Prof. Dr. H. Brenner **Osnabrück WS 2022/2023**

Maß- und Integrationstheorie

Vorlesung 22

Orthonormalsysteme

DEFINITION 22.1. Es sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle -,-\rangle$. Eine Familie von Vektoren v_i, $i \in I$, von V heißt *Orthonormalsystem*, wenn

$$\langle v_i, v_i\rangle = 1 \text{ für alle } i \in I \text{ und } \langle v_i, v_j\rangle = 0 \text{ für } i \neq j$$

gilt.

Zu einem gegebenen Orthonormalsystem v_i, $i \in I$, und einem Vektor $v \in V$ spielen die Koeffizienten $\langle v, v_i, \rangle$ eine wichtige Rolle, man spricht von den *Fourierkoeffizienten* des Vektors bezüglich des Systems, wobei diese Sprechweise insbesondere im Kontext von Fourierreihen verwendet wird. Eine wichtige Frage ist, in welcher Beziehung v zu $\sum_{i \in I} \langle v, v_i, \rangle v_i$ steht, wobei bei I unendlich zuerst zu klären ist, in welchem Sinne eine solche unendlich Summe verstanden werden kann. Im endlichen Fall haben wir folgende Beschreibung, auf die man weitere Resultate zurückführen kann.

LEMMA 22.2. *Es sei V ein $\mathbb{K}$-Vektorraum mit einem Skalarprodukt $\langle -,-\rangle$, sei v_i, $i \in E$, ein endliches Orthonormalsystem mit dem davon erzeugten Untervektorraum U. Dann gilt für die orthogonale Projektion*

$$p_U(v) = \sum_{i \in E} \langle v, v_i\rangle v_i .$$

Beweis. Es sei

$$w = v - \sum_{i \in E} \langle v, v_i\rangle v_i ,$$

es ist nach Korollar 21.8 lediglich zu zeigen, dass w orthogonal zu den v_j ist. Dies ergibt sich direkt aus

$$\begin{aligned} \langle w, v_j\rangle &= \left\langle v - \sum_{i \in E} \langle v, v_i\rangle v_i, v_j \right\rangle \\ &= \langle v, v_j\rangle - \sum_{i \in E} \langle \langle v, v_i\rangle v_i, v_j\rangle \\ &= \langle v, v_j\rangle - \langle v, v_j\rangle \\ &= 0. \end{aligned}$$

□

Beispiel 22.3. Es sei E eine endliche Menge und

$$V = \mathbb{K}^E \cong \mathbb{K}^{\#(E)}$$

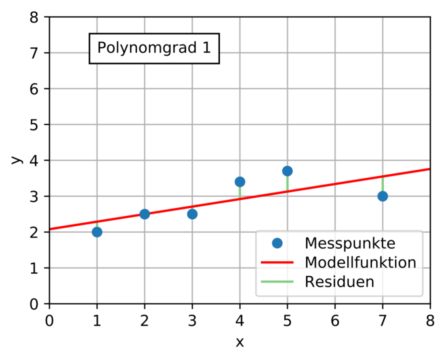
die Menge der $\mathbb{K}$-wertigen Funktionen auf E, versehen mit dem Standardskalarprodukt. Eine Funktion $f \in V$ kann einfach durch eine vollständige Wertetabelle beschrieben werden. Es kann aber auch sinnvoll sein, die Funktion f durch eine Funktion g aus einem vorgegebenen Untervektorraum $U \subseteq V$ zu approximieren. Dabei liefert das Skalarprodukt und die zugehörige orthogonale Projektion auf U ein naheliegendes Hilfsmittel, um eine optimale Approximation zu finden. Nach Korollar 21.7 ist $p_U(f)$ diejenige Funktion, die unter allen Funktionen aus U zu f den minimalen Abstand besitzt, wobei der Abstand zu f über das Skalarprodukt gegeben ist, also durch

$$\|g - f\|^2 = \sum_{i \in E} |g_i - f_i|^2 .$$

Wenn g_j, $j \in J$, eine Orthonormalbasis von U ist, so ist

$$g = p_U(f) = \sum_{j \in J} \langle f, g_j \rangle \, g_j$$

nach Lemma 22.2 die beste Approximation. Das so bestimmte g minimiert also die Summe der einzelnen Differenzquadrate, man spricht von der *Methode der kleinsten Fehlerquadrate*.



Eine typische Anwendung ist, wenn E Messtellen repräsentiert, etwa $E \subseteq \mathbb{R}^d$, und f_e Messergebnisse, die eventuell fehlerhaft sein können. Man weiß aus physikalischen Gründen, dass die Abhängigkeit einer gewissen Gesetzmäßigkeit gehorchen muss, beispielsweise ein linearer Zusammenhang sein muss oder als Flugbahn eines Planeten eine Ellipse sein muss oder ähnliches. Diese Gesetzmäßigkeit legt den (typischerweise niedrigdimensionalen) Untervektorraum U fest, in dem nach einer optimalen Approximation gesucht wird, das den Messergebnissen möglichst nahe kommt.

Beispiel 22.4. Von einer unbekannten Funktion $f\colon \mathbb{R} \to \mathbb{R}$ sei der Datensatz $(1,11)$, $(4,20)$, $(6,23)$ gegeben und es sei bekannt, dass f eine affin-lineare Funktion sein muss. Der Datensatz beruht auf Messungen, in denen Fehler und Ungenauigkeiten vorkommen können, die drei Punkte liegen nicht wirklich auf einer Geraden. Es wird nach der affin-linearen Funktion $ax+b$ gesucht, die gut zu den Daten passt. Wir betrachten die Abbildung

$$\Psi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3,\ (a,b) \longmapsto (a+b, 4a+b, 6a+b),$$

die einem Parameterpaar (a,b), das die affin-lineare Funktion $ax+b$ repräsentiert, die Auswertung an den drei Punkten $(1,4,6)$ zuordnet. Dabei ist Ψ eine injektive lineare Abbildung und das Bild $U = \operatorname{bild} \varphi$ ist ein zweidimensionaler Untervektorraum von $\mathbb{R}^3$. Diese Ebene steht senkrecht zum Vektor $\begin{pmatrix} -2 \\ 5 \\ -3 \end{pmatrix}$, eine Basis ist durch $\begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix}$ gegeben (die unter Ψ von der Basis $\begin{pmatrix} -1 \\ 6 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ des $\mathbb{R}^2$ herrührt). Die optimale Approximation (im Sinne der euklidischen Norm bzw. im Sinne der kleinsten Fehlerquadrate) ist die orthogonale Projektion des Wertetupels $\begin{pmatrix} 11 \\ 20 \\ 23 \end{pmatrix}$ auf die Ebene. Dies führt zum linearen Gleichungssystem

$$\begin{pmatrix} 11 \\ 20 \\ 23 \end{pmatrix} = \alpha \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} + \gamma \begin{pmatrix} -2 \\ 5 \\ -3 \end{pmatrix}$$

mit den Lösungen $\alpha = \frac{218}{95}$, $\beta = \frac{901}{190}$ und $\gamma = \frac{9}{38}$. Daher ist

$$\begin{aligned} p_U \begin{pmatrix} 11 \\ 20 \\ 23 \end{pmatrix} &= \frac{218}{95} \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} + \frac{901}{190} \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} \\ &= \frac{1}{190} \begin{pmatrix} 2180 \\ 3575 \\ 4505 \end{pmatrix} \\ &= \frac{1}{38} \begin{pmatrix} 436 \\ 715 \\ 901 \end{pmatrix}. \end{aligned}$$

Der entsprechende Punkt auf dem $\mathbb{R}^2$ ist

$$\begin{aligned} \frac{218}{95} \begin{pmatrix} -1 \\ 6 \end{pmatrix} + \frac{901}{190} \begin{pmatrix} 1 \\ -1 \end{pmatrix} &= \frac{1}{190} \begin{pmatrix} -436+901 \\ 2616-901 \end{pmatrix} \\ &= \frac{1}{190} \begin{pmatrix} 465 \\ 1715 \end{pmatrix} \\ &= \frac{1}{38} \begin{pmatrix} 93 \\ 343 \end{pmatrix}. \end{aligned}$$

Die beste Approximation ist also

$$f(x) = \frac{93}{38}x + \frac{343}{38}.$$

Es ist $f(1) = \frac{436}{38} \sim 11,473$, $f(4) = \frac{715}{38} \sim 18,815$ und $f(6) = \frac{901}{38} \sim 23,710$.

Satz 22.5. *Es seien $x_1, \ldots, x_n$ verschiedene reelle Zahlen, $n \geq 2$, und $y_1, \ldots, y_n$ reelle Zahlen. Es sei*[1] $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ *und* $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$. *Dann ist die affin-lineare Funktion $ax + b$ mit*

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

und

$$b = \bar{y} - a\bar{x}$$

die optimale lineare Approximation für den Datensatz

$$f(x_i) = y_i$$

im Sinne der minimalen Fehlerquadrate. D.h. die Summe der Fehlerquadrate $\sum_{i=1}^n (ax_i + b - y_i)^2$ wird für die angegebenen Koeffizienten a und b minimal.

Beweis. Wir betrachten die Abbildung

$$\psi \colon \mathbb{R}^2 \longrightarrow \mathbb{R}^n,\ (a, b) \longmapsto (ax_1 + b, \ldots, ax_n + b).$$

Diese Abbildung ist linear und injektiv, da

$$\psi(1, 0) = (x_1, \ldots, x_n) =: v_1$$

und

$$\psi(0, 1) = (1, \ldots, 1) =: v_2$$

linear unabhängig sind. Es sei

$$U = \operatorname{bild} \psi = \langle v_1, v_2 \rangle \subseteq \mathbb{R}^n.$$

Es geht darum, die orthogonale Projektion von $y = (y_1, \ldots, y_n)$ auf U zu bestimmen. Der Vektor

$$u_2 = \frac{v_2}{\| v_2 \|} = \left(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}} \right)$$

ist normiert. Wegen

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \bar{x} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\rangle = 0$$

bildet

$$u_1 := \frac{v_1 - \bar{x} v_2}{\| v_1 - \bar{x} v_2 \|} = \frac{1}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \begin{pmatrix} x_1 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}$$

[1] $\bar{x}$ bzw. $\bar{y}$ bezeichnen also den Durchschnitt der Messstellen bzw. der Messwerte.

zusammen mit u_2 eine Orthonormalbasis von U. Es entspricht u_2 der konstanten Funktion $\frac{1}{\sqrt{n}}$ und u_1 der affin-linearen Funktion $\frac{1}{\sqrt{\sum_{i=1}^n (x_i-\bar{x})^2}}(x-\bar{x})$. Nach Lemma 22.2 ist

$$p_U(y) = \langle y, u_1\rangle u_1 + \langle y, u_2\rangle u_2,$$

dabei ist

$$\langle y, u_1\rangle = \frac{\sum_{i=1}^n y_i(x_i-\bar{x})}{\sqrt{\sum_{i=1}^n (x_i-\bar{x})^2}}$$

und

$$\langle y, u_2\rangle = \frac{\sum_{i=1}^n y_i}{\sqrt{n}}.$$

Somit ist die optimale affin-lineare Funktion gleich

$$\frac{\sum_{i=1}^n y_i(x_i-\bar{x})}{\sqrt{\sum_{i=1}^n (x_i-\bar{x})^2}} \cdot \frac{1}{\sqrt{\sum_{i=1}^n (x_i-\bar{x})^2}}(x-\bar{x}) + \frac{\sum_{i=1}^n y_i}{\sqrt{n}} \cdot \frac{1}{n},$$

also ist

$$a = \frac{\sum_{i=1}^n y_i(x_i-\bar{x})}{\sum_{i=1}^n (x_i-\bar{x})^2}$$

und

$$b = \bar{y} - a\bar{x}.$$

□

Den Graphen der approximierenden affin-linearen Funktion im vorstehenden Satz nennt man *Ausgleichsgerade*.

Beispiel 22.6. In der Situation von Beispiel 22.4 kommt man mit Satz 22.5 deutlich schneller ans Ziel. Es ist

$$\bar{x} = \frac{1+4+6}{3} = \frac{11}{3}$$

und daher

$$\begin{aligned} a &= \frac{\left(1-\frac{11}{3}\right)\cdot 11 + \left(4-\frac{11}{3}\right)\cdot 20 + \left(6-\frac{11}{3}\right)\cdot 23}{\left(1-\frac{11}{3}\right)^2 + \left(4-\frac{11}{3}\right)^2 + \left(6-\frac{11}{3}\right)^2} \\ &= \frac{-\frac{8}{3}\cdot 11 + \frac{1}{3}\cdot 20 + \left(\frac{7}{3}\right)\cdot 23}{\left(-\frac{8}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{7}{3}\right)^2} \\ &= 3\cdot\frac{-88+20+161}{64+1+49} \\ &= 3\cdot\frac{93}{114} \\ &= \frac{93}{38} \end{aligned}$$

und

$$b = \bar{y} - a\bar{x} = 18 - \frac{93}{38}\cdot\frac{11}{3} = \frac{2052-1023}{114} = \frac{1029}{114} = \frac{343}{38}.$$

Vollständige Orthonormalsysteme

DEFINITION 22.7. Ein Orthonormalsystem v_i, $i \in I$, in einem $\mathbb{K}$-Vektorraum V mit Skalarprodukt heißt *vollständig* oder eine *Hilbertbasis*, wenn der von den v_i erzeugte Untervektorraum dicht in V ist.

Die folgende Aussage heißt *Besselsche Abschätzung.*

LEMMA 22.8. *Es sei V ein $\mathbb{K}$-Vektorraum mit einem Skalarprodukt $\langle -, - \rangle$ und sei v_i, $i \in I$, ein Orthonormalsystem. Dann ist für jeden Vektor $v \in V$ die Familie $|\langle v, v_i \rangle|^2$, $i \in I$, summierbar und es gilt*

$$\sum_{i \in I} |\langle v, v_i \rangle|^2 \, \leq \, \|v\|^2 .$$

Beweis. Für jede endliche Teilmenge $E \subseteq I$ schreiben wir

$$v \, = \, \sum_{i \in E} \langle v, v_i \rangle \, v_i + w$$

(dabei ist w orthogonal zu $\langle v_i, i \in E \rangle$ und hängt von E ab) und erhalten aufgrund der Orthogonalitätsbeziehungen

$$\begin{aligned} \sum_{i \in E} |\langle v, v_i \rangle|^2 \, &= \, \sum_{i \in E} |\langle v, v_i \rangle|^2 \, \langle v_i, v_i \rangle \\ &\leq \, \sum_{i \in E} \langle \langle v, v_i \rangle \, v_i, \langle v, v_i \rangle \, v_i \rangle + \langle w, w \rangle \\ &= \, \left\langle \sum_{i \in E} \langle v, v_i \rangle \, v_i + w, \sum_{i \in E} \langle v, v_i \rangle \, v_i + w \right\rangle \\ &= \, \langle v, v \rangle \\ &= \, \|v\|^2 . \end{aligned}$$

Nach Aufgabe 17.14 (Analysis (Osnabrück 2021-2023)) ist die Familie summierbar und ihre Summe ist durch $\|v\|^2$ beschränkt. □

KOROLLAR 22.9. *Es sei V ein $\mathbb{K}$-Hilbertraum und sei v_i, $i \in I$, ein Orthonormalsystem. Dann ist zu einem Vektor $v \in V$ die Vektorenfamilie $\langle v, v_i \rangle \, v_i$, $i \in I$, summierbar.*

Beweis. Für jede endliche Teilmenge $E \subseteq I$ ist

$$\| \sum_{i \in E} \langle v, v_i \rangle \, v_i \|^2 \, = \, \sum_{i \in E} \| \langle v, v_i \rangle \, v_i \|^2 \, = \, \sum_{i \in E} |\langle v, v_i \rangle|^2 \, ,$$

was nach Lemma 22.8 durch $\|v\|^2$ beschränkt ist. Daher ist die Familie eine Cauchy-Familie und somit wegen der Vollständigkeit des Raumes nach Lemma 19.3 summierbar. □

Im Allgemeinen gibt es keinen direkten Zusammenhang zwischen v und $\sum_{i \in I} \langle v, v_i \rangle \, v_i$, man denke etwa an kleine Orthonormalsysteme. Der folgende Satz charakterisiert die vollständigen Orthonormalsysteme.

Satz 22.10. *Es sei V ein $\mathbb{K}$-Vektorraum mit einem Skalarprodukt $\langle -, - \rangle$ und sei v_i, $i \in I$, ein Orthonormalsystem. Dann sind folgende Aussagen äquivalent.*

(1) *Die Familie ist vollständig.*
(2) *Für jedes $v \in V$ gilt*

$$v = \sum_{i \in I} \langle v, v_i \rangle v_i .$$

(3) *Für jedes $v \in V$ gilt*

$$\| v \|^2 = \sum_{i \in I} | \langle v, v_i \rangle |^2 .$$

Beweis. Von (1) nach (2). Die Vollständigkeit des Orthonormalsystems bedeutet, dass es zu jedem Vektor $v \in V$ und jedem $\epsilon > 0$ ein Koeffiziententupel a_i mit einer endlichen Trägermenge $E \subseteq I$ mit

$$\| v - \sum_{i \in E} a_i v_i \| \leq \epsilon$$

gibt. Nach Lemma 22.2 erfüllt erst recht $\sum_{i \in E} \langle v, v_i \rangle v_i$ diese Eigenschaft. Dies heißt aber, dass die Summe $\sum_{i \in I} \langle v, v_i \rangle v_i$ gleich v ist. Von (2) nach (1) ergibt sich aus Lemma 19.4.

Zum Nachweis der Äquivalenz von (2) und (3) ziehen wir für eine endliche Teilmenge $E \subseteq I$ die Gleichung

$$\| v - \sum_{i \in E} \langle v, v_i \rangle v_i \|^2 = \| v \|^2 - \sum_{i \in E} | \langle v, v_i \rangle |^2$$

heran. (2) bedeutet, dass die linke Seite beliebig klein wird, (3) bedeutet, dass die rechte Seite beliebig klein wird, daher sind die Eigenschaften äquivalent. $\square$

Die Gleichung in (3) des vorstehenden Satzes nennt man auch *Parsevalsche Gleichung*.

Lemma 22.11. *Es sei v_i, $i \in I$, ein Orthonormalsystem in einem $\mathbb{K}$-Hilbertraum V. Dann kann man das System zu einem vollständigen Orthonormalsystem ergänzen.*

Beweis. Siehe Aufgabe 22.12. $\square$

Bemerkung 22.12. In einem $\mathbb{K}$-Vektorraum V mit Skalarprodukt kann man ein gegebenes System von linear unabhängigen Vektoren v_n, $n \in \mathbb{N}$, mit Hilfe des Orthonomalisierungsverfahrens im endlichdimensionalen Fall in ein abzählbar unendliches Orthonormalsystem überführen. Speziell kann man in einem separablen Hilbertraum aus jeder linear unabhängigen Familie, die einen dichten Untervektorraum erzeugt, ein abzählbares vollständiges Orthonormalsystem gewinnen.

Definition 22.13. Es sei V ein $\mathbb{K}$-Hilbertraum und sei v_i, $i \in I$, ein vollständiges Orthonormalsystem. Dann nennt man zu $v \in V$ die Darstellung

$$v = \sum_{i \in I} \langle v, v_i \rangle \, v_i$$

die *Fourierentwicklung* von v und die rechte Seite eine *Fouriersumme*. Die Koeffizienten $\langle v, v_i \rangle$ heißen *Fourierkoeffizienten*.

Im separablen Fall, wenn das vollständige Orthonormalsystem abzählbar ist und durch v_n, $n \in \mathbb{N}$, (oder $\mathbb{Z}$ als geordnete Indexmenge) gegeben ist, so nennt man die Darstellung

$$v = \sum_{n \in \mathbb{N}} \langle v, v_n \rangle \, v_n$$

auch die Fourierreihe zu v bezüglich des gegebenen Systems. Die Sprechweise wird insbesondere bei periodischen Funktionen der Periodenlänge 1 mit dem trigonometrischen Orthonormalsystem verwendet, siehe insbesondere Satz 23.6. Bei anderer Periodenlänge ist der Sprachgebrauch nicht einheitlich.

Abbildungsverzeichnis