5 Innere-Produkt-Räume und Fourier-Approximation

In diesem Kapitel betrachten wir immer Vektorräume über dem Körper der reellen Zahlen $\mathbb{R}$ oder dem Körper der komplexen Zahlen $\mathbb{C}$.

5.1 Was ist ein IP-Raum ?

Definition 5.1.1: *Sei V ein Vektorraum über $\mathbb{R}$. Ein* **inneres Produkt** *oder* **Skalarprodukt** *auf V ist eine binäre Abbildung vom Vektorraum V in den Grundkörper $\mathbb{R}$, $f : V^2 \to \mathbb{R}$, geschrieben als $f(x, y) = \langle x|y\rangle$, sodass für alle $x, x', y \in V$ und alle $\alpha \in \mathbb{R}$ folgendes gilt:*

(1) $\langle x + x'|y\rangle = \langle x|y\rangle + \langle x'|y\rangle$;

(2) $\langle \alpha x|y\rangle = \alpha\langle x|y\rangle$;

(3) $\langle y|x\rangle = \langle x|y\rangle$;

(4) $\langle x|x\rangle \geq 0$, *mit Gleichheit genau dann, wenn* $x = 0$.

Unter einem **reellen inneren Produktraum** *verstehen wir einen Vektorraum V über $\mathbb{R}$, auf welchem ein inneres Produkt $\langle .|.\rangle$ erklärt ist.*
In analoger Weise können wir ein inneres Produkt auf einem Vektorraum V über $\mathbb{C}$ betrachten, also eine binäre Abbildung von V in den Grundkörper $\mathbb{C}$; wir müssen nur die Eigenschaft (3) ersetzen durch die Eigenschaft ($\overline{3}$):

($\overline{3}$) $\langle y|x\rangle = \overline{\langle x|y\rangle}$,

wobei $\overline{c}$ die komplex konjugierte Zahl zu c ist (man beachte, dass daraus wegen $\langle x|x\rangle = \overline{\langle x|x\rangle}$ folgt $\langle x|x\rangle \in \mathbb{R}$, also die Bedingung (4) sinnvoll ist).
Unter einem **komplexen inneren Produktraum** *verstehen wir einen Vektorraum V über $\mathbb{C}$, auf welchem ein inneres Produkt $\langle .|.\rangle$ erklärt ist.*
Abkürzend nennen wir einen inneren Produktraum einfach **IP-Raum**. □

Wir erinnern daran, dass für eine komplexe Zahl $c = a+bi$ die die komplex konjugierte Zahl die Gestalt $\overline{c} = a - bi$ hat. Für eine reelle Zahl r gilt $\overline{r} = r$. Wir können also die Eigenschaft ($\overline{3}$) auch für einen reellen Vektorraum mit innerem Produkt fordern, sie wird in diesem Fall einfach zu (3). Wir werden also meist einfach von (3) sprechen, im komplexen Fall aber darunter ($\overline{3}$) verstehen.

Lemma 5.1.2: *Sei V ein (reeller oder komplexer) IP-Raum mit innerem Produkt $\langle .|.\rangle$, und K der Grundkörper, also $K = \mathbb{R}$ oder $K = \mathbb{C}$. Dann gelten auch die folgenden Beziehungen für alle $x, y, y' \in V$ und $\alpha \in K$:*

(5) $\langle x|y + y'\rangle = \langle x|y\rangle + \langle x|y'\rangle$,

(6) $\langle x|\alpha y\rangle = \overline{\alpha}\langle x|y\rangle$,

(7) $\langle x|0_V\rangle = 0 = \langle 0_V|y\rangle$.

Beweis: zu (5): wegen (1) und (3) haben wir

$$\langle x|y+y'\rangle = \overline{\langle y+y'|x\rangle} = \overline{\langle y|x\rangle} + \overline{\langle y'|x\rangle} = \langle x|y\rangle + \langle x|y'\rangle \ .$$

zu (6): wegen (3) und (2) haben wir

$$\langle x|\alpha y\rangle = \overline{\langle \alpha y|x\rangle} = \overline{\alpha\langle y|x\rangle} = \overline{\alpha}\overline{\langle y|x\rangle} = \overline{\alpha}\langle x|y\rangle \ .$$

zu (7): $\langle x|0_V\rangle = \langle x|0\cdot 0_V\rangle = 0\cdot\langle x|0_V\rangle = 0$. Analog für $\langle 0_V|y\rangle$. □

Beispiel 5.1.3: Auf dem reellen Vektorraum $\mathbb{R}^n$ ist die Abbildung

$$\begin{array}{ll} \langle .|.\rangle : & \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \\ & \langle (x_1,\ldots,x_n)|(y_1,\ldots,y_n)\rangle = \sum_{i=1}^n x_i y_i \end{array}$$

offensichtlich ein inneres Produkt. Es heisst das **kanonische innere Produkt** oder **kanonische Skalarprodukt** auf $\mathbb{R}^n$.
Auf dem komplexen Vektorraum $\mathbb{C}^n$ ist die Abbildung

$$\begin{array}{ll} \langle .|.\rangle : & \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} \\ & \langle (z_1,\ldots,z_n)|(w_1,\ldots,w_n)\rangle = \sum_{i=1}^n z_i \overline{w_i} \end{array}$$

offensichtlich ein inneres Produkt. Es heisst das **kanonische innere Produkt** oder **kanonische Skalarprodukt** auf $\mathbb{C}^n$. □

Beispiel 5.1.4: (a) Seien $a,b\in\mathbb{R}$ mit $a<b$ und sei $V = C([a,b],\mathbb{R})$ der reelle Vektorraum der stetigen Funktionen von $[a,b]$ nach $\mathbb{R}$. Aus elementaren Eigenschaften des Integrals leitet man her, dass die Abbildung von $V\times V$ nach $\mathbb{R}$, definiert durch

$$(f,g)\mapsto \langle f|g\rangle = \int_a^b fg \ ,$$

ein inneres Produkt auf $V = C([a,b],\mathbb{R})$ ist.
(b) Ebenso ist auf dem reellen Vektorraum der Polynome vom Grad kleiner oder gleich n, also $\mathrm{Pol}_n(\mathbb{R})$, die Abbildung

$$(p,q)\mapsto \langle p|q\rangle = \int_0^1 pq$$

ein inneres Produkt. □

Beispiel 5.1.5: Für $n\times n$ Matrizen A ist die **Spur** (**trace**) definiert als

$$\mathrm{spur}(A) = \sum_{i=1}^n a_{ii} \ .$$

Mittels der Abbildung

$$\langle A|B\rangle = \mathrm{spur}(B^T A) = \sum_{i,j=1}^n a_{ij} b_{ij}$$

wird der reelle Vektorraum $\mathrm{Mat}_{n\times n}(\mathbb{R})$ zu einem inneren Produktraum.
Mittels der Abbildung

$$\langle A|B\rangle = \mathrm{spur}(\overline{B^T}A)$$

wird der komplexe Vektorraum $\mathrm{Mat}_{n\times n}(\mathbb{C})$ zu einem inneren Produktraum. □

Definition 5.1.6: *Sei V ein (reeller oder komplexer) IP-Raum. Für jedes $x \in V$ definieren wir die* **Norm** *von x als die nicht-negative reelle Zahl*

$$\| x \| = \sqrt{\langle x|x\rangle} \ .$$

Wir nennen ein $x \in V$ **normiert**, *wenn $\| x \| = 1$.*
Für $x, y \in V$ definieren wir den **Abstand** *(***Distanz***) zwischen x und y als*

$$d(x,y) \ = \| x - y \| \ .$$ □

Aus der Eigenschaft (4) folgt sofort: $\| x \| = 0 \iff x = 0_V$.

Beispiel 5.1.7: Sei $\mathbb{R}^2$ der reelle IP-Raum unter dem kanonischen inneren Produkt. Dann gilt offensichtlich für $x = (x_1, x_2)$

$$\| x \| = \sqrt{x_1^2 + x_2^2} \ ,$$

also $\| x \|$ ist der Abstand von x zum Ursprung $(0,0)$.
Ebenso gilt für $y = (y_1, y_2)$

$$\| x - y \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \ ,$$

also $\| x - y \|$ ist der Abstand zwischen x und y. □

Satz 5.1.8: *Sei V ein IP-Raum. Dann gilt für alle $x, y \in V$ und alle Skalare λ:*

(i) $\| \lambda x \| = |\lambda| \, \| x \|$;

(ii) (Cauchy-Schwarz-Ungleichung) [1] $\quad |\langle x|y\rangle| \le \| x \| \cdot \| y \|$;

(iii) (Dreiecksungleichung) $\quad \| x + y \| \le \| x \| + \| y \|$.

Beweis: ad (i): $\| \lambda x \|^2 = \langle \lambda x|\lambda x\rangle = \lambda\overline{\lambda}\langle x|x\rangle = |\lambda|^2 \| x \|^2$.
ad (ii): Die Beziehung ist trivial für $x = 0_V$. Sei also nun $x \neq 0$, und daher auch $\| x \| \neq 0$.
Sei weiters

$$z := y - \frac{\langle y|x\rangle}{\| x \|^2} x \ .$$

Dann ist zunächst

$$\langle z|x\rangle \ = \langle y|x\rangle - \frac{\langle y|x\rangle}{\| x \|^2}\langle x|x\rangle \ = \ 0 \ .$$

Daraus folgt

$$\begin{aligned} 0 \ \le \| z \|^2 = \langle z|z\rangle &= \langle z|y\rangle \\ &= \langle y|y\rangle - \tfrac{\langle y|x\rangle}{\|x\|^2}\langle x|y\rangle \\ &= \langle y|y\rangle - \tfrac{|\langle x|y\rangle|^2}{\|x\|^2} \end{aligned}$$

und somit offensichtlich die Behauptung.

[1] Augustin L. Cauchy (1789–1857), Hermann A. Schwarz (1843–1921)

ad (iii): es gilt

$$\begin{aligned}
\| x+y \|^2 &= \langle x+y|x+y\rangle \\
&= \langle x|x\rangle + \langle x|y\rangle + \langle y|x\rangle + \langle y|y\rangle \\
&= \| x \|^2 + \langle x|y\rangle + \overline{\langle x|y\rangle} + \| y \|^2 \\
&= \| x \|^2 + 2\mathrm{re}\langle x|y\rangle + \| y \|^2 \\
&\leq \| x \|^2 + 2|\langle x|y\rangle| + \| y \|^2 \\
&\leq \| x \|^2 + 2 \| x \| \| y \| + \| y \|^2 \\
&= (\| x \| + \| y \|)^2 \ .
\end{aligned}$$

Dabei haben wir für die zweite Ungleichung die Cauchy-Schwarz-Ungleichung benutzt. Daraus folgt nun sofort die Dreiecksungleichung. □

Beispiel 5.1.9: (a) Im inneren Produktraum $\mathbb{R}^n$ unter dem kanonischen inneren Produkt bedeutet die Cauchy-Schwarz-Ungleichung:

$$|\sum_{i=1}^n x_i y_i| \ \leq \ \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2} \ .$$

(b) Im inneren Produktraum $C([a,b],\mathbb{R})$ mit dem Integral als innerem Produkt (vgl. Beispiel 5.1.4(a)) bedeutet die Cauchy-Schwarz-Ungleichung

$$|\int_a^b f\cdot g| \ \leq \ \sqrt{\int_a^b f^2} \cdot \sqrt{\int_a^b g^2} \ ,$$

also eine Integralabschätzung. □

Definition 5.1.10: *Sei V ein innerer Produktraum. $x, y \in V$ heissen* **orthogonal** *genau dann, wenn $\langle x|y\rangle = 0$. Wir schreiben dafür $x \perp y$.*
Eine Teilmenge S von $V \setminus \{0\}$ heisst eine **orthogonale Teilmenge** *bzw. ein* **Orthogonalsystem** *genau dann, wenn jedes Paar verschiedener Elemente von S orthogonal ist.*
Eine **orthonormale Teilmenge** *bzw. ein* **Orthonormalsystem** *von V ist ein Orthogonalsystem S, sodass $\| x \| = 1$ für jedes $x \in S$.*

Beispiel 5.1.11: (a) Bzgl. des kanonischen inneren Produkts bilden die kanonischen Basen von $\mathbb{R}^n$ und $\mathbb{C}^n$ orthonormale Teilmengen. In $\mathbb{R}^2$ etwa sind zwei Vektoren $x = (x_1, x_2), y = (y_1, y_2)$ orthogonal genau dann, wenn $x_1y_1 + x_2y_2 = 0$. Geometrisch bedeutet das, dass die Gerade vom Nullpunkt durch x senkrecht liegt zur Geraden vom Nullpunkt durch y.
(b) Die Matrizen $E_{i,j}$ bilden eine orthonormale Teilmenge des inneren Produktraums der $n \times n$ Matrizen aus Bsp. 5.1.5. □

Satz 5.1.12: *Sei V ein innerer Produktraum und S eine orthogonale Teilmenge, welche nicht den Nullvektor enthält. Dann ist S linear unabhängig.*

Beweis: Sei also $S \subseteq V \setminus \{0\}$ orthogonal und seien $x_1, \ldots, x_n \in S$. Aus

$$\sum_{i=1}^n \lambda_i x_i \ = \ 0$$

folgt für jedes $j \in \{1, \ldots, n\}$

$$\lambda_j \| x_j \|^2 = \lambda_j \langle x_j|x_j\rangle = \sum_{i=1}^n \lambda_i \langle x_i|x_j\rangle = \langle \sum_{i=1}^n \lambda_i x_i | x_j\rangle = \langle 0|x_j\rangle = 0 \ .$$

Somit muss für alle j der Linearkoeffizient $\lambda_j = 0$ sein, d.h. S ist linear unabhängig. $\square$

Satz 5.1.13: *Sei V ein innerer Produktraum und sei $\{b_1, \ldots, b_n\}$ eine orthonormale Teilmenge. Dann gilt für alle $x \in V$:*

$$\sum_{i=1}^{n} |\langle x|b_i\rangle|^2 \leq \| x \|^2 \quad . \qquad (\textbf{Bessel} - \textbf{Ungleichung})$$

Beweis: Sei $z := x - \sum_{i=1}^{n}\langle x|b_i\rangle b_i$. Dann gilt

$$\begin{array}{rcl} 0 & \leq & \langle z|z\rangle \\ & = & \langle x|x - \sum_{i=1}^{n}\langle x|b_i\rangle b_i\rangle \; - \; \sum_{i=1}^{n}\langle x|b_i\rangle \cdot \langle b_i|x - \sum_{j=1}^{n}\langle x|b_j\rangle b_j\rangle \\ & = & \langle x|x\rangle \; - \; \sum_{i=1}^{n} \overline{\langle x|b_i\rangle}\langle x|b_i\rangle \; - \; \sum_{i=1}^{n}\langle x|b_i\rangle \cdot (\overline{\langle x|b_i\rangle} - \overline{\langle x|b_i\rangle} \cdot \langle b_i|b_i\rangle) \\ & = & \langle x|x\rangle \; - \; \sum_{i=1}^{n}\langle x|b_i\rangle\overline{\langle x|b_i\rangle} \\ & = & \| x \|^2 \; - \; \sum_{i=1}^{n} |\langle x|b_i\rangle|^2 \; . \end{array}$$

$\square$

Satz 5.1.14: *Sei V ein innerer Produktraum, sei $B = \{b_1, \ldots, b_n\}$ eine orthonormale Teilmenge, und sei $W = \mathrm{span}(B)$. Dann sind die folgenden Aussagen äquivalent:*

(1) $x \in W$;

(2) $\sum_{i=1}^{n} |\langle x|b_i\rangle|^2 = \| x \|^2$;

(3) $x = \sum_{i=1}^{n}\langle x|b_i\rangle b_i$;

(4) $\langle x|y\rangle = \sum_{i=1}^{n}\langle x|b_i\rangle\langle b_i|y\rangle$ *für alle* $y \in V$. *(***Parseval-Identität***)*

Beweis: (2) $\Longrightarrow$ (3): sei z so wie im Beweis von Satz 5.1.13. Gilt die Gleichheit in (2), dann ist $z = 0$. Das ist gleichbedeutend mit (3).
(3) $\Longrightarrow$ (4): ist $x = \sum_{i=1}^{n}\langle x|b_i\rangle b_i$, dann gilt für alle $y \in V$

$$\langle x|y\rangle = \langle\sum_{i=1}^{n}\langle x|b_i\rangle b_i|y\rangle = \sum_{i=1}^{n}\langle x|b_i\rangle\langle b_i|y\rangle \; .$$

(4) $\Longrightarrow$ (2): folgt sofort aus (4) für $y = x$.
(3) $\Longrightarrow$ (1): offensichtlich.
(1) $\Longrightarrow$ (3): ist $x = \sum_{i=1}^{n} \lambda_i b_i$, dann gilt für jedes $j \in \{1, \ldots, n\}$

$$\lambda_j = \sum_{i=1}^{n} \lambda_i\langle b_i|b_j\rangle = \langle\sum_{i=1}^{n} \lambda_i b_i|b_j\rangle = \langle x|b_j\rangle.$$

D.h. es gilt (3). $\square$

Ein Skalarprodukt erzeugt also eine Norm auf einem reellen oder komplexen Vektorraum. Wir können aber auch die wesentlichen Eigenschaften einer solchen Norm zur abstrakten Definition eines normierten Vektorraums heranziehen.

Definition 5.1.15: *Sei V ein reeller oder komplexer Vektorraum, also $K = \mathbb{R}$ oder $K = \mathbb{C}$. Eine Abbildung $\| \; \|: V \to \{x \in \mathbb{R}|x \geq 0\}$ heisst eine* **Norm** *auf V genau dann, wenn für alle $x, x' \in V$, $\lambda \in K$ gilt*

(1) $\| x \| \geq 0$, *mit Gleichheit genau dann, wenn* $x = 0$,

(2) $\| \lambda x \| = |\lambda| \cdot \| x \|$,

(3) $\| x + x' \| \leq \| x \| + \| x' \|$ *(Dreiecksungleichung).*

Ist auf V *eine Norm definiert, so heisst* V *ein* **normierter Vektorraum**.

Beispiel 5.1.16: Wir haben oben schon gesehen, dass jedes Skalarprodukt eine Norm induziert. Aber nicht jede Norm ist von einem Skalarprodukt induziert. Ein Beispiel dafür ist die **Maximumsnorm** etwa auf $\mathbb{R}^2$:

$$\| (x, y) \|_{\max} := \max(|x|, |y|).$$

Es gibt kein Skalarprodukt auf $\mathbb{R}^2$, welches die Maximumsnorm induziert. □

Mittels eines Skalarprodukts können wir auch vom Winkel zwischen zwei Elementen eines Vektorraums sprechen. Dieser so eingeführte Begriff des Winkels stimmt im Fall von $\mathbb{R}^2$ mit dem allgemein üblichen Begriff überein.

Satz 5.1.17: *Sei* V *ein reeller IP-Raum. Dann gilt für alle* $x, y \in V \setminus \{0\}$

$$-1 \leq \frac{\langle x|y \rangle}{\| x \| \cdot \| y \|} \leq 1 .$$

Beweis: das folgt unmittelbar aus der Cauchy-Schwarz-Ungleichung, Satz 5.1.8(ii). □

Definition 5.1.18: *Sei* V *ein reeller IP-Raum und seien* x, y *zwei vom Nullvektor verschiedene Elemente von* V*. Dann heisst der in* $[0, \pi]$ *eindeutig bestimmte Winkel* α *mit*

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

der **Winkel zwischen** x **und** y*, geschrieben als* $\alpha = \angle(x, y)$.

Satz 5.1.19: *Der Winkel zwischen zwei Vektoren in einem reellen IP-Raum ist unabhängig von der "Länge" des Vektors. Insbesondere können wir die Vektoren normieren vor der Bestimmung des Winkels:*

$$\angle(x, y) = \angle\left(\frac{x}{\| x \|}, \frac{y}{\| y \|}\right) \quad \textit{für alle } x, y \in V,\ x \neq 0 \neq y.$$

Beweis: Aufgrund der Gesetzmässigkeiten in einem IP-Raum gilt

$$\frac{\langle \frac{x}{\|x\|} | \frac{y}{\|y\|} \rangle}{\| \frac{x}{\|x\|} \| \cdot \| \frac{y}{\|y\|} \|} = \frac{\frac{1}{\|x\| \cdot \|y\|} \cdot \langle x|y \rangle}{\frac{\|x\|}{\|x\|} \cdot \frac{\|y\|}{\|y\|}} = \frac{\langle x|y \rangle}{\| x \| \cdot \| y \|} .$$

Daraus folgt unmittelbar die Behauptung. □

Satz 5.1.20: *In* $\mathbb{R}^2$ *mit dem kanonischen Skalarprodukt ist der Winkel zwischen zwei Vektoren unabhängig von Rotation. Dazu schreiben wir (in isomorpher Weise) die Vektoren als Spaltenvektoren. Wir betrachten für zwei vom Nullvektor verschiedene Vektoren* $x = (x_1, x_2)^T$ *und* $y = (y_1, y_2)^T$ *die um den Winkel* θ *rotierten Vektoren (vgl. Kap. 2.1)*

$$x_\theta := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad , \quad y_\theta := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$

Dann gilt $\angle(x,y) = \angle(x_\theta, y_\theta)$.

Beweis: Aus den trigonometrischen Rechengesetzen sieht man

$$\langle x_\theta | y_\theta \rangle = x_1 y_1 + x_2 y_2 = \langle x | y \rangle \qquad \text{und} \qquad \| x_\theta \| = \| x \| \ , \ \| y_\theta \| = \| y \| \ .$$

Daraus folgt unmittelbar die Behauptung. □

Aufgrund dieser Invarianzen reicht es aus, in $\mathbb{R}^2$ etwa den Winkel zwischen einem Vektor $x = (x_1, x_2)$ der Länge (Norm) 1 und dem Vektor $e_1 = (1, 0)$ zu betrachten. Durch Skizzierung auf dem Einheitskreis sieht man sofort, dass in der klassischen analytischen Geometrie für den Winkel α zwischen x und e_1 gelten sollte

$$\cos \alpha = x_1.$$

Das ergibt sich offensichtlich auch aus obiger allgemeiner Definition angewandt auf diesen Spezialfall:

$$\cos \alpha = \frac{\langle x | e_1 \rangle}{\| x \| \cdot \| e_1 \|} = \langle x | e_1 \rangle = x_1.$$

5.2 Orthonormalbasen und Fourier-Approximation

Definition 5.2.1: *Sei V ein innerer Produktraum. Eine* **orthonomale Basis** *oder* **Orthonormalbasis** *von V ist eine orthonormale Teilmenge, welche eine Basis von V darstellt.*

Beispiel 5.2.2: Die kanonischen Basen in $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathrm{Mat}_{n\times n}(\mathbb{R})$, $\mathrm{Mat}_{n\times n}(\mathbb{C})$ sind orthonormale Basen. □

Definition 5.2.3: *Sei $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ oder $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. Die Matrix A heisst* **orthogonal**, *falls die Zeilen von A bzgl. des kanonischen Skalarprodukts eine Orthonormalbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ bilden. Orthogonale Matrizen über dem Komplexen nennt man auch* **unitär**. *(Vgl. dazu Def. 2.1.26.)* □

Man beachte, dass für orthogonale Matrizen nicht nur die Orthogonalität der Zeilen gefordert wird, sondern deren Orthonormalität.

Beispiel 5.2.4: Rotationsmatrizen (vgl. Kap. 2.1)

$$R_\theta \;=\; \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

sind orthogonale Matrizen.
Die Matrix

$$S_x \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

bewirkt eine Spiegelung um die x-Achse:

$$S_x \cdot \begin{pmatrix} x \\ y \end{pmatrix} \;=\; \begin{pmatrix} x \\ -y \end{pmatrix} \;.$$

Solche Spiegelungsmatrizen sind ebenfalls orthogonale Matrizen. □

Jeder endlichdimensionale innere Produktraum besitzt eine orthonormale Basis. Eine solche orthonormale Basis kann auch aus einer gegebenen Basis konstruiert werden durch den sogenannten Gram-Schmidt Orthonormalisierungsprozess.

Satz 5.2.5 (Gram-Schmidt Orthonormalisierungsprozess): *Sei V ein innerer Produktraum. Für jeden Vektor $x \neq 0$ verwenden wir die Bezeichnung $x^* = x/\parallel x \parallel$. Sei $X = \{x_1, \ldots, x_k\}$ eine k-elementige linear unabhängige Teilmenge von V, und seien die Vektoren y_i $(1 \le i \le k)$ definiert als*

$$\begin{array}{rcl} y_1 & := & x_1^*, \\ y_2 & := & (x_2 - \langle x_2|y_1\rangle y_1)^*, \\ & \vdots & \\ y_k & := & (x_k - \sum_{i=1}^{k-1}\langle x_k|y_i\rangle y_i)^* \end{array}$$

(vergleiche Satz 5.2.8 und Def. 5.2.9).
Dann ist $Y = \{y_1, \ldots, y_k\}$ eine orthonormale Basis für $\mathrm{span}(X)$.

Beweis: Per Induktion über j, $1 \le j \le k$, zeigt man leicht:

$$\mathrm{span}(y_1, \ldots, y_j) = \mathrm{span}(x_1, \ldots, x_j) \;.$$

Die Induktionsbasis $j = 1$ ist klar. $\sum_{i=1}^{j-1}\langle x_j|y_i\rangle y_i$ ist laut Induktionshypothese in $\mathrm{span}(x_1,\ldots,x_{j-1})$. Also ist wegen der linearen Unabhängigkeit der $x_1,\ldots,x_k$ der Vektor

$$y_j \;=\; (x_j - \sum_{i=1}^{j-1}\langle x_j|y_i\rangle y_i)^*$$

ungleich dem Nullvektor und in $\mathrm{span}(x_1,\ldots,x_j) \setminus \mathrm{span}(x_1,\ldots,x_{j-1})$. Es gilt also $\mathrm{span}(x_1,\ldots,x_j) = \mathrm{span}(y_1,\ldots,y_j)$.
Offensichtlich ist aufgrund der Konstruktion jedes y_i normiert.
Es bleibt also nur mehr zu zeigen, dass Y ein Orthogonalsystem ist. Wir beweisen das mittels Induktion über k.
Für $k = 1$ ist die Behauptung trivial.
Als Induktionshypothese nehmen wir nun an, dass für ein $r \geq 2$ die Menge $\{y_1,\ldots,y_{r-1}\}$ orthogonal ist.
Setzen wir nun

$$\alpha_r \;:=\; \parallel x_r - \sum_{i=1}^{r-1}\langle x_r|y_i\rangle y_i \parallel \,,$$

dann sehen wir aus der Definition von y_r

$$\alpha_r y_r \;=\; x_r - \sum_{i=1}^{r-1}\langle x_r|y_i\rangle y_i \;.$$

Daher gilt für jedes $j < r$

$$\alpha_r\langle y_r|y_j\rangle \;=\; \langle x_r|y_j\rangle - \sum_{i=1}^{r-1}\langle x_r|y_i\rangle\langle y_i|y_j\rangle \;=\; \langle x_r|y_j\rangle - \langle x_r|y_j\rangle \;=\; 0 \;.$$

Nun ist aber $\alpha_r \neq 0$, und daher ist $\langle y_r|y_j\rangle = 0$.
Somit ist $\{y_1,\ldots,y_r\}$ orthogonal. □

Korollar: *Jeder endlichdimensionale innere Produktraum hat eine orthonormale Basis.*

Beweis: Man wende den Gram-Schmidt-Orthonormalisierungsprozess an. □

Beispiel 5.2.6: (a) Im inneren Produktraum $\mathbb{R}^3$ bzgl. des kanonischen inneren Produkts gehen wir aus von der Basis

$$x_1 = (0,1,1), \quad x_2 = (1,0,1), \quad x_3 = (1,1,0).$$

Daraus wollen wir eine orthonormale Basis $\{y_1,y_2,y_3\}$ konstruieren mittels des Gram-Schmidt-Orthonormalisierungsprozesses.
Zunächst setzen wir

$$y_1 := \frac{x_1}{\parallel x_1 \parallel} = \frac{1}{\sqrt{2}}(0,1,1).$$

Danach bestimmen wir

$$\begin{aligned} x_2 - \langle x_2|y_1\rangle y_1 &= (1,0,1) - \tfrac{1}{\sqrt{2}}\langle(1,0,1)|(0,1,1)\rangle\tfrac{1}{\sqrt{2}}(0,1,1) \\ &= (1,0,1) - \tfrac{1}{2}(0,1,1) \\ &= \tfrac{1}{2}(2,-1,1)\,, \end{aligned}$$

und erhalten durch Normalisierung

$$y_2 := \frac{1}{\sqrt{6}}(2,-1,1).$$

Schliesslich bestimmen wir

$$x_3 - \langle x_3|y_2\rangle y_2 - \langle x_3|y_1\rangle y_1 = \frac{2}{3}(1,1,-1),$$

und erhalten durch Normalisierung

$$y_3 := \frac{1}{\sqrt{3}}(1,1,-1).$$

Somit haben wir eine orthonormale Basis $\{y_1, y_2, y_3\}$ gefunden.
(b) Sei $F = \{f_1 = 1, f_2 = x, f_3 = x^2\}$ Basis eines Teilraums W des reellen Vektorraums $V = C([0,1],\mathbb{R})$ bzgl. des Skalarprodukts

$$\langle f|g\rangle \ = \ \int_0^1 fg\,dx \ .$$

Wir wollen mithilfe des Gram-Schmidt Orthonormalisierungsprozesses eine orthonormale Basis G für W berechnen:

$$\begin{aligned} g_1 &= f_1^* \ = \ 1 \ , \\ g_2 &= (f_2 \ - \ \langle f_2|g_1\rangle \cdot g_1)^* \ = \ 2\sqrt{3}(x-\tfrac{1}{2}) \ , \\ g_3 &= (f_3 \ - \ \langle f_3|g_1\rangle \cdot g_1 - \ \langle f_3|g_2\rangle \cdot g_2)^* \ = \ 6\sqrt{5}(x^2 - x + \tfrac{1}{6}) \ . \end{aligned}$$

□

Satz 5.2.7 (Pythagoras): (Pythagoras von Samos, ca 580–500 v.Chr.) *Sei V ein IP-Raum und x, y orthogonale Elemente in $V \setminus \{0\}$, also $x \perp y$. Dann gilt*

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

Beweis: $\| \ x+y \ \|^2 = \langle x+y|x+y\rangle = \langle x|x\rangle + \langle x|y\rangle + \langle y|x\rangle + \langle y|y\rangle = \| \ x \ \|^2 + \| \ y \ \|^2$. □

Im folgenden geben wir zu einem Vektor $v \in V$ und dem von einer Orthonormalbasis aufgespannten Teilraum W die "beste" Approximation in W zu v an.

Satz 5.2.8: *Sei $S = \{s_1, \ldots, s_m\}$ ein Orthonormalsystem im IP-Raum V. Sei $v \in V$ und sei $v_S := \sum_{i=1}^m \langle v|s_i\rangle s_i$. Dann gilt für alle $t \in \operatorname{span}(S)$:*

(i) *$(v - v_S) \perp t$,*

(ii) *$\| \ v - v_S \ \| \leq \| \ v - t \ \|$.*

Beweis: zu (i): Sei $t = \sum_{i=1}^m \lambda_i s_i$. Dann gilt

$$\begin{aligned} \langle v - v_S|t\rangle &= \langle v|t\rangle - \langle v_S|t\rangle \\ &= \textstyle\sum_{i=1}^m \overline{\lambda_i}\langle v|s_i\rangle - \sum_{i=1}^m \langle v|s_i\rangle\langle s_i|t\rangle \\ &= \textstyle\sum_{i=1}^m \overline{\lambda_i}\langle v|s_i\rangle - \sum_{i=1}^m \langle v|s_i\rangle\langle s_i|\lambda_i s_i\rangle \\ &= \textstyle\sum_{i=1}^m \overline{\lambda_i}\langle v|s_i\rangle - \sum_{i=1}^m \overline{\lambda_i}\langle v|s_i\rangle \\ &= 0 \ . \end{aligned}$$

zu (ii): $t - v_S \in \mathrm{span}(S)$, also wegen (i) ist

$$v - v_S \ \perp \ t - v_S \ .$$

Weiters haben wir

$$v - t \ = \ (v - v_S) + (v_S - t) \ .$$

Wegen Satz 5.2.7 (Pythagoras) gilt also

$$\| \ v - t \ \|^2 = \| \ v - v_S \ \|^2 + \| \ v_S - t \ \|^2 \ ,$$

und daher

$$\| \ v{-}t \ \| \geq \| \ v{-}v_S \ \| \qquad \square.$$

Bemerkung: Wenn wir also den Gram-Schmidt Orthogonalisierungsprozess noch einmal betrachten, so sehen wir, dass wir jeweils die besten Approximationen von x_i auf den bisher erzeugten Teilraum $\mathrm{span}\{y_1, \ldots, y_{i-1}\}$ betrachten, und y_i dann die normierte Version von x_i minus diese beste Approximation ist.

Definition 5.2.9: *Sei $S = \{s_1, \ldots, s_m\}$ ein Orthonormalsystem im IP-Raum V. Sei $v \in V$ und sei $v_S := \sum_{i=1}^{m} \langle v|s_i \rangle s_i$. Der Vektor v_S heisst die* **orthogonale Projektion** *von v auf $\mathrm{span}(S)$ oder die* **Fourier-Approximation** *(Jean-Baptiste Fourier, 1768–1830) von v bzgl. S. Wie wir oben gesehen haben ist v_S die beste Approximation zu v in $\mathrm{span}(S)$. Für $x \in \mathrm{span}(S)$ heisst die Identität (3) in Satz 5.1.14*

$$x \ = \ \sum_{i=1}^{n} \langle x|s_i \rangle s_i$$

die **Fourier-Entwicklung** *von x bzgl. S.*
Die Skalare $\langle x|s_i \rangle$ heissen dabei die **Fourier-Koeffizienten** *von x.*

Beispiel 5.2.6 (Fortsetzung): Wie in 5.2.6(b) betrachten wir den Teilraum W von V. Nun wollen wir für

$$h \ = \ x^4$$

die Fourierentwicklung bzgl. der Basis G von W, also die beste Approximation h' zu h in W berechnen:

$$\begin{aligned} h' \ &= \ \langle h|g_1 \rangle \cdot g_1 + \langle h|g_2 \rangle \cdot g_2 + \langle h|g_3 \rangle \cdot g_3 \\ &= \ \tfrac{1}{5} g_1 + \tfrac{2\sqrt{3}}{15} g_2 + \tfrac{2\sqrt{5}}{35} g_3 \\ &= \ \tfrac{12}{7} x^2 - \tfrac{32}{35} x + \tfrac{3}{35} \end{aligned} \qquad \square$$

Satz 5.2.10: *Sei V ein endlichdimensionaler innerer Produktraum der Dimension n. Zu jeder orthonormalen Teilmenge $\{x_1, \ldots, x_k\}$ gibt es $x_{k+1}, \ldots, x_n \in V$, sodass $\{x_1, \ldots, x_n\}$ eine orthonormale Basis von V ist.*

Beweis: Man erweitert $\{x_1, \ldots, x_k\}$ zu einer Basis von V und wendet darauf Gram-Schmidt an. Dabei bleiben die Elemente $x_1, \ldots, x_k$ unverändert. $\square$

Definition 5.2.11: *Seien V und W innere Produkträume über demselben Körper. Ein Vektorraumisomorphismus $f : V \to W$ ist ein* **innerer Produktisomorphismus** *genau dann, wenn f innere Produkte erhält, also wenn gilt*

$$\langle f(x)|f(y)\rangle = \langle x|y\rangle \quad \text{für alle } x, y \in V.$$

Satz 5.2.12: *Seien V und W endlichdimensionale innere Produkträume über demselben Körper mit $\dim(V) = n = \dim(W)$, sei $B = \{b_1, \ldots, b_n\}$ eine orthonormale Basis von V, und sei $f \in \mathrm{Hom}(V, W)$ eine lineare Abbildung.*
Dann ist f ein innerer Produktisomorphismus genau dann, wenn $f(B) = \{f(b_1), \ldots, f(b_n)\}$ eine orthonormale Basis von W ist.

Beweis: "$\Longrightarrow$": Ist f ein innerer Produktisomorphismus, dann ist $f(B)$ sicherlich eine Orthonormalbasis von W.
"$\Leftarrow$": Sei also nun $f(B)$ eine Orthonormalbasis von W. Dann ist f auf jeden Fall einmal ein Vektorraumisomorphismus von V nach W. Für alle $x \in V$ wenden wir die Fourier-Entwicklung bzgl. der Orthonormalbasis B an und erhalten

$$\begin{aligned}
\langle f(x)|f(b_j)\rangle &= \langle f\big(\textstyle\sum_{i=1}^n \langle x|b_i\rangle b_i\big)|f(b_j)\rangle \\
&= \langle \textstyle\sum_{i=1}^n \langle x|b_i\rangle f(b_i)|f(b_j)\rangle \\
&= \textstyle\sum_{i=1}^n \langle x|b_i\rangle \langle f(b_i)|f(b_j)\rangle \\
&= \langle x|b_j\rangle \; .
\end{aligned}$$

Ebenso erhalten wir $\langle f(b_j)|f(x)\rangle = \langle b_j|x\rangle$. Wendet man nun Satz 5.1.14(4) (Parseval-Identität) sowohl in V als auch in W an, so ergibt sich

$$\begin{aligned}
\langle f(x)|f(y)\rangle &= \textstyle\sum_{j=1}^n \langle f(x)|f(b_j)\rangle \langle f(b_j)|f(y)\rangle \\
&= \textstyle\sum_{j=1}^n \langle x|b_j\rangle \langle b_j|y\rangle \\
&= \langle x|y\rangle \; .
\end{aligned}$$

Somit ist f also ein innerer Produktisomorphismus. □

Beispiel 5.2.13: Wir betrachten noch ein Beispiel zur Fourierentwicklung. Als IP-Raum nehmen wir $C([0, 2\pi], \mathbb{R})$ mit dem in Beispiel 5.1.4(a) definierten inneren Produkt

$$\langle f|g\rangle = \int_0^{2\pi} fg \; .$$

Die Funktionen $e(x), s_k(x), c_k(x)$ seien erklärt als

$$e(x) = 1, \quad s_k(x) = \sin(kx), \quad c_k(x) = \cos(kx) \quad \text{für } k \in \mathbb{N} \setminus \{0\} \; .$$

Dann ist

$$\tilde{S} = \{e\} \;\cup\; \{s_k | 1 \le k\} \;\cup\; \{c_k | 1 \le k\}$$

eine orthogonale Teilmenge von $C([0,2\pi],\mathbb{R})$ (Details siehe Übungen).
Normieren wir die Elemente in $\tilde{S}$, so erhalten wir das Orthonormalsystem

$$S = \{\frac{1}{\sqrt{2\pi}}\} \ \cup\ \{\frac{1}{\sqrt{\pi}}s_k \mid 1 \leq k\} \ \cup\ \{\frac{1}{\sqrt{\pi}}c_k \mid 1 \leq k\}$$

(man beachte, dass die Elemente von S Funktionen von $[0,2\pi]$ nach $\mathbb{R}$ sind).
Sei nun W_n der $(2n+1)$-dimensionale Teilraum von $C([0,2\pi],\mathbb{R})$, welcher aufgespannt wird durch die Orthonormalbasis

$$S_n = \{\frac{1}{\sqrt{2\pi}}\} \ \cup\ \{\frac{1}{\sqrt{\pi}}s_k \mid 1 \leq k \leq n\} \ \cup\ \{\frac{1}{\sqrt{\pi}}c_k \mid 1 \leq k \leq n\}\ .$$

Dann erhalten wir die beste Approximation zu einer Funktion $f(x) \in C([0,2\pi],\mathbb{R})$ als das Element

$$f_n(x) \ = \ \lambda_0\frac{1}{\sqrt{2\pi}} + \sum_{k=1}^{n}\lambda_{k,1}\frac{1}{\sqrt{\pi}}\sin(kx) + \sum_{k=1}^{n}\lambda_{k,2}\frac{1}{\sqrt{\pi}}\cos(kx)$$

wobei

$$\lambda_0 = \langle f(x)|\tfrac{1}{\sqrt{2\pi}}\rangle = \tfrac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x)dx\ ,$$

$$\lambda_{k,1} = \langle f(x)|\tfrac{1}{\sqrt{\pi}}\sin(kx)\rangle = \tfrac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\sin(kx)dx \quad \forall 1 \leq k \leq n\ ,$$

$$\lambda_{k,2} = \langle f(x)|\tfrac{1}{\sqrt{\pi}}\cos(kx)\rangle = \tfrac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\cos(kx)dx \quad \forall 1 \leq k \leq n\ .$$

Ist $f(x)$ unendlich oft differenzierbar, dann ist $(f_n(x))_{n\geq 1}$ eine Cauchy-Folge mit Limes $f(x)$. Wir können also dann schreiben

$$f(x) \ = \ \lambda_0\frac{1}{\sqrt{2\pi}} + \sum_{k\geq 1}\lambda_{k,1}\frac{1}{\sqrt{\pi}}\sin(kx) + \sum_{k\geq 1}\lambda_{k,2}\frac{1}{\sqrt{\pi}}\cos(kx)\ .$$

Das ist die **Darstellung der Funktion $f(x)$ als Fourier-Reihe**. □