8 Euklidische und unitäre Vektorräume

In diesem Kapitel werden nur endlich dimensionale Vektorräume über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ betrachtet. Der Querstrich bezeichnet die komplexe Konjugation ($z = x + iy$, $\overline{z} = x - iy$). Wenn der zugrunde liegende Vektorraum reell ist, so hat er keine Bedeutung.

8.1 Skalarprodukte Sei V ein linearer Raum über $\mathbb{K}$. Eine Abbildung $(\cdot, \cdot) : V \times V \to \mathbb{K}$ heißt *inneres Produkt* oder *Skalarprodukt* in V, wenn die folgenden Bedingungen

(a) $(\alpha_1 x_1 + \alpha_2 x_2, x_3) = \alpha_1 (x_1, x_3) + \alpha_2 (x_2, x_3)$ (Linearität),

(b) $(x_1, x_2) = \overline{(x_2, x_1)}$ (Antisymmetrie),

(c) $(x, x) > 0$ für $x \neq 0$ (Definitheit),

erfüllt sind, wobei $\alpha_i \in \mathbb{K}$ und $x_i \in V$.

Wegen (b) ist $(x, x) \in \mathbb{R}$. Aus (a) und (b) folgt, dass das innere Produkt eine *Sesquilinearform* ist, d.h. es ist linear in der ersten Komponente und *antilinear* in der zweiten,

$$(x_1, \alpha_2 x_2 + \alpha_3 x_3) = \overline{(\alpha_2 x_2 + \alpha_3 x_3, x_1)} = \overline{\alpha_2}(x_1, x_2) + \overline{\alpha_3}(x_1, x_3).$$

Im Fall reeller Räume ist das innere Produkt eine Bilinearform.

Mit Hilfe des Skalarprodukts definieren wir später Schnittwinkel zweier sich schneidender Geraden sowie Abstände zwischen zwei Punkten des Vektorraums. Mit

$$\|x\| = (x, x)^{1/2}$$

können wir die „Entfernung" des Punktes x zum Nullpunkt definieren.

Im $\mathbb{R}^n$ ist das Standardprodukt

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k, \quad |x| = \|x\| = \Big(\sum_{k=1}^{n} |x_k|^2 \Big)^{1/2}.$$

Nach dem Satz des Pythagoras ist $|x|$ gerade die Länge des Vektors x. Man beachte die Notation: $(\cdot, \cdot)$ ist ein allgemeines Skalarprodukt, $\langle \cdot, \cdot \rangle$ ist reserviert für das Standardprodukt im $\mathbb{K}^n$, das im Fall $\mathbb{K} = \mathbb{C}$ so ausschaut:

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \overline{y_k}.$$

Im Komplexen wird in der zweiten Komponente des Produkts komplex konjugiert, damit $\langle x, x \rangle$ reell und ≥ 0 ist.

Ein Vektorraum mit Skalarprodukt heißt *euklidischer Vektorraum* ($\mathbb{K} = \mathbb{R}$) bzw. *unitärer Vektorraum* ($\mathbb{K} = \mathbb{C}$)). Wir sprechen von einem Raum mit Skalarprodukt, wenn wir es offen lassen, ob der Raum reell oder komplex ist.

Lemma 8.1 (Cauchy-Ungleichung) *In einem Raum mit Skalarprodukt gilt für alle x, y*

$$|(x, y)| \leq \|x\| \, \|y\|.$$

Beweis: Aus den Axiomen für das innere Produkt erhalten wir

$$\begin{aligned} 0 \leq (\alpha x + y, \alpha x + y) &= |\alpha|^2 \|x\|^2 + (\alpha x, y) + (y, \alpha x) + \|y\|^2 \\ &= |\alpha|^2 \|x\|^2 + 2\mathrm{Re}\big(\alpha (x, y)\big) + \|y\|^2. \end{aligned}$$

Wir können $x \neq 0$ voraussetzen und wählen $\alpha = -\overline{(x,y)}/\|x\|^2$, also

$$0 \leq \|\alpha x + y\|^2 = \|y\|^2 - \frac{|(x,y)|^2}{\|x\|^2}.$$

Damit ist die Ungleichung bewiesen. □

Lemma 8.2 $\|x\| = (x,x)^{1/2}$ *ist eine* Norm *auf* V, *sie besitzt die Eigenschaften*

(a) $\|x\| > 0$ *für* $x \neq 0$ *(Definitheit),*

(b) $\|\alpha x\| = |\alpha|\,\|x\|$ *für alle* $\alpha \in \mathbb{K}$ *(positive Homogenität),*

(c) $\|x+y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung).*

Beweis: Die beiden ersten Normaxiome folgen direkt aus der Definition der Sesquilinearform, die Dreiecksungleichung beweist man mit Hilfe der Cauchy-Ungleichung

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + (x,y) + (y,x) + \|y\|^2 \|x\|^2 + 2\mathrm{Re}\,(x,y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Aus der Dreiecksungleichung folgt die *umgekehrte Dreiecksungleichung*

$$\big|\|x\| - \|y\|\big| \leq \|x-y\|.$$

Dies folgt aus

$$\|x\| = \|x - y + y\| \leq \|x-y\| + \|y\|.$$

Die andere Richtung beweist man, indem man die Rollen von x und y vertauscht.

8.2 Orthogonalität Sei V ein Vektorraum mit Skalarprodukt. Zwei Vektoren $x, y \in V$ heißen *orthogonal*, wenn $(x,y) = 0$. Wir schreiben dafür $x \perp y$.

Satz 8.3 (Pythagoras) (a) *In einem euklidischen oder unitären Vektorraum gilt*

$$x \perp y \;\Rightarrow\; \|x\|^2 + \|y\|^2 = \|x+y\|^2.$$

(b) *In einem euklidischen Vektorraum gilt auch die Umkehrung:*

$$\|x\|^2 + \|y\|^2 = \|x+y\|^2 \;\Rightarrow\; x \perp y.$$

Beweis: (a) $\|x+y\|^2 = (x+y, x+y) = \|x\|^2 + (x,y) + (y,x) + \|y\|^2 = \|x\|^2 + \|y\|^2$.

(b) Im euklidischen Fall gilt in der letzten Formel $(x,y) + (y,x) = 2(x,y)$, so dass wir auf $x \perp y$ schließen können. Dagegen ist bei unitären Räumen $(x,y) + (y,x) = (x,y) + \overline{(x,y)} = 2\mathrm{Re}\,(x,y)$ und wir erhalten in diesem Fall nur, dass (x,y) rein imaginär ist. □

Eine Menge von Vektoren $x_1, \ldots, x_k$ heißt *Orthogonalsystem*, wenn die Vektoren nicht verschwinden und paarweise orthogonal sind, also $(x_i, x_j) = 0$ für $i \neq j$ erfüllt ist. Ein Orthogonalsystem heißt *Orthonormalsystem*, wenn zusätzlich $\|x_i\| = 1$ für $i = 1, \ldots, k$ erfüllt ist. Aus einem Orthogonalsystem $x_1, \ldots, x_k$ erhalten wir mit der Normierung $y_i = x_i/\|x_i\|$ ein Orthonormalsystem $y_1, \ldots, y_k$.

Die Vektoren in einem Orthogonalsystem sind linear unabhängig, denn in

$$\alpha_1 x_1 + \ldots + \alpha_k x_k = 0$$

können wir von rechts mit x_j multiplizieren und die Linearität des Skalarprodukts ausnutzen,

$$0 = (\alpha_1 x_1 + \ldots + \alpha_k x_k, x_j) = \alpha_1(x_1, x_j) + \ldots + \alpha_k(x_k, x_j) = \alpha_j(x_j, x_j).$$

Es folgt $\alpha_j = 0$.

Nun wollen wir eine l.u. Menge von Vektoren $u_1, \ldots, u_k$ so linear kombinieren, dass eine Orthogonalsystem $x_1, \ldots, x_k$ entsteht mit $\operatorname{span}\{u_1, \ldots, u_i\} = \operatorname{span}\{x_1, \ldots, x_i\}$, $1 \le i \le k$. Wir setzen $x_1 = u_1$. Anschließend bestimmen wir $\alpha \in \mathbb{K}$ so, dass

$$\alpha x_1 + u_2 \perp x_1 \;\Rightarrow\; \alpha = -(u_2, x_1)/\|x_1\|^2.$$

Mit diesem α ist dann $x_2 = \alpha x_1 + u_2 \perp x_1$. Allgemeiner verwenden wir den folgenden

Satz 8.4 (Orthogonalisierungsverfahren von Gram-Schmidt) *Sei V ein Vektorraum mit Skalarprodukt $(\cdot,\cdot)$ und sei $u_1, \ldots, u_k$ eine l.u. Menge von Vektoren in V. Dann erhält man durch*

$$x_1 = u_1, \quad x_{i+1} = u_{i+1} - \sum_{j=1}^{i} \frac{(u_{i+1}, x_j)}{\|x_j\|^2} x_j \text{ für } i = 1, \ldots, k-1 \tag{8.1}$$

ein Orthogonalsystem mit

$$\operatorname{span}\{u_1, \ldots, u_i\} = \operatorname{span}\{x_1, \ldots, x_i\} \text{ für } 1 \le i \le k. \tag{8.2}$$

Insbesondere sind die Vektoren $x_i \ne 0$ und können mit $y_i = x_i/\|x_i\|$ zu einem Orthonormalsystem $y_1, \ldots, y_k$ normiert werden.

Beweis: Wir zeigen die Orthogonalität der Vektoren $x_1, \ldots, x_k$ sowie (8.2) mit Hilfe von (8.1) durch Induktion über k. Für $k = 1$ ist $x_1 = u_1$ und (8.2) erfüllt. Sei also die Behauptung für k erfüllt, insbesondere dürfen wir (8.2) für dieses k verwenden sowie die Orthogonalität der Vektoren $x_1, \ldots, x_k$. Mit dem Ansatz

$$x_{k+1} = u_{k+1} + \alpha_1 x_1 + \ldots + \alpha_k x_k \tag{8.3}$$

erhalten wir aus der Multiplikation mit dem Vektor x_i

$$(x_{k+1}, x_i) = (u_{k+1}, x_i) + \alpha_i(x_i, x_i),$$

denn wegen der Induktionsvoraussetzung gilt $(x_j, x_i) = 0$ für $j \ne i$. Es gilt daher $(x_{k+1}, x_i) = 0$ genau dann, wenn wir $\alpha_i = -(u_{k+1}, x_i)/(x_i, x_i)$ wählen, das ist gerade (8.1). Wäre $x_{k+1} = 0$, so $u_{k+1} \in \operatorname{span}\{x_1, \ldots, x_k\} = \operatorname{span}\{u_1, \ldots, u_k\}$ im Widerspruch zur vorausgesetzten linearen Unabhängigkeit der $u_1, \ldots, u_{k+1}$. $\operatorname{span}\{x_1, \ldots, x_{k+1}\} = \operatorname{span}\{u_1, \ldots, u_{k+1}\}$ folgt aus (8.3. $\square$

Für die Orthogonalisierung mit einem Computerprogramm ist das hier vorgestellte Verfahren die denkbar schlechteste Möglichkeit, weil Rundungsfehler die Orthogonalität stören und einmal gestörte Orthogonalität zu größeren Fehlern in den folgenden Schritten führt (=Aufschaukelung von Rundungsfehlern). Besser ist daher das *modifizierte Gram-Schmidt-Verfahren* oder das *Householder-Verfahren.*

Sei V ein Vektorraum mit Skalarprodukt $(\cdot,\cdot)$ und U ein Unterraum von V. Die Menge

$$U^\perp = \{x \in V : (x, u) = 0 \text{ für alle } u \in U\}$$

heißt *orthogonales Komplement* von U in V. Gilt für einen Vektor $x \in V$, dass $(x, u) = 0$ für alle $u \in U$, so sagen wir, dass x *senkrecht auf* U steht und schreiben $x \perp U$.

Als kleines Beispiel betrachten wir den $\mathbb{R}^2$ mit den drei prinzipiellen Unterräumen $\{0\}, g, \mathbb{R}^2$, wobei g eine Gerade durch den Nullpunkt in Richtung $x \in \mathbb{R}^2$ bezeichnet. Es gilt dann $\{0\}^\perp = \mathbb{R}^2$,

$\mathbb{R}^{2\perp} = \{0\}$. Alle Vektoren, die auf x senkrecht stehen, bilden das orthogonale Komplement von g. Mit $y \perp x$, $y \neq 0$, gilt dann $g^\perp = \{\alpha y : \alpha \in \mathbb{R}\}$.

Allgemeiner gilt in einem beliebigen Vektorraum mit Skalarprodukt $\{0\}^\perp = V$, $V^\perp = \{0\}$. Zur Berechnung von $U^\perp$ für einen nichttrivialen Unterraum U von V mit $\dim V = n$ wählen wir eine Basis $u_1, \ldots, u_r$ von U und ergänzen sie nach dem Basisergänzungssatz 5.9 mit $u_{r+1}, \ldots, u_n$ zu einer Basis von V. In dieser Reihenfolge der Vektoren wenden wir den Gram-Schmidt-Algorithmus 8.1 an und normieren die erhaltenen Vektoren, was zu einem Orthonormalsystem $x_1, \ldots, x_n$ von V führt. Wegen (8.2) gilt $U = \operatorname{span}\{x_1, \ldots, x_r\}$ und die Vektoren in $U' = \operatorname{span}\{x_{r+1}, \ldots, x_n\}$ stehen senkrecht auf U. Daher ist $U' \subset U^\perp$. Jeder Vektor aus V lässt sich als eine Linearkombination $v = \sum_{i=1}^n \alpha_i x_i$ schreiben. Sei $u = \sum_{i=1}^r \alpha_i x_i$. Ist $u = 0$, so ist $u \in U'$, andernfalls ist $u \in U$ mit $(u, u) > 0$. Damit ist $U' = U^\perp$ gezeigt.

Aus dieser Konstruktion lassen sich alle wichtigen Eigenschaften von $U^\perp$ ablesen:

Satz 8.5 (a) *$U^\perp$ ist Unterraum von V,*

(b) $U \cap U^\perp = \{0\}$,

(c) $\dim V = \dim U + \dim U^\perp$.

Beispiele 8.6 (i) Sei $V = \mathbb{R}^3$ mit dem Standard-Skalarprodukt $\langle \cdot, \cdot \rangle$ versehen. Sei $U = \{(x, y, z)^T : 2x + 3y + 4z = 0\}$ eine Ebene. Dann ist $U^\perp = \operatorname{span}\{(2, 3, 4)^T\}$ wegen

$$\langle (2, 3, 4)^T, (x, y, z)^T \rangle = 0 \;\Leftrightarrow\; 2x + 3y + 4z = 0.$$

(ii) Allgemeiner nennen wir einen Unterraum U eines endlich dimensionalen Vektorraums V *Hyperebene*, wenn $\dim U = \dim V - 1 > 0$. Da hier das Skalarprodukt nicht eingeht, gilt diese Definition auch in allgemeinen Vektorräumen über beliebigen Körpern. Im Falle von euklidischen oder unitären Vektoräumen gestatten diese Hyperebenen eine einfache Darstellung mit Hilfe des Skalarprodukts. Wie in der Konstruktion des orthogonalen Komplements beschrieben erhalten wir $U^\perp = \operatorname{span}\{x\}$ mit $x \neq 0$. Dann gilt

$$U = \{u \in V : \langle x, u \rangle = 0\} \;\Leftrightarrow\; U = \{u \in V : x_1 u_1 + \ldots + x_n u_n = 0\}.$$

Man nennt dies die *Hessesche Normalenform* der Hyperebene U. Anders ausgedrückt: Die Hyperebene kann charakterisiert werden durch einen beliebigen Vektor $x \neq 0$, der senkrecht auf U steht und Normale von U genannt wird.

(iii) Sei $V = \mathbb{C}^3$ versehen mit dem Standard-Skalarprodukt $\langle \cdot, \cdot \rangle$. Für $U = \operatorname{span}\{(1, i, 0)^T, (0, 0, 1)^T\} = \operatorname{span}\{x_1, x_2\}$ wollen wir das orthogonale Komplement bestimmen. Die beiden erzeugenden Vektoren sind bereits orthogonal. Durch Probieren finden wir heraus, dass $e_2 = (0, 1, 0)^T$ nicht im Bild dieser beiden Vektoren ist. Nach (8.1) erhalten wir

$$\begin{aligned} x_3 &= e_2 - \frac{\langle e_2, x_1 \rangle}{\langle x_1, x_1 \rangle} x_1 - \frac{\langle e_2, x_2 \rangle}{\langle x_2, x_2 \rangle} x_2 \\ &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \right\rangle} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{-i}{2} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

Damit ist $U^\perp = \operatorname{span}\{(i, 1, 0)^T\}$.

(iv) Ist $x = (x_1, x_2)^T \in \mathbb{K}^2$, so gilt für $x^\perp = (-\overline{x_2}, \overline{x_1})^T$, dass $\langle x, x^\perp \rangle = 0$. □

Für einen Unterraum U des Raums V hatten wir mit Hilfe einer Orthonormalbasis $x_1, \ldots, x_n$ von V mit $U = \operatorname{span}\{x_1, \ldots, x_r\}$ den Unterraum $U^\perp = \operatorname{span}\{x_{r+1}, \ldots, x_n\}$ konstruiert. Entwickeln wir ein beliebiges $v \in V$ nach dieser Basis, $v = \sum_{i=1}^n \alpha_i x_i$, so erhalten wir mit

$$u = \sum_{i=1}^{r} \alpha_i x_i, \quad u^\perp = \sum_{i=r+1}^{n} \alpha_i x_i \tag{8.4}$$

eine Zerlegung

$$v = u + u^\perp \text{ mit } u \in U, u^\perp \in U^\perp$$

u und $u^\perp$ sind nach Konstruktion eindeutig bestimmt.

Die *Orthogonalprojektion* von V auf U ist die Abbildung

$$p_U : V \to U \subset V, \quad v = u + u^\perp \mapsto u.$$

Satz 8.7 *Für die Orthogonalprojektion p_U eines Vektorraums V auf einen Unterraum U gilt:*

(a) *p_U ist linear mit $p_U^2 = p_U \circ p_U = p_U$.*

(b) $\operatorname{Bild} p_U = U$, $\operatorname{Kern} p_U = U^\perp$.

(c) *Es gilt $\|p_U v\| \leq \|v\|$.*

Beweis: Die Eigenschaften folgen aus (8.4). □

Die Berechnung der Orthogonalprojektion erfolgt ebenfalls über (8.4). Es gilt

$$(u, x_j) = (\sum_{i=1}^{n} \alpha_i x_i, x_j) = \alpha_j (x_j, x_j) = \alpha_j.$$

Damit können wir durch einfaches multiplizieren mit x_j das α_j rekonstruieren. Daher

$$p_U(v) = \sum_{i=1}^{r} (u, x_i) x_i. \tag{8.5}$$

Beispiel 8.8 Sei $V = \mathbb{R}^4$ versehen mit dem Standard-Produkt $\langle \cdot, \cdot \rangle$. Sei

$$U = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad v = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}.$$

Gesucht ist die Orthogonalprojektion von v auf U. Wir bestimmen eine Orthonormalbasis von U:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |v_1| = \sqrt{3},$$

$$v_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |v_2| = \sqrt{3},$$

$$v_3 = \begin{pmatrix}0\\0\\1\\1\\0\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\-1\\1\\0\\0\end{pmatrix} = \frac{2}{3}\begin{pmatrix}-1\\0\\1\\1\\0\end{pmatrix}, \quad |v_3| = \frac{2}{\sqrt{3}}.$$

Damit erhalten wir die Orthonormalbasis von U

$$x_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\0\\1\\0\end{pmatrix}, \quad x_2 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\\0\\0\end{pmatrix}, \quad x_3 = \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\0\\1\\1\\0\end{pmatrix}.$$

Gemäß (8.5) folgt

$$p_U(v) = \langle v, x_1\rangle x_1 + \langle v, x_2\rangle x_2 + \langle v, x_3\rangle x_3 = \frac{1}{3}(3, 5, 8, 13, 0)^T.$$

□

8.3 Orthogonale und unitäre Matrizen In diesem Abschnitt betrachten wir nur die Vektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$ versehen mit dem zugehörigen Standard-Produkt.

Eine reelle bzw. komplexe $(n \times n)$-Matrix heißt *orthogonal* bzw. *unitär*, wenn

$$A^T A = E_n \quad \text{bzw.} \quad \overline{A}^T A = E_n.$$

Dies bedeutet, dass A regulär ist mit $A^{-1} = A^T$ bzw. $A^{-1} = \overline{A}^T$. Damit gilt auch $A\overline{A}^T = E_n$ (im Reellen hat der Querstrich wie immer keine Bedeutung). Wir bezeichnen mit a_i die Spaltenvektoren von A, $A = (a_1| \dots |a_n)$. Dann bedeutet $A^T A = E_n$ im Reellen, dass

$$\langle a_i, a_j\rangle = \delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}.$$

Die Spaltenvektoren der Matrix bilden damit ein Orthonormalsystem. Interpretieren wir $AA^T = E_n$ auf die gleiche Weise, kommen wir zur analogen Schlussfolgerung, dass auch die Zeilenvektoren ein Orthonormalsystem bilden.

Im Komplexen können wir genauso folgern wegen

$$(\overline{A}^T A)_{ij} = \sum_k \overline{a}_{ki} a_{kj} = \langle a_j, a_i\rangle.$$

Wir formulieren diese Ergebnisse nur für den komplexen Fall, im Reellen gilt der folgende Satz völlig analog.

Satz 8.9 *Sei $A \in \mathbb{C}^{n\times n}$. Die folgenden Aussagen sind äquivalent:*

(a) *A ist eine unitäre Matrix.*

(b) *A ist regulär mit $A^{-1} = \overline{A}^T$.*

(c) *Die Spaltenvektoren (bzw. Zeilenvektoren) bilden eine Orthonormalbasis des $\mathbb{C}^n$ bezüglich des Standard-Produkts.*

Beispiele 8.10 (i) Im $\mathbb{R}^2$ sind die Drehmatrizen mit Winkel ω

$$A = \begin{pmatrix}\cos\omega & -\sin\omega\\ \sin\omega & \cos\omega\end{pmatrix}$$

offenbar orthogonal.

(ii) Im $\mathbb{R}^n$ ist eine Hyperebene durch einen Vektor $w \in \mathbb{R}^n \setminus \{0\}$ gegeben: $U = \{x : \langle w, x\rangle = 0\}$ (s. Beispiel 8.6 (ii)). Hier können wir $|w| = 1$ voraussetzen. Dann besitzt die Spiegelung an dieser Hyperebene die Darstellungsmatrix

$$S = E_n - 2ww^T.$$

Dabei ist $A = ww^T$ die $(n \times n)$-Matrix mit Einträgen $a_{ij} = w_i w_j$. Ist $x \in \mathbb{R}^n$, so gilt $x = z + \alpha w$ mit $z \in U$, denn $U = \operatorname{span}\{w\}^\perp$. Die Spiegelung an U muss diesen Vektor abbilden auf $Sx = z - \alpha w$ und das ist der Fall:

$$\begin{aligned} Sx &= (E_n - 2ww^T)(z + \alpha w) = z + \alpha w - 2(ww^T)(z + \alpha w) \\ &= z + \alpha w - 2w(w^T z) - 2\alpha w(w^T w). \end{aligned}$$

Im Reellen gilt für Spaltenvektoren x, y, dass $x^T y = \langle x, y\rangle$. Damit ist $w^T z = 0$ wegen $w \perp z$ und $w^T w = 1$ wegen $|w| = 1$. Insgesamt erhalten wir $Sx = z - \alpha w$ wie behauptet. Man rechnet leicht nach $S^2 = E_n$ sowie $S = S^T$. Damit ist S orthogonal. □

Für beliebige reelle $(n \times n)$-Matrizen A gilt

$$\langle Ax, y\rangle = \langle x, A^T y\rangle \quad \text{für alle } x, y \in \mathbb{R}^n$$

wegen

$$\langle Ax, y\rangle = \sum_{k=1}^{n} (Ax)_k y_k = \sum_{k=1}^{n}\sum_{j=1}^{n} a_{kj} x_j y_k = \sum_{k=1}^{n}\sum_{j=1}^{n} a^T_{jk} x_j y_k = \langle x, A^T y\rangle.$$

Damit gilt für eine orthogonale $(n \times n)$-Matrix A

$$\langle Ax, Ay\rangle = \langle x, A^T Ay\rangle = \langle x, y\rangle,$$

insbesondere auch für $y = x$: $|Ax| = |x|$. Damit erhält eine orthogonale Matrix das Skalarprodukt und damit auch die Längen von Vektoren. Die zugehörigen orthogonalen Selbstabbildungen $f(x) = Ax$ erhalten damit alle Strukturen, die in einem euklidischen Vektorraum vorhanden sind. Die ganze Herleitung gilt auch im unitären Raum $\mathbb{C}^n$.

Wir zeigen: Hat $(n \times n)$-Matrix A die Eigenschaft

$$|Ax| = |x| \quad \text{für alle } x \in \mathbb{K}^n,$$

so ist sie bereits orthogonal bzw. unitär. Im Reellen gilt

$$|Ax + Ay|^2 = |x + y|^2 \;\Rightarrow\; (Ax, Ay) = (x, y) \text{ wegen } |Ax| = |x|,\ |Ay| = |y|,$$

woraus $(x, A^T Ay) = (x, y)$ folgt. Wir können hier für x die kanonischen Einheitsvektoren einsetzen und erhalten $A^T Ay = y$ für alle y und damit $A^T A = E_n$. Im Komplexen folgt mit gleicher Überlegung nur $\operatorname{Re}(Ax, Ay) = \operatorname{Re}(x, y)$. Wir können hier aber x durch ix ersetzen und erhalten dann auch $\operatorname{Im}(Ax, Ay) = \operatorname{Im}(x, y)$. Der Rest verläuft genauso wie zuvor.