Approximationstheorie

Inhaltsverzeichnis

1 Einführung **1**

2 Approximation in normierten Räumen **7**

3 Approximation im Hilbertraum **14**

3.1 Best-Approximation und Orthogonal-Projektion . . . 14

3.2 Der Hilbertraum $L^2(0,1)$ . . . 18

4 Approximation in $C_*[0,1]$ **28**

4.1 Jackson-Sätze . . . 31

4.2 Umkehrsätze . . . 40

4.3 Weitere Aussagen . . . 45

5 Approximation in $C[a,b]$ **47**

5.1 Timan-Sätze . . . 48

5.2 Sätze von Ditzian und Totik . . . 53

5.3 Aussagen zum Bernstein-Operator . . . 54

5.4 Markoff- und Bernstein-Ungleichungen . . . 55

6 Charakterisierung der Bestapproximation **57**

6.1 Charakterisierung der Bestapproximation in $L^p(D)$ . . . 61

6.2 Charakterisierung der Bestapproximation in $C(K)$ . . . 65

6.3 Bestapproximation mit Haarschen Räumen . . . 69

6.3.1 Haarsche Räume und der Eindeutigkeitssatz von Haar . . . 70

6.3.2 Existenz positiver Funktionen . . . 76

6.3.3 Charakterisierung des Proximums . . . 80

6.3.4 Remez-Algorithmus . . . 84

7 Der Bernstein-Operator und die Sätze von Weierstraß und Korovkin **95**

1 Einführung

Die Approximationstheorie hat das Ziel, zu gegebenen Funktionen (oder diskreten Daten oder ganzen Funktionsklassen) Methoden zur Konstruktion einfacher Funktionen anzugeben, die die gegebenen Funktionen gut approximieren und eventuell weitere Eigenschaften, wie z.B. Monotonie, erhalten.

1.1 Beispiel. Als einfaches Beispiel zur Einführung wollen wir drei verschiedene Methoden zur Approximation der Funktion

$$f : [-1, 1] \longrightarrow \mathbb{R}, \quad f(x) = \frac{1}{a - x} \qquad \text{(mit festem } a > 1\text{)}$$

durch Polynome vom Grad n angeben.

a) **Taylorentwicklung**.
f besitzt im Intervall $[-1, 1] \subset (-a, a)$ die Taylorentwicklung (mit Entwicklungspunkt $x_0 = 0$)

$$f(x) = \frac{1}{a - x} = \frac{1}{a\,(1 - \frac{x}{a})} = \frac{1}{a} \sum_{k=0}^{\infty} \left(\frac{x}{a}\right)^k = \sum_{k=0}^{\infty} \frac{x^k}{a^{k+1}}.$$

Wegen $|x/a| < 1$ konnte hierbei die geometrische Reihe zu Hilfe genommen werden, um die Taylorreihe von f hinzuschreiben. Die Partialsumme der Taylorreihe ist das n-te Taylorpolynom

$$\frac{1}{a - x} \approx \sum_{k=0}^{n} \frac{x^k}{a^{k+1}} =: p_n^{(T)}(f, x_0, x). \tag{1.1}$$

Der *Approximationsfehler* wird berechnet als die Maximumsnorm von $f - p_n^{(T)}$ auf dem Intervall $[-1, 1]$, also

$$\|f - p_n^{(T)}\|_\infty = \max_{x \in [-1,1]} \left| \sum_{k=n+1}^{\infty} \frac{x^k}{a^{k+1}} \right| = \sum_{k=n+1}^{\infty} \frac{1}{a^{k+1}} = \frac{a^{-n-1}}{a - 1} =: E_n^{(T)}(f). \tag{1.2}$$

b) **Orthogonalentwicklung nach Chebyshev-Polynomen**.
Die Chebyshev-Polynome (1. Art) sind definiert durch

$$T_n(x) = \cos(n\ \arccos(x)), \quad n \in \mathbb{N}_0,\ x \in [-1, 1].$$

Sie können auch rekursiv bestimmt werden durch

$$T_0(x) = 1, \qquad T_1(x) = x, \qquad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1,\ x \in \mathbb{R}.$$

An der Rekursion erkennt man besser, dass T_k ein Polynom vom exakten Grad k ist. Also sind $T_0, \dots, T_n$ eine Basis des Polynomraums $\mathcal{P}_n$. Die trigonometrische Form verrät dagegen die Maximumsnorm auf dem Intervall $[-1, 1]$,

$$\|T_k\|_\infty = \max_{x \in [-1,1]} |T_k(x)| = 1 = T_k(1).$$

Die besondere Bedeutung der T_k rührt von der Orthogonalitätsbeziehung

$$\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} T_n(x)\, T_m(x)\, dx = \begin{cases} \pi/2 & \text{für } n = m, \\ 0 & \text{für } n \neq m, \end{cases}$$

her, das heißt die Chebyshev-Polynome sind paarweise orthogonal bezüglich des Skalarprodukts

$$\langle f, g \rangle_C = \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} f(x)\, g(x)\, dx, \qquad f, g \in C[-1, 1].$$

Die Orthogonalprojektion von $C[-1, 1]$ auf den Raum der Polynome vom Grad $\leq n$ (bzgl. dieses Skalarprodukts) wird in der Numerik als Gauß-Chebyshev-Approximation bezeichnet. Weil die T_k, $k = 0, \dots, n$, eine Orthogonalbasis bilden, ist die Orthogonalprojektion von f gegeben durch

$$p_n^{(C)} = \sum_{k=0}^{n} c_k T_k \quad \text{mit} \quad c_k = \frac{2}{\pi} \langle f, T_k \rangle_C .$$

Für das gegebene f erhalten wir $p_n^{(C)}$ mit dem folgenden Trick, ohne die Skalarprodukte auszurechnen. Man findet im "Handbook of Mathematical Functions" von Abramowitz und Stegun (Seite 783) die erzeugende Funktion der Chebyshev-Polynome

$$\frac{1-z^2}{1-2xz+z^2} \quad = \quad 1 + 2 \sum_{k=1}^{\infty} z^k\, T_k(x) \tag{1.3}$$

für alle $|x| \leq 1$, $z \in \mathbb{C}$ mit $|z| < 1$. Durch eine spezielle Wahl von z erhalten wir auf der linken Seite ein Vielfaches von $f(x) = \frac{1}{a-x}$: dazu setzen wir

$$z = \beta := a - \sqrt{a^2 - 1},$$

erhalten aus $a > 1$ sofort $z \in (0, 1)$ und rechnen einfach nach, dass $1 - 2az + z^2 = 0$ gilt. Hieraus folgt dann

$$1 + z^2 = 2az, \quad 1 - z^2 = 2 - 2az, \quad \frac{1 - a\beta}{\beta(a - x)} = 1 + 2 \sum_{k=1}^{\infty} \beta^k\, T_k(x). \tag{1.4}$$

Dies ergibt die Orthogonalprojektion von f

$$\frac{1}{a-x} \approx p_n^{(C)}(x) = \frac{\beta}{1-a\beta}\left(1 + 2\sum_{k=1}^{n} \beta^k\, T_k(x)\right). \tag{1.5}$$

Der Fehler lässt sich sowohl in der vom Skalarprodukt induzierten Norm als auch in der Maximumsnorm angeben. Beachtet man noch

$$\frac{\beta}{1-a\beta} = \frac{1}{\sqrt{a^2-1}},$$

so ergibt der "Satz des Pythagoras" für die paarweise orthogoanlen T_k

$$\begin{aligned}
\|f - p_n^{(C)}\|_C &= \frac{1}{\sqrt{a^2-1}}\left\|2\sum_{k=n+1}^{\infty} \beta^k\, T_k\right\|_C \\
&= \frac{1}{\sqrt{a^2-1}}\left(4\sum_{k=n+1}^{\infty} \beta^{2k}\, \underbrace{\|T_k\|_C^2}_{=\pi^2/4}\right)^{1/2} \\
&= \frac{\pi}{\sqrt{a^2-1}}\,\frac{\beta^{n+1}}{\sqrt{1-\beta^2}}.
\end{aligned}$$

(Hinweis: $\|f - p_n^{(C)}\|_C = \inf_{p\in\mathcal{P}_n} \|f-p\|_C$, also ist $p_n^{(C)}$ die Bestapproximation bzgl. dieser Norm.)

In der Maximumsnorm auf dem Intervall $[-1,1]$ erhalten wir

$$\begin{aligned}
\|f - p_n^{(C)}\|_\infty &= \max_{x\in[-1,1]}\left|\frac{2}{\sqrt{a^2-1}}\sum_{k=n+1}^{\infty} \beta^k\, T_k(x)\right| \\
&= \frac{2}{\sqrt{a^2-1}}\sum_{k=n+1}^{\infty} \beta^k \\
&= \frac{2\beta^{n+1}}{\sqrt{a^2-1}\,(1-\beta)} =: E_n^{(C)}(f).
\end{aligned} \tag{1.6}$$

c) **Bestapproximation von f auf $[-1,1]$ bzgl. der Maximumsnorm**

Die Bestapproximation an f ist definiert als Polynom $p_n^* \in \mathcal{P}_n$ mit

$$\|f - p_n^*\|_\infty = \inf_{p\in\mathcal{P}_n} \|f-p\|_\infty.$$

Wir werden in den folgenden Kapiteln viel zur Lösung dieses Approximationsproblems lernen. Mit Hilfe des Alternantensatzes in Kapitel 6.3 lässt sich für das gegebene f beweisen (siehe G. Meinardus, "Approxmationstheorie"), dass

$$\|f - p_n^*\|_\infty = \frac{\beta^n}{a^2-1} =: E_n^*(f), \quad \beta = a - \sqrt{a^2-1}. \tag{1.7}$$

Insgesamt gilt also

$$E_n^*(f) < E_n^{(C)}(f) = \underbrace{\frac{2\beta\sqrt{a^2-1}}{1-\beta}}_{=1+\sqrt{\frac{a+1}{a-1}}} E_n^*(f) < E_n^{(T)}(f).$$

Die Orthogonalentwicklung nach Chebyshev-Polynomen ist also auch bzgl. der Maximumsnorm auf $C[-1,1]$ fast so gut wie die zu dieser Norm gebildete Bestapproximation p_n^*.

Wir halten die folgenden Beobachtungen fest:

1. Die Methoden in Beispiel 1.1 lassen sich auf beliebige Funktionen in $C[-1,1]$ (bzw. $C^n[-1,1]$) anwenden. Dabei sind a) und b) lineare Methoden, d.h. für $c_1, c_2 \in \mathbb{R}$ und $f_1, f_2 \in C[-1,1]$ ist $p_n^{(C)}(c_1 f_1 + c_2 f_2) = c_1 p_n^{(C)}(f_1) + c_2 p_n^{(C)}(f_2)$. Hingegen ist die Methode c) der Bestapproximation in der Maximumsnorm nicht-linear!

2. Bei allen drei Methoden liegt lineare Konvergenzordnung vor, d.h. für $n \to \infty$ gilt jeweils $E_n = O(q^n)$ mit einem $0 < q < 1$. Der Wert von q ist bei b) und c) der gleiche, nämlich $q = \beta = a - \sqrt{a^2-1}$, bei a) ist $q = 1/a$. Abbildung 1.1 zeigt die schnellere Konvergenz (kleineres q) bei b) und c) gegenüber der Taylorentwicklung a).

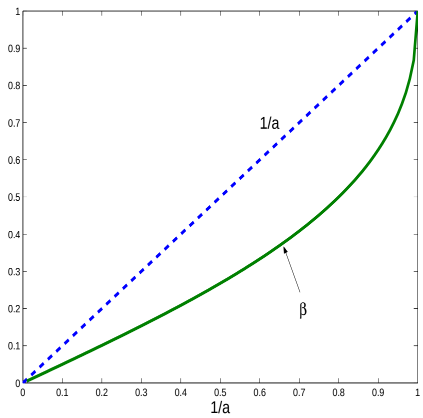

Abbildung 1.1: Vergleich der Konvergenzfaktoren $1/a$ bzw. β

Wozu sollte man die einfache rationale Funktion $f(x) = \frac{1}{a-x}$ überhaupt approximieren? Eine praktische Anwendung ergibt sich für die Invertierung von Matrizen.

1.2 Anwendung zur Matrixinversion. Es sei $A \in \mathbb{R}^{m\times m}$ symmetrisch mit Eigenwerten in $[-1,1]$. Gesucht ist die Inverse $(a\,\mathrm{Id} - A)^{-1}$ für $a > 1$.

Aus der Linearen Algebra wissen wir, dass A eine Spektralzerlegung

$$A = \sum_{k=1}^{m} \lambda_k \, v_k \, v_k^\top$$

besitst, wobei $(v_1, \ldots, v_m)$ eine Orthonormalbasis des $\mathbb{R}^m$ aus Eigenvektoren von A ist. Mit Hilfe dieser Zerlegung lassen sich grundlegende Operationen (z.B. Potenzbildung A^k oder Einsetzen von A in die Variable x eines Polynoms) auf die Eigenwerte reduzieren: für jedes Polynom $p(x) = \sum_{j=0}^{n} c_j x^j$ gilt

$$p(A) = \sum_{j=0}^{n} c_j A^j = \sum_{k=1}^{m} p(\lambda_k) \, v_k \, v_k^\top .$$

Auch für die gesuchte Inverse gilt

$$(a \operatorname{Id} - A)^{-1} = \sum_{k=1}^{m} \frac{1}{a - \lambda_k} \, v_k \, v_k^\top .$$

Natürlich kennt man in der Regel die Eigenwerte λ_k von A nicht. Weiss man nur, dass alle Eigenwerte im Intervall $[-1, 1]$ liegen (z.B. durch Betrachtung der Gerschgorin-Kreise), so erhält man durch die Matrizen

$$p_n^{(T)}(A), \quad p_n^{(C)}(A), \quad p_n^*(A)$$

jeweils Approximationen von $(a \operatorname{Id} - A)^{-1}$, weil ja z.B. $p_n^{(C)}(\lambda_k)$ eine gute Approximation von $\frac{1}{a-\lambda_k}$ ist. Für numerische Zwecke wird gern die Matrix

$$p_n^{(C)}(A) = \frac{1}{\sqrt{a^2 - 1}} \left(1 + 2 \sum_{k=1}^{n} \beta^k \, T_k(A) \right)$$

verwendet, weil die Berechnung rekursiv mit Hilfe der Rekursion der Chebyshev-Polynome erfolgen kann. Dies führt zu sogenannten "semi-iterativen" Methoden zur Lösung von linearen Gleichungssystemen in der Numerik. Deren Qualität ist manchmal vergleichbar mit dem meist bevorzugten Verfahren der konjugierten Gradienten, vgl. Numerik I-II.

Zentrale Fragestellungen der Approximationstheorie sind: die Existenz und Eindeutigkeit (vgl. Kapitel 2, 3 und Abschiutt 6.3.1), die Charakterisierung der Bestapproximation (vgl. Kapitel 6) sowie Abschätzungen des Approximationsfehlers für bestimmte Klassen von Funktionen (vgl. Kapitel 4 und 5).

Literaturangaben zur Vorlesung Approximationstheorie

1. E. W. Cheney, Introduction to Approximation Theory, 2nd edition, Chelsea, New York, 1982.

2. P. J. Davis, Interpolation and Approximation, Blaisdell, New York, 1963. Reprint: Dover, New York.

3. R. A. DeVore, G. G. Lorentz, Constructive Approximation, Springer-Verlag, New York, 1993.

4. M. W. Müller, Approximationstheorie, Akad. Verl.-Ges., Wiesbaden, 1978.

5. M. J. D. Powell, Approximation Theory and Methods, Cambridge University Press, 1981

6. T. Sauer, Approximationstheorie (Vorlesungsskript), Universität Passau, `http://www.fim.uni-passau.de/digitale-bildverarbeitung/lehre/`

2 Approximation in normierten Räumen

Die folgenden Begriffe werden in den Grundvorlesungen der Linearen Algebra, Analysis und Numerik eingeführt.

2.1 Definition (normierter Vektorraum über $\mathbb{R}$ bzw. $\mathbb{C}$). *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ bzw. $\mathbb{K} = \mathbb{C}$. Dann heißt eine Abbildung $\|.\| : V \longrightarrow \mathbb{R}$ Norm, falls*

1. *$\|v\| \geq 0$ für alle $v \in V$ und ($\|v\| = 0 \Longleftrightarrow v = 0$)* *(Definitheit)*
2. *$\|\alpha\, v\| = |\alpha|\, \|v\|$ für alle $v \in V$ und $\alpha \in \mathbb{K}$* *(Homogenität)*
3. *$\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$.* *(Dreiecksungleichung)*

*Ist $(V, \|.\|)$ **vollständig** (d.h. alle Cauchy-Folgen in V besitzen einen Grenzwert in V), so heißt $(V, \|.\|)$ **Banachraum**.*

2.2 Beispiel. a) Es sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall. Der Vektorraum der stetigen Funktionen

$$C(I) = \{f : [a, b] \to \mathbb{K} : f \text{ stetig}\}$$

ist versehen mit der Maximumsnorm

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$$

ein Banachraum. Die Konvergenz bzgl. diesr Norm ist die gleichmäßige Konvergenz, die wir aus der Analysis kennen.

b) Sei $1 \leq p < \infty$ und

$$\mathcal{L}^p(a, b) := \{f : (a, b) \longrightarrow \mathbb{K}\text{Lebesgue-messbar, } \int_a^b |f(x)|^p\, dx < \infty\}.$$

Ein wichtiger Schritt zur Definition der L^p-Räume ist es, Funktionen f und g zu identifizieren (sie sind "das gleiche Element", bilden mathematisch genauer eine Äquivalenzklasse), wenn $f(x) = g(x)$ für fast alle $x \in (a, b)$ gilt; der Begriff "für fast alle $x \in (a, b)$" oder kurz "fast überall" besagt, dass $f(x) = g(x)$ außerhalb einer Menge vom Lebesgue-Maß 0 gilt. All diese Äquivalenzklassen bilden den normierten Vektorraum $L^p(a, b)$ mit der L^p-Norm

$$\|f\|_p := \left(\int_a^b |f(x)|^p\, dx \right)^{\frac{1}{p}}.$$

$L^p(a, b)$ ist ein Banachraum.

Beispiel: Die Funktion $f \in \mathcal{L}^1(0,1)$ mit $f(x) = x^{-1/2}$ hat die Norm $\|f\|_1 = \int_0^1 x^{-1/2}\,dx = 2$. Sie wird mit der Funktion $g \in \mathcal{L}^1(0,1)$ mit

$$g(x) = \begin{cases} x^{-1/2} & \text{für } x \in (0,1) \cap \mathbb{R}, \\ 0 & \text{für } x \in (0,1) \cap \mathbb{Q} \end{cases}$$

identifiziert, beide Funktionen f und g repräsentieren also die gleiche "Funktion" in $L^1(0,1)$.

c) Der Vektorraum $C[a,b]$ mit der L^2-Norm ist ein normierter Raum, aber **kein** Banachraum. (Übungen)

2.3 Bemerkung (2. Dreiecksungleichung)**.** Die Normabbildung $\|.\| : V \longrightarrow \mathbb{R}$ ist stetig, und sogar Lipschitzstetig, denn es gilt als Folgerung aus der Dreiecksungleichung auch

$$\Big| \|v\| - \|w\| \Big| \leq \|v - w\|.$$

2.4 Bemerkung (Endlich-dimensionale Vektorräume)**.** In einem endlich-dimensionalen Vektorraum V sind alle Normen **äquivalent**, das heißt es gibt $\alpha, \beta > 0$, mit

$$\alpha \, \|v\|_I \leq \|v\|_{II} \leq \beta \, \|v\|_I$$

für alle $v \in V$, wobei $\|.\|_I, \|.\|_{II}$ beliebige Normen auf V sind.
Insbesondere ist jeder n-dimensionale normierte Vektorraum V als topologischer Raum isomorph zu $(\mathbb{K}^n, \|.\|_2)$.
Der Satz von Alaoglu besagt: Der normierte Vektorraum V ist genau dann endlich-dimensional, wenn seine Einheitskugel $B = \{v \in V : \|v\| \leq 1\}$ kompakt ist.

2.5 Beispiel. Der Folgenraum $V = \ell^2(\mathbb{N})$ der quadrat-summierbaren Folgen mit dem Skalarprodukt

$$\langle\, a, b \,\rangle = \sum_{k=1}^{\infty} a_k \, b_k, \quad a = (a_k)_{k\in\mathbb{N}}, \quad b = (b_k)_{k\in\mathbb{N}}$$

ist unendlich-dimensional. Wir zeigen, dass die Einheitskugel nicht kompakt ist.
Für $j \in \mathbb{N}$ ist das Element $e_j = (\delta_{jk})_{k\in\mathbb{N}}$ ein Element der Einheitskugel. Die Folge dieser Elemente $(e_j)_{j\in\mathbb{N}}$ ist eine beschränkte Folge in V, besitzt aber keine konvergente Teilfolge wegen $\|e_i - e_j\| = 2$ für alle $i \neq j$. Daran erkennt man, dass die Einheitskugel von $\ell^2(\mathbb{N})$ nicht kompakt ist.

Wir beschreiben nun das Problem der **linearen Bestapproximation**, das in dieser Vorlesung intensiv behandelt wird.

2.6 Definition. *Sei V ein normierter Vektorraum und $M \subseteq V$ ein Teilraum. Zu $f \in V$ heißt*

- $E_M(f) := \inf\limits_{v \in M} \|f - v\|$ *der* ***Approximationsgrad*** *von f (bzgl. M),*
- *jedes Element $v^* \in M$ mit $\|f - v^*\| = E_M(f)$* ***Proximum*** *(oder* ***Bestapproximation****) von f in M,*
- $\Pi_M(f) \subset M$ *die Menge aller Proxima von f in M und*
- $\Pi_M : V \longrightarrow \mathcal{P}(M)$, $f \longmapsto \Pi_M(f)$ *die* ***metrische Projektion*** *von V auf M. Hierbei ist $\mathcal{P}(M)$ die Potenzmenge von M.*

Der Teilraum $M \subseteq V$ heißt

- ***Existenzmenge****, wenn zu jedem $f \in V$ mindestens ein Proximum $v^* \in M$ existiert,*
- ***Eindeutigkeitsmenge****, wenn zu jedem $f \in V$ höchstens ein Proximum $v^* \in M$ existiert,*
- ***Chebyshev-Menge****, wenn zu jedem $f \in V$ genau ein Proximum $v^* \in M$ existiert.*

2.7 Bemerkung. a) Im Allgemeinen kann $\Pi_M(f)$ für ein f leer sein, für ein anderes f unendlich viele Elemente enthalten. Allgemeine Aussagen zu Existenz- und Eindeutigkeitsmengen werden wir in diesem Kapitel kennenlernen.

b) Falls M eine Chebyshev-Menge ist, so ist die metrische Projektion auch als Abbildung von V nach M zu verstehen: $\Pi_M : V \longrightarrow M$, $\Pi_M(f) = v^*$ mit $\|f - v^*\| = E_M(f)$ ist dann wohldefiniert.

Achtung: Im Allgemeinen ist diese Abbildung nicht linear.

2.8 Lemma. *Für $f \in V$ ist die Menge $\Pi_M(f)$ konvex und beschränkt (oder leer).*

Beweis. Es sei $\Pi_M(f) \neq \emptyset$.
1. Konvexität: Wähle $v_1, v_2 \in \Pi_M(f)$ und setze $w := \alpha\, v_1 + (1-\alpha)\, v_2 \in M$ mit $\alpha \in (0,1)$. Dann gilt

$$\|f - w\| = \|\alpha\,(f - v_1) + (1-\alpha)\,(f - v_2)\| \leq \alpha\, \underbrace{\|f - v_1\|}_{=E_M(f)} + (1-\alpha)\, \underbrace{\|f - v_2\|}_{=E_M(f)} = E_M(f).$$

Also erzielt auch das Element $w \in M$ den Approximationsgrad $E_M(f)$, d.h. $w \in \Pi_M(f)$.

2. Beschränktheit: Für jedes Proximum v^* gilt

$$\begin{aligned}\|v^*\| &= \|v^* - f + f\| \leq \|v^* - f\| + \|f\| \\ &= \|f\| + \inf_{v \in M} \|v - f\| \\ &\leq \|f\| + \|0 - f\| = 2\,\|f\|.\end{aligned}$$

□

Als Hilfsmittel zur Untersuchung von Existenzmengen benötigen wir den folgenden Begriff.

2.9 Definition (Minimalfolge). *Gegeben sei ein normierter Vektorraum V, ein Teilraum $M \subseteq V$ sowie eine Funktion $f \in V$. Eine Folge $(v_n)_{n\in\mathbb{N}}$ von Elementen $v_n \in M$ heißt* ***Minimalfolge****, falls*

$$\lim_{n\to\infty} \|f - v_n\| = \inf_{v \in M} \|f - v\| = E_M(f).$$

Die Definition der Minimalfolge lässt sofort den folgenden Schluss zu.

2.10 Lemma. *Zu jedem $f \in V$ existiert eine Minimalfolge.*

Als erste Anwendung der Minimalfolge können wir die Stetigkeit der Abbildung E_M nachweisen.

2.11 Lemma. *Der Approximationsgrad $E_M : V \longrightarrow \mathbb{R}$ ist Lipschitzstetig: Es gilt*

$$|E_M(f) - E_M(g)| \leq \|f - g\|$$

für alle $f, g \in V$.

Beweis. Seien $f, g \in V$ und $\varepsilon > 0$ beliebig. Wir wählen Minimalfolgen (v_n) zu f und (w_n) zu g, sowie $n_0 \in \mathbb{N}$ mit

$$\begin{aligned}E_M(f) &\leq \|f - v_n\| < E_M(f) + \varepsilon, \\ E_M(g) &\leq \|g - w_n\| < E_M(g) + \varepsilon\end{aligned}$$

für alle $n \geq n_0$. Mit der 2. Dreiecksungleichung ergibt sich

$$\begin{aligned}\|f - g\| &\geq \|f - w_n\| - \|w_n - g\| \geq E_M(f) - E_M(g) - \varepsilon \\ \|f - g\| &\geq \|g - v_n\| - \|v_n - f\| \geq E_M(g) - E_M(f) - \varepsilon.\end{aligned}$$

Also ist

$$|E_M(f) - E_M(g)| \leq \|f - g\| + \epsilon$$

für alle $\epsilon > 0$ und damit folgt die Behauptung. □

Der folgende Satz ist der zentrale Existenzsatz zur Bestapproximation im Banachraum. Wesentliche Voraussetzung ist die Endlich-Dimensionalität von M.

2.12 Satz (Existenzsatz)**.** *Sei* $(V, \|.\|)$ *ein normierter Vektorraum und* $M \subseteq V$ *ein endlich-dimensionaler Teilraum. Dann gilt* $\Pi_M(f) \neq \emptyset$ *für jedes* $f \in V$*.*

Beweis. Sei (v_n) eine Minimalfolge zu f, also

$$\lim_{n\to\infty} \|f - v_n\| = E_M(f). \tag{2.1}$$

Wir betrachten zunächst die skalare Folge der Normen $(\|v_n\|)_{n\in\mathbb{N}}$. Wegen

$$\|v_n\| \leq \|v_n - f\| + \|f\|$$

und der Konvergenz der skalaren Folge $(\|v_n - f\|)$ gegen $E_M(f)$ ist auch $(\|v_n\|)_{n\in\mathbb{N}}$ beschränkt.

Das heißt genau, dass die Folge $(v_n)_{n\in\mathbb{N}}$ in M beschränkt ist. Weil M endlich-dimensional ist, existiert eine konvergente Teilfolge $(v_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert $v^* := \lim\limits_{k\to\infty} v_{n_k} \in M$. Die Stetigkeit der Norm erlaubt den Grenzübergang

$$\|f - v^*\| = \|f - \lim_{k\to\infty} v_{n_k}\| = \lim_{k\to\infty} \|f - v_{n_k}\| = E_M(f),$$

also ist v^* Proximum. □

2.13 Bemerkung. Ein Beispiel in den Übungen zeigt, dass auf die Endlich-Dimensionalität im Allgemeinen nicht verzichtet werden kann. In Kapitel 3 wird jedoch für Hilberträume eine viel schwächere Voraussetzung verlangt: dort muss M ein abgeschlossener Teilraum sein, egal ob endlich- oder unendlich-dimensional.

Zur Diskussion der Eindeutigkeitsmengen betrachten wir eine geometrische Eigenschaft des Banachraums, die durch die Norm induziert wird.

2.14 Definition (Strikt konvexer Raum)**.** *Ein normierter Raum* $(V, \|.\|)$ *heißt* ***strikt konvex****, wenn für alle* $f, g \in V$ *mit* $\|f\| = \|g\| = 1$ *und* $f \neq g$ *sowie für alle* $0 < \alpha < 1$ *die Abschätzung*

$$\|\alpha\, f + (1-\alpha)\, g\| < 1$$

gilt; geometrisch bedeutet dies: die Elemente auf der Verbindungsstrecke zwischen zwei Randpunkten der Einheitskugel liegen im Inneren der Kugel.

2.15 Beispiel. a) Die euklidische Norm im $\mathbb{R}^n$ ist strikt konvex. Veranschaulichung am Einheitskreis, Rechnung für $\|x\|_2 = \|y\|_2 = 1$:

$$\|\alpha x + (1-\alpha)y\|_2^2 = \alpha^2 + (1-\alpha)^2 + 2\alpha(1-\alpha)\langle x, y\rangle = 1 - 2\alpha(1-\alpha)(1 - \langle x, y\rangle).$$

Für $\alpha \in (0,1)$ ist der letzte Ausdruck nur dann 1, wenn $\langle x, y\rangle = 1$ gilt; dies ist auf Grund der Cauchy-Schwarz Ungleichung nur dann der Fall, wenn $x = y$ ist.

b) Die Maximums-Norm im $\mathbb{R}^n$ ist nicht strikt konvex.

2.16 Satz (Eindeutigkeitssatz)**.** *Ist $(V, \|.\|)$ strikt konvex und $M \subseteq V$ ein linearer Teilraum, so besitzt jedes $f \in V$ höchstens ein Proximum $v^* \in M$.*

Beweis. 1. Fall: $E_M(f) = 0$, $v_1, v_2 \in \Pi_M(f)$. Dann ist $\|f - v_1\| = 0 = \|f - v_2\|$, also $v_1 = v_2 = f$.
2. Fall: $E_M(f) > 0$, $v_1, v_2 \in \Pi_M(f)$, $v_1 \neq v_2$. Wähle $0 < \alpha < 1$ beliebig. Dann ist für $w = \alpha v_1 + (1-\alpha) v_2 \in M$

$$\|f - w\| = \|\alpha(f - v_1) + (1-\alpha)(f - v_2)\| \overset{2.14}{<} E_M(f),$$

weil $\|f - v_1\| = \|f - v_2\| = E_M(f)$ gilt. Dies ist ein Widerspruch zur Definition von $E_M(f)$. □

Wir haben bereits ganz allgemein gezeigt, dass der Approximationsgrad $E_M(f)$ stetig von $f \in V$ abhängt. Für Chebyshev-Mengen M ist eine viel wichtigere Frage, ob die metrische Projektion $\Pi_M : V \to M$ ebenfalls stetig von f abhängt. Nur dann sind Algorithmen zur Bestapproximation, also zur Bestimmung von $v^* = \Pi_M(f)$, sinnvoll einsetzbar. Die folgende Aussage liefert die Stetigkeit für endlich-dimensionale Chebyshev-Mengen. Später werden wir für Spezialfälle sogar eine lokale Form der Lipschitz-Stetigkeit zeigen können (vgl. Satz von Freud für Haarsche Räume in Abschnitt 6.3).

2.17 Satz. *Sei $(V, \|.\|)$ ein normierter Vektorraum, $M \subseteq V$ eine endlich-dimensionale Chebyshev-Menge. D ann ist die metrische Projektion $\Pi_M : V \longrightarrow M$, $f \longmapsto \Pi_M(f)$ stetig. (Beachte: Π_M ist im Allgemeinen nicht linear)*

Beweis. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge in V mit $f = \lim_{n\to\infty} f_n \in V$. Weiter seien $v_n^* := \Pi_M(f_n)$ und $v^* = \Pi_M(f)$ die Proxima. Zu zeigen ist $v^* = \lim_{n\to\infty} v_n^*$. Dazu verwenden wir den folgenden Konvergenzsatz: Es seien $(a_n)_{n\in\mathbb{N}}$ eine Folge im normierten Raum V und $a \in V$. Falls **jede Teilfolge** $(a_{n_k})_{k\in\mathbb{N}}$ **eine gegen** $a \in V$ **konvergierende Teilfolge** $(a_{n_{k_j}})_{j\in\mathbb{N}}$ **besitzt**, dann konvergiert die Folge $(a_n)_{n\in\mathbb{N}}$ gegen a.

Schritt 1. Jede Teilfolge von (v_n^*) besitzt eine konvergente Teilfolge:
Dies folgt aus der Beschränktheit und $\dim(M) < \infty$: $\|v_n^*\| \leq 2\,\|f_n\|$ für alle $n \in \mathbb{N}$, und (f_n) ist beschränkt, weil konvergent.

Schritt 2. Jede konvergente Teilfolge $(v_{n_k}^*)$ hat den Grenzwert v^*:

Sei dazu $w = \lim_{k\to\infty} v_{n_k}^* \in M$. Dann gilt

$$\begin{aligned}
E_M(f) &\leq \|f-w\| = \lim_{k\to\infty} \|f - v_{n_k}^*\| \leq \limsup_{k\to\infty} (\underbrace{\|f-f_{n_k}\|}_{\longrightarrow 0} + \underbrace{\|f_{n_k} - v_{n_k}^*\|}_{=E_M(f_{n_k})}) \\
&= \limsup_{k\to\infty} \underbrace{(E_M(f_{n_k}) - E_M(f))}_{\overset{2.11}{\leq} \|f_{n_k}-f\| \longrightarrow 0} + E_M(f) \\
&= E_M(f).
\end{aligned}$$

In dieser Ungleichungskette muss also überall Gleichheit gelten, insbesondere $E_M(f) = \|f-w\|$, also $w = \Pi_M(f) = v^*$ aufgrund der Eindeutigkeit des Proximums an f.

$\square$