

17. Orthogonalsysteme

17.1. Winkel und Orthogonalität

Vorbemerkung: Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und zugehöriger Norm $\|\cdot\|$, dann gilt nach Cauchy-Schwarz:

$$\forall x, y \in V \setminus \{0\} : \frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \leq 1$$

Definition: (a) Sei $\mathbb{K} = \mathbb{R}$. Für $x, y \in V \setminus \{0\}$ sei $\phi = \angle(x, y) \in [0, \pi]$ diejenige (eindeutig bestimmte) Zahl mit

$$\cos \phi = \frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|}$$

ϕ heißt der **Winkel** zwischen x und y .

(b) x, y heißen **orthogonal** oder senkrecht zueinander, falls $\langle x, y \rangle = 0$ gilt.
Schreibe: $x \perp y$.

(c) Teilmengen $M, N \subseteq V$ heißen orthogonal, falls gilt:

$$\forall x \in M \forall y \in N : x \perp y$$

Schreibe: $M \perp N$.

(d) Eine Teilmenge $B \subseteq V$ heißt **Orthogonalsystem** (OGS), falls für $x, y \in B$ gilt:

$$x \neq y \implies x \perp y$$

(e) Ein Orthogonalsystem B heißt **Orthonormalsystem** (ONS) wenn gilt:

$$\forall x \in B : \|x\| = 1$$

(f) Eine Basis B von V heißt **Orthogonalbasis** (OGB), bzw. **Orthonormalbasis** (ONB), falls B ein Orthogonalsystem, bzw. Orthonormalsystem, ist.

Beispiel: (1) Sei $V = \mathbb{K}^n$ mit Standardskalarprodukt $\langle \cdot, \cdot \rangle$ und $S := \{e_1, \dots, e_n\}$ Standardbasis.

Dann ist S eine Orthonormalbasis und jede Teilmenge $T \subseteq S$ ist ein Orthonormalsystem.

(2) Sei $I := [a, b]$ ein Intervall.

Sei $V := \{p \in \text{Abb}(I, \mathbb{C}) \mid \exists P \in \mathbb{C}[T] : p(t) = P(t)\}$.

$w : I \rightarrow \mathbb{R}_{\geq 0}$ sei stetig und mit der Eigenschaft $w(t) = 0$ nur für endlich viele $t \in I$.

Wir erhalten ein Skalarprodukt auf V :

$$\langle p, q \rangle_w := \int_I w(t) p(t) \overline{q(t)} dt$$

Eine Basis von V ist $\{p_n(t) =: t^n \mid n \in \mathbb{N}_0\}$.

Gesucht ist eine Orthonormalbasis und ein Verfahren zu ihrer Bestimmung.

Bemerkung: Jedes Orthogonalsystem B mit $0 \notin B$ ist linear unabhängig.

Beweis: Es ist

$$\sum_{b \in B} \alpha_b \cdot b = 0$$

Dann gilt für alle $c \in B$:

$$\begin{aligned} 0 &= \langle 0, c \rangle \\ &= \left\langle \sum_b \alpha_b b, c \right\rangle \\ &= \sum_b \alpha_b \langle b, c \rangle \\ &\stackrel{b=0 \forall b \neq c}{=} \alpha_c \underbrace{\langle c, c \rangle}_{\neq c} \\ &\implies \alpha_c = 0 \end{aligned}$$

■

17.2. Das E. Schmidtsche Orthogonalisierungsverfahren

Satz 13:

Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und sei $M := \{x_0, x_1, \dots\}$ eine abzählbare Teilmenge von V .

(1) Es existiert ein Orthogonalsystem $\{y_0, y_1, \dots\}$ derart, dass gilt:

$$\forall n : \quad \langle y_0, y_1, \dots, y_n \rangle = \langle x_0, \dots, x_n \rangle \quad (\text{gleiche lineare Hülle}) \quad (17.1)$$

(2) Falls M linear unabhängig ist, so sind alle $y_i \neq 0$ und $B := \{z_0, z_1, \dots\}$ mit

$$z_i := \frac{1}{\|y_i\|} y_i$$

ist ein Orthonormalsystem mit

$$\forall n : \quad \langle z_0, z_1, \dots, z_n \rangle = \langle x_0, x_1, \dots, x_n \rangle$$

Beweis: (1) Wir beschreiben einen Algorithmus zum Auffinden der y_n .

Start: $y_0 := x_0$. Angenommen: alle y_m für $m < n$ sind bereits gefunden. Setze

$$y_n := x_n - \sum_{i=0}^{n-1} \frac{\langle x_n, y_i \rangle}{\langle y_i, y_i \rangle} y_i$$

(nur über i mit $y_i \neq 0$ summieren).

Damit folgt:

$$\begin{aligned} y_n &\in \underbrace{\langle y_0, y_1, \dots, y_{n-1}, x_n \rangle}_{=\langle x_0, \dots, x_{n-1} \rangle} = \langle x_0, \dots, x_n \rangle \\ x_n &\in \langle y_0, y_1, \dots, y_n \rangle \end{aligned}$$

Daraus folgt (17.1).

Rest: Für alle $m < n$: $y_n \perp y_m$. Damit:

$$\begin{aligned} \langle y_n, y_m \rangle &= \langle x_n, y_m \rangle - \sum_{i=0}^{n-1} \frac{\langle x_n, y_i \rangle}{\langle y_i, y_i \rangle} \langle y_i, y_m \rangle \\ &= \langle x_n, y_m \rangle - \sum_{i=0}^{n-1} \frac{\langle x_n, y_i \rangle}{\langle y_i, y_i \rangle} \delta_{im} \langle y_i, y_i \rangle \\ &= \langle x_n, y_m \rangle - \langle x_n, y_m \rangle \\ &= 0 \end{aligned}$$

(2) ✓ Leicht selbst zu verifizieren. ■

Beispiel: Sei $V = \mathbb{R}^3$ mit dem Standardskalarprodukt, $M := \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$. Dann:

$$y_0 = x_0, \quad \|y_0\| = \sqrt{2}$$

$$y_1 = x_1 - \frac{\langle x_1, y_0 \rangle}{\langle y_0, y_0 \rangle} y_0 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \|y_1\| = \sqrt{\frac{3}{2}}$$

$$y_2 = x_2 - \frac{\langle x_2, y_0 \rangle}{\langle y_0, y_0 \rangle} y_0 - \frac{\langle x_2, y_1 \rangle}{\langle y_1, y_1 \rangle} y_1 = \frac{2}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Auswirkungen (des Orthogonalisierungsverfahrens) auf Matrizen:

Sei $V \cong \mathbb{K}^n$ mit Standardskalarprodukt s und Basis $B := \{b_1, \dots, b_n\}$. Das Orthogonalisierungsverfahren liefert eine Orthogonalbasis $C = \{c_1, \dots, c_n\}$.

$$c_\nu = b_\nu - \sum_{i=1}^{n-1} \frac{\langle b_\nu, c_i \rangle}{\langle c_i, c_i \rangle} c_i = \dots = \sum_{i=1}^{\nu} \alpha_{i\nu} b_i \longrightarrow A = M_{BC} = (\alpha_{i\nu}) = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

$$\frac{c_\nu}{\|c_\nu\|} = \sum_{i=1}^{\nu} \frac{\alpha_{i\nu}}{\|c_i\|} b_i = z_\nu \longrightarrow M_{BZ} = \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$$

Erinnere (Darstellungsmatrix):

$$D_{BB}(s) = (s(b_\nu, b_\mu)) \in \mathbb{K}^n$$

Da C eine Orthogonalbasis ist, folgt $s(c_\nu, c_\mu) = \delta_{\nu\mu} \|c_\nu\|^2$, also

$$D_{CC}(s) = \text{diag}(\dots, \|c_\nu\|^2, \dots)$$

Falls Z eine Orthonormalbasis ist, so folgt $D_{ZZ}(s) = I$.

Generell für beliebige Basen B, C und $A = M_{BC}$:

$$\begin{aligned} D_{CC}(s) &= (s(c_\nu, c_\mu)) \\ &= \left(s \left(\sum_{i=1}^n \alpha_{i\nu}, \sum_j \alpha_{j\mu} b_j \right) \right) \\ &= \left(\sum_i \sum_j \alpha_{i\nu} \overline{\alpha_{j\mu}} \cdot s(b_i, b_j) \right) \\ &= A^\top (s(b_i, b_j)) \overline{A} \end{aligned}$$

Definition: Für beliebige $D \in \mathbb{K}^{n \times n}$ setze $D^* := \overline{D}^\top$ (die sogenannte **Adjungierte**).

$$D_{CC}(s) = A^\top D_{BB}(s) \overline{A}$$

Speziell für jede Orthonormalbasis C :

$$D_{CC}(s) = I,$$

das folgt wegen $M_{BC}^{-1} = M_{CB}$.

Es ist

$$D_{BB}(s) = D^* D$$

für $D := \overline{M}_{CB}$, wobei D obere Dreiecksmatrix ist.

Satz 14:

Für $P \in \mathbb{K}^{n \times n}$ ist äquivalent:

- (1) P ist hermitesch (symmetrisch) und positiv definit
- (2) Es gibt ein $A \in \text{GL}_n(\mathbb{K})$ mit $P = A^* A$.

Beweis: Der Beweis erfolgt durch Ringschluss:

- (1) \implies (2) Sei $V = \mathbb{K}^n$, $B = \{e_1, \dots, e_n\}$ die Standardbasis.
 P definiert ein Skalarprodukt:

$$\begin{aligned} s(x, y) &= (\overline{\eta}_1, \dots, \overline{\eta}_n) \cdot P \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \\ P &= (s(e_i, e_j)) = D_{BB}(s) \end{aligned}$$

Das Schmidtsche Orthogonalisierungsverfahren liefert eine Orthonormalbasis und damit $P = D^*D$.

(2) \implies (1) $P = A^*A$ ist hermitesch; zu zeigen: $P^* = P$

$$\begin{aligned}
 P^* &= (A^* \cdot A)^* \\
 &= \overline{(A^* \cdot A)}^\top \\
 &= \overline{(A^\top \cdot A)}^\top \\
 &= (A^\top \cdot \overline{A})^\top \\
 &= \overline{A}^\top \cdot A \\
 &= A^*A \\
 &= P
 \end{aligned}$$

Daraus folgt: $s(x, y)$ ist hermitesche Form. Für alle $x \in \mathbb{K}^n$ gilt:

$$\begin{aligned}
 s(x, x) &= \overline{x}^\top \cdot P \cdot x \\
 &= \overline{x}^\top \left(\overline{A}^\top \cdot A \right) x \\
 &= (\overline{Ax})^\top Ax \\
 &= s_0(Ax, Ax) \\
 &\geq 0
 \end{aligned}$$

Weiterhin:

$$\begin{aligned}
 s_0(Ax, Ax) &= 0 \\
 \iff Ax &= 0 \\
 \stackrel{A \text{ inv.}}{\iff} x &= 0 \\
 \implies s &\text{ positiv definit}
 \end{aligned}$$

■

Falls speziell B und C Orthonormalbasen sind, folgt:

$$D_{BB}(s) = I = D_{CC}(s)$$

und $D := \overline{M}_{CB}$.

Folgerung: Die Basiswechselmatrix $A = M_{BC}$ einer Orthonormalbasis C in eine andere Orthonormalbasis B gehört zur **orthogonalen Gruppe**

$$O(n) := \left\{ A \in \text{GL}_n(\mathbb{R}) \mid A^\top \cdot A = I \right\} \quad \text{für } \mathbb{K} = \mathbb{R}$$

beziehungsweise zur **unitären Gruppe**

$$U(n) := \left\{ A \in \text{GL}_n(\mathbb{C}) \mid A^\top \cdot \overline{A} = I \right\} \quad \text{für } \mathbb{K} = \mathbb{C}$$

Bemerkung: $O(n)$, beziehungsweise $U(n)$, ist eine Untergruppe von $\text{GL}_n(\mathbb{R})$, beziehungsweise $\text{GL}_n(\mathbb{C})$.

Folgerung (Iwasawa-Zerlegung): Jede Matrix $g \in \mathrm{GL}_n(\mathbb{R})$ hat eine eindeutige Produktzerlegung

$$g = k \cdot b$$

mit $k \in O(n)$ und $b \in B(n) := \left\{ \begin{pmatrix} \beta_1 & & * \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix} \mid \beta_\nu > 0 \right\}$. Das heißt:

$$\mathrm{GL}_n(\mathbb{R}) = O(n) \cdot B(n)$$

Analog gilt:

$$\mathrm{GL}_n(\mathbb{C}) = U(n) \cdot B(n)_{\mathbb{C}}$$

mit

$$B(n)_{\mathbb{C}} := \left\{ \begin{pmatrix} \beta_1 & & * \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix} \in \mathrm{GL}_n(\mathbb{C}) \mid \beta_\nu \in \mathbb{R}_{>0} \right\}$$

Beweis: Für $\mathbb{K} = \mathbb{R}$ $g \in \mathrm{GL}_n(\mathbb{R})$ folgt: Die Spalten b_1, \dots, b_n sind eine Basis von \mathbb{R}^n . Das Schmidtsche Orthogonalisierungsverfahren liefert eine Orthonormalbasis $\{c_1, \dots, c_n\}$ mit Übergangsmatrix $A = M_{BC} \in B(n)$. Denn:

$$c_\nu := \sum_{i=1}^{\nu} \alpha_{i\nu} b_i$$

besagt $g \cdot A = (c_1, \dots, c_n)$ und $k = (c_1, \dots, c_n) \in O(n)$, da $k^\top \cdot k = (\langle c_i, c_j \rangle) = I$. ■

17.3. Orthogonale Projektion und orthogonales Komplement

Satz 15 (Satz von Pythagoras):

Für $x, y \in V$ mit $x \perp y$ gilt:

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

Beweis:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \underbrace{\langle x, y \rangle}_{=0} + \underbrace{\langle y, x \rangle}_{=0} + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 \end{aligned}$$

Ist $\{x_1, \dots, x_N\}$ ein Orthogonalsystem, so folgt

$$\left\| \sum_{\nu=1}^N x_\nu \right\|^2 = \sum_{\nu=1}^N \|x_\nu\|^2$$

Der Beweis folgt leicht mit vollständiger Induktion. ■

Satz 16:

Sei $U \leq V$ mit $\dim V < \infty$.

- (1) Für alle $x \in V$ existiert genau ein $y \in U$ mit $d := \|x - y\| = \min\{\|x - u\| \mid u \in U\}$.
- (2) Dieses $y \in U$ ist auch charakterisiert durch: $(x - y) \perp U$.
Schreibe: $y =: \Pi_U(x)$.
- (3) Die Abbildung $\Pi_U \in \text{End}(V)$ ist stetig; es gilt $\Pi_U^2 = \Pi_U$ und $\|\Pi_U(x)\| \leq \|x\|$.
 d heißt **Abstand** von x und U , $y = \Pi_U(x)$ die **orthogonale Projektion** von x auf U , $z := x - y$ heißt **Lot** von x auf U , y **Lotfußpunkt**.

Beweis: (1) Wähle eine Orthonormalbasis $S = \{e_i \mid i = 1, \dots, r\}$ in U . Setze $y = \Pi_U(x) := \sum_{i=1}^r \langle x, e_i \rangle e_i$.

Behauptung: $\forall u' \in U : \quad x - y \perp y - u'$

$$\begin{aligned} \langle x - y, y - u' \rangle &= \underbrace{\langle x, y \rangle - \langle y, y \rangle}_{\stackrel{!}{=} 0} + \underbrace{\langle y, u' \rangle - \langle x, u' \rangle}_{\stackrel{!}{=} 0} \\ \langle x, y \rangle &= \left\langle x, \sum_i \langle x, e_i \rangle e_i \right\rangle = \sum_i \overline{\langle x, e_i \rangle} \langle x, e_i \rangle \end{aligned}$$

u' in Basisdarstellung: Mit $u' = \sum_j \alpha_j e_j$ folgt

$$\begin{aligned} \langle y, u' \rangle &= \left\langle \sum_i \langle x, e_i \rangle e_i, \sum_j \alpha_j e_j \right\rangle \\ &= \sum_i \langle x, e_i \rangle \overline{\alpha_i} \end{aligned}$$

Weiterhin gilt:

$$\begin{aligned} \langle x, u' \rangle &= \left\langle x, \sum_j \alpha_j e_j \right\rangle \\ &= \sum_j \overline{\alpha_j} \langle x, e_j \rangle \end{aligned}$$

Mit Pythagoras folgt:

$$\begin{aligned} \|x - u'\|^2 &= \|x - y + y - u'\|^2 \\ &= \|x - y\|^2 + \underbrace{\|y - u'\|^2}_{\geq 0} \\ &\geq \|x - y\|^2 \end{aligned}$$

Es ist also $\|x - u'\| \geq \|x - y\|$, wobei Gleichheit genau für $y - u' = 0$ gilt. Damit folgt die Eindeutigkeit von y .

- (2) Sei $y \in U$ und $x - y \perp U$. Dann gilt $\langle x, e_i \rangle = \langle y, e_i \rangle$ für alle i .
Es folgt:

$$\begin{aligned} y &= \sum_i \langle y, e_i \rangle e_i \\ &= \sum_i \langle x, e_i \rangle e_i \\ &= \Pi_U(x) \end{aligned}$$

- (3) Aus $x - y \perp y$ folgt mit Pythagoras:

$$\begin{aligned} \|x\|^2 &= \|x - y\|^2 + \|y\|^2 \\ &\geq \|y\|^2 \\ &= \|\Pi_U(x)\|^2 \end{aligned}$$

Es folgt: Π_U ist (Lipschitz-)stetig.

$\Pi_U^2 = \Pi_U$ ist leicht selbst zu verifizieren. ■

Definition: Sei $M \subseteq V$ Teilmenge. Der Vektorraum

$$M^\perp := \{y \in V \mid y \perp M\}$$

heißt **Orthogonalraum** oder **orthogonales Komplement** von M .

Lemma:

(1) $M_1 \subseteq M_2 \implies M_1^\perp \supseteq M_2^\perp$

(2) $\langle M \rangle^\perp = M^\perp$

- (3) Aus $M_i \subseteq V$, ($i = 1, \dots, n$) folgt

$$\left(\bigcup_{i=1}^n M_i \right)^\perp = \bigcap_{i=1}^n M_i^\perp$$

- (4) Aus $U_i \leq V$ (Teilräume) folgt

$$\left(\bigcap_{i=1}^n U_i \right)^\perp \supseteq \sum_{i=1}^n (U_i^\perp)$$

(5) $\langle M \rangle \leq (M^\perp)^\perp$ und $M^\perp = ((M^\perp)^\perp)^\perp$.

- (6) Im Spezialfall $\dim V < \infty$ gilt:

(a) Mit $U \leq V$ folgt $V = U \oplus U^\perp$ (insbesondere $\dim U + \dim U^\perp = \dim V$) und $(U^\perp)^\perp = U$

(b) Mit $U_i \leq V$ folgt $(\bigcap_{i=1}^n U_i)^\perp = \sum_{i=1}^n (U_i^\perp)$

Beweis: Übung!

■

