

Kapitel 1

Bestapproximation

1.1 Einführung

Bemerkung 1.1 *Motivation.* Seien $[a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall und $f : [a, b] \rightarrow \mathbb{R}$ eine reellwertige Funktion. Von dieser Funktion sei es im Laufe einer numerischen Berechnung notwendig, Funktionswerte an (vielen) unterschiedlichen Stellen $x \in [a, b]$ zu berechnen, wobei die Stellen vor der Rechnung nicht bekannt sind. Ist $f(x)$ eine „einfache“ Funktion, zum Beispiel ein Polynom, bereitet die Berechnung der Funktionswerte keine Schwierigkeiten. Bei Polynomen verwendet man dafür zweckmäßigerweise das Horner-Schema.

Aber schon bei anderen elementaren Funktionen, wie $e^x, \sin(x), \ln(x)$ ist die Berechnung von Funktionswerten für beliebige Argumente x aus dem Definitionsbereich der jeweiligen Funktion schwierig. In diesem Falle ist es zweckmäßig, den gesuchten Funktionswert mit einer vorgegebenen Genauigkeit ε zu approximieren. Dabei ersetzt man die betrachtete Funktion $f(x)$ durch eine Funktion $\varphi(x)$, welche einfacher berechenbar ist (etwa ein Polynom), und deren Werte sich in $[a, b]$ nicht um mehr als ε von den Werten von $f(x)$ unterscheiden. \square

Bemerkung 1.2 *Bestapproximation – abstrakte Aufgabenstellung.* Sei V ein normierter Raum reellwertiger Funktion, welche über $[a, b]$ definiert sind, zum Beispiel $V = C([a, b])$. Gegeben sei $f \in V$. Weiter sei U eine Menge reellwertiger Funktionen über $[a, b]$, welche nur aus „einfach berechenbaren“ Funktionen besteht. Des Weiteren bezeichne $\|\cdot\|$ eine Norm, welche auf $V \cap U$ definiert ist. Dann hat die Aufgabe der Bestapproximation die Form: Finde $u \in U$ so dass

$$\|f - u\| \leq \|f - v\| \quad \forall v \in U. \quad (1.1)$$

Bei dieser Aufgabenstellung sind die Wahl von U und die Wahl der Norm $\|\cdot\|$ noch frei. \square

Definition 1.3 **Tschebyscheff-Approximation.** Betrachtet man (1.1) für die Maximumsnorm

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

so spricht man von Tschebyscheff¹-Approximation. \square

Beispiel 1.4 *Tschebyscheff-Approximation.* Seien $[a, b] = [0, \pi]$ und $f(x) = \sin(x)$. Gesucht ist die Tschebyscheff-Approximation in der Menge der konstanten Funktionen über $[a, b]$, $U = P_0([a, b])$. Es sind

$$\min_{x \in [a,b]} f(x) = 0, \quad \max_{x \in [a,b]} f(x) = 1.$$

¹Pafnuti Lwowitsch Tschebyscheff (1821 – 1894)

Die konstante Funktion, deren maximaler Abstand zu beiden Extremwerten minimal ist, ist $u(x) = 1/2$. Dies ist die Tschebyscheff-Approximation. Dann ist

$$\|\sin(x) - u\|_\infty = \frac{1}{2}.$$

□

Definition 1.5 Der Raum $L^2(a, b)$. Der Raum $L^2(a, b)$ besteht aus allen Funktionen $f : (a, b) \rightarrow \mathbb{R}$, welche in (a, b) quadratisch (Lebesgue-) integrierbar sind

$$L^2(a, b) = \left\{ f : \int_a^b (f(x))^2 dx < \infty \right\}.$$

Der Raum ist ausgestattet mit der Norm

$$\|f\|_{L^2} = \left(\int_a^b (f(x))^2 dx \right)^{1/2} = \left(\int_a^b f^2(x) dx \right)^{1/2}. \quad (1.2)$$

□

Beispiel 1.6 Bestapproximation in $\|\cdot\|_{L^2}$. Betrachte die gleiche Ausgangssituation wie im Beispiel 1.4. Nun ist aber diejenige konstante Funktion $u \in P_0([a, b])$ gesucht, für die $\|f - u\|_{L^2}$ minimal wird. Diese Norm ist genau dann minimal, wenn das Quadrat der Norm minimal ist. Dies folgt aus der strengen Monotonie der Wurzelfunktion. Mit binomischer Formel und der Eigenschaft, dass u eine Konstante ist, erhält man

$$\begin{aligned} \int_0^\pi (\sin(x) - u)^2 dx &= \int_0^\pi \sin^2(x) dx - 2u \int_0^\pi \sin(x) dx + \pi u^2 \\ &= \frac{\pi}{2} - 4u + \pi u^2. \end{aligned}$$

Dies ist eine quadratische Funktion in u , deren Bild (Parabel) nach oben geöffnet ist. Demzufolge besitzt sie ein Minimum, welches man wie üblich berechnet

$$u = \frac{2}{\pi} \approx 0.63661977236758134308,$$

siehe Abbildung 1.1. Man erhält

$$\|\sin(x) - u\|_{L^2} = \left(\int_0^\pi (\sin(x) - u)^2 dx \right)^{1/2} = \left(\frac{\pi^2 - 8}{2\pi} \right)^{1/2} \approx 0.5454876555. \quad (1.3)$$

Der Vergleich mit Beispiel 1.4 zeigt, dass die Wahl unterschiedlicher Normen zu unterschiedlichen Ergebnissen führen kann. □

Bemerkung 1.7 Stückweise definierte Funktionen. In den Beispielen 1.4 und 1.6 wurden Bestapproximierte gesucht, die im gesamten Intervall die gleich Form besitzen. Die zu approximierende Funktion $f(x)$ kann sich aber in unterschiedlichen Teilintervallen von $[a, b]$ unterschiedlich verhalten. Dann ist es zweckmäßig, sie durch eine stückweise definierte Funktion zu approximieren. □

Definition 1.8 Der Raum S_n oder P_1 . Sei das Intervall $[a, b]$ durch ein Gitter mit n Gitterzellen trianguliert

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = b.$$

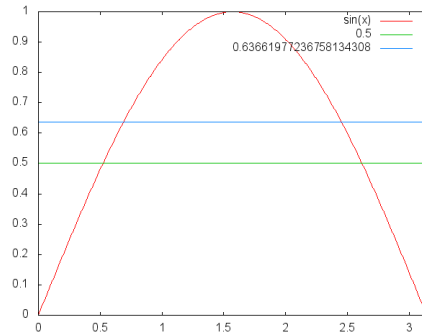


Abbildung 1.1: Bestapproximation von $\sin(x)$ in $[0, \pi]$: Tschebyscheff-Approximation und Approximation bezüglich $\|\cdot\|_{L^2}$.

Dann besteht der Raum S_n aus allen stetigen und stückweise linearen Funktionen (Polygonzügen)

$$S_n = \{f : f \in C([a, b]), f|_{[x_{i-1}, x_i]} \text{ ist linear } \forall i = 1, \dots, n\}.$$

Soll nicht die Anzahl der Gitterzellen sondern die stückweise Linearität des Raumes hervorgehoben werden, so wird er im Allgemeinen mit P_1 bezeichnet. \square

Beispiel 1.9 *Approximation in S_2 mit $\|\cdot\|_{L^2}$.* Betrachte wiederum die Situation von Beispiel 1.4. Man kann zeigen, siehe Beispiel 1.40, dass die Bestapproximation durch eine lineare Funktion $\alpha x + \beta$ in $[a, b]$ gegeben ist durch $\alpha = 0, \beta = 2/\pi$. Das bedeutet, die beste lineare Approximation ist gerade die Konstante aus Beispiel 1.6 und der Approximationsfehler ist durch (1.3) gegeben.

Betrachtet man ein Gitter aus zwei Gitterzellen mit $x_1 = \pi/2$ und den Polygonzug

$$p_2 = \begin{cases} \frac{2}{\pi}x & x \in \left[0, \frac{\pi}{2}\right], \\ -\frac{2}{\pi}x + 2 & x \in \left(\frac{\pi}{2}, \pi\right] \end{cases} \in S_2,$$

so erhält man

$$\begin{aligned} \|\sin(x) - p_2\|_{L^2} &= \left(\int_0^{\pi/2} \left(\sin(x) - \frac{2}{\pi}x \right)^2 dx + \int_{\pi/2}^{\pi} \left(\sin(x) + \frac{2}{\pi}x - 2 \right)^2 dx \right)^{1/2} \\ &= 0.2674224922. \end{aligned}$$

Das ist schon nur noch etwa der halbe Fehler im Vergleich mit der Bestapproximation mit einer globalen linearen Funktion. Im Raum S_2 ist es jedoch noch nicht der bestmögliche Wert. Dessen Berechnung wird in Beispiel 1.43 erklärt. \square

1.2 Bestapproximation in normierten Räumen und Prä-Hilbert-Räumen

Bemerkung 1.10 *Inhalt.* In diesem Abschnitt werden zunächst Aussagen zur Existenz einer Lösung des Problems der Bestapproximierenden und deren Eindeutigkeit im Rahmen von allgemeinen normierten Räumen vorgestellt. (Prä-)Hilbert-Räume sind spezielle normierte Räume, für die man die Bestapproximierende genauer charakterisieren kann. \square

Definition 1.11 Normierter Raum. Ein linearer Raum V heißt normiert, wenn es eine Abbildung $\|\cdot\|_V : V \rightarrow \mathbb{R}$ gibt, die folgenden Bedingungen genügt:

- i) $\|v\|_V \geq 0$ für alle $v \in V$, wobei $\|v\|_V = 0$ genau dann, wenn $v = 0$,
- ii) $\|\alpha v\|_V = |\alpha| \|v\|_V$ für alle $v \in V$ und für alle $\alpha \in \mathbb{R}$,
- iii) Dreiecksungleichung: $\|v_1 + v_2\|_V \leq \|v_1\|_V + \|v_2\|_V$ für alle $v_1, v_2 \in V$.

□

Bemerkung 1.12 Bestapproximations–Aufgabe. Die allgemeine Aufgabe der Bestapproximation (1.1) wird etwas konkretisiert. Anstelle einer beliebigen Menge U wird nun der Fall betrachtet, dass $U \subset V$ ein endlich–dimensionaler Teilraum ist. Ansonsten ändert sich die Bestapproximations–Aufgabe nicht. Sei $f \in V$ gegeben, finde $u \in U$ so dass

$$\|f - u\|_V \leq \|f - v\|_V \quad \forall v \in U. \quad (1.4)$$

□

Satz 1.13 Existenz einer Lösung der Bestapproximations–Aufgabe. Die Bestapproximations–Aufgabe (1.4) besitzt eine Lösung.

Beweis: Definiere die Abbildung $g : U \rightarrow \mathbb{R}$ durch $g(v) = \|f - v\|_V$ für alle $v \in U$. Die Funktion $g(v)$ ist wegen $g(v) \geq 0$ nach unten beschränkt. Mit Hilfe der Dreiecksungleichung erhält man

$$|g(v) - g(w)| = \left| \|f - v\|_V - \|f - w\|_V \right| \leq \|v - w\|_V \quad \forall v, w \in U, \quad (1.5)$$

da beispielsweise gilt

$$\begin{aligned} \|f - v\|_V - \|f - w\|_V &= \|(f - w) - (v - w)\|_V - \|f - w\|_V \\ &\leq \|f - w\|_V + \|v - w\|_V - \|f - w\|_V = \|v - w\|_V, \end{aligned}$$

und analog der andere Fall. Da die Norm eine stetige Funktion ist, folgt aus (1.5), dass auch $g(v)$ eine stetige Abbildung ist.

Betrachte nun eine Kugel, welche nur Funktionen aus U bis zu einer bestimmten Norm enthält

$$B = \{v \in U : \|v\|_V \leq 2\|f\|_V\}.$$

Da U ein Unterraum ist, ist $0 \in U$ und offenbar auch $0 \in B$. Zunächst wird gezeigt, dass kein Minimum von $g(v)$ außerhalb von B liegen kann. Betrachte dazu ein $v \in U$ mit $v \notin B$, also $\|v\|_V > 2\|f\|_V$. Dann folgt mit Dreiecksungleichung

$$g(v) = \|f - v\|_V \geq \|v\|_V - \|f\|_V > 2\|f\|_V - \|f\|_V = \|f\|_V = g(0).$$

Es ist auch anschaulich klar, dass der Mittelpunkt der Kugel eine bessere Approximation als v ist, falls v eine Norm hat, die den doppelten Abstand von f zum Mittelpunkt übersteigt. Somit reduziert sich die Bestapproximations–Aufgabe zu: Finde $u \in B$, so dass

$$g(u) \leq g(v) \quad \forall v \in B.$$

Die Kugel B ist abgeschlossen, da der Rand $\|v\|_V = 2\|f\|_V$ mit zu B gehört, und beschränkt. Nun benötigt man die Eigenschaft, dass U endlich–dimensional ist. In endlich–dimensionalen Räumen ist eine abgeschlossene und beschränkte Menge auch kompakt. Nach dem Satz von Weierstraß nimmt eine stetige Funktion auf einer kompakten Menge ihre Extremwerte an. Damit existiert ein $u \in B$, so dass

$$g(u) = \inf_{v \in B} g(v) = \min_{v \in B} g(v)$$

ist. ■

Bemerkung 1.14 Zum Beweis. Der Beweis ist nicht konstruktiv, das heißt, es wird nicht angegeben, wie eine Lösung konstruiert werden kann. □

Beispiel 1.15 *Nichteindeutigkeit der Bestapproximation.* Seien $V = C([-1, 1])$ ausgestattet mit $\|\cdot\|_V = \|\cdot\|_\infty$, $f(x) = 1$ und $U = \{p_\alpha : p_\alpha(x) = \alpha(1+x), \alpha \in \mathbb{R}\}$. Dann ist die Differenz zwischen $f(x)$ und einer beliebigen Funktion aus U gegeben durch

$$\|f - p_\alpha\|_\infty = \max_{x \in [-1, 1]} |1 - \alpha(1+x)| = \max_{x \in [-1, 1]} |-\alpha x + (1-\alpha)| = g_\alpha(x).$$

Im Betrag steht eine lineare Funktion. Solche Funktionen nehmen ihren betragsmäßig größten Wert in einem der Intervallenden an. Es gelten

$$g_\alpha(-1) = 1, \quad g_\alpha(1) = |1 - 2\alpha|.$$

Damit folgt

$$\|f - p_\alpha\|_\infty = \begin{cases} 1, & \alpha \in [0, 1], \\ |1 - 2\alpha|, & |\alpha - 0.5| > 0.5. \end{cases}$$

Es gilt $|1 - 2\alpha| > 1$ für $|\alpha - 0.5| > 0.5$. Damit sind alle Geraden $p_\alpha(x)$ mit $\alpha \in [0, 1]$ eine Bestapproximation. **Bild** \square

Definition 1.16 **Streng normierter Raum.** Ein linearer normierter Raum V heißt streng normiert, wenn aus $\|v + w\|_V = \|v\|_V + \|w\|_V$ folgt, dass $v = \alpha w$ für ein $\alpha \in \mathbb{R}$. \square

Satz 1.17 **Eindeutigkeit der Lösung der Bestapproximations-Aufgabe in streng normierten Räumen.** Sei V ein streng normierter Raum. Dann existiert genau eine Lösung der Bestapproximations-Aufgabe (1.4).

Beweis: Die Existenz der Lösung wurde bereits in Satz 1.13 gezeigt. Es bleibt, die Eindeutigkeit zu beweisen.

Ist $f \in U$, so ist $u = f$ die eindeutige Lösung, denn dann gilt wegen der ersten Normeigenschaft

$$0 = \|f - u\|_V < \|f - v\|_V \quad \forall v \in V, v \neq u.$$

Indirekter Beweis. Sei $f \notin U$ und seien $u_1 \neq u_2$ zwei Lösungen von (1.4). Dann folgt mit Dreiecksungleichung

$$\left\| f - \frac{1}{2}(u_1 + u_2) \right\|_V \leq \left\| \frac{1}{2}f - \frac{1}{2}u_1 \right\|_V + \left\| \frac{1}{2}f - \frac{1}{2}u_2 \right\|_V = \frac{1}{2} \|f - u_1\|_V + \frac{1}{2} \|f - u_2\|_V.$$

Da sowohl u_1 als auch u_2 Bestapproximierende sind, sind die Normen auf der rechten Seite der Ungleichung gleich und es folgt

$$\left\| f - \frac{1}{2}(u_1 + u_2) \right\|_V \leq \|f - u_1\|_V = \|f - u_2\|_V.$$

Da u_1 nach Voraussetzung eine Bestapproximierende ist, kann nur das Gleichheitszeichen gelten und demzufolge ist auch $\frac{1}{2}(u_1 + u_2)$ eine Bestapproximierende. Es folgt also

$$\|f - u_1\|_V + \|f - u_2\|_V = 2 \left\| f - \frac{1}{2}(u_1 + u_2) \right\|_V = \|(f - u_1) + (f - u_2)\|_V.$$

Da V streng normiert ist, gibt es nun ein $\alpha \in \mathbb{R}$ mit $f - u_1 = \alpha(f - u_2)$. Das ist äquivalent zu

$$(1 - \alpha)f = u_1 - \alpha u_2 \in U. \tag{1.6}$$

Da $u_1, u_2 \in U$ ist auch die rechte Seite dieser Gleichung Element von U . Da $f \notin U$ ist und U ein Raum ist, kann (1.6) nur gelten, wenn $1 - \alpha = 0$, also $\alpha = 1$. Dann folgt

$$0 = u_1 - u_2 \iff u_1 = u_2.$$

Das ist ein Widerspruch zur Annahme, dass u_1 und u_2 verschieden sind. Demzufolge ist die Lösung von (1.4) eindeutig. \blacksquare

Beispiel 1.18 *Raum mit strenger Norm.* Seien $V = \mathbb{R}^N$ und $\|\cdot\|_V = \|\cdot\|_2$ die Euklidische Norm. Dann folgt aus $\|\mathbf{x} + \mathbf{y}\|_2 = \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$ durch quadrieren

$$\|\mathbf{x} + \mathbf{y}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 + 2\|\mathbf{x}\|_2\|\mathbf{y}\|_2.$$

Mit der Definition der Euklidischen Norm mit dem Skalarprodukt von Vektoren erhält man aber auch

$$\|\mathbf{x} + \mathbf{y}\|_2^2 = (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 + 2(\mathbf{x}, \mathbf{y}).$$

Also gilt $(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|_2\|\mathbf{y}\|_2$. Nun gilt für den Winkel γ zwischen den Vektoren \mathbf{x} und \mathbf{y}

$$\cos(\gamma) = \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2} = 1.$$

Also ist $\gamma = 0$ und es folgt $\mathbf{x} = \alpha\mathbf{y}$ für ein $\alpha \in \mathbb{R}$. □

Definition 1.19 **Skalarprodukt, Prä-Hilbert-Raum.** Sei V ein reeller linearer Raum. Eine Abbildung $(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ heißt Skalarprodukt, wenn für alle $u, v, w \in V$ und für alle $\alpha, \beta \in \mathbb{R}$ die folgenden Bedingungen erfüllt sind:

- i) Symmetrie: $(u, v) = (v, u)$,
- ii) Linearität: $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$,
- iii) positive Definitheit: $(v, v) \geq 0$ und $(v, v) = 0$ genau dann wenn $v = 0$.

Wird der Raum V mit der Norm

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

versehen, so wird er Prä-Hilbert-Raum genannt. □

Beispiel 1.20 *Prä-Hilbert-Raum.* Sei $V = C([a, b])$. Dann definiert die Abbildung

$$(u, v) = \int_a^b u(x)v(x) dx \tag{1.7}$$

ein Skalarprodukt of V , was man einfach nachrechnen kann. Demzufolge ist $(V, \|\cdot\|_{L^2})$ ein Prä-Hilbert-Raum. Dieser Raum ist aber nicht vollständig in dieser Norm. Betrachte zum Beispiel das Intervall $[-1, 1]$ und die Funktionen

$$f_\varepsilon(x) = \begin{cases} 0, & x \in [0, -\varepsilon), \\ 1/(2\varepsilon)x + 1/2, & x \in [-\varepsilon, \varepsilon), \\ 1, & x \in [\varepsilon, 1], \end{cases} \quad \text{mit } \varepsilon > 0.$$

Es gilt

$$\int_{-1}^1 f_\varepsilon^2(x) dx = \frac{2\varepsilon}{3} + (1 - \varepsilon) = 1 - \frac{\varepsilon}{3}.$$

Demzufolge gilt für alle Funktionen $f_\varepsilon(x) \in (V, \|\cdot\|_{L^2})$. Sei nun $0 < \varepsilon_1 < \varepsilon_2$. Dann ist

$$\|f_{\varepsilon_1} - f_{\varepsilon_2}\|_{L^2}^2 = \int_{-1}^1 (f_{\varepsilon_1} - f_{\varepsilon_2})^2 dx = \int_{-\varepsilon_2}^{\varepsilon_2} (f_{\varepsilon_1} - f_{\varepsilon_2})^2 dx \leq 2\varepsilon_2,$$

da $(f_{\varepsilon_1} - f_{\varepsilon_2})^2 \leq 1$. Damit ist $\{f_\varepsilon\}$ eine Cauchy²-Folge in $(V, \|\cdot\|_{L^2})$ für $\varepsilon \rightarrow 0$, aber die Grenzfunktion ist eine unstetige Funktion.

Vervollständigt man $(V, \|\cdot\|_{L^2})$, so erhält man $L^2(-1, 1)$. □

Definition 1.21 **Hilbert-Raum.** Ein vollständiger Prä-Hilbert-Raum wird Hilbert-Raum genannt. □

²Augustin Louis Cauchy (1789 – 1857)

Beispiel 1.22 Hilbert–Raum. Hilbert–Räume sind zum Beispiel:

- \mathbb{R}^N ausgestattet mit dem Euklidischen Skalarprodukt,
- $L^2(a, b)$ ausgestattet mit dem Skalarprodukt (1.7).

□

Satz 1.23 Prä–Hilbert–Raum ist streng normiert. *Jeder Prä–Hilbert–Raum V ist streng normiert.*

Beweis: Analog zu Beispiel 1.18 erhält man für $u, v \in V$ mit $\|u + v\|_V = \|u\|_V + \|v\|_V$ dass

$$(u, v) = \|u\|_V \|v\|_V.$$

gilt. Dass man aus dieser Aussage folgern kann, dass u ein Vielfaches von v ist folgt aus einer Eigenschaft, welche man beim Beweis der Cauchy–Schwarz³–Ungleichung

$$(u, v) \leq \|u\|_V \|v\|_V \quad \forall u, v \in V \quad (1.8)$$

als Nebenprodukt erhält, siehe Literatur oder unten.

Für Interessenten: Beweis von (1.8). Ist $(u, v) = 0$, dann gilt (1.8) offensichtlich. Sei nun $(u, v) \neq 0$. Dann ist $\beta = (u, v)/|(u, v)|$ wohldefiniert. Nach Eigenschaft iii) des Skalarproduktes gilt für beliebiges $\lambda \in \mathbb{R}$

$$\begin{aligned} 0 &\leq (\beta u + \lambda v, \beta u + \lambda v) = \beta^2(u, u) + 2\lambda\beta(u, v) + \lambda^2(v, v) \\ &= (u, u) + 2\lambda\beta(u, v) + \lambda^2(v, v), \end{aligned} \quad (1.9)$$

da $\beta^2 = 1$. Das ist eine quadratische Funktion in λ , welche keine negativen Werte annehmen darf. Das ist genau dann der Fall, wenn die Diskriminante nicht-positiv ist

$$\frac{\beta^2(u, v)^2}{(v, v)^2} - \frac{(u, u)}{(v, v)} \leq 0 \iff (u, v)^2 \leq (u, u)(v, v) \iff |(u, v)| \leq \|u\|_V \|v\|_V.$$

Gleichheit gilt genau dann, wenn eine Nullstelle angenommen wird. Nach (1.9) muss dann

$$0 = (\beta u + \lambda v, \beta u + \lambda v) = \|\beta u + \lambda v\|_V$$

gelten, woraus nach einer Normeigenschaft

$$\beta u + \lambda v = 0 \iff u = -\frac{\lambda}{\beta}v$$

folgt. Also ist v ein Vielfaches von u . ■

Folgerung 1.24 Eindeutigkeit der Lösung der Bestapproximations–Aufgabe im Prä–Hilbert–Raum. *Im Prä–Hilbert–Raum besitzt die Bestapproximations–Aufgabe (1.4) eine eindeutige Lösung.*

Beweis: Die Aussage folgt sofort durch Kombination der Sätze 1.17 und 1.23. ■

Bemerkung 1.25 Charakterisierung der Bestapproximierenden in Prä–Hilbert–Räumen. In Prä–Hilbert–Räumen kann die Bestapproximierende mit Hilfe des Skalarproduktes charakterisiert werden. Diese Charakterisierung ist auch für die praktische Berechnung der Bestapproximierenden nützlich. □

Satz 1.26 Charakterisierung der Bestapproximierenden in Prä–Hilbert–Räumen *Sei V ein Prä–Hilbert–Raum. Dann ist die Bestapproximations–Aufgabe (1.4) äquivalent zur Lösung der Aufgabe: Finde $u \in U$, so dass*

$$(f - u, v) = 0 \quad \forall v \in U. \quad (1.10)$$

³Hermann Amandus Schwarz (1843 – 1921)

Beweis: Zum Beweis des Satzes müssen zwei Aussagen gezeigt werden.

i) Zu zeigen: Ist $u \in U$ die Lösung von (1.4), dann löst u auch (1.10). Sei $\alpha \neq 0$ beliebig aber fest gewählt und sei $v \in U$ beliebig gewählt. Da u die Lösung von (1.4) ist, folgt

$$\|f - u\|_V^2 \leq \|f - (u + \alpha v)\|_V^2 = \|(f - u) - \alpha v\|_V^2 = \|f - u\|_V^2 - 2\alpha(f - u, v) + \alpha^2 \|v\|_V^2.$$

Demzufolge ist

$$0 \leq -2\alpha(f - u, v) + \alpha^2 \|v\|_V^2.$$

Division durch $\alpha > 0$ und Grenzübergang $\alpha \rightarrow 0$ liefert

$$0 \leq -2(f - u, v).$$

Führt man dieselben Operationen für $\alpha < 0$ durch, erhält man

$$0 \geq -2(f - u, v).$$

Die beiden Ungleichungen können nur gemeinsam gelten, wenn

$$0 = (f - u, v).$$

Da $v \in U$ beliebig gewählt war, löst u (1.10).

ii) Zu zeigen: Ist $u \in U$ die Lösung von (1.10), dann löst u auch (1.4). Sei $v \in U$, $v \neq 0$, beliebig aber fest gewählt. Dann folgt

$$\|f - (u + v)\|_V^2 = \|f - u\|_V^2 - 2(f - u, v) + \|v\|_V^2 = \|f - u\|_V^2 + \|v\|_V^2 > \|f - u\|_V^2.$$

Damit löst u die Bestapproximationsaufgabe (1.4). ■

Definition 1.27 Orthogonalität. Sei V ein Prä-Hilbert-Raum. Die Elemente $u, v \in V$ heißen zueinander orthogonal, wenn $(u, v) = 0$. In Anlehnung an endlich-dimensionale Räume sagt man auch, dass u und v senkrecht aufeinander stehen. □

Bemerkung 1.28 *Berechnung der Bestapproximierenden mit Hilfe von (1.10).* Aus (1.10) folgt, dass der Raum U , und insbesondere die Bestapproximierende u , und der Fehler $f - u$ orthogonal zueinander sind.

Die Gleichung (1.10) ist der Ausgangspunkt bei der praktischen Berechnung der Bestapproximierenden. Sei U ein n -dimensionaler Raum. Dann wird U mit einer Basis $\{\varphi_1, \dots, \varphi_n\}$ ausgestattet. Damit existiert die eindeutige Darstellung

$$u = \sum_{j=1}^n u_j \varphi_j$$

mit zu berechnenden Koeffizienten $u_j \in \mathbb{R}$. Man kann zeigen (Übungsaufgabe) dass (1.10) genau dann für alle $v \in U$ erfüllt ist, wenn es für alle Basisfunktionen von U erfüllt ist. Damit erhält man n Gleichungen

$$(u, \varphi_i) = (f, \varphi_i) \implies \sum_{j=1}^n (\varphi_j, \varphi_i) u_j = (f, \varphi_i), \quad i = 1, \dots, n. \quad (1.11)$$

Setzt man

$$\begin{aligned} A &= (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}, a_{ij} = (\varphi_j, \varphi_i), \\ \mathbf{b} &= (b_i)_{i=1}^n \in \mathbb{R}^n, b_i = (f, \varphi_i), \\ \mathbf{u} &= (u_i)_{i=1}^n \in \mathbb{R}^n, \end{aligned} \quad (1.12)$$

so lässt sich (1.11) als lineares Gleichungssystem

$$A\mathbf{u} = \mathbf{b} \quad (1.13)$$

schreiben. Dieses System wird Normalgleichungen genannt. Die in (1.12) definierte Matrix A wird Gram⁴sche Matrix genannt. Sie ist symmetrisch und positiv definit (Übungsaufgabe). \square

Bemerkung 1.29 *Zum Normalgleichungssystem* (1.13). Für die praktische Rechnung wäre es besonders einfach, wenn A eine Diagonalmatrix ist. Das ist genau dann der Fall, wenn

$$(\varphi_i, \varphi_j) = 0 \quad \forall i, j = 1, \dots, n, \quad i \neq j.$$

In diesem Falle nennt man $\{\varphi_i\}_{i=1}^n$ eine Orthogonalbasis.

Aus einer vorgegebenen Basis kann man mit Hilfe des Gram–Schmidt⁵schen Orthogonalisierungsverfahrens eine Orthogonalbasis berechnen. Die Basisvektoren einer Orthogonalbasis sind nur bis auf einen Faktor bestimmt. \square

Beispiel 1.30 *Orthogonalbasis*. Seien $V = C([-1, 1])$ ausgestattet mit dem Skalarprodukt $(u, v) = \int_{-1}^1 u(x)v(x) dx$ und $U = P_2$ der Raum der quadratischen Polynome auf $[-1, 1]$. Eine Basis von P_2 ist sicherlich

$$\{\psi_1 = 1, \psi_2 = x, \psi_3 = x^2\}.$$

Diese Basis ist aber nicht orthogonal, da

$$(\psi_1, \psi_3) = \int_{-1}^1 \psi_1(x)\psi_3(x) dx = \int_{-1}^1 x^2 dx = \frac{2}{3}.$$

Beim Gram–Schmidtschen Orthogonalisierungsverfahrens fängt man mit einer Basisfunktion an und nimmt dann Schritt für Schritt weitere Funktionen zur Orthogonalbasis hinzu. Diese Funktionen erhält man durch Orthogonalisierung der ursprünglichen Basisfunktionen zu den bereits in der Orthogonalbasis befindlichen Funktionen. Setze $\varphi_1 = \psi_1$. Es gilt

$$\int_{-1}^1 \varphi_1(x)\psi_2(x) dx = \int_{-1}^1 x dx = 0.$$

Deswegen kann man $\varphi_2 = \psi_2$ wählen. Zur Berechnung von φ_3 werden die Anteile von ψ_3 abgezogen, die schon in dem von $\{\varphi_1, \varphi_2\}$ aufgespannten Raum liegen

$$\varphi_3 = \psi_3 - \frac{(\psi_3, \varphi_2)}{(\varphi_2, \varphi_2)}\varphi_2 - \frac{(\psi_3, \varphi_1)}{(\varphi_1, \varphi_1)}\varphi_1.$$

Man rechnet direkt nach, dass mit diesem Ansatz $(\varphi_3, \varphi_1) = (\varphi_3, \varphi_2) = 0$ gelten. Setzt man die Funktionen ein, so ergibt sich

$$\varphi_3 = x^2 - \frac{\int_{-1}^1 x^3 dx}{\int_{-1}^1 x^2 dx}x - \frac{\int_{-1}^1 x^2 dx}{\int_{-1}^1 dx} = x^2 - \frac{1}{3}.$$

Diese Herangehensweise kann man auf einen Polynomraum P_n beliebigen Grades erweitern. Orthogonale Polynome zum Skalarprodukt $(u, v) = \int_{-1}^1 u(x)v(x) dx$ haben dann die Form

$$\varphi_i(x) = \frac{i!}{(2i)!} \frac{d^i}{dx^i} (x^2 - 1)^i, \quad i = 0, \dots, n.$$

⁴Jørgen Pedersen Gram (1850 – 1916)

⁵Erhard Schmidt (1876 – 1959)

Bei dieser Orthogonalbasis besitzen die Basisfunktionen den Koeffizienten Eins vor dem Term höchsten Grades. Mit der Normierung

$$\varphi_i(x) = \frac{1}{2^i i!} \frac{d^i}{dx^i} (x^2 - 1)^i, \quad i = 0, \dots, n,$$

werden die Polynome Legendre⁶-Polynome genannt. \square

Satz 1.31 Drei-Term-Rekursion für orthogonale Polynome. *Zu jedem Skalarprodukt (\cdot, \cdot) auf P_n gibt es eindeutig bestimmte Orthogonalpolynome $\varphi_i \in P_n$, $i = 0, \dots, n$, wobei alle Basisfunktionen den Koeffizienten Eins vor dem Term höchster Ordnung besitzen. Diese Polynome genügen einer Drei-Term-Rekursion*

$$\varphi_i(x) = (x + \alpha_i)\varphi_{i-1}(x) + \beta_i\varphi_{i-2}(x), \quad i = 1, \dots, n,$$

mit den Anfangswerten

$$\varphi_{-1}(x) = 0, \varphi_0(x) = 1, \beta_1 = 0$$

und den Koeffizienten

$$\alpha_i = -\frac{(x\varphi_{i-1}, \varphi_{i-1})}{(\varphi_{i-1}, \varphi_{i-1})}, \beta_i = -\frac{(\varphi_{i-1}, \varphi_{i-1})}{(\varphi_{i-2}, \varphi_{i-2})}.$$

Beweis: Siehe Literatur, zum Beispiel [Deuffhard and Hohmann, 2008, Satz 6.2] oder [Beresin and Shidkow, 1970, Kapitel 5.4.2]. Man stellt $x\varphi_i(x)$ als Linearkombination der Basispolynome dar und erhält die Aussagen durch Koeffizientenvergleiche. \blacksquare

Bemerkung 1.32 Zur Bedeutung orthogonaler Polynome. Orthogonale Polynome und ihre Eigenschaften sind beispielsweise bei der Herleitung von Quadraturformeln optimaler Ordnung von Bedeutung, siehe Kapitel 4. \square

1.3 Tschebyscheff-Approximation mittels Polynomen

Bemerkung 1.33 Motivation. Dieser Abschnitt betrachtet den Spezialfall $V = C([a, b])$, $\|\cdot\|_V = \|\cdot\|_\infty$ und $U = P_n$. Man kann einfach Beispiele konstruieren, die zeigen, dass in diesem Fall kein streng normierter Raum vorliegt. (Übungsaufgabe) Demzufolge ist mit den bisherigen Erkenntnissen zwar die Existenz der Lösung der Bestapproximations-Aufgabe: finde $u \in P_n$ so dass

$$\|f - u\|_\infty \leq \|f - v\|_\infty \quad \forall v \in P_n, \quad (1.14)$$

gesichert, jedoch nicht ihre Eindeutigkeit.

In dem Spezialfall, dass $U = P_n$ ist, kann man jedoch auch Eigenschaften von Polynomen nutzen und zeigen, dass (1.14) eine eindeutige Lösung besitzt. Die zugehörige Mathematik ist etwas umfangreicher und soll hier nur angedeutet werden. \square

Definition 1.34 Haar-Raum und Tschebyscheff-System. Seien $v_1, \dots, v_n \in C([a, b])$ gegeben. Dann spannen diese Funktionen einen endlich-dimensionalen Unterraum

$$V_n = \text{span} \{v_1, \dots, v_n\} \subset C([a, b])$$

auf. Dieser Unterraum wird Haar⁷-Raum genannt, wenn für jedes $v \in V_n \setminus \{0\}$ gilt, dass es höchstens n verschiedene Nullstellen auf $[a, b]$ besitzt. Eine Basis eines Haar-Raumes wird als Tschebyscheff-System bezeichnet. \square

⁶Adrien-Marie Legendre (1752 – 1833)

⁷Alfréd Haar (1885 – 1933)

Beispiel 1.35 *Tschebyscheff-System.*

- Die Monome $\{1, x, x^2, \dots, x^n\}$ bilden ein Tschebyscheff-System der Dimension $n + 1$ in $C([a, b])$.
- Die Funktionen $\{1, e^x, e^{2x}, \dots, e^{nx}\}$ bilden ein Tschebyscheff-System der Dimension $n + 1$ in $C([a, b])$.
- Die Funktionen $\{1, \cos(x), \sin(x), \dots, \cos(nx), \sin(nx)\}$ bilden ein Tschebyscheff-System der Dimension $2n + 1$ in $C([0, 2\pi])$.

Die Richtigkeit des ersten Beispiels folgt daraus, dass jede Linearkombination der Basisfunktionen ein Polynom höchstens n -ten Grades ist. Nach dem Fundamentalsatz der Algebra besitzt solch ein Polynom höchstens n Nullstellen in \mathbb{R} .

Für die anderen Beispiele sei auf die Literatur verwiesen. \square

Satz 1.36 **Eindeutigkeit der Lösung der Bestapproximations-Aufgabe** (1.14). *Die Bestapproximations-Aufgabe (1.14) besitzt eine eindeutige Lösung.*

Beweis: Die Aussage des Satzes gilt nicht nur für Polynom-Räume sondern für Haarsche Räume. Der Beweis nutzt den sogenannten Tschebyscheffschen Alternantensatz. Für den detaillierten Beweis sei auf die Literatur verwiesen, beispielsweise auf [Beresin and Shidkow, 1970, Kap. 4.2.2]. \blacksquare

Beispiel 1.37 *Tschebyscheff-Approximation durch einen Polynomraum.* Seien $V = C([-1, 1])$, $f(x) = x^{n+1}$ und $U = P_n$. Dann lautet (1.14): finde $u \in P_n$ so dass

$$\|x^{n+1} - u\|_\infty \leq \|x^{n+1} - v\|_\infty \quad \forall v \in P_n.$$

Es ist $x^{n+1} - v$ ein Polynom vom Grad $(n + 1)$ auf $[-1, 1]$, dessen Betrag möglichst klein sein soll. Das Finden dieses Polynomes ist eine klassische Aufgabenstellung in der Analysis. Als Ergebnis erhält man

$$x^{n+1} - u = \frac{1}{2^n} T_{n+1}(x) \quad \implies \quad u = x^{n+1} - \frac{1}{2^n} T_{n+1}(x), \quad (1.15)$$

wobei $T_{n+1}(x)$ das Tschebyscheff-Polynom 1. Art vom Grad $(n + 1)$ auf $[-1, 1]$ ist. Die Tschebyscheff-Polynome sind beispielsweise gegeben durch, [Beresin and Shidkow, 1970, Kapitel 2.3.2] oder [Hanke-Bourgeois, 2006, Kap. 32],

$$T_n(x) = \cos(n \arccos(x)), \quad n = 0, 1, 2, \dots$$

Tschebyscheff-Polynome 1. Art sind orthogonal bezüglich des Skalarprodukts

$$(u, v) = \int_{-1}^1 \frac{u(x)v(x)}{\sqrt{1-x^2}} dx.$$

Nach Satz 1.31 genügen sie einer Drei-Term-Rukursion. Diese ist durch

$$T_0(x) = 1, T_1(x) = x, T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), n = 1, 2, \dots$$

gegeben.

Betrachtet man die Bestapproximations-Aufgabe für $n = 0$, so erhält man mit (1.15)

$$u = x - \frac{1}{2^0} T_1(x) = x - x = 0.$$

Das ist auch anschaulich klar. Für $n = 1$ erhält man mit (1.15)

$$u = x^2 - \frac{1}{2^1} T_2(x) = x^2 - \frac{1}{2} (2x^2 - 1) = \frac{1}{2}.$$

Man wird allgemein feststellen, dass man für (1.15) eine Lösung in P_{n-1} erhält. Das liegt daran, dass falls x^{n+1} eine gerade Funktion ist, die Bestapproximation auch eine gerade Funktion ist und analog falls x^{n+1} eine ungerade Funktion ist. \square

1.4 Approximation stetiger Funktionen in der Norm $\|\cdot\|_{L^2}$

Bemerkung 1.38 *Bekanntes.* Aus Beispiel 1.20 ist bekannt, dass $C([a, b])$ ausgestattet mit dem Skalarprodukt (1.7) und der dadurch induzierten Norm ein Prä-Hilbert-Raum ist. Des Weiteren ist bekannt, dass die Bestapproximations-Aufgabe (1.4) eine eindeutige Lösung besitzt und dass sie äquivalent zur Aufgabe (1.10) ist. Die prinzipielle Herangehensweise zur Lösung von (1.10) ist in Bemerkung 1.28 beschrieben.

Der einzige offene Punkt bei der Berechnung der Bestapproximation ist die Wahl der Basis des Raumes U . Dieser Punkt wird in diesem Abschnitt für $U = P_n$ und $U = S_n$ diskutiert. \square

Bemerkung 1.39 *Der Fall $U = P_n$.* Betrachte den Fall, dass Unterraum, in dem die Bestapproximierende gesucht ist, der Raum $U = P_n$ ist. Dann kann man die Bestapproximierende durch Aufstellung des Systems der Normalgleichungen (1.13) und dessen Lösung bestimmen. Für die konkrete Rechnung muss man den Raum P_n mit einer Basis ausstatten. Betrachte der Einfachheit halber das Intervall $[0, 1]$.

Eine naheliegende Wahl für eine Basis sind die Monome

$$\varphi_i(x) = x^i, \quad i = 0, \dots, n.$$

Dann besitzen die Einträge der Gramschen Matrix die Gestalt

$$a_{ij} = (\varphi_j, \varphi_i) = \int_0^1 x^j x^i dx = \int_0^1 x^{i+j} dx = \frac{1}{i+j+1} x^{i+j+1} \Big|_{x=0}^{x=1} = \frac{1}{i+j+1}.$$

Die Matrix A hat also die Gestalt

$$A = \begin{pmatrix} 1 & 1/2 & 1/3 & \dots \\ 1/2 & 1/3 & 1/4 & \dots \\ 1/3 & 1/4 & 1/5 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Das ist eine sogenannte Hilbert-Matrix. Die Kondition der Matrix ist verhält sich wie $\mathcal{O}((1 + \sqrt{2})^{4n} / \sqrt{n})$, sie wächst also exponentiell in n . Damit muss man bei der Lösung des Normalgleichungssystems schon für relativ kleine Werte von n mit großen Rundungsfehlern rechnen. Die Monome sind als Basis ungeeignet!

Wie bereits in Bemerkung 1.29 festgehalten, ist eine orthogonale Basis ideal für U . Im Beispiel 1.30 wurde die Konstruktion solch einer Basis für das Intervall $[-1, 1]$ demonstriert. Man erhält die Legendre-Polynome. Analog kann man natürlich auch für das Intervall $[0, 1]$ vorgehen. Auch diese Polynome werden Legendre-Polynome genannt. Man kann die Legendre-Polynome in $[0, 1]$ aus den Monomen analog zum Beispiel 1.30 berechnen. Alternativ kann man die Legendre-Polynome in $[0, 1]$ aus den Legendre-Polynomen in $[-1, 1]$ durch die Variablen-Transformation $x \mapsto 2x - 1$ bestimmen. Damit erhält man für die ersten drei Polynome

$$\begin{aligned} 1 &\rightarrow 1, \\ x &\rightarrow 2x - 1, \\ \frac{1}{2}(3x^2 - 1) &\rightarrow \frac{1}{2}(3(2x - 1)^2 - 1) = 6x^2 - 6x + 1. \end{aligned}$$

\square

Beispiel 1.40 *Bestapproximation in P_n in der Norm $\|\cdot\|_{L^2}$.* Es wird nun eine Erweiterung von Beispiel 1.6 betrachtet. Jetzt sollen die Bestapproximationen im Raum P_1 und P_2 gefunden werden.

Nach Bemerkung 1.39 ist es günstig, als Basis die Legendre–Polynome in $[0, \pi]$ zu wählen. Man erhält diese beispielsweise aus den Legendre–Polynomen in $[-1, 1]$ durch die Variablentransformation $x \mapsto \frac{2}{\pi}x - 1$

$$\begin{aligned} 1 &\rightarrow 1 = p_0, \\ x &\rightarrow \frac{2}{\pi}x - 1 = p_1, \\ \frac{1}{2}(3x^2 - 1) &\rightarrow \frac{1}{2}\left(3\left(\frac{2}{\pi}x - 1\right)^2 - 1\right) = \frac{6x^2 - 6\pi x + \pi^2}{\pi^2} = p_2. \end{aligned}$$

Betrachte zunächst die Bestapproximation in P_1 , das heißt es ist eine Bestapproximation der Gestalt $u = u_0p_0 + u_1p_1$ gesucht. Für das Normalgleichungssystem (1.13) erhält man

$$\mathbf{A}\mathbf{u} = \begin{pmatrix} (p_0, p_0) & (p_1, p_0) \\ (p_0, p_1) & (p_1, p_1) \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = \begin{pmatrix} \pi & 0 \\ 0 & \frac{\pi}{3} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = \begin{pmatrix} (\sin(x), p_0) \\ (\sin(x), p_1) \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Damit ist die Bestapproximation in P_1 gegeben durch $u = \frac{2}{\pi}$. Man erhält also eine Konstante. Der Fehler ist

$$\|\sin(x) - u\|_{L^2} = \left(\frac{\pi^2 - 8}{2\pi}\right)^{1/2} \approx 0.5454876555.$$

Nun wird die Bestapproximation in P_2 betrachtet. Man kann erwarten, dass man den Sinus in $[0, \pi]$ durch eine nach unten geöffnete Parabel schon recht gut approximieren kann. Die Bestapproximation hat die Gestalt $u = u_0p_0 + u_1p_1 + u_2p_2$ und für die Normalgleichungen erhält man

$$\mathbf{A}\mathbf{u} = \begin{pmatrix} \pi & 0 & 0 \\ 0 & \frac{\pi}{3} & 0 \\ 0 & 0 & \frac{\pi}{5} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ \frac{2\pi^2 - 24}{\pi^2} \end{pmatrix} = \mathbf{b}.$$

Damit hat die Bestapproximierende die Gestalt

$$\begin{aligned} u &= \frac{2}{\pi} + \frac{10\pi^2 - 120}{\pi^3}p_2 = \left(\frac{60}{\pi^3} - \frac{720}{\pi^5}\right)x^2 + \left(\frac{720}{\pi^4} - \frac{60}{\pi^2}\right)x + \left(\frac{12}{\pi} - \frac{120}{\pi^3}\right) \\ &\approx -0.41769x^2 + 1.13122x - 0.05046, \end{aligned}$$

siehe Abbildung 1.2. Der Fehler ist

$$\|\sin(x) - u\|_{L^2} = 0.0305999\dots$$

□

Bemerkung 1.41 *Der Fall $U = S_n$.* Der Raum S_n besteht aus Polygonzügen über $[a, b]$, welche über einem Gitter $a = x_0 < x_1 < \dots < x_n = b$ definiert sind, siehe Definition 1.8. Zur Aufstellung des Normalgleichungssystems (1.13) muss man eine günstige Basis in S_n wählen. Hierbei bietet es sich an, möglichst einfache Polygone zu nehmen. Das führt zur sogenannten Knotenbasis $\{\varphi_i\}_{i=0}^n$, welche durch die Eigenschaft

$$\varphi_i \in S_n : \varphi_i(x_k) = \delta_{ik}, \quad k = 0, \dots, n, \quad (1.16)$$

charakterisiert ist. Hierbei ist δ_{ik} das Kronecker⁸–Delta. Die Basisfunktion φ_i hat im Knoten i den Wert Eins und sie verschwindet in allen anderen Knoten.

⁸Leopold Kronecker (1823 – 1891)

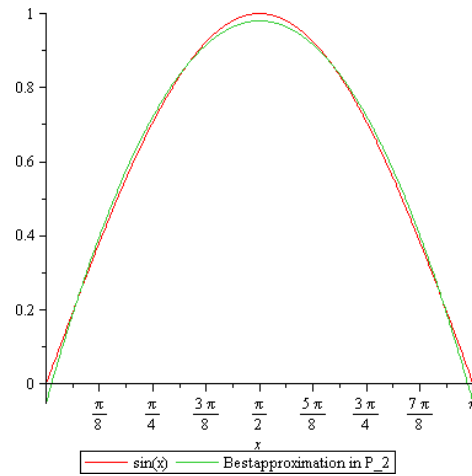


Abbildung 1.2: Bestapproximation von $\sin(x)$ in $[0, \pi]$ bezüglich $\|\cdot\|_{L^2}$ mit einem quadratischen Polynom.

Setzt man $h_i = x_i - x_{i-1}$, $i = 1, \dots, n$, so erhält man folgende Darstellungen der Basisfunktionen

$$\begin{aligned} \varphi_0(x) &= \begin{cases} 1 - \frac{1}{h_1}(x - x_0) & \text{für } x \in [x_0, x_1], \\ 0 & \text{sonst,} \end{cases} \\ \varphi_i(x) &= \begin{cases} 1 + \frac{1}{h_i}(x - x_i) & \text{für } x \in [x_{i-1}, x_i], \\ 1 - \frac{1}{h_{i+1}}(x - x_i) & \text{für } x \in [x_i, x_{i+1}], \\ 0 & \text{sonst,} \end{cases} \quad i = 1, \dots, n-1, \\ \varphi_n(x) &= \begin{cases} 1 + \frac{1}{h_n}(x - x_n) & \text{für } x \in [x_{n-1}, x_n], \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Auf Grund ihrer Gestalt, nennt man diese Basisfunktionen auch Hütchenfunktionen. Diese Funktionen spielen bei der numerischen Lösung von Randwertproblemen partieller Differentialgleichungen eine Rolle und sie werden dort linear Finite Elemente genannt.

Mit Hilfe der Knotenbasis hat die gesuchte Funktion u die Gestalt

$$u(x) = \sum_{j=0}^n u_j \varphi_j(x)$$

mit unbekanntem Koeffizienten u_j , $j = 1, \dots, n$. Dieser Polygonzug ist eindeutig bestimmt, wenn man seine Werte in allen Knoten kennt. Man erhält für den Knoten x_i , $i = 0, \dots, n$, wegen der Eigenschaft der Knotenbasis (1.16)

$$u(x_i) = \sum_{j=0}^n u_j \varphi_j(x_i) = u_i.$$

Das heißt, aus den Koeffizienten sieht man sofort die Werte des Polygonzuges in den Knoten. \square

Bemerkung 1.42 *Berechnung der Gramschen Matrix im Fall $U = S_n$.* Die Funktionen der Knotenbasis besitzen einen lokalen Träger. Sie sind höchstens in zwei nebeneinander liegenden Intervallen von Null verschieden. Damit folgt, dass

$$(\varphi_i, \varphi_j) = 0 \quad \text{falls} \quad |i - j| \geq 2,$$

weil dann $\varphi_i(x)\varphi_j(x) \equiv 0$ in $[a, b]$ ist. Damit ist klar, dass die Gramsche Matrix höchstens Einträge in der Diagonalen und den beiden Nebendiagonalen besitzen kann. Sie ist also eine Tridiagonalmatrix.

Zur Berechnung der Matrixeinträge muss man Integrale von Funktionen berechnen, die höchstens in zwei nebeneinander liegenden Teilintervallen von Null verschieden sind. In den Teilintervallen, in denen die Funktionen nicht verschwinden, sind sie quadratisch. Man erhält für $i = 1, \dots, n$,

$$\begin{aligned} a_{ii} &= (\varphi_i, \varphi_i) = \int_{x_{i-1}}^{x_{i+1}} \varphi_i^2(x) dx \\ &= \int_{x_{i-1}}^{x_i} \left(1 + \frac{1}{h_i}(x - x_i)\right)^2 dx + \int_{x_i}^{x_{i+1}} \left(1 - \frac{1}{h_{i+1}}(x - x_i)\right)^2 dx \\ &= \frac{h_i}{3} \left(1 + \frac{1}{h_i}(x - x_i)\right)^3 \Big|_{x_{i-1}}^{x_i} + \frac{(-h_{i+1})}{3} \left(1 - \frac{1}{h_{i+1}}(x - x_i)\right)^3 \Big|_{x_i}^{x_{i+1}} \\ &= \frac{h_i}{3} + \frac{h_{i+1}}{3} = \frac{1}{3}(h_i + h_{i+1}). \end{aligned}$$

Analog berechnet man

$$\begin{aligned} a_{00} &= \frac{h_1}{3}, \\ a_{nn} &= \frac{h_n}{3}, \\ a_{i,i+1} = a_{i-1,i} &= \frac{h_i}{6}, \quad i = 1, \dots, n. \end{aligned}$$

Definiert man

$$u_{-1} = u_{n+1} = h_0 = h_{n+1} = 0,$$

dann haben die Normalgleichungen die Gestalt

$$\frac{h_i}{6}u_{i-1} + \frac{h_i + h_{i+1}}{3}u_i + \frac{h_{i+1}}{6}u_{i+1} = (f, \varphi_i), \quad i = 0, \dots, n.$$

Oft werden die Gleichungen mit $6/(h_i + h_{i+1})$ durchmultipliziert. Dann erhält man das Normalgleichungssystem

$$\mathbf{A}\mathbf{u} = \begin{pmatrix} 2 & \lambda_0 & & & & & \\ \mu_1 & 2 & \lambda_1 & & & & \\ & \mu_2 & 2 & \lambda_2 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \lambda_{n-1} & \\ & & & & \mu_n & 2 & \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix} = \frac{6}{h_i + h_{i+1}} \begin{pmatrix} (f, \varphi_0) \\ (f, \varphi_1) \\ (f, \varphi_2) \\ \vdots \\ (f, \varphi_{n-1}) \\ (f, \varphi_n) \end{pmatrix} \quad (1.17)$$

mit

$$\mu_i = \frac{h_i}{h_i + h_{i+1}}, \quad \lambda_i = \frac{h_{i+1}}{h_i + h_{i+1}}.$$

Man kann zeigen, dass man die LU-Zerlegung einer Tridiagonalmatrix mit $\mathcal{O}(n)$ floating point Operationen berechnen kann, siehe [Kielbasiński and Schwetlick, 1988,

Kap. 6.4]. Damit ist das Normalgleichungssystem (1.17) mit optimalem Aufwand lösbar. Eine Abschätzung der Kondition dieser Matrix wird in Satz 1.44 bewiesen. \square

Beispiel 1.43 *Bestapproximation in S_2 in der Norm $\|\cdot\|_{L^2}$.* Nun wird Beispiel 1.9 fortgesetzt. Die Basisfunktionen von S_2 mit dem Knoten $x_1 = \pi/2$ haben die Form

$$\begin{aligned}\varphi_0(x) &= \begin{cases} 1 - \frac{2}{\pi}x & \text{für } x \in [0, \pi/2], \\ 0 & \text{sonst,} \end{cases} \\ \varphi_1(x) &= \begin{cases} 1 + \frac{2}{\pi}\left(x - \frac{\pi}{2}\right) & \text{für } x \in [0, \pi/2], \\ 1 - \frac{2}{\pi}\left(x - \frac{\pi}{2}\right) & \text{für } x \in (\pi/2, \pi], \\ 0 & \text{sonst,} \end{cases} \\ \varphi_2(x) &= \begin{cases} 1 + \frac{2}{\pi}(x - \pi) & \text{für } x \in [\pi/2, \pi], \\ 0 & \text{sonst.} \end{cases}\end{aligned}$$

Der Ansatz für die Bestapproximierende ist

$$u(x) = u_0\varphi_0(x) + u_1\varphi_1(x) + u_2\varphi_2(x).$$

Für das Normalgleichungssystem erhält man

$$\begin{aligned}\begin{pmatrix} (\varphi_0, \varphi_0) & (\varphi_1, \varphi_0) & (\varphi_2, \varphi_0) \\ (\varphi_0, \varphi_1) & (\varphi_1, \varphi_1) & (\varphi_2, \varphi_1) \\ (\varphi_0, \varphi_2) & (\varphi_1, \varphi_2) & (\varphi_2, \varphi_2) \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} &= \begin{pmatrix} \frac{\pi}{6} & \frac{\pi}{12} & 0 \\ \frac{\pi}{12} & \frac{\pi}{3} & \frac{\pi}{12} \\ 0 & \frac{\pi}{12} & \frac{\pi}{6} \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} \\ &= \begin{pmatrix} (\sin(x), \varphi_0) \\ (\sin(x), \varphi_1) \\ (\sin(x), \varphi_2) \end{pmatrix} = \begin{pmatrix} 1 - \frac{2}{\pi} \\ \frac{4}{\pi} \\ 1 - \frac{2}{\pi} \end{pmatrix}.\end{aligned}$$

Die Lösung dieses Systems ist

$$\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} = \frac{1}{\pi^2} \begin{pmatrix} 8(\pi - 3) \\ -4\pi + 24 \\ 8\pi - 3 \end{pmatrix} \approx \begin{pmatrix} 0.11477 \\ 1.15846 \\ 0.11477 \end{pmatrix}.$$

Damit hat die Bestapproximierende die Gestalt

$$u = 0.11477\varphi_0(x) + 1.15846\varphi_1(x) + 0.11477\varphi_2(x),$$

siehe Abbildung 1.3. Der Fehler ist

$$\|\sin(x) - u\|_{L^2} = \sqrt{\frac{1}{2} \left(\pi - \frac{32}{\pi} + \frac{192}{\pi^2} - \frac{384}{\pi^3} \right)} \approx 0.11125.$$

\square

Satz 1.44 *Kondition der Systemmatrix des Normalgleichungssystems (1.17). Es gilt*

$$\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty \leq 3,$$

wobei $\|\cdot\|_\infty$ die Zeilensummennorm einer Matrix ist

$$\|A\|_\infty = \max_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{i=1, \dots, m} \sum_{j=1}^n |a_{ij}|, \quad A \in \mathbb{R}^{m \times n}. \quad (1.18)$$

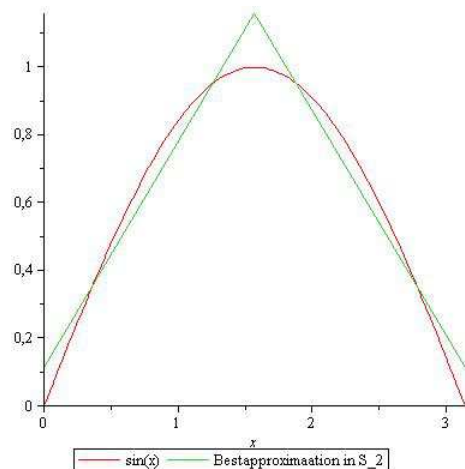


Abbildung 1.3: Bestapproximation von $\sin(x)$ in $[0, \pi]$ bezüglich $\|\cdot\|_{L^2}$ mit einem Polygon aus S_2 .

Beweis: Da bekannt ist, dass (1.17) für alle rechte Seiten eine eindeutige Lösung besitzt, Folgerung 1.24, muss A eine nicht-singuläre Matrix sein und A^{-1} existiert und ist ebenfalls nicht-singulär.

In jeder Zeile von A gibt es höchstens drei Einträge. Es gilt

$$|\mu_i| + |a_{ii}| + |\lambda_i| \leq \frac{h_i}{h_i + h_{i+1}} + 2 + \frac{h_{i+1}}{h_i + h_{i+1}} = 3, \quad i = 0, \dots, n.$$

Also folgt $\|A\|_\infty = 3$.

Der Term $\|A^{-1}\|_\infty$ wird abgeschätzt. Sei $\mathbf{v} \in \mathbb{R}^{n+1}$ beliebig aber fest gewählt. Es gilt

$$(A\mathbf{v})_i = \mu_i v_{i-1} + 2v_i + \lambda_i v_{i+1} \iff v_i = \frac{1}{2} ((A\mathbf{v})_i - \mu_i v_{i-1} - \lambda_i v_{i+1}),$$

$i = 0, \dots, n$. Dann folgt mit Dreiecksungleichung

$$\begin{aligned} |v_i| &\leq \frac{1}{2} |(A\mathbf{v})_i| + \frac{1}{2} \mu_i |v_{i-1}| + \frac{1}{2} \lambda_i |v_{i+1}| \\ &\leq \frac{1}{2} \max_{i=0, \dots, n} |(A\mathbf{v})_i| + \frac{1}{2} \mu_i \max_{i=0, \dots, n} |v_i| + \frac{1}{2} \lambda_i \max_{i=0, \dots, n} |v_i| \\ &= \frac{1}{2} \|A\mathbf{v}\|_\infty + \frac{1}{2} (\mu_i + \lambda_i) \|\mathbf{v}\|_\infty \\ &= \frac{1}{2} \|A\mathbf{v}\|_\infty + \frac{1}{2} \|\mathbf{v}\|_\infty, \quad i = 0, \dots, n. \end{aligned}$$

Diese Ungleichung gilt für alle Indizes, also gilt sie auch für das Maximum

$$\|\mathbf{v}\|_\infty = \max_{i=0, \dots, n} |v_i| \leq \frac{1}{2} \|A\mathbf{v}\|_\infty + \frac{1}{2} \|\mathbf{v}\|_\infty \iff \|\mathbf{v}\|_\infty \leq \|A\mathbf{v}\|_\infty. \quad (1.19)$$

Da A^{-1} nicht-singulär ist, bildet A^{-1} den Raum \mathbb{R}^{n+1} auf sich selber ab. Damit gibt es einen Vektor $\mathbf{w} \in \mathbb{R}^{n+1}$ mit $\mathbf{v} = A^{-1}\mathbf{w}$ für jeden Vektor $\mathbf{v} \in \mathbb{R}^{n+1}$. Man erhält mit der Definition der Norm (1.18) und (1.19)

$$\|A^{-1}\|_\infty = \max_{\mathbf{w} \in \mathbb{R}^{n+1} \setminus \{0\}} \frac{\|A^{-1}\mathbf{w}\|_\infty}{\|\mathbf{w}\|_\infty} = \max_{\mathbf{v} \in \mathbb{R}^{n+1} \setminus \{0\}} \frac{\|\mathbf{v}\|_\infty}{\|A\mathbf{v}\|_\infty} \leq 1.$$

Damit ist die Aussage des Satzes bewiesen. ■

Bemerkung 1.45 *Zum Satz 1.44.* Die Matrix ist also sehr gut konditioniert. Da man zudem nur relativ wenige Matrixeinträge berechnen muss, deren Berechnung leicht ist und sich das entstehende lineare System mit optimalem Aufwand lösen lässt, ist die Knotenbasis (1.16) eine gute Wahl als Basis von S_n . □