Musterlösung Serie 11

Äusseres Produkt, Singulärwertzerlegung

1. Seien $F\colon V \to W$ und $G\colon W \to U$ lineare Abbildungen von endlichdimensionalen K-Vektorräumen und $r \geqslant 0$. Zeigen Sie:

   (a) Die Zuordnung

   $$\Lambda^r F\colon \Lambda^r V \to \Lambda^r W\,, v_1 \wedge \cdots \wedge v_r \mapsto F(v_1) \wedge \cdots \wedge F(v_r)$$

   definiert eine lineare Abbildung.

   (b) Es gilt die Funktorialität $\Lambda^r(G \circ F) = \Lambda^r G \circ \Lambda^r F$.

   (c) F ist injektiv genau dann, wenn $\Lambda^r F$ ungleich Null ist für $r = \dim(V)$.

   (d) F ist surjektiv genau dann, wenn $\Lambda^r F$ ungleich Null ist für $r = \dim(W)$.

   Lösung: In der nachfolgenden Lösung für (c) und (d) kommen zwei verschiedene Ideen zum Tragen, die auch auf die jeweils andere Teilaufgabe angewendet werden können.

   (a) Man rechnet direkt nach, dass die Zuordnung

   $$(v_1, \ldots, v_r) \mapsto F(v_1) \wedge \cdots \wedge F(v_r)$$

   eine alternierende multilineare Abbildung definiert und daher nach universeller Eigenschaft eine lineare Abbildung

   $$\Lambda^r F\colon \Lambda^r V \to \Lambda^r W$$

   induziert.

   (b) Die Aussage ist offensichtlich für Elemente der Form $v_1 \wedge \cdots \wedge v_r \in \Lambda^r V$. Das genügt bereits, da der Vektorraum $\Lambda^r V$ von solchen Elementen erzeugt wird.

   (c) Es sei $r = \dim(V)$ und $b_1, \ldots, b_r$ eine Basis von V. Wir wissen, dass dann $\Lambda^r V$ eindimensional ist und von dem Element $b_1 \wedge \ldots \wedge b_n$ erzeugt wird. Nach Definition von $\Lambda^r f$ haben wir

   $$\Lambda^r f(b_1 \wedge \ldots \wedge b_n) = f(b_1) \wedge \ldots \wedge f(b_n).$$

   Somit ist $\Lambda^r f$ ungleich Null genau dann, wenn $f(b_1) \wedge \ldots \wedge f(b_n) \neq 0$ ist. Das ist äquivalent dazu, dass die Vektoren $f(b_1), \ldots, f(b_n)$ linear unabhängig sind, was wiederum äaquivalent dazu ist, dass f injektiv ist. Damit ist die Behauptung bewiesen.

(d) Es sei $r = \dim(W)$. Dann ist $\Lambda^r W$ eindimensional. Wir beweisen die beiden Richtungen der Behauptung separat.

Nehmen wir zunächst an, dass f surjektiv ist. Dann existiert eine lineare Abbildung $g\colon W \to V$ mit $f \circ g = \mathrm{id}_W$. Nach der Funktorialität der äusseren Potenz gilt folglich

$$\Lambda^r f \circ \Lambda^r g \;=\; \Lambda^r(f \circ g) \;=\; \Lambda^r(\mathrm{id}_W) \;=\; \mathrm{id}_{\Lambda^r W}.$$

Wegen $\Lambda^r W \neq 0$ ist diese Abbildung ungleich Null und daher auch $\Lambda^r f$.

Nehmen wir umgekehrt an, dass f nicht surjektiv ist. Dann ist $W' := \mathrm{Im}(f)$ ein Unterraum von W der Dimension $< r$. Somit ist $\Lambda^r W' = 0$. Für die Inklusionsabbildung $i\colon W' \subset W$ ist daher auch die induzierte lineare Abbildung $\Lambda^r i\colon \Lambda^r W' \to \Lambda^r W$ gleich Null. Ausserdem ist $f = i \circ f'$ für einen Homomorphismus $f'\colon V \to W'$. Nach der Funktorialität des äusseren Potenz ist daher

$$\Lambda^r f \;=\; \Lambda^r(i \circ f') \;=\; \Lambda^r i \circ \Lambda^r f',$$

und aus $\Lambda^r i = 0$ folgt nun auch $\Lambda^r f = 0$, wie gewünscht.

2. Ein Element $\alpha \in \Lambda^k V$ heißt *rein*, falls Vektoren $w_1, \ldots, w_k \in V$ existieren mit $\alpha = w_1 \wedge \ldots \wedge w_k$. Zeigen Sie:

   (a) Für jedes $\alpha \in \Lambda^2 V$ existiert eine Basis $\{b_i\}_{i \in I}$ von V und ein $k \geqslant 0$ mit $\{1, \ldots, 2k\} \subset I$ und $\alpha = b_1 \wedge b_2 + \ldots + b_{2k-1} \wedge b_{2k}$.

   (b) Ist $2 \neq 0$ in K, so ist $\alpha \in \Lambda^2 V$ genau dann rein, wenn $\alpha \wedge \alpha = 0$ ist in $\Lambda^4 V$.

Lösung:

(a) Es sei $\alpha \in \Lambda^2 V$ und $\{b_i\}_{i \in I}$ eine Basis von V. Dann ist α eine endliche Linearkombination von Vektoren der Form $b_i \wedge b_j$. Nach geeigneter Wahl der Indexmenge I können wir ohne Beschränkung der Allgemeinheit annehmen, dass

$$\alpha \;=\; \sum_{1 \leqslant i < j \leqslant n} a_{ij}\, b_i \wedge b_j \tag{1}$$

ist für ein $0 \leqslant n < \infty$ und für Koeffizienten $a_{ij} \in K$. Es genügt also zu zeigen:

Behauptung. Für jede Wahl linear unabhängiger Vektoren $b_1, \ldots, b_n \in V$ und für jedes $\alpha \in \Lambda^2 V$ der Form (1) existiert eine Basis $(\tilde b_1, \ldots, \tilde b_n)$ des Unterraums $\langle b_1, \ldots, b_n \rangle$ und ein $0 \leqslant k \leqslant n/2$ mit

$$\alpha \;=\; \tilde b_1 \wedge \tilde b_2 + \ldots + \tilde b_{2k-1} \wedge \tilde b_{2k}\,.$$

Beweis. Wir verwenden Induktion über n. Im Fall $n = 0$ oder $\alpha = 0$ gilt die Aussage mit $k = 0$. Sei andernfalls $n > 0$ und $\alpha \neq 0$, und die Behauptung gelte für alle $n' < n$. Nach einer etwaigen Umordnung der Vektoren $b_1, \ldots, b_n$ können wir $a_{12} \neq 0$ annehmen. Setze

$$\begin{aligned} \tilde{b}_1 &:= a_{12} \cdot (b_1 - \textstyle\sum_{2<j\leqslant n} \frac{a_{2j}}{a_{12}} \cdot b_j) \qquad \text{und} \\ \tilde{b}_2 &:= b_2 + \textstyle\sum_{2<j\leqslant n} \frac{a_{1j}}{a_{12}} \cdot b_j. \end{aligned}$$

Dann gilt

$$\begin{aligned} \tilde{b}_1 \wedge \tilde{b}_2 &= a_{12} b_1 \wedge b_2 + \sum_{2<j\leqslant n} a_{1j} b_1 \wedge b_j - \sum_{2<j\leqslant n} a_{2j} b_j \wedge b_2 - \sum_{2<j,j'\leqslant n}^{n} \frac{a_{2j} a_{1j'}}{a_{12}} b_j \wedge b_{j'} \\ &= \sum_{2\leqslant j\leqslant n} a_{1j} b_1 \wedge b_j + \sum_{2<j\leqslant n} a_{2j} b_2 \wedge b_j - \sum_{2<j,j'\leqslant n}^{n} \frac{a_{2j} a_{1j'}}{a_{12}} \cdot b_j \wedge b_{j'}. \end{aligned}$$

Folglich gibt es Koeffizienten $c_{ij} \in K$ mit

$$\alpha = \tilde{b}_1 \wedge \tilde{b}_2 + \sum_{2<i<j\leqslant n} c_{ij} \cdot b_i \wedge b_j.$$

Durch Anwenden der Induktionsvoraussetzung auf die linear unabhängigen Vektoren $b_3, \ldots, b_n$ und das Element $\sum_{2<i<j\leqslant n} c_{ij} \cdot b_i \wedge b_j$ erhalten wir eine Basis $\tilde{b}_3, \ldots, \tilde{b}_n$ von $\langle b_3, \ldots, b_n \rangle$ und eine Zahl $1 \leqslant k \leqslant n/2$ mit

$$\sum_{2<i<j\leqslant n} c_{ij} \cdot b_i \wedge b_j = \tilde{b}_3 \wedge \tilde{b}_4 + \ldots + \tilde{b}_{2k-1} \wedge \tilde{b}_{2k}.$$

Damit hat

$$\alpha = \tilde{b}_1 \wedge \tilde{b}_2 + \tilde{b}_3 \wedge \tilde{b}_4 + \ldots + \tilde{b}_{2k-1} \wedge \tilde{b}_{2k}$$

die gesuchte Form. Ausserdem zeigt die Konstruktion von $\tilde{b}_1$ und $\tilde{b}_2$, dass $(\tilde{b}_1, \tilde{b}_2, b_3, \ldots, b_n)$ eine Basis von $\langle b_1, \ldots, b_n \rangle$ ist, und somit gilt dasselbe auch für $(\tilde{b}_1, \tilde{b}_2, \tilde{b}_3, \ldots, \tilde{b}_n)$. Damit ist die Behauptung bewiesen. $\square$

(b) Falls α rein ist, so gibt es Vektoren $w_1, w_2 \in V$ mit $\alpha = w_1 \wedge w_2$. Nach der Definition von alternierenden Abbildungen ist dann

$$\alpha \wedge \alpha = w_1 \wedge w_2 \wedge w_1 \wedge w_2 = 0.$$

Umgekehrt sei $\alpha \in \Lambda^2 V$ mit $\alpha \wedge \alpha = 0$. Nach (a) existiert eine Basis $\{b_i\}_{i\in I}$ von V mit

$$\alpha = b_1 \wedge b_2 + \ldots + b_{2k-1} \wedge b_{2k}$$

für ein $k \geqslant 0$. Dann ist

$$0 = \alpha \wedge \alpha = \sum_{i=1}^{k} \sum_{j=1}^{k} b_{2i-1} \wedge b_{2i} \wedge b_{2j-1} \wedge b_{2j}.$$

Hier ist der Summand $b_{2i-1} \wedge b_{2i} \wedge b_{2j-1} \wedge b_{2j}$ gleich Null für $i = j$, weil er dann zweimal denselben Faktor b_{2i} enthält. Andernfalls ist er invariant unter Vertauschung von i und j, weil diese Vertauschung durch die Transpositionen $2i \leftrightarrow 2j$ und $2i-1 \leftrightarrow 2j-1$ erreicht wird, die beide ein Vorzeichenwechsel bewirken. Insgesamt folgt daraus

$$0 \; = \; 2 \cdot \sum_{0<i<j\leqslant k} b_{2i-1} \wedge b_{2i} \wedge b_{2j-1} \wedge b_{2j}.$$

Da $2 \neq 0$ ist und die Vektoren $b_{2i-1} \wedge b_{2i} \wedge b_{2j-1} \wedge b_{2j}$ für alle $i < j$ linear unabhängig sind, muss die Summe leer, also $k \leqslant 1$ sein. Im Fall $k = 0$ ist $\alpha = 0 \wedge 0$ und im Fall $k = 1$ ist $\alpha = b_1 \wedge b_2$; in beiden Fällen ist α also rein.

3. Zeigen Sie, dass für jedes $\varepsilon \in \{0, 1\}$ durch die Zuordnung

$$\begin{aligned} \rho\colon S_n &\to \mathrm{End}(\textstyle\bigotimes^n V)\,, \\ \sigma &\mapsto \big(v_1 \otimes \cdots \otimes v_n \mapsto \mathrm{sign}(\sigma)^\varepsilon v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}\big) \end{aligned}$$

eine Darstellung von S_n auf dem Vektorraum $\bigotimes^n V$ definiert wird.

Lösung:

Da $\mathrm{sign(id)} = 1$ ist in beiden Fällen $\rho(\mathrm{id}) = \mathrm{id}$. Seien $\tau, \sigma \in S_n$ und $v_1, \ldots, v_n \in V$. Weiter sei $u_j = v_{\sigma^{-1}j}$, dann ist

$$\begin{aligned} \rho(\tau)\big(\rho(\sigma)(v_1 \otimes \cdots \otimes v_n)\big) &= \rho(\tau)\big(\mathrm{sign}(\sigma)(u_1 \otimes \cdots \otimes u_n)\big) \\ &= \mathrm{sign}(\tau)\mathrm{sign}(\sigma)(u_{\tau^{-1}1} \otimes \cdots \otimes u_{\tau^{-1}n}) \\ &= \mathrm{sign}(\tau \circ \sigma)(v_{\sigma^{-1}\tau^{-1}1} \otimes \cdots \otimes v_{\sigma^{-1}\tau^{-1}n}) \\ &= \mathrm{sign}(\tau \circ \sigma)(v_{(\tau\circ\sigma)^{-1}1} \otimes \cdots \otimes v_{(\tau\circ\sigma)^{-1}n}) \\ &= \rho(\tau \circ \sigma)(v_1 \otimes \cdots \otimes v_n) \end{aligned}$$

und daher $\rho(\tau \circ \sigma) = \rho(\tau) \circ \rho(\sigma)$.

4. Bestimmen Sie die Singulärwertzerlegung $A = UDV^\dagger$ der komplexen Matrix

$$A = \frac{1}{2}\begin{pmatrix} 3+2i & i & 0 \\ -i & 3+2i & 0 \\ 0 & 0 & 2i \end{pmatrix},$$

wobei $(\;)^\dagger = \overline{(\;)}^T$.

Lösung: Wir berechnen $A^\dagger A$ und erhalten

$$A^\dagger A = \frac{1}{4}\begin{pmatrix} 14 & 6i & 0 \\ -6i & 14 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Wegen des Eintrags unten rechts besitzt $A^\dagger A$ einen Eigenwert 1. Um die anderen Eigenwerte zu bestimmen berechnen wir das charakteristische Polynom der oberen linken 2×2-Matrix. Dieses ist $(X - \frac{14}{4})(X - \frac{14}{4}) + \frac{(6i)^2}{16} = X^2 - 7X + 10 = (X-5)(X-2)$. Also hat $A^\dagger A$ die Eigenwerte 5, 2 und 1. Daraus schliessen wir, dass A die Singulärwerte $\sqrt{5}$, $\sqrt{2}$ und 1 besitzt und die Matrix D gleich

$$D = \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ist. Die normierten Eigenvektoren von $A^\dagger A$ zu den Eigenwerten 5 bzw. 2 bzw. 1 sind

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \text{ bzw. } u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \text{ bzw. } u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Diese Eigenvektoren sind die Spalten der Matrix V, also ist

$$V^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Matrix U finden wir durch Lösen der Matrixgleichung $AV = UD$. Wir finden

$$U = \begin{pmatrix} \frac{2+i}{\sqrt{10}} & \frac{1+i}{2} & 0 \\ \frac{1-2i}{\sqrt{10}} & \frac{i-1}{2} & 0 \\ 0 & 0 & i \end{pmatrix}.$$

Die Singulärwertzerlegung von A ist demnach

$$A = \frac{1}{2} \begin{pmatrix} 3+2i & i & 0 \\ -i & 3+2i & 0 \\ 0 & 0 & 2i \end{pmatrix} = \begin{pmatrix} \frac{2+i}{\sqrt{10}} & \frac{1+i}{2} & 0 \\ \frac{1-2i}{\sqrt{10}} & \frac{i-1}{2} & 0 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

5. Sei $A \in \mathrm{M}(n \times n, \mathbb{C})$ und $A = UDV^\dagger$ die Singulärwertzerlegung. Zeigen Sie, dass $UV^\dagger \in U(n)$ die in der Frobeniusnorm optimale Approximation von A durch eine unitäre Matrix ist, also für alle $X \in U(n)$ gilt:

$$\|X - A\|_F \geqslant \|UV^\dagger - A\|_F.$$

Lösung:

Der Beweis folgt dem reellen Fall (siehe Satz 2.8 Zusatzskript). Zunächst bemerkt man, dass wir $A = D$ annehmen können, da

$$\|X - A\|_F = \|U^\dagger (X - A) V\|_F = \|X' - D\|_F,$$

mit $X' = U^\dagger X V \in U(n)$. Wir zeigen also, dass $E \in U(n)$ die beste Approximation von D unter allen $X \in U(n)$ ist. Sei $Re(U)$ der Realteil der Matrix U. Die Frobeniusnorm von $X - D$ berechnet sich als (Erinnerung: D ist reell)

$$\begin{aligned} \|X - D\|_F &= \mathrm{tr}(X^\dagger X) - \mathrm{tr}(DX^\dagger) - \mathrm{tr}(XD) + \mathrm{tr}(D^2) \\ &= n - 2\mathrm{tr}(Re(U)D) + \mathrm{tr}(D^2)\,, \end{aligned}$$

wobei wir $\mathrm{tr}(DX^\dagger) = \mathrm{tr}(\overline{X}D)$ verwendet haben. Wie im rellen Fall verwenden wir nun, dass für alle $X \in U(n)$ und jeden Eintrag x_{ij} von $Re(U)$ stets $x_{ij} \leqslant 1$ ist (aus der Dreiecksungleichung). Folglich ist

$$\mathrm{tr}(Re(U)D) \leqslant \mathrm{tr}(D)\,,$$

und somit $\|X - D\|_F \geqslant \|E - D\|_F$.