8. Tschebyscheff-Approximation: Theorie

Im Folgenden untersuchen wir Bestapproximationen bezüglich der Maximumsnorm. Die Wurzeln dieser Theorie gehen auf Pafnuti Lwowitsch Tschebyscheff (1821–1894) zurück. Tschebyscheff untersuchte polynomiale und rationale Bestapproximationen für stetige Funktionen $f \in \mathrm{C}[a,b]$ bezüglich $\|\cdot\|_\infty$. Die zentrale Aussagen dieser Theorie, der so genannte Alternantensatz (auch Satz von Tschebyscheff genannt) wurde jedoch von Blichfeld (1901) und Kirchfelder (1902) bewiesen. Tschebyscheff zeigte die folgende schwächere Aussage:

Ist $p \in \Pi_n$ eine Bestapproximation zu $f \in \mathrm{C}^1[a,b]$, so gibt es wenigstens $(n+2)$ kritische Punkte, das sind Randpunkte des Intervalls oder stationäre Punkte der Fehlerfunktion $e := f - p$.

Haarsche Räume.

Es sei $B \subset \mathbb{R}$ eine nichtleere und kompakte Menge und $R := \mathrm{C}(B)$ ausgestattet mit der Maximumsnorm $\|f\|_\infty := \max_{x\in B} |f(x)|$.

Natürlich lässt sich auch eine komplexe Variante der obigen Voraussetzungen formulieren, also $B \subset \mathbb{C}$, und die meisten der folgenden Aussagen bleiben unverändert oder mit geringen Modifikationen gültig. Anders sieht es jedoch beim mehrdimensionalen Fall, also $B \subset \mathbb{R}^m$, bzw. $B \subset \mathbb{C}^m$ mit $m > 1$ aus, wie wir sogleich sehen werden.

Gegeben sei wieder ein endlich dimensionaler, linearer Teilraum, $V \subset R$. Typische Beispiele sind mit $B = [a,b]$ die Polynomräume, $V := \Pi_n[a,b]$ mit $\dim V = n+1$, die trigonometrischen Polynome $V = T_n[a,b]$ mit $\dim V = 2n+1$, oder die Spline-Räume $V = S_m(t_0,\ldots,t_n)$ mit $\dim V = m+n$.

Definition und Satz (8.1)

Sei $V = V_n$ ein Teilraum von R der Dimension $\dim V = n+1$. Dann sind die folgenden drei Eigenschaften von V äquivalent:

(H1) Jedes Element $p \in V$, $p \neq 0$, hat höchstens n Nullstellen.

(H2) Zu $(n+1)$ Stützstellen $(t_j, f_j) \in B \times \mathbb{R}$, $j = 0,1,\ldots,n$, mit paarweise verschiedenen t_j gibt es genau eine interpolierende Funktion $p \in V$.

(H3) Ist $(h_0,\ldots,h_n)$ irgendeine Basis von V und sind $t_0,\ldots,t_n \in B$ paarweise verschieden, so ist die Matrix

$$D(t_0,\ldots,t_n) \;:=\; \begin{pmatrix} h_0(t_0) & \ldots & h_n(t_0) \\ \vdots & & \vdots \\ h_0(t_n) & \ldots & h_n(t_n) \end{pmatrix}$$

regulär in $\mathbb{R}^{(n+1,n+1)}$.

Erfüllt V diese Eigenschaften, so heißt V ein *Haarscher Raum*, benannt nach Alfred Haar (1885 – 1933). Eine beliebige Basis eines Haarschen Raumes heißt auch ein *Haarsches System* oder ein *Tschebyscheff-System* in $\mathrm{C}(B)$.

Beweis: Die Äquivalenz von (H2) und (H3) ist aus der Numerik wohlbekannt: Die eindeutige Existenz einer interpolierenden Funktion $p = \sum \alpha_j h_j \in V$ ist äquivalent zur eindeutigen Lösbarkeit des linearen Gleichungssystems

$$\begin{pmatrix} h_0(t_0) & \ldots & h_n(t_0) \\ \vdots & & \vdots \\ h_0(t_n) & \ldots & h_n(t_n) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} f_0 \\ \vdots \\ f_n \end{pmatrix}$$

und damit zur Regularität der Koeffizientenmatrix.

Ist (H3) nicht erfüllt, so gibt es eine Basis $(h_0,\ldots,h_n)$ von V und paarweise verschiedene Punkte $t_0,\ldots,t_n \in B$, so dass das obige lineare Gleichungssystem für $f_j = 0$ eine Lösung $\alpha \neq 0$ besitzt. $p := \sum \alpha_j h_j$ ist damit eine Element aus $V\setminus\{0\}$ mit wenigstens $(n+1)$ Nullstellen. Damit ist auch (H1) nicht erfüllt. Umgekehrt: Gilt (H3), so hat das homogene lineare Gleichungssystem nur die triviale Lösung $\alpha \neq 0$. Jedes $p \in V$, $p \neq 0$, hat daher höchstens n Nullstellen. □

Beispiel (8.2)

Die Monome $(1, t, \ldots, t^n)$ bilden ein Haarsches System in $\mathrm{C}[a,b]$, $a < b$.

Die Aussage (H1) hängt dabei mit dem Fundamentalsatz der Algebra zusammen, (H2) entspricht der Existenz und Eindeutigkeit des Problems der Interpolation durch Polynome, (H3) entspricht schließlich der Regularität der Vandermonde-Matrix.

Das Beispiel lässt sich unmittelbar auf $\mathbb{C}$ übertragen: $(1, z, \ldots, z^n)$ ist ein Haarsches System in $\mathrm{C}(B;\mathbb{C})$, wobei $B \subset \mathbb{C}$ wenigstens $(n+1)$ Punkte enthalten muss.

Beispiel (8.3)

Die trigonometrischen Funktionen $(1, \cos t, \sin t, \ldots, \cos(nt), \sin(nt))$ bilden ein Haarsches System in $\mathrm{C}[0, 2\pi[$.

Beweis: Man sieht dies anhand der komplexen Darstellung

$$p(t) \;=\; \frac{a_0}{2} \;+\; \sum_{k=1}^{n} [a_k \cos(kt) + b_k \sin(kt)] \;=\; \sum_{k=-n}^{n} \gamma_k \mathrm{e}^{ikt} \;=\; \mathrm{e}^{-int} \sum_{k=0}^{2n} \gamma_{k-n} z^k$$

mit $\gamma_0 = a_0/2$, $\gamma_k = (a_k - ib_k)/2$, $\gamma_{-k} = (a_k + ib_k)/2$ $(k > 0)$ und $z = \mathrm{e}^{it}$.

Zu $(2n+1)$ verschiedenen Punkten in $[0, 2\pi[$ sind auch die $z_j := \mathrm{e}^{it_j}$ paarweise verschieden. Daher gibt es zu den Knoten $(z_j, \mathrm{e}^{int_j} f_j)$, $j = 0, \ldots, 2n$, ein eindeutig bestimmtes Interpolationspolynom $q(z) = \sum_{k=0}^{2n} \gamma_{k-n} z^k$. Wegen

$$f_j = \mathrm{e}^{-int_j} \sum_{k=0}^{2n} \gamma_{k-n} z_j^k = \sum_{k=0}^{2n} \gamma_{k-n} \mathrm{e}^{i(k-n)t_j} = \sum_{k=0}^{2n} \overline{\gamma_{k-n}} \mathrm{e}^{i(n-k)t_j} = \mathrm{e}^{-int_j} \sum_{\ell=0}^{2n} \overline{\gamma_{n-\ell}} z_j^\ell$$

gilt $\gamma_{k-n} = \overline{\gamma_{n-k}}$, $k = 0, \ldots, 2n$. Damit ist das zugehörige interpolierende trigonometrische Polynom $p(t)$ – mit *reellen* Koeffizienten a_k, b_k – ebenfalls eindeutig bestimmt.
$\square$

Durch gerade bzw. ungerade Fortsetzung einer Funktion aus $\mathrm{C}([0, \pi[)$ bzw. $\mathrm{C}(]0, \pi[)$ ergibt sich:

$(1, \cos t, \ldots, \cos(nt))$ ist ein Haarsches System in $\mathrm{C}([0, \pi[)$, $(\sin t, \ldots, \sin(nt))$ ist ein Haarsches System in $\mathrm{C}(]0, \pi[)$.

Beispiel (8.4)

Für $\lambda_0 < \lambda_1 < \ldots < \lambda_n$ bilden die Funktionen $(\mathrm{e}^{\lambda_0 t}, \ldots, \mathrm{e}^{\lambda_n t})$ ein Haarsches System in $\mathrm{C}[a, b]$.

Beweis: (per Induktion über n)
Für $n = 0$ ist die Gültigkeit von (H1) klar. Zum Induktionsschritt: Hat $p(t) = \sum_{k=0}^{n} a_k \mathrm{e}^{\lambda_k t}$ $(n+1)$ Nullstellen in $[a, b]$, so hat $q(t) := (\mathrm{e}^{-\lambda_0 t} p(t))'$ nach dem Satz von Rolle dort wenigstens n Nullstellen. Da aber $q \in \mathrm{Spann}(\mathrm{e}^{(\lambda_1 - \lambda_0)t}, \ldots, \mathrm{e}^{(\lambda_n - \lambda_0)t})$ folgt aus der Induktionsvoraussetzung, dass $q = 0$ ist. Damit ergibt sich aber auch $p = 0$.
$\square$

Eine Variante des obigen Beweises zeigt weiter: Für $\lambda_0 < \lambda_1 < \ldots < \lambda_n$ und $m_i \in \mathbb{N}_0$, $i = 0, \ldots, n$, ist $(\mathrm{e}^{\lambda_0 t}, t\,\mathrm{e}^{\lambda_0 t}, \ldots, t^{m_0} \mathrm{e}^{\lambda_0 t}, \ldots, t^{m_n} \mathrm{e}^{\lambda_n t})$ ein Haarsches System in $\mathrm{C}[a, b]$, $a < b$.

Ferner folgt mittels Substitution $\mathrm{e}^t \to x$:

Für $\lambda_0 < \lambda_1 < \ldots < \lambda_n$ ist $(x^{\lambda_0}, \ldots, x^{\lambda_n})$ ein Haarsches System in $\mathrm{C}[a, b]$, $0 < a < b$.

Beispiel (8.5) (Mairhuber, Proc. AMS 7, 1956)

Die folgende Konstruktion von Mairhuber zeigt, dass im (reellen) mehrdimensionalen Fall, $\mathrm{C}(B)$, $B \subset \mathbb{R}^m$, $m \geq 2$, i. Allg. kein Haarsches System existiert.

Es sei $m = 2$ und die Menge $B \subset \mathbb{R}^2$ enthalte eine y-förmige Teilmenge. Legt man die Punkte $z_0, \ldots, z_n$ wie in Abb. 8.1 und nimmt an, dass $h_0, \ldots, h_n$ ein Haarsches System ist, so darf $d(z_0, \ldots, z_n) := \det D(z_0, \ldots, z_n)$ nicht verschwinden. d hängt dabei stetig von den $z_0, \ldots, z_n$ ab.

Wir führen nun folgende Punktverschiebung durch: z_0 wandere über Position a zur Position b, anschließend wandere z_1 über die Position a zur alten Position von z_0 und sodann z_0 von Position b über a zur alten Position von z_1.

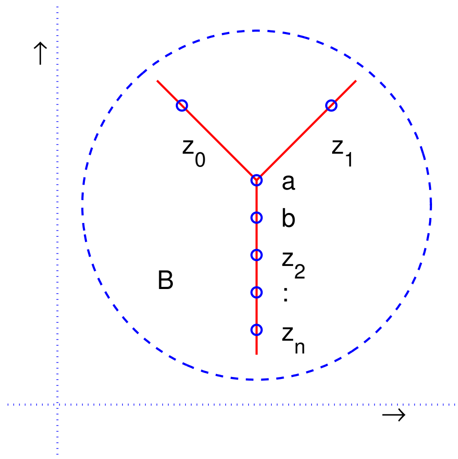

Abb. 8.1: Beispiel von Mairhuber

Bei diesem ganzen Prozeß bleiben die Punkte jeweils paarweise verschieden, so dass $d(z_0, \ldots, z_n)$ sein Vorzeichen nicht wechselt. Andererseits entsteht die Endposition durch Vertauschung der ersten beiden Zeilen in der Matrix $D(z_0, \ldots, z_n)$. Damit folgt aber $d(z_0, z_1, \ldots, z_n) = -d(z_1, z_0, \ldots, z_n)$ und wir erhalten einen Widerspruch! □

Für den eindimensionalen, reellen Fall mit $B = [a, b]$ sind die folgenden weiteren Eigenschaften Haarscher Räume hilfreich.

Satz (8.6)

Sei V ein $(n+1)$ dimensionaler Haarscher Teilraum von $\mathrm{C}[a, b]$. Dann gelten

(H4) Hat $p \in V \setminus \{0\}$ im Intervall $[a, b]$ m Nullstellen, von denen k Nullstellen ohne Vorzeichenwechsel sind (diese liegen dann im offenen Intervall $]a, b[$), so gilt $m + k \leq n$.

(H5) Zu $k \leq m \in \mathbb{N}_0$, $m + k = n$, und paarweise verschiedenen Punkten $t_1, \ldots, t_k \in]a, b[$, $t_{k+1}, \ldots, t_m \in [a, b]$ existiert eine Funktion $p \in V \setminus \{0\}$, die genau die Nullstellen $t_1, \ldots, t_m$ besitzt und für die $t_1, \ldots, t_k$ Nullstellen ohne Vorzeichenwechsel sind.

(H6) Zu $m \leq n$ vorgegebenen paarweise verschiedenen Punkten $t_1, \ldots, t_m \in]a, b[$, existiert eine Funktion $p \in V \setminus \{0\}$, die genau die Nullstellen $t_1, \ldots, t_m$ besitzt und diese sämtlich Nullstellen mit Vorzeichenwechsel sind.

Beweis: Für den Beweis dieser Aussagen sei auf das Lehrbuch von Powell (Anhang) verwiesen. □

Das Kolmogoroff-Kriterium.

Es sei wieder $B \subset \mathbb{R}$ eine kompakte Menge, $R = \mathrm{C}(B)$ ausgestattet mit der Maximumsnorm $\|\cdot\|_\infty$. Weiter sei $f \in \mathrm{C}(B)$ und V ein $(n+1)$-dimensionaler linearer Teilraum von $\mathrm{C}(B)$.

Definition (8.7)

Zu $p \in V$ heißt $e := f - p$ die *Fehlerfunktion* der Approximation p von f. Die Menge

$$A \;=\; A(f,p) \;:=\; \{t \in B: \; |e(t)| \;=\; \|e\|_\infty\}$$

heißt *Menge der Extremalpunkte* von p bezüglich f.

Es gibt nun die folgende Charakterisierung einer Bestapproximation durch ihre Extremalpunkte (nach Andrey Nikolaevich Kolmogoroff; 1903 – 1987)

Satz (8.8) (Kolmogoroff, 1948)

$p \in V$ ist genau dann eine Bestapproximation von $f \in \mathrm{C}(B)$ bezüglich V, $\|\cdot\|_\infty$, wenn das folgende Kriterium gilt

$$\forall\, q \in V: \quad \min\{(f(t) - p(t))\, q(t): \; t \in A(f,p)\} \;\leq\; 0.$$

Beweis:

⇒: Angenommen, das Kolmogoroff-Kriterium gilt nicht, d.h. es gibt ein $q \in V$ und ein $\varepsilon > 0$, so dass

$$\forall\, t \in A(f,p): \quad (f(t) - p(t))\, q(t) \;>\; 2\,\varepsilon.$$

Da $A(f,p)$ kompakt ist, gibt es eine in B offene Menge $U \supset A(f,p)$ mit

$$\forall\, t \in U: \quad (f(t) - p(t))\, q(t) \;>\; \varepsilon.$$

Mit $M := \|q\|_\infty$, $p_1 := p + \lambda q$, $\lambda > 0$ folgt dann für $t \in U$:

$$\begin{aligned}
(f(t) - p_1(t))^2 &= ((f(t) - p(t)) - \lambda q(t))^2 \\
&= (f(t) - p(t))^2 \;-\; 2\,\lambda\,(f(t) - p(t))\, q(t) \;+\; \lambda^2 (q(t))^2 \\
&< \|e\|_\infty^2 \;-\; 2\,\lambda\,\varepsilon \;+\; \lambda^2\, M^2 \\
&< \|e\|_\infty^2 \;-\; \lambda\,\varepsilon, \qquad \text{für} \quad 0 < \lambda < \varepsilon/M^2.
\end{aligned}$$

Da $B \setminus U$ kompakt ist und $A(f,p) \subset U$ existiert ein $\delta > 0$ mit $\forall\, t \in B \setminus U$: $|f(t) - p(t)| < \|e\|_\infty - \delta$.

Für $\lambda < \delta/(2M)$ und $t \in B \setminus U$ folgt:

$$\begin{aligned}
|f(t) - p_1(t)| &\leq |f(t) - p(t)| \;+\; \lambda\,|q(t)| \\
&\leq \|e\|_\infty \;-\; \delta \;+\; \delta/(2M)\, M \;=\; \|e\|_\infty \;-\; \delta/2.
\end{aligned}$$

Insgesamt ist damit gezeigt: $\|f - p_1\|_\infty \;<\; \|f - p\|_\infty$, im Widerspruch zur Minimaleigenschaft von p.

$\Leftarrow$: Gelte das Kolmogoroff-Kriterium und sei $p_1 \in V$, $q := p_1 - p$.
Dann existiert ein $t_0 \in A(f,p)$ mit $(f(t_0) - p(t_0))\, q(t_0) \le 0$. Damit folgt

$$\begin{aligned}
(f(t_0) - p_1(t_0))^2 &= ((f(t_0) - p(t_0)) - q(t_0))^2 \\
&= (f(t_0) - p(t_0))^2 \;-\; 2\,(f(t_0) - p(t_0))\, q(t_0) \;+\; q(t_0)^2 \\
&\ge (f(t_0) - p(t_0))^2 \;=\; \|f - p\|^2.
\end{aligned}$$

Es folgt $\|f - p_1\|_\infty \;\ge\; \|f - p\|_\infty$, $\forall\, p_1 \in V$. Damit ist p Bestapproximation von f aus V. $\square$

Bemerkung (8.9)

Für den komplexen Fall $B \subset \mathbb{C}$ (kompakt) lautet das Kolmogoroff-Kriterium

$$\forall\, q \in V: \quad \min\{\, \mathrm{Re}\,[(f(t) - p(t))\, \overline{q(t)}\,] : \; t \in A(f,p)\,\} \;\le\; 0.$$

Der Beweis erfolgt analog zu dem von Satz (8.8).

Beispiel (8.10)

Sei $R := \mathrm{C}[1,2]$, $V := \Pi_1[1,2]$ und $f(t) = t^2$. Zunächst bestimmen wir analog zu den Beispielen 1.3b) und 1.6c) einen Kandidaten $p(t) = p^*(t) = a_0 + a_1\, t$ aus dem nichtlinearen Gleichungssystem

$$\begin{aligned}
f(1) \;-\; p(1) &= \delta, \\
f(\tau) \;-\; p(\tau) &= -\delta, \\
f(2) \;-\; p(2) &= \delta, \\
f'(\tau) \;-\; p'(\tau) &= 0.
\end{aligned}$$

Eine einfache Rechnung liefert die (eindeutige) Lösung $a_0 = -17/8$, $a_1 = 3$ und $\tau = 1.5$. Damit ist

$$\|f - p^*\|_\infty \;=\; \delta \;=\; 1/8, \quad A(f,p^*) \;=\; \{1, 1.5, 2\}.$$

Ist p^* nun tatsächlich eine Bestapproximation? Das Kolmogoroff-Kriterium besagt, dass für alle $q \in \Pi_1$ gelten muss:

$$\min\{\delta\, q(1),\; -\delta\, q(3/2),\; \delta\, q(2)\} \;\le\; 0.$$

Wäre dies nicht der Fall, so müsste $q(1) > 0$, $q(3/2) < 0$ und $q(2) > 0$ sein, im Widerspruch zu $q \in \Pi_1$. Damit ist gezeigt, dass das Kolmogoroff-Kriterium erfüllt ist; p^* ist also tatsächlich Bestapproximation von f aus V!

Eine Folgerung aus dem Kolmogoroff-Kriterium ist der folgende Eindeutigkeitssatz für $\|\cdot\|_\infty$–Bestapproximationen.

Satz (8.11) (Haarscher Eindeutigkeitssatz)

Sei wieder $B \subset \mathbb{R}$ kompakt, $R := \mathrm{C}(B)$ ausgestattet mit $\|\cdot\|_\infty$. Ferner sei $V \subset R$ ein $(n+1)$–dimensionaler Haarscher Teilraum von R. Dann gelten

a) Ist $p \in V$ Bestapproximation von $f \in R \setminus V$ bezüglich V, so enthält $A(f,p)$ wenigstens $(n+2)$ Punkte.

b) Zu jedem $f \in R$ gibt es genau eine Bestappoximation aus V.

Beweis: zu a) Nehmen wir an, es gäbe $(n+1)$ paarweis verschiedene Punkte $t_0, \ldots, t_n \in B$ mit $A(f,p) \subset \{t_0, \ldots, t_n\}$. Nach (H2) existiert dann $q \in V$ mit $q(t_j) = e(t_j) := f(t_j) - p(t_j)$. Für alle $t_j \in A(f,p)$ gilt somit

$$(f(t_j) - p(t_j))\, q(t_j) \;=\; e(t_j)^2 \;=\; \|f - p\|_\infty^2 \;>\; 0.$$

Das Kolmogoroff-Kriterium ist damit *nicht* erfüllt und somit p keine Bestapproximation. Widerspruch!

zu b) Seien p_1, p_2 Bestapproximationen an f aus V und gelte o.B.d.A. $f \in \mathrm{C}(B) \setminus V$. Dann ist auch $p = (p_1 + p_2)/2$ eine Bestapproximation.

Nach Teil a) existieren wenigstens $(n+2)$ Extremalpunkte $t_0, \ldots, t_{n+1} \in A(f,p)$. Es gilt also

$$f(t_j) - p(t_j) \;=\; \sigma_j\, d_V(f), \quad j = 0, \ldots, n+1; \quad |\sigma_j| \;=\; 1,$$

$$\Rightarrow \; \left|\frac{1}{2}\,(f(t_j) - p_1(t_j)) \;+\; \frac{1}{2}\,(f(t_j) - p_2(t_j))\right| \;=\; d_V(f).$$

Da aber zugleich $|f(t_j) - p_k(t_j)| \le d_V(f)$ gilt, folgt hiermit

$$f(t_j) - p_1(t_j) \;=\; f(t_j) - p_2(t_j) \;=\; \sigma_j\, d_V(f), \qquad j = 0, \ldots, n+1.$$

Damit ist aber $(p_2 - p_1)(t_j) = 0$, $j = 0, \ldots, n+1$, und damit nach (H1): $p_2 = p_1$. □

Bemerkungen (8.12)

a) Der Beweis des Haarschen Eindeutigkeitssatzes gilt analog im komplexen Fall $(B \subset \mathbb{C})$. Zum Beweisteil b) beachte man, dass $(\mathbb{C}, |\cdot|)$ strikt normiert ist.

b) Es gilt in gewissem Sinn die Umkehrung des Haarschen Eindeutigkeitssatzes: Ist V ein endlich dimensionaler Teilraum und erfüllt V nicht die Haarsche Bedingung, so existiert ein $f \in \mathrm{C}(B)$ zu dem es mehrere Bestapproximationen gilt; vgl. Schönhage, Satz 6.4.

Die Aussagen (8.12)b) und (8.11) lassen sich wie folgt zusammenfassen

Folgerung (8.13)

Für einen linearen Teilraum $V \subset \mathrm{C}(B)$ mit $\dim V = n+1$ sind äquivalent

a) Jedes $p \in V \setminus \{0\}$ hat höchstens n Nullstellen,

b) Zu jedem $f \in \mathrm{C}(B)$ gibt es genau eine Bestapproximation von f aus V.

Neben der Frage der Eindeutigkeit einer Bestapproximation ist auch die Frage nach der *strikten* Eindeutigkeit von Interesse, gemeint ist damit ein mindestens lineares Anwachsen des Fehlers mit dem Abstand von der Bestapproximation.

Definition (8.14)

$p \in V$ heißt eine *strikt eindeutige* Bestapproximation von f bzgl. V, falls es ein $\gamma > 0$ gibt mit

$$\forall\, q \in V: \ \|f - q\|_\infty \ \geq \ \|f - p\|_\infty + \gamma\, \|p - q\|_\infty.$$

Im reellen Fall lässt sich tatsächlich aus der Haarschen Bedingung die Existenz einer strikt eindeutigen Bestapproximation folgern, vgl. Nürnberger; Theorem 3.18.

Satz (8.15)

Für einen endlich dimensionalen linearen Teilraum $V \subset \mathrm{C}[a,b]$ sind äquivalent

a) V ist Haarscher Teilraum,

b) Zu jedem $f \in \mathrm{C}[a,b]$ gibt es eine strikt eindeutige Bestapproximation von f aus V.

Alternanten.

Im Folgenden sei $B \subset [a,b]$ kompakt, $R = \mathrm{C}(B)$, V sei ein $(n+1)$–dimensionaler Teilraum von $\mathrm{C}[a,b]$. Ferner enthalte B wenigstens $(n+2)$ Punkte.

Satz (8.16) (de la Vallee-Pouissin)

Erfüllt V die Haarsche Bedingung bzgl. $\mathrm{C}[a,b]$ und gibt es zu $f \in \mathrm{C}(B)$ und $p \in V$ Punkte $t_0 < t_1 < \ldots < t_{n+1} \in B$, so dass mit einem $\sigma \in \{-1, 1\}$ gilt

$$\mathrm{sign}((f-p)(t_j)) \ = \ \sigma\,(-1)^j, \quad j = 0, \ldots, n+1,$$

so folgt $$\min_j |(f-p)(t_j)| \ \leq \ d_V(f) \ = \ \inf_{q \in V} \max_{t \in B} |f(t) - q(t)|.$$

Gleichheit kann hierbei nur für $|(f-p)(t_j)| = |(f-p)(t_k)|, \ \forall\, j,k$ auftreten.

Beweis:

Sei $(h_0, \ldots, h_n)$ eine Basis von V. Dann hat die folgende Matrix aufgrund der Haarschen Bedingung maximalen Rang ($= n+1$)

$$D(t_0, \ldots, t_{n+1}) \ = \ \begin{pmatrix} h_0(t_0) & \ldots & h_n(t_0) \\ \vdots & & \vdots \\ h_0(t_{n+1}) & \ldots & h_n(t_{n+1}) \end{pmatrix} \ \in \ \mathbb{R}^{(n+2,n+1)}.$$

Es gibt somit einen nichtverschwindenden Vektor $(\lambda_0, \ldots, \lambda_{n+1}) \in \mathbb{R}^{n+2} \setminus \{0\}$ mit den Eigenschaften

$$\begin{array}{lll} \text{(a)} & \displaystyle\sum_{j=0}^{n+1} \lambda_j \, h_k(t_j) \;=\; 0, \qquad k \;=\; 0,1,\ldots,n \\ \text{(b)} & \displaystyle\sum_{j=0}^{n+1} |\,\lambda_j\,| \;=\; 1. \end{array} \tag{8.17}$$

Da ferner *jede* quadratische Teilmatrix von $D(t_0, \ldots, t_{n+1})$ aus $n+1$ Zeilen nach der Haarschen Bedingung regulär ist, verschwindet auch keins der λ_j.

Für das lineare Funktional $\ell \in R^*$ mit $\ell(q) := \sum_0^{n+1} \lambda_j \, q(t_j)$ gilt somit $\ell \in V^\perp$ und $\|\,\ell\,\| = 1$.

Aufgrund des Dualitätsprinzips, Satz (2.27), folgt $|\,\ell(f)\,| \leq d_V(f)$.

Seien nun $q_k \in V$, $k = 0, 1, \ldots, n$, bestimmt durch die Interpolationsbedingungen

$$q_k(t_j) \;=\; 0, \;\; j \in \{0, \ldots, n+1\} \setminus \{k, k+1\}, \qquad q_k(t_k) \;=\; 1.$$

Nach (H2) sind die q_k hiermit eindeutig bestimmt, $q_k \neq 0$ und, da q_k nach (H1) höchstens n Nullstellen besitzen kann, ist auch $q_k(t_{k+1}) \;>\; 0$.

Damit folgt

$$0 \;=\; \ell(q_k) \;=\; \lambda_k \cdot 1 \;+\; \lambda_{k+1}\; q_k(t_{k+1})$$

$$\Rightarrow \;\; \operatorname{sign} \lambda_{k+1} \;=\; -\operatorname{sign} \lambda_k, \qquad \lambda_j \;\neq\; 0.$$

und somit

$$\begin{aligned} d_V(f) \;&\geq\; |\,\ell(f)\,| \;=\; |\,\ell(f-p)\,| \;=\; \Big|\,\sum_{j=0}^{n+1} \lambda_j\,(f-p)(t_j)\,\Big| \\ &=\; \sum_{j=0}^{n+1} |\,\lambda_j\,|\,|\,(f-p)(t_j)\,| \;\geq\; \Big(\sum_{j=0}^{n+1} |\,\lambda_j\,|\Big) \; \min_j |\,(f-p)(t_j)\,| \\ &=\; \min_j |\,(f-p)(t_j)\,|\,. \end{aligned}$$

Gleichheit kann höchstens dann vorliegen, falls alle $|\,(f-p)(t_j)\,|$ gleich sind. □

Folgerung (8.18)

Gilt für $(n+2)$ Extremalpunkte $t_0 < t_1 < \ldots < t_{n+1}$ von p bezüglich f mit einem festen $\sigma \in \{-1, 1\}$

$$(f-p)(t_j) \;=\; \sigma\,(-1)^j\,\|\,f-p\,\|_\infty, \quad j = 0, \ldots, n+1,$$

so ist p Bestapproximation von f aus V.

Hierzu gilt nun auch die Umkehrung:

Satz (8.19) (Alternantensatz)

Sei wieder $f \in R = \mathrm{C}(B)$, $B \subset [a, b]$ eine kompakte Menge, die wenigstens $(n+2)$ Punkte enthält. Ferner sei V ein $(n+1)$–dimensionaler Haarscher Teilraum von $\mathrm{C}[a, b]$.
Ein Element $p \in V$ ist genau dann Bestapproximation von f aus V, wenn es $(n+2)$ Punkte $t_0 < t_1 < \ldots < t_{n+1}$ aus B gibt mit

$$(f - p)(t_j) \;=\; \sigma\,(-1)^j\,\|\,f - p\,\|_\infty, \quad j = 0, \ldots, n+1, \qquad \sigma \in \{-1, 1\}. \tag{8.20}$$

Das Tupel $(t_0, \ldots, t_{n+1})$ heißt dann eine *Alternante* der Fehlerfunktion $e = f - p$.

Beweis:

Nach (8.18) genügt es zu zeigen, dass es zu jeder Bestapproximation $p \in V$ eine Alternante der Fehlerfunktion gibt.

Sei also $p \in V$ Bestapproximation von $f \in \mathrm{C}(B)$ und nehmen wir an, dass es keine Alternante zu p gäbe. Dann existiert eine Unterteilung

$$a \;=\; \tau_0 \;<\; \tau_1 \;<\; \ldots \;<\; \tau_m \;<\; \tau_{m+1} \;=\; b,$$

wobei $m \leq n$ ist, sowie ein $\delta > 0$ und ein $\sigma \in \{-1, 1\}$, so dass für die Fehlerfunktion $e := f - p$ und $t \in B \cap [\tau_j, \tau_{j+1}]$, $j = 0, 1, \ldots, m$ gilt

$$\sigma\,(-1)^j \;=\; +1 \qquad \Rightarrow \quad e(t) \;\in\;]\,-\,\|\,e\,\| + \delta\,,\,\|\,e\,\|\,]$$

$$\sigma\,(-1)^j \;=\; -1 \qquad \Rightarrow \quad e(t) \;\in\; [\,-\,\|\,e\,\|\,,\,\|\,e\,\| - \delta\,[\,.$$

Wir wenden nun (H6) an, vgl. (8.6). Demnach gibt es eine Funktion $q \in V$, die genau die Nullstellen $\tau_1, \ldots, \tau_m$ in $[a, b]$ besitzt und dazwischen die folgende Vorzeichenverteilung besitzt

$$\operatorname{sign} q(t) \;=\; \sigma\,(-1)^j, \qquad t \in]\tau_j, \tau_{j+1}[\,.$$

Für hinreichend kleines $\varepsilon > 0$ erfüllt $\tilde{e} := e - \varepsilon\, q$ damit die Bedingung $\|\,\tilde{e}\,\|_\infty < \|\,e\,\|_\infty$. Das bedeutet aber, dass $p + \varepsilon q \in V$ eine bessere Approximation von f liefert als p. Widerspruch! $\square$

Beispiel (8.21)

Die Funktion $f := \sin$ sei auf einem Intervall $[a, b]$ durch eine Konstante $p \in \Pi_0$ zu approximieren. Enthält das Intervall nun wenigstens zwei Punkte der Form $t_k = (2k+1)\,\pi/2$, $k \in \mathbb{Z}$, so ist $p^* = 0$ Bestapproximation. Mit zwei benachbarten Punken t_k und t_{k+1} ist ja bereits eine Alternante der Länge zwei gegeben.

Enthält das Intervall $[a, b]$ sogar drei Punkte der obigen Form, so ist $p^* = 0$ sogar Bestapproximation bezüglich Π_1.

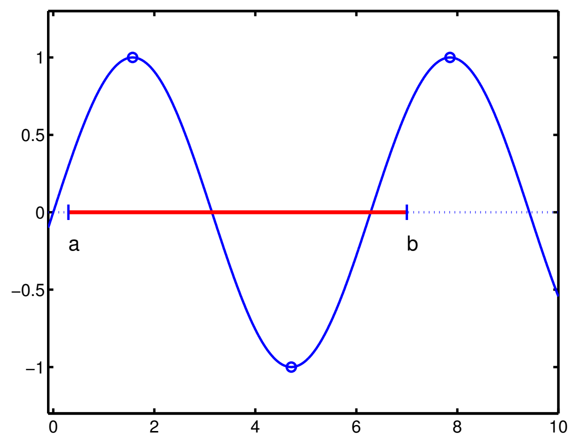

Abb. 8.2 Alternante zu Beispiel (8.21)

Beispiel (8.22)

Die Funktion $f := \sin$ soll im Intervall $[0, \pi/2]$ durch eine Parabel der Form $p(t) = a_0\, t + a_1\, t^2$ approximiert werden.

Es ist zu beachten, dass der lineare Raum V der Polynome dieser Form *keinen* Haarschen Raum über $[0, \pi/2]$ bildet. Man kann sich damit behelfen, dass man 0 nicht zur Alternante hinzunimmt und das Problem auf einem Intervall $[a, \pi/2]$ mit kleinem $a > 0$ betrachtet (dort ist V ein Haarscher Raum).

Die numerische Berechnung ergibt die Alternante

$$t_0 \doteq 0.28373316, \quad t_1 \doteq 1.10612446, \quad t_2 \doteq 1.57079633,$$

die Bestapproximation: $p^*(t) \doteq 1.13662336\, t - 0.31121899\, t^2$

und die Minimalabweichung: $\|f - p^*\|_\infty \doteq 0.017501718$.

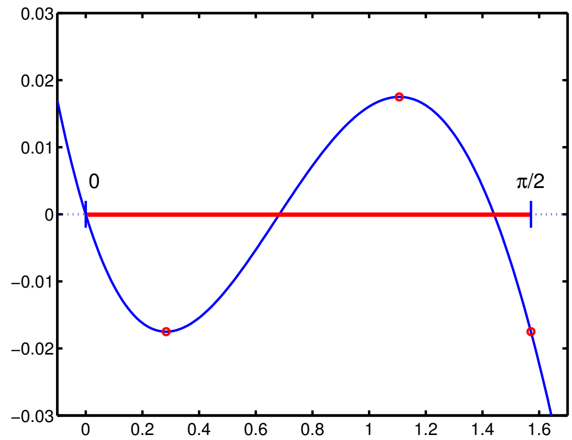

Abb. 8.3 Fehlerfunktion zu Beispiel (8.22)

Beispiel (8.23)

Die Funktion $f(t) := 1/(1+t)$, $0 \le t \le 1$, ist durch ein Polynom $p \in \Pi_2[0,1]$ zu approximieren. Eine numerische Berechnung ergibt die Alternante

$$t_0 \;=\; 0, \quad t_1 \doteq 0.20710678, \quad t_2 \doteq 0.70710678, \quad t_3 \;=\; 1,$$

die Bestapproximation $p^*(t) \;=\; a_0 \,+\, a_1\, t \,+\, a_2\, t^2$ mit

$$a_0 \doteq 0.99264069, \;\; a_1 \doteq -0.82842712, \;\; a_2 \doteq 0.34314575,$$

die Minimalabweichung lautet $\|f - p^*\|_\infty \doteq 0.735931288 \times 10^{-2}$.

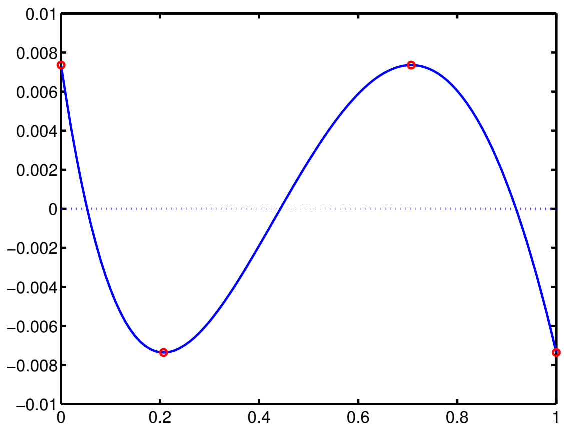

Abb. 8.4 Fehlerfunktion zu Beispiel (8.23)

Satz (8.24)

Sei $R = \mathrm{C}[a,b]$ und V ein $(n+1)$-dimensionaler Teilraum von R, der die konstanten Funktionen enthält. Schließlich sei $f \in R \setminus V$, so dass sowohl V, wie auch $\mathrm{Spann}(V \cup \{f\})$ Haarsche Teilräume sind.
Für die Bestapproximation p von f aus V gilt dann: Die Fehlerfunktion $e := f - p$ besitzt genau $(n+2)$ Extremalpunkte $a = t_0 < \ldots < t_{n+1} = b$. Zwischen benachbarten Extremalpunkten ist e streng monoton. Insbesondere ist die Alternante gemäß (8.19) eindeutig bestimmt und enthält die Randpunkte des Intervalls.

Beweis:

Für ein beliebiges $c \in \mathbb{R}$ ist $f - p - c \in \mathrm{Spann}(V \cup \{f\})$. Daher hat $f - p - c$ höchstens $n+1$ Nullstellen in $[a,b]$. Die Behauptung ergibt sich hieraus mit dem Alternantensatz und dem Zwischenwertsatz. $\square$

Für das Beispiel (8.23) sind die Voraussetzungen des Satzes (8.24) erfüllt. Die Alternante ist also eindeutig bestimmt und enthält die Randpunkte $t_0 = 0$ und $t_4 = 1$.

Beispiel (8.25)

Sei $B := \{t_0, \dots, t_{n+1}\} \subset [a,b]$, V_n ein $(n+1)$–dimensionaler Haarscher Teilraum von $\mathrm{C}[a,b]$. $h_{n+1} \in \mathrm{C}[a,b] \setminus V$ bezeichne ein neues Basiselement, so dass $V_{n+1} := \mathrm{Spann}(V_n \cup \{h_{n+1}\})$ ebenfalls ein Haarscher Teilraum von $\mathrm{C}[a,b]$ ist.

Zu $f \in \mathrm{C}(B)$ konstruieren wir $g_1,\, g_2 \in V_{n+1}$ durch die Interpolationsbedingungen

$$g_1(t_j) \;=\; f(t_j), \quad g_2(t_j) \;=\; (-1)^j, \quad j = 0,\, 1,\, \dots\,,\, n+1.$$

Ferner bestimmen wir $\mu \in \mathbb{R}$ durch die Forderung $p := g_1 - \mu\, g_2 \;\in\; V_n$. Damit stellen wir fest

$$(f-p)(t_j) \;=\; \mu\,(-1)^j, \quad j = 0, 1, \dots, n+1.$$

Da es in B keine weiteren Punkte gibt, ist $(t_0, \dots, t_{n+1})$ eine Alternante zur Fehlerfunktion $e := f - p$. Damit ist p Bestapproximation von $f \in \mathrm{C}(B)$ bezüglich V_n.

Beispiel (8.26)

Die Bestapproximation der Funktion $f(t) := t^{n+1} \in \mathrm{C}[-1,1]$ bezüglich $\Pi_n[-1,1]$ ist gegeben durch $p(t) := t^{n+1} \;-\; 2^{-n}\, T_{n+1}(t)$ mit $d_{\Pi_n}(f) = 2^{-n}$.

Die Begründung ergibt sich unmittelbar aus den Eigenschaften der Tschebyscheff–Polynome, vgl. Satz (3.13). Zunächst ist p tatsächlich ein Polynom n-ten Grades, ferner hat die Fehlerfunktion $e \;=\; f - p \;=\; 2^{-n} T_{n+1}$ die Alternante

$$t_k^E \;=\; \cos\Big(\frac{n+1-k}{n+1}\,\pi\Big), \;\; k = 0,\, \dots\,,\, n+1. \qquad \square$$

Wir formulieren noch die Folgerung aus dem Alternantensatz, die sich für die Tschebyscheff-Approximation von $2\,\pi$–periodischen Funktionen durch trigonometrische Polynome ergibt.

Satz (8.27)

Sei $f \in \mathrm{C}_{2\,\pi}$ und $n \in \mathbb{N}$. Dann gibt es genau eine Bestapproximation von f aus T_n bezüglich $\|\cdot\|_\infty$. Diese ist charakterisiert durch eine Alternante der Länge $2\,n + 2$ im halboffenen Intervall $[0, 2\,\pi[$.

Beweis: Anwendung des Alternantensatzes für das Intervall $[0, b]$ und Betrachtung des Grenzübergangs $b \uparrow 2\,\pi$. $\square$

Satz (8.28) (Fehlerdarstellung)

Sei $f \in \mathrm{C}^{n+1}[-1,1]$ und $E_n(f) := d_{\Pi_n[-1,1]}(f)$ (bzgl. $\|\cdot\|_\infty$).

a) Gilt für eine Vergleichsfunktion $f_0 \in \mathrm{C}^{n+1}[-1,1]$ die Abschätzung $|\, f^{(n+1)}(t)\,| \;\le\; f_0^{(n+1)}(t), \;\; \forall\, t \in [-1,1]$, so folgt $E_n(f) \;\le\; E_n(f_0)$.

b) Es gibt ein $\tau \in [-1,1]$ mit

$$E_n(f) \;\leq\; \frac{|\, f^{(n+1)}(\tau)\,|}{2^n\,(n+1)!}. \tag{8.29}$$

Beweis:

zu a): Ist p Bestapproximation von f aus $\Pi_n[-1,1]$, so existiert eine Alternante $(t_0,\ldots,t_{n+1})$ mit

$$(f-p)(t_j) \;=\; (-1)^j\,\mu, \quad |\mu| \;=\; E_n(f).$$

Es sei nun $p_0 \in \Pi_n$ die Bestapproximation von f_0 auf $B := \{t_0,\ldots,t_{n+1}\}$. Dann gilt analog

$$(f_0-p_0)(t_j) \;=\; (-1)^j\,\mu_0, \quad |\mu_0| \;\leq\; E_n(f_0).$$

Für $F \;:=\; \mu_0\,(f-p) \;-\; \mu\,(f_0-p_0)$ gilt dann $F(t_j)=0,\ j=0,\ldots,n+1$ und somit nach dem Satz von Rolle

$$\exists\,\tau \;\in\; [-1,1]: \;\; F^{(n+1)}(\tau) \;=\; \mu_0\, f^{(n+1)}(\tau) \;-\; \mu\, f_0^{(n+1)}(\tau) \;\;=\;\; 0.$$

Gilt nun $|\, f^{(n+1)}(t)\,| \;<\; f_0^{(n+1)}(t)$, für alle $t \in [-1,1]$, so folgt

$$E_n(f) \;=\; |\mu| \;=\; |\mu_0|\,\frac{|\, f^{(n+1)}(\tau)\,|}{|\, f_0^{(n+1)}(\tau)\,|} \;\leq\; |\mu_0| \;\leq\; E_n(f_0).$$

Gilt dagegen nur $|\, f^{(n+1)}(t)\,| \;\leq\; f_0^{(n+1)}(t)$, für alle $t \in [-1,1]$, so wende man die obige Überlegung auf $\widetilde{f_0}(t) \;:=\; f_0(t) \;+\; \varepsilon\, t^{n+1}$, $\varepsilon > 0$, an.

Es ist dann $\widetilde{f}_0^{(n+1)}(t) \;=\; f_0^{(n+1)}(x) \;+\; (n+1)!\,\varepsilon \;>\; f_0^{(n+1)}(t)$ und somit

$$E_n(f) \;\leq\; E_n(\widetilde{f_0}) \;\leq\; E_n(f_0) \;+\; 2^{-n}\,\varepsilon.$$

Für $\varepsilon \downarrow 0$ folgt die Behauptung.

zu b): Man setze $f_0(t) \;:=\; c\,t^{n+1}/(n+1)!$.

Für die Minimalabweichung gilt nach (8.26): $E_n(f_0) \;=\; c\;2^{-n}/(n+1)!$.

Sei nun $\tau \in [-1,1]$ mit

$$c \;:=\; |\, f^{(n+1)}(\tau)\,| \;=\; \max\{\,|\, f^{(n+1)}(t)\,|: \; -1 \leq t \leq 1\,\}$$

Dann folgt $|\, f^{(n+1)}(t)\,| \;\leq\; c \;=\; f_0^{(n+1)}(t)$ und damit nach a)

$$E_n(f) \;\leq\; \frac{|\, f^{(n+1)}(\tau)\,|}{2^n\,(n+1)!} \qquad \square$$