Musterlösung Serie 19

SCALAR PRODUCTS, BILINEAR FORMS

1. Für welche Werte $a \in \mathbb{R}$ ist für $x = \binom{x_1}{x_2}$ und $y = \binom{y_1}{y_2}$ durch

$$\langle x, y\rangle := x_1y_1 + ax_1y_2 + ax_2y_1 + 7x_2y_2$$

ein Skalarprodukt auf $\mathbb{R}^2$ definiert?

Lösung: Man prüft direkt, dass $\langle\cdot,\cdot\rangle$ eine symmetrische Bilinearform auf $\mathbb{R}^2$ ist; nämlich mit der symmetrischen Darstellungsmatrix $\left(\begin{smallmatrix}1 & a\\ a & 7\end{smallmatrix}\right)$. Sodann berechnen wir:

$$\begin{aligned}\langle x, x\rangle &= x_1^2 + 2ax_1x_2 + 7x_2^2\\ &= x_1^2 + 2ax_1x_2 + a^2x_2^2 - a^2x_2^2 + 7x_2^2\\ &= (x_1 + ax_2)^2 + (7 - a^2)x_2^2.\end{aligned}$$

Ist $\langle\cdot,\cdot\rangle$ positiv definit, so folgt mit $x = \binom{-a}{1} \neq 0$, dass $7 - a^2 > 0$ sein muss; also $|a| < \sqrt{7}$. Ist umgekehrt $|a| < \sqrt{7}$, so gilt wegen der obigen Rechnung $\langle x, x\rangle > 0$ für alle $x \neq 0$ und daher ist $\langle\cdot,\cdot\rangle$ positiv definit. Somit ist $\langle\cdot,\cdot\rangle$ ein Skalarprodukt ist genau dann, wenn $|a| < \sqrt{7}$ gilt.

2. Sei V der Vektorraum der reellen Polynome vom Grad $\leqslant n$.

   (a) Zeige, dass durch

   $$\langle p, q\rangle := \int_0^\infty p(t)q(t)e^{-t}\,dt$$

   ein Skalarprodukt auf V definiert wird.

   (b) Bestimme die Matrix des Skalarprodukts bezüglich der Basis $1, x, \ldots, x^n$.

Lösung:

   (a) Zunächst wissen wir aus der Analysis, dass das uneigentliche Integral konvergiert. Sodann zeigt man direkt aus den Linearitätseigenschaften des Integrals, dass $\langle\cdot,\cdot\rangle$ eine Bilinearform ist. Offensichtlich ist sie symmetrisch. Sei nun $p \in V \smallsetminus \{0\}$ beliebig. Wähle einen Punkt $x_0 > 0$ mit $p(x_0) \neq 0$. Da p eine stetige Funktion induziert, gilt dann $|p(x)| \geqslant c := \frac{1}{2}|p(x_0)| > 0$ auf einem ganzen Intervall $[a, b] \subset \mathbb{R}^{\geqslant 0}$ mit $x_0 \in [a, b]$. Da die Funktion $t \mapsto e^{-t}$ streng monoton fallend und $e^{-t} > 0$ für alle t ist, folgt

   $$\langle p, p\rangle \;=\; \int_0^\infty p(t)^2e^{-t}\,dt \;\geqslant\; \int_a^b p(t)^2e^{-t}\,dt \;\geqslant\; c\cdot e^{-b}\cdot(b-a) \;>\; 0.$$

   Damit ist $\langle\cdot,\cdot\rangle$ positiv definit und daher ein Skalarprodukt.

(b) Für alle $k \in \mathbb{Z}^{\geqslant 0}$ sei

$$a(k) := \int_0^\infty t^k e^{-t}\, dt.$$

Dann ist

$$a(0) = \int_0^\infty e^{-t}\, dt = -e^{-t}\Big|_0^\infty = 1,$$

und für $k \geqslant 1$ gilt nach partieller Integration

$$a(k) = \int_0^\infty t^k e^{-t}\, dt = -t^k e^{-t}\Big|_0^\infty - \int_0^\infty kt^{k-1}(-e^{-t})\, dt = k \cdot a(k-1).$$

Durch Induktion über k folgt daraus $a(k) = k!$. Sei schliesslich $A := (a_{ij})_{i,j}$ die Darstellungsmatrix des Skalarproduktes $\langle \cdot, \cdot \rangle$ bezüglich der geordneten Basis $(1, x, \ldots, x^n)$. Dann ist

$$a_{ij} = \langle x^{i-1}, x^{j-1} \rangle = a(i+j-2) = (i+j-2)!\,.$$

3. Sei $V = \mathbb{R}^2$ mit dem Standard-Skalarprodukt ausgestattet und für $i = 1, 2$ sei $v_i \in V \setminus \{0\}$. Zeigen Sie, dass die Formel

$$\langle v_1, v_2 \rangle = ||v_1||\, ||v_2|| \cos(\widehat{v_1, v_2}),$$

die den Cosinus eines Winkels definiert, drehinvariant ist. Anders ausgedrückt, zeigen Sie, dass für jede Drehung der Ebene $R : V \to V$ gilt

$$\cos(\widehat{v_1, v_2}) = \cos(\widehat{Rv_1, Rv_2}),$$

was wir von einer guten Definition des Winkels zwischen 2 Vektoren erwarten würden.

Lösung: Sei $v_1 = (x_1, y_1)^T$, $v_2 = (x_2, y_2)^T$. Schreibe θ für den Winkel von v_1 zur horizontalen Achse. Wir erhalten

$$\begin{aligned} \cos(\theta) &= \frac{x_1}{||v_1||} \\ \sin(\theta) &= \frac{y_1}{||v_1||}. \end{aligned}$$

Die Matrix $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$ rotiert v_1 auf den Vektor $(1, 0)$ und rotiert v_2 um den gleichen Winkel. Dies erhält den Winkel $\widehat{v_1, v_2}$ und lässt das Skalarprodukt unverändert. Tatsächlich ist

$$\begin{aligned}
&\left\langle \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} v_1, \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} v_2 \right\rangle \\
=&\left(\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \cdot v_1 \right)^T \cdot \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \cdot v_2 \\
=&v_1^T \cdot \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}^T \cdot \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \cdot v_2 \\
=&v_1^T \cdot I_2 \cdot v_2 \\
=&\langle v_1, v_2 \rangle .
\end{aligned}$$

Somit können wir ohne Beschränkung der Allgemeinheit $v_1 = (1, 0)$ annehmen. In diesem Fall gilt einerseits

$$\left\langle v_1, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right\rangle = x_2$$

und andererseits

$$\left|\left| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|\right| \, ||v_2|| \cos(\widehat{v_1, v_2}) = ||v_2|| \, \frac{x_2}{||v_2||} = x_2.$$

Daraus folgt die gewünschte Formel.

4. Sei $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. Zeige:

   (a) Die Matrix $A^T A$ ist symmetrisch.

   (b) Die Matrix $A^T A$ ist positiv definit genau dann, wenn A invertierbar ist.

   (c) Es gilt $\mathrm{Rang}(A^T A) = \mathrm{Rang}(A)$.

Lösung: Aussage (a) folgt aus der Rechnung $(A^T A)^T = A^T (A^T)^T = A^T A$.

Für den Rest berechnen wir zur Vorbereitung für jedes $v \in \mathbb{R}^n$

$$(*) \qquad v^T \cdot (A^T A) \cdot v \;=\; (v^T A^T) \cdot Av \;=\; (Av)^T Av \;=\; \|Av\|^2,$$

wobei $|| \; ||$ die übliche euklidische Norm auf $\mathbb{R}^n$ ist.

Ist nun A invertierbar, so folgt $Av \neq 0$ und somit $v^T \cdot (A^T A) \cdot v = \|Av\|^2 > 0$; also ist $A^T A$ positiv definit. Ist umgekehrt $A^T A$ positiv definit, so folgt $\|Av\|^2 = v^T \cdot (A^T A) \cdot v > 0$ und daraus $Av \neq 0$. Also hat die lineare Abbildung $L_A \colon \mathbb{R}^n \to \mathbb{R}^n$, $v \mapsto Av$ den Kern Null, und daher ist A invertierbar. Damit ist (b) bewiesen.

Für (c) behaupten wir zunächst: $\mathrm{Kern}(L_A) = \mathrm{Kern}(L_{A^T A})$.

Beweis: Für alle $v \in \mathbb{R}^n$ mit $Av = 0$ gilt auch $A^T Av = A^T 0 = 0$; also haben wir die Inklusion „$\subset$“. Umgekehrt sei $A^T Av = 0$. Aus $(*)$ folgt dann $\|Av\|^2 =$

$v^T \cdot (A^T A) \cdot v = v^T 0 = 0$. Da die euklidische Norm positiv definit ist, folgt daraus $Av = 0$, also $v \in \mathrm{Kern}(L_A)$. Dies zeigt die Inklusion „$\supset$“, und wir sind fertig. *q.e.d.*

Aus der Behauptung folgt schliesslich

$$\begin{aligned} \mathrm{Rang}(A) &= \dim \mathrm{Bild}(L_A) = n - \dim \mathrm{Kern}(L_A) \\ &= n - \dim \mathrm{Kern}(L_{A^T A}) = \dim \mathrm{Bild}(L_{A^T A}) = \mathrm{Rang}(A^T A). \end{aligned}$$

5. (a) Sei $\| \cdot \|$ eine Norm auf dem $\mathbb{R}$-Vektorraum V. Zeige, dass die Norm genau dann von einem Skalarprodukt $\langle \cdot, \cdot \rangle$ auf V induziert wird, wenn sie für alle $x, y \in V$ die *Parallelogrammidentität*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

erfüllt.

(b) Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum. Betrachte die folgende Abbildung:

$$\begin{array}{rccc} ||\cdot||_1 : & V & \to & \mathbb{R}_{\geqslant 0} \\ & v = (v_1, v_2, \ldots, v_n) & \mapsto & \sum_{i=1}^n |v_i| \end{array}$$

Zeige, dass $||\cdot||_1$ eine Norm auf V definiert, und beweise, dass sie nicht von einem Skalarprodukt herrührt.

Lösung:

(a) Im Fall $||x||^2 = \langle x, x \rangle$ gilt

$$\begin{aligned} ||x + y||^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &= ||x||^2 + 2\langle x, y \rangle + ||y||^2 \\ ||x - y||^2 &= \langle x - y, x - y \rangle \\ &= \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle \\ &= ||x||^2 - 2\langle x, y \rangle + ||y||^2 , \end{aligned}$$

also ist

$$||x + y||^2 + ||x - y||^2 = 2\,||x||^2 + 2\,||y||^2 \ .$$

Sei umgekehrt $||\cdot||$ eine Norm auf V, welche die Parallelogrammidentität erfüllt. Definiere

$$\langle x, y \rangle := \frac{1}{2}\big(||x + y||^2 - ||x||^2 - ||y||^2\big) = \frac{1}{4}\big(||x + y||^2 - ||x - y||^2\big),$$

wobei die letzte Gleichung aus der Parallelogrammidentität (PI) folgt. Diese Abbildung erfüllt $\langle x, y \rangle = \langle y, x \rangle$ und $||x||^2 = \langle x, x \rangle$ für alle $x, y \in V$. Da $||\cdot||$ bereits positiv definit ist, bleibt also nur noch zu zeigen, dass $\langle \cdot, \cdot \rangle$ bilinear

ist. Aufgrund der Symmetrie genügt es zu zeigen, dass die Abbildung linear in der ersten Variablen ist.
Für beliebige $x, x', y \in V$ berechnen wir

$$
\begin{aligned}
\langle x + x', y\rangle &= \frac{1}{4}(||(x+y) + x'||^2 - ||x + x' - y||^2) \\
&\overset{\text{(PI)}}{=} \frac{1}{4}(2\,||x+y||^2 + 2\,||x'||^2 - ||x + y - x'||^2 - ||x + x' - y||^2) \\
&\overset{\text{(PI)}}{=} \frac{1}{4}(2\,||x+y||^2 + 2\,||x'||^2 - 2\,||x||^2 - 2\,||x' - y||^2) \\
&= \frac{1}{4}(2\,||x+y||^2 - 2\,||x||^2 - 2\,||x'||^2 + 4\,||x'||^2 - 2\,||x' - y||^2) \\
&\overset{\text{(PI)}}{=} \frac{1}{4}(2\,||x+y||^2 - 2\,||x||^2 - 2\,||x'||^2 + 2\,||x' + y||^2 - 4\,||y||^2) \\
&= \frac{1}{2}(||x+y||^2 - ||x||^2 - ||y||^2 + ||x' + y||^2 - ||x'||^2 - ||y||^2) \\
&= \langle x, y\rangle + \langle x', y\rangle.
\end{aligned}
$$

Also ist die Abbildung additiv in der ersten Variablen. Das Verhalten unter skalarer Multiplikation können wir nur indirekt erschliessen. Zunächst zeigen wir die folgenden Aussagen für alle $n \in \mathbb{Z}^{>0}$ und alle $x, y \in V$:

(i) Wegen
$$0 = \langle 0, y\rangle = \langle x + (-x), y\rangle = \langle x, y\rangle + \langle -x, y\rangle$$
gilt $\langle -x, y\rangle = -\langle x, y\rangle$.

(ii) Aus der Additivität folgt durch Induktion $\langle nx, y\rangle = n\langle x, y\rangle$.

(iii) Aus (ii) mit $\frac{1}{n}x$ anstelle von x folgt
$$\langle x, y\rangle = \langle n\frac{1}{n}x, y\rangle = n\langle \frac{1}{n}x, y\rangle,$$
und daher $\langle \frac{1}{n}x, y\rangle = \frac{1}{n}\langle x, y\rangle$.

Aus allen drei Aussagen zusammen folgt nun
$$\langle \frac{p}{q} \cdot x, y\rangle = \frac{p}{q} \cdot \langle x, y\rangle$$
für alle $\frac{p}{q} \in \mathbb{Q}$ und alle $x, y \in V$.
Nun fixieren wir beliebige $x, y \in V$. Der von diesen erzeugte Unterraum U ist dann isomorph zu $\mathbb{R}^n$ für ein $n \leqslant 2$. Die Einschränkung von $||\cdot||$ auf U ist wieder eine Norm und entspricht daher einer Norm auf $\mathbb{R}^n$. Wie in der Vorlesung erklärt, ist diese Norm nun aber eine Lipschitz-stetige Funktion auf $\mathbb{R}^n$. Insbesondere ist sie also stetig. Daraus folgt, dass auch die Einschränkung von $\langle\cdot,\cdot\rangle$ auf $U \times U$ einer stetigen Funktion auf $\mathbb{R}^n \times \mathbb{R}^n$ entspricht. Dies impliziert wiederum, dass die Funktion
$$\mathbb{R} \to \mathbb{R},\ t \mapsto \langle tx, y\rangle - t\langle x, y\rangle$$

stetig ist. Nun zeigt aber die obige Gleichung, dass diese Funktion auf $\mathbb{Q}$ verschwindet. Da sie stetig ist, verschwindet sie daher auf ganz $\mathbb{R}$. Es gilt also

$$\langle tx, y\rangle = t\langle x, y\rangle$$

für alle $t \in \mathbb{R}$. Also ist $\langle\cdot,\cdot\rangle$ linear in der ersten Variable, und wir sind fertig.

(b) Um zu zeigen, dass $||\cdot||_1$ nicht durch ein Skalarprodukt induziert wird, genügt es ein Paar von Vektoren zu finden, für welches $||\cdot||_1$ nicht die Parallelogrammgleichung erfüllt. Seien $x = (1, 0, 0, \ldots, 0)^T$, $y = (0, 1, 0, \ldots, 0)^T$. Dann gilt

$$\begin{aligned} ||x + y||_1^2 &= (1 + 1)^2 = 4 \\ ||x - y||_1^2 &= (1 + 1)^2 = 4 \\ 2\,||x||_1^2 &= 2 \\ 2\,||y||_1^2 &= 2. \end{aligned}$$

Also ist die Parallelogrammgleichung verletzt.

6. Sei $K = \mathbb{R}$, $V = M_{n\times n}(K)$ und betrachte die Abbildung

$$\begin{array}{ccc} V \times V & \to & K \\ (A, B) & \mapsto & \mathrm{Tr}(A^T B). \end{array}$$

Zeige, dass sie ein Skalarprodukt auf V definiert und finde eine orthonormale Basis bezüglich dieses Skalarprodukts. Die induzierte Norm wird als *Hilbert-Schmidt-Norm* bezeichnet. Gib eine Formel für die Norm einer Matrix $A \in V$ in Bezug auf ihre Einträge an.

Lösung: Die Grundregeln der Matrixmultiplikation implizieren, dass die Abbildung $(A, B) \mapsto A^T B$ bilinear ist. Da die Spurabbildung linear ist, ist die angegebene Abbildung also bilinear. Wegen

$$\langle A, B\rangle = \mathrm{Tr}(A^T B) = \mathrm{Tr}((A^T B)^T) = \mathrm{Tr}(B^T (A^T)^T) = \mathrm{Tr}(B^T A) = \langle B, A\rangle$$

ist sie ausserdem symmetrisch. Sodann schreiben wir $A = (v_1, \ldots, v_n)$ mit Spaltenvektoren v_i. Dann ist A^T die Matrix mit den Zeilen $v_1^T, \ldots, v_n^T$ und folglich

$$A^T A \;=\; \big(v_i^T v_j\big)_{i,j=1,..,n}.$$

Die Spur ist definiert als die Summe der Diagonaleinträge; darum gilt also

$$\langle A, A\rangle \;=\; \mathrm{Tr}(A^T A) \;=\; \sum_{i=1}^{n} v_i^T v_i.$$

Hier ist jeder Summand $v_i^T v_i$ das Betragsquadrat von v_i bezüglich des Standardskalarprodukts auf $\mathbb{R}^n$ und folglich $\geqslant 0$. Also ist $\langle A, A\rangle \geqslant 0$. Für $A \neq 0$ ist ausserdem mindestens ein $v_i \neq 0$, also mindestens ein Summand > 0 und daher $\langle A, A \rangle > 0$. Somit ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt.

Offenbar ist die Menge aller $n \times n$-Elementarmatrizen eine Basis von $\mathrm{Mat}_{n\times n}(\mathbb{R})$. Durch direkte Rechnung verifizieren wir

$$\langle E_{ij}, E_{k\ell} \rangle = \begin{cases} 1 & \text{falls } (i,j) = (k,\ell), \\ 0 & \text{sonst,} \end{cases}$$

also ist dies eine Orthonormalbasis.

Aliter: Wir identifizieren $\mathrm{Mat}_{n\times n}(\mathbb{R})$ mit $\mathbb{R}^{n^2}$, indem wir die Koeffizienten einer Matrix in irgendeiner fest gewählten Reihenfolge auflisten. Dies ist ein Vektorraum-Isomorphismus. Für zwei $n \times n$-Matrizen $A = (a_{ij})_{i,j}$ und $B = (b_{ij})_{i,j}$ zeigt nun eine direkte Rechnung die Gleichung

$$\langle A, B \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}.$$

Über den genannten Isomorphismus entspricht $\langle \cdot, \cdot \rangle$ also dem Standardskalarprodukt auf $\mathbb{R}^{n^2}$; somit ist auch diese ein Skalarprodukt. Ausserdem entsprechen die $n \times n$-Elementarmatrizen genau den Standardbasisvektoren von $\mathbb{R}^{n^2}$; und da diese eine Orthonormalbasis für das Standardskalarprodukt bilden, gilt das Entsprechende für die $n \times n$-Elementarmatrizen.

Wir bezeichnen die induzierte Norm mit $||\cdot||$. Seien $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$, $A^T = (\tilde{a}_{ij})_{1 \leqslant i,j \leqslant n}$ und $A^T A = (c_{ij})_{1 \leqslant i,j \leqslant n}$. Es gilt

$$c_{ij} = \sum_{k=1} \tilde{a}_{ik} a_{kj} = \sum_{k=1} a_{ki} a_{kj}.$$

Also ist $c_{ii} = \sum_{k=1}^{n} a_{ki}^2$ und

$$||A||^2 = \mathrm{Tr}(A^T A) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki}^2$$

ist einfach die euklidische Norm auf A gesehen als ein Element von $\mathbb{R}^{n^2}$.