 D(\varphi_1), \ldots, r_q D(\varphi_q)).$$

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

6. Ist $A \in \mathrm{M}_n(\mathbb{R})$ selbstadjungiert, so gibt es eine orthogonale Matrix $O \in \mathrm{O}(n)$, so dass OAO^{-1} in Diagonalgestalt ist. Die Diagonaleinträge sind dabei bis auf Permutation eindeutig bestimmt.

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

$$OAO^{-1} = \operatorname{block}(1, \ldots, 1, -1, \ldots, -1, D(\varphi_1), \ldots, D(\varphi_r)).$$

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

Beweis. Wir betrachten den Fall, dass $A \in \mathrm{M}_n(\mathbb{C})$ normal ist. Es sei $\mathcal{B} = (e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{C}^n$ und $f\colon V \to V$ der eindeutige Endomorphismus mit $\mathrm{M}_{\mathcal{B}}(f) = A$. 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 $U \coloneqq T^{\mathcal{B}}_{\mathcal{C}}$ gilt nun, dass

$$UAU^{-1} = T^{\mathcal{B}}_{\mathcal{C}} \operatorname{M}_{\mathcal{B}}(f) T^{\mathcal{C}}_{\mathcal{B}} = \operatorname{M}_{\mathcal{C}}(f)$$

eine Diagonalmatrix ist. Die Spalten der Matrix $U^{-1} = T^{\mathcal{C}}_{\mathcal{B}}$ 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} ist somit unitär. Deshalb ist auch U unitär.

Das zeigt die erste Aussage. Die anderen Aussagen ergeben sich analog über die jeweilige Normalform der entsprechenden Endomorphismen. □

Im Folgenden seien $I, J \in \operatorname{M}_2(\mathbb{C})$ mit

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

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

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

$$\exp(\log(r)I + \theta J) = \exp(\log(r)I)\exp(\theta J) = e^{\log(r)} I \exp(\theta J) = r \exp(\theta J).$$

Aus $J^2 = -I$ ergibt sich 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 $\exp(\log(r)I + \theta J) = r\exp(\theta J) = rD(\theta)$. □

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

$$C \coloneqq \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \,\middle|\, a, b \in \mathbb{R} \right\} \subseteq \operatorname{M}_2(\mathbb{R}).$$

Die Abbildung $\Phi \colon \mathbb{C} \to C$ 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. für alle $z_1, z_2 \in \mathbb{C}$ gilt $\Phi(z_1 + z_2) = \Phi(z_1) + \Phi(z_2)$, $\Phi(z_1 \cdot z_2) = \Phi(z_1) \cdot \Phi(z_2)$ und $\Phi(1) = I$.) 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 ist $\Phi(a + ib) = aI + bJ$ für alle $a, b \in \mathbb{R}$ und $\Phi(re^{i\varphi}) = rD(\varphi)$ für alle $r \geq 0$ und $\varphi \in \mathbb{R}$.

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 auf Aussagen über die Exponentialabbildung auf $\mathbb{C}$ zurückführen.

Inbesondere übersetzt sich das Problem, einen Logarithmus einer Matrix $rD(\varphi) \in C$ zu finden, dazu, einen Logarithmus einer komplexen Zahl $re^{i\varphi}$ zu finden. Konkret ergibt sich aus $\exp(\log(r) + i\varphi) = re^{i\varphi}$, dass

$$rD(\varphi) = \Phi(re^{i\varphi}) = \Phi(\exp(\log(r) + i\varphi)) = \exp(\Phi(\log(r) + i\varphi)) = \exp(\log(r)I + \varphi J).$$

Man bemerke, dass unter diesem Blickwinkel der obige Beweis von Lemma 3 eine Übersetzung des üblichen Beweises für $\exp(i\varphi) = \cos\varphi + i\sin\varphi$ ist.

Der Vorteil an unitären (und damit auch orthogonalen) Basiswechselmatrizen besteht darin, dass sie mit dem Matrixadjungieren verträglich sind:

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

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

Korollar 6. *Es sei $A \in \mathrm{M}_n(\mathbb{C})$ und $U \in \mathrm{U}(n)$.*

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.

Allgemeiner gilt für $S \in \mathrm{GL}_n(\mathbb{C})$ genau dann $(SAS^{-1})^* = SA^*S^{-1}$ für alle $A \in \mathrm{M}_n(\mathbb{C})$, wenn $S = \lambda U$ für ein eine unitäre Matrix $U \in \mathrm{U}(n)$ und einen invertierbaren Skalar $\lambda \in \mathbb{C}^\times$. Durch passende Wahl von U lässt sich dabei λ als $\lambda = \sqrt{\mathrm{tr}(S^*S)/n}$ wählen, also als positiver reeller Skalar.

2 Lösungen zu Aufgabe 1

Es sei $A \in \mathrm{GL}_n(\mathbb{R})$.

a)

Angenommen, es ist $A = \exp(B)$ für eine selbstadjungierte Matrix $B \in \mathrm{M}_n(\mathbb{R})$. Da B selbstadjungiert ist, gibt es nach Theorem 2 eine orthogonale Basiswechselmatrix $O \in \mathrm{O}(n)$, so dass $OBO^{-1} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Da B selbstadjungiert ist, ist auch $A = \exp(B)$ selbstadjungiert, also symmetrisch. Außerdem ist

$$OAO^{-1} = O \exp(B) O^{-1} = \exp(OBO^{-1}) = \exp(\mathrm{diag}(\lambda_1, \dots, \lambda_n)) = \mathrm{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$$

mit $\exp(\lambda_1), \dots, \exp(\lambda_n) > 0$. Also sind alle Eigenwerte von A positiv.

Angenommen, A ist symmetrisch mit positiven Eigenwerten. Da A symmetrisch und reell ist, ist A selbstadjungiert. Da A selbstadjungiert ist, gibt es nach Theorem 2 eine orthogonale Basiswechselmatrix $O \in \mathrm{O}(n)$, sodass $OAO^{-1} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Nach Annahme ist dabei $\lambda_1, \dots, \lambda_n > 0$. Für $D \coloneqq \mathrm{diag}(\log(\lambda_1), \dots, \log(\lambda_n))$ ist

$$OAO^{-1} = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = \exp(\mathrm{diag}(\log(\lambda_1), \dots, \log(\lambda_n))) = \exp(D),$$

und somit $A = O^{-1} \exp(D) O = \exp(O^{-1} D O)$. Da D als reelle Diagonalmatrix selbstadjungiert ist und O normal ist, ist nach Korollar 6 auch $O^{-1} D O \in \mathrm{M}_n(\mathbb{R})$ selbstadjungiert.

b)

Angenommen, es ist $A = \exp(B)$ für antiselbstadjungiertes $B \in \mathrm{M}_n(\mathbb{R})$. Da B reell und antiselbstadjungiert ist, ist $\exp(A)$ reell und unitär, also orthogonal. Da B reell und antiselbstadjungiert ist, ist B schiefsymmetrisch. Deshalb sind sind alle Diagonaleinträge von B Null, weshalb $\mathrm{tr}\, B = 0$ und deshalb $\det A = \det \exp(B) = \exp(\mathrm{tr}\, B) = \exp(0) = 1$.

Angenommen, A ist orthogonal mit $\det A = 1$. Da A orthogonal ist, gibt es nach Theorem 2 eine orthogonale Matrix $O \in \mathrm{O}(n)$, so dass

$$OAO^{-1} = \mathrm{block}(1, \dots, 1, \underbrace{-1, \dots, -1}_{r}, D(\varphi_1), \dots, D(\varphi_r)),$$

Die Vielfachheit r des Einträges -1 muss gerade sein, da

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

Für $s = r/2$ ist deshalb

$$OAO^{-1} = \mathrm{block}(1, \dots, 1, \underbrace{-I, \dots, -I}_{s}, D(\varphi_1), \dots, D(\varphi_r)).$$

Für die Matrix

$$B \coloneqq \mathrm{block}\Big(0, \dots, 0, \underbrace{\frac{\pi}{2} J, \dots, \frac{\pi}{2} J}_{s}, \varphi_1 J, \dots, \varphi_r J\Big)$$

gilt nach Lemma 3, dass $OAO^{-1} = \exp(B)$, und somit $A = O^{-1} \exp(B) O = \exp(O^{-1} B O)$. Da B antiselbstadjungiert ist (denn J ist antiselbstadjungiert) und O orthogonal ist, ist nach Korollar 6 auch OBO^{-1} antiselbstadjungiert.

c)

Angenommen, es ist $A = \exp(B)$ für normales $B \in \mathrm{M}_n(\mathbb{R})$. Als normale Matrix ist B über $\mathbb{C}$ diagonalisierbar, d.h. es gibt $S \in \mathrm{GL}_n(\mathbb{C})$ mit $SBS^{-1} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. (Nach Theorem 2 lässt sich S orthogonal wählen, dies ist hier aber nicht nötig.) Da B eine reelle Matrix ist, ist für jeden nicht reellen Eigenwert λ von B auch $\overline{\lambda}$ ein Eigenwert von B, und λ und $\overline{\lambda}$ haben die gleichen algebraischen (und geometrischen) Vielfachheiten. Durch passende Wahl von S ist deshalb o.B.d.A.

$$SBS^{-1} = \mathrm{diag}(\mu_1, \dots, \mu_r, \lambda_1, \overline{\lambda_1}, \dots, \lambda_s, \overline{\lambda_s})$$

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

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

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 $\lambda = \overline{e^{\lambda_j}} = e^{\overline{\lambda_j}}$. Also ist λ ein Eigenwert mit gerader Vielfachheit.

Angenommen, $A \in \mathrm{M}_n(\mathbb{R})$ ist normal und invertierbar, so dass alle negativen reellen Eigenwerte von A gerade Vielfachheit haben. Da A reell und normal ist, gibt es nach Theorem 2 eine orthogonale Matrix $O \in \mathrm{O}(n)$ mit

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

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

$$\begin{aligned}
B := \mathrm{block}\Big(&\log(\lambda_1), \dots, \log(\lambda_p), \log(|\mu_1|)I + \pi J, \dots, \log(|\mu_q|)I + \pi J, \\
&\log(r_1)I + \varphi_1 J, \dots, \log(r_s)I + \varphi_s J\Big)
\end{aligned}$$

gilt nach Lemma 3, dass $OAO^{-1} = \exp(B)$ und somit $A = O^{-1}\exp(B)O = \exp(O^{-1}BO)$.

Die Matrix B ist normal: Es ist

$$\begin{aligned}
B^* = B^T = \mathrm{block}\Big(&\log(\lambda_1), \dots, \log(\lambda_p), \log(|\mu_1|)I - \pi J, \dots, \log(|\mu_q|)I - \pi J, \\
&\log(r_1)I - \varphi_1 J, \dots, \log(r_s)I - \varphi_s J\Big),
\end{aligned}$$

und es genügt zu überprüfen, dass die einzelnen Blöcke von B und B^* jeweils miteinander kommutieren.

Für die (1×1)-Blöcke $\log(\lambda_j)$ mit $j = 1, \dots, p$ ist dies klar. Für die (2×2)-Blöcke $\log(|\mu_j|)I + \pi J$ und $\log(|\mu_j|)I - \pi J$ mit $j = 1, \dots, q$, sowie auch für die 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 nach Korollar 6 damit auch $O^{-1}BO$ normal.