10. L_1–Approximation

Problemstellung.

Wir betrachten den reellen Vektorraum $R = \mathrm{C}[a,b]$, $a < b$, sowie die zugehörige L_1–Norm

$$\| f \|_1 \;:=\; \int_a^b |f(t)|\,dt \tag{10.1}$$

Weiter sei $V \subset R$ ein $(n+1)$-dimensionaler Teilraum und $f \in R \setminus V$. Wir suchen eine L_1–Bestapproximation $p^* \in V$, also

$$\forall\, p \in V: \quad \|f - p^*\|_1 \;\le\; \|f - p\|_1.$$

Wie bisher wird zu einer Approximation $p \in V$ mit $e := f - p$ die *Fehlerfunktion* bezeichnet.
Zur Berechnung von $\|e\|_1$ definieren wir die *Vorzeichenfunktion* $s: [a,b] \to \mathbb{R}$ gemäß

$$s(t) \;:=\; \text{sign } e(t) \;:=\; \begin{cases} 1, & \text{falls } e(t) > 0, \\ 0, & \text{falls } e(t) = 0, \\ -1, & \text{falls } e(t) < 0, \end{cases} \tag{10.2}$$

Damit folgt $\|f - p\|_1 \;=\; \|e\|_1 \;=\; \int_a^b s(t)\,e(t)\,dt.$

Beispiel (10.3)

Sei $f \in \mathrm{C}^1[a,b]$ streng monoton wachsend, $f' > 0$ und $V = \Pi_0[a,b]$.

Wir haben nun ein $\tau \in \,]a,b[$ zu bestimmen, so dass für $p_\tau(t) := f(\tau)$ gilt $\|f - p_\tau\|_1$ minimal!

Für die Funktion

$$\Phi(\tau) \;:=\; \|f - p_\tau\|_1 \;=\; \int_a^\tau (f(\tau) - f(t))\,dt \;+\; \int_\tau^b (f(t) - f(\tau))\,dt$$

findet man

$$\begin{aligned} \Phi'(\tau) \;&=\; (f(\tau) - f(\tau)) \;+\; \int_a^\tau f'(\tau)\,dt \;-\; (f(\tau) - f(\tau)) \;-\; \int_\tau^b f'(\tau)\,dt \\ &=\; f'(\tau)\,(2\,\tau - a - b) \end{aligned}$$

Φ' hat also die einzige Nullstelle $\tau^* = (a+b)/2$. Wegen $f' > 0$ ist ferner $\Phi'(t) < 0$ für $t \in [a, \tau^*[$ und $\Phi'(t) > 0$ für $t \in]\tau^*, b]$.

Damit ist τ^* also ein striktes globales Minimum von Φ.

Man beachte die bemerkenswerte Eigenschaft, dass τ^* unabhängig von f ist.

Charakterisierung der Bestapproximation.

Wir nehmen an, dass die Nullstellenmenge

$$Z \;:=\; \{t \in [a,b] :\; e(t) = 0\} \tag{10.4}$$

aus endlich vielen kompakten Teilintervall von $[a,b]$ besteht. Diese können auch einpunktig sein.

Satz (10.5) (Kripke, Rivlin, 1965)

p ist genau dann L_1–Bestapproximation von f aus V, wenn gilt

$$\forall\, q \in V : \quad \Big| \int_a^b s(t)\, q(t)\, dt \Big| \;\leq\; \int_Z |q(t)|\, dt\,.$$

Bemerkungen (10.6)

a) Für das Beispiel (10.3) ist $Z = \{\tau\}$ und $s(t) \;=\; \operatorname{sign}(t-\tau)$. Nach (10.5) ist τ so zu bestimmen, dass

$$\forall\, q \in \Pi_0 : \quad \int_a^b s(t)\, q(t)\, dt \;=\; 0.$$

Hieraus folgt sofort $\tau = (a+b)/2$.

b) Besteht Z nur aus endlich vielen Punkten, so besagt der Charakterisierungssatz

$$p \quad L_1 - \text{Bestapproximation} \quad \Leftrightarrow \quad \forall\, q \in V :\; \langle \operatorname{sign}(f-p), q\rangle \;=\; 0$$

Man vergleiche dies auch mit den Charakterisierungen

$$p \quad L_2 - \text{Bestapproximation} \quad \Leftrightarrow \quad \forall\, q \in V :\; \langle f-p, q\rangle \;=\; 0$$

$$p \quad L_\infty - \text{Bestapproximation} \quad \Leftrightarrow \quad \forall\, q \in V :\; \min\nolimits_{A(f,p)} (f-p)\, q \;\leq\; 0$$

Beweis zu (10.5)

⇒: Sei $p \in V$ Bestapproximation, so dass die Charakterisierung

$$\forall\, q \in V : \quad \Big| \int_a^b s(t)\, q(t)\, dt \Big| \;\leq\; \int_Z |q(t)|\, dt\,.$$

nicht erfüllt ist. Dann existiert ein $q \in V$ mit

$$\eta \;:=\; \Big| \int_a^b s\,q \,\Big| \;-\; \int_Z |\,q\,| \;>\; 0. \tag{$*$}$$

O.B.d.A. können wir annehmen, dass $\int sq \;>\; 0$ und $\|q\|_\infty \;=\; 1$ gelten.

Wir erweitern Z durch die Nachbarmenge $(\Theta > 0)$

$$Z_\Theta \;:=\; \{t \in [a,b] : \;\; 0 < |f(t) - p(t)| \le \Theta\}.$$

Wegen der Voraussetzung an Z und der Stetigkeit von $|e|$ ist Z_Θ meßbar und für hinreichend kleines $\Theta > 0$ gilt

$$\int_{Z_\Theta} dt \;<\; \eta/2.$$

Schließlich setzen wir $Z_R := [a,b] \setminus (Z \cup Z_\Theta)$.

Zur Abschätzung von $\|f - (p + \Theta q)\|_1$ zerlegen wir das Integral über $[a,b]$ in die drei Teilintegrale über Z, Z_Θ und Z_R.

(i) Für $t \in Z$ hat man $|\,f - p - \Theta\,q\,| \;=\; \Theta\,|\,q\,|$. Damit gilt

$$\int_Z |f - p - \Theta\,q| \;\;=\;\; \Theta \int_Z |q|.$$

(ii) Für $t \in Z_\Theta$ gilt

$$|f - p - \Theta\,q| \;\le\; |f - p\,| \;+\; \Theta\,|q| \;\le\; |f - p\,| \;+\; \Theta\,(2 - sq)$$

Die letzte Abschätzung gilt, da $\|q\|_\infty = 1$, also $2 - sq \ge 1$. Somit

$$\int_{Z_\Theta} |f - p - \Theta\,q| \;\le\; \int_{Z_\Theta} |f - p| \;+\; \Theta \int_{Z_\Theta} (2 - sq).$$

(iii) Für $t \in Z_R$ hat man $|f - p| > \Theta$ und damit $\operatorname{sign}(f - p - \Theta q) = \operatorname{sign}(f - p) = s$. Hiermit ergibt sich schließlich

$$|f - p - \Theta\,q| \;\;=\;\; |f - p| \;-\; \Theta\,s\,q,$$

also

$$\int_{Z_R} |f - p - \Theta\,q| \;\le\; \int_{Z_R} |f - p| \;-\; \Theta \int_{Z_R} s\,q.$$

Insgesamt folgt nunmehr

$$\begin{aligned}
\|f - (p + \Theta q)\|_1 \;&\le\; \|f - p\|_1 \;+\; \Theta \int_Z |\,q\,| \;+\; \Theta \int_{Z_\Theta} (2 - sq) \;-\; \Theta \int_{Z_R} s\,q \\
&\le\; \|f - p\|_1 \;+\; \Theta \int_Z |\,q\,| \;+\; 2\,\Theta \int_{Z_\Theta} dt \;-\; \Theta \int_a^b s\,q
\end{aligned}$$

$$\leq \quad \|f-p\|_1 \;+\; 2\,\Theta \int\limits_{Z_\Theta} dt \;-\; \Theta \left(\int\limits_a^b s\,q \;-\; \int\limits_Z |\,q\,| \right)$$

$$<_{(*)} \quad \|f-p\|_1 \;+\; \Theta\,\eta \;-\; \Theta\,\eta \;=\; \|f-p\|_1.$$

Damit ist aber p keine L_1–Bestapproximation von f.

$\Leftarrow$: Wir haben zu zeigen, dass aus der Charakterisierung

$$\forall\, q \in V: \quad |\int\limits_a^b s(t)\,q(t)\,dt\,| \;\leq\; \int\limits_Z |q(t)|\,dt\,.$$

folgt, dass p Bestapproximation ist. Dazu schätzen wir ab:

$$\begin{aligned}
\|\,f-p+q\,\|_1 \;&=\; \int\limits_a^b |\,f(t)-p(t)+q(t)\,|\,dt \\
&=\; \int\limits_a^b |\,s\,(f-p+q)\,| \;+\; \int\limits_Z |\,f-p+q\,| \\
&\geq\; \int\limits_a^b s\,(f-p+q) \;+\; \int\limits_Z |\,f-p+q\,| \\
&=\; \int\limits_a^b s\,(f-p) \;+\; \int\limits_a^b s\,q \;+\; \int\limits_Z |\,q\,| \\
&\geq\; \int\limits_a^b s\,(f-p) \;=\; \|\,f-p\,\|_1 \qquad \square
\end{aligned}$$

L_1–Approximation für Haarsche Räume.

Satz (10.7)

Ist V ein Haarscher Teilraum von $C[a,b]$, $p \in V$ L_1–Bestapproximation und hat $e := f - p$ nur endlich viele Nullstellen, so wechselt e wenigstens $(n+1)$ mal das Vorzeichen.

Beweis:

Hat e genau m Nullstellen mit Vorzeichenwechsel und ist $m \leq n$, so existiert nach (H6) eine Funktion $q \in V$, die genau diese m Nullstellen besitzt und dort das Vorzeichen wechselt. O.B.d.A. gilt somit

$$\forall\, t \in [a,b] \setminus Z: \quad s(t)\,q(t) \;>\; 0.$$

Damit folgt aber

$$\int_a^b s(t)\,q(t)\,dt \;>\; 0, \qquad \int_Z |\,q(t)\,|\,dt \;=\; 0,$$

im Widerspruch zur Charakterisierung (10.5). □

Satz (10.8) (Eindeutigkeit)

Ist V ein Haarscher Teilraum, so gibt es zu jedem $f \in \mathrm{C}[a,b]$ höchstens eine L_1–Bestapproximation (und damit auch genau eine nach Satz (2.2)).

Beweis: Sind p_1, p_2 Bestapproximationen von f, so folgt für $p := (p_1 + p_2)/2$

$$|\,f(t) - p(t)\,| \;\le\; \frac{1}{2}\,|\,f(t) - p_1(t)\,| \;+\; \frac{1}{2}\,|\,f(t) - p_2(t)\,|\,.$$

Somit ist auch p eine L_1–Bestapproximation und die Integrale über die obige Ungleichung müssen jeweils gleich sein, damit muss aus Stetigkeitsgründen auch punktweise Gleichheit gelten

$$\forall\, t \in [a,b]: \quad |\,f(t) - p(t)\,| \;=\; \frac{1}{2}\,|\,f(t) - p_1(t)\,| \;+\; \frac{1}{2}\,|\,f(t) - p_2(t)\,|\,.$$

Nach (10.7) hat $(f - p)$ aber $(n+1)$ Nullstellen in $[a,b]$. Diese müssen nach Obigem zuleich Nullstellen von $(f - p_1)$ und von $(f - p_2)$ sein. Damit hat aber auch $(p_2 - p_1) \in V$ diese $(n+1)$ Nullstellen. Da V ein Haarscher Raum ist, folgt $p_1 = p_2$. □

Satz (10.9) (L_1–Knoten)

Ist V ein Haarscher Teilraum, $p \in V$ L_1–Bestapproximation von $f \in R$ und hat die Fehlerfunktion $e := f - p$ genau $(n+1)$ Nullstellen, so hängen diese nicht von f ab.

Beweis: Die Fehlerfunktion $e := f - p$ habe genau die Nullstellen $\tau_0 < \ldots < \tau_n$ in $[a,b]$.

Zu $g \in R$ sei weiter $q \in V$ L_1-Bestapproximation, so dass die Fehlerfunktion $\widetilde{e} := g - q$ genau die Nullstellen $\sigma_0 < \ldots < \sigma_n$ in $[a,b]$ besitzt.

Nach (10.7) haben alle Nullstellen Vorzeichenwechsel, liegen also insbesondere im offenen Intervall $]a,b[$. Die Anfangwerte $e(a)$ und $\widetilde{e}(a)$ verschwinden nicht und o.E.d.A. nehmen wir an, dass $e(a)\,\widetilde{e}(a) > 0$ und $\tau_0 < \sigma_0$ gelten.

Da V ein Haarscher Raum ist, gibt es ein $h \in V$, das genau die Nullstellen $\tau_1, \ldots, \tau_n$ besitzt und dort jeweils das Vorzeichen wechselt. Wieder gelte o.E.d.A. $e(a)h(a) < 0$.

Nach dem Charakterisierungssatz (10.5) gilt nun mit $s(t) := \operatorname{sign} e(t)$, $\widetilde{s}(t) := \operatorname{sign} \widetilde{e}(t)$:

$$\int_a^b s(t)\,h(t)\,dt \;=\; \int_a^b \widetilde{s}(t)\,h(t)\,dt \;=\; 0$$

und somit auch

$$\int_a^b (\widetilde{s}(t) \;-\; s(t))\; h(t)\, dt \;=\; 0. \qquad (*)$$

Nach Konstruktion ist $s(t) = \widetilde{s}(t)$ für $t \in [a, \tau_0[$ und $s(t)\, h(t) > 0, \;\; \forall\, t \in\,]\tau_0, b] \setminus \{\tau_1, \ldots, \tau_n\}$. Hieraus lässt sich nun unmittelbar ableiten

$$\forall\, t \in [a,b] : \;\; [s(t) - \widetilde{s}(t)]\, h(t) \;\geq\; 0.$$

Zusammen mit $(*)$ folgt somit

$$\forall\, t \in [a,b] : \;\; [s(t) - \widetilde{s}(t)]\, h(t) \;=\; 0.$$

Dies kann aber nur gelten, wenn alle Nullstellen übereinstimmen, also $\forall\, j : \; \sigma_j = \tau_j$ gilt. □

Die Nullstellen der Fehlerfunktion $e = f - p$ sind also (sofern es nur $(n+1)$ Nullstellen gibt) unabhängig von f. Sie heißen die *zu V gehörigen L_1–Knoten* $\tau_0 < \ldots < \tau_n$. Sind diese Knoten bekannt, so lässt sich zu vorgegebenem $f \in R$ die L_1–Bestapproximation p als Lösung der folgenden Interpolationsaufgabe ermitteln

$$p \;\in\; V, \;\; \text{mit } p(\tau_j) \;=\; f(\tau_j), \;\; j = 0, \ldots, n.$$

Satz (10.10) (Polynomräume)

Die L_1–Knoten für $V = \Pi_n[-1,1]$ sind die inneren Extremalstellen des Tschebyscheff-Polynoms T_{n+2}:

$$\tau_k \;=\; \cos\left[\frac{n+1-k}{n+2}\,\pi\right], \qquad k = 0, 1, \ldots, n.$$

Beweis:

Die τ_k seien wie oben gegeben, ferner sei $\tau_{-1} := -1$ und $\tau_{n+1} := 1$. Die Vorzeichenfunktion lautet also

$$s(t) \;=\; \begin{cases} (-1)^j & : \;\; \tau_{j-1} < t < \tau_j \\[1ex] 0 & : \;\; t \;=\; \tau_j \end{cases}$$

und es ist zu zeigen, dass für alle $k = 0, 1, \ldots, n$ gilt

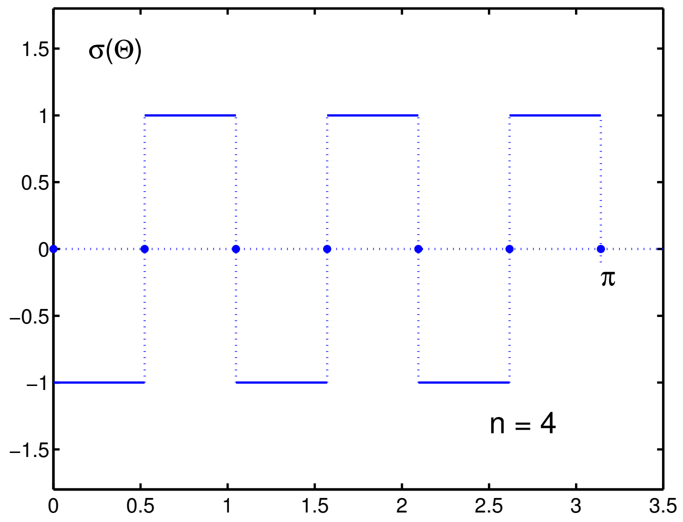
$$\int_{-1}^{1} s(t)\, T_k(t)\, dt \;=\; 0. \qquad (10.11)$$

Die Substitution $t = \cos\Theta$, $dt = -\sin\Theta\, d\Theta$ ergibt die zu (10.11) äquivalente Bedingung

$$\int_0^{\pi} \sigma(\Theta)\; \cos(k\Theta)\; \sin(\Theta)\, d\,\Theta \;=\; 0, \qquad k = 0, \ldots, n, \qquad (10.12)$$

wobei

$$\sigma(\Theta) \;:=\; s(\cos\Theta) \;=\; \begin{cases} (-1)^{n-j} & : \; \dfrac{j-1}{n+2}\,\pi < \Theta < \dfrac{j}{n+2}\,\pi \\[2ex] 0 & : \; \Theta \;=\; \dfrac{j}{n+2}\,\pi . \end{cases} \tag{10.13}$$



Wir setzen σ zu einer ungeraden und $2\,\pi$–periodischen Funktion auf $\mathbb{R}$ fort. Damit gilt

$$\forall\, \Theta \in \mathbb{R} : \quad \sigma(\Theta \,+\, \frac{\pi}{n+2}) \;=\; -\sigma(\Theta) \tag{10.14}$$

und es folgt

$$\begin{aligned} \int_0^\pi \sigma(\Theta)\, \cos(k\Theta)\, \sin\Theta\, d\Theta \;&=\; \frac{1}{2} \int_0^\pi \sigma(\Theta)\, [\, \sin((k+1)\Theta) - \sin((k-1)\Theta)\,]\, d\Theta \\ &=\; \frac{1}{4} \int_{-\pi}^{\pi} \sigma(\Theta)\, [\, \sin((k+1)\Theta) - \sin((k-1)\Theta)\,]\, d\Theta. \end{aligned}$$

Wir setzen nun

$$I_m \;:=\; \int_{-\pi}^{\pi} \sigma(\Theta)\, \sin(m\,\Theta)\, d\Theta \tag{10.15}$$

Verschiebung des $2\,\pi$–periodischen Integranden mit $\Theta \;=\; \varphi + \dfrac{\pi}{n+2}$ ergibt unter Verwendung von (10.14)

$$
\begin{aligned}
I_m &= \int_{-\pi}^{\pi} \sigma(\varphi + \frac{\pi}{n+2})\ \sin(m\,[\varphi + \frac{\pi}{n+2}])\, d\varphi \\
&= -\int_{-\pi}^{\pi} \sigma(\varphi)\ \Big\{ \sin(m\,\varphi)\ \cos(\frac{m\pi}{n+2})\ +\ \cos(m\varphi)\ \sin(\frac{m\pi}{n+2}) \Big\}\ d\varphi \\
&= -\ \cos(\frac{m\pi}{n+2})\ I_m,
\end{aligned}
$$

wobei sich die letzte Gleichheit dadurch ergibt, dass $\sigma(\varphi)$ eine ungerade, $\cos(m\,\varphi)$ jedoch eine gerade Funktion ist.

Hieraus folgt insbesondere $I_m = 0$ für alle $m = 0, 1, \ldots, n+1$ und somit auch

$$\int_0^{\pi} \sigma(\Theta)\ \cos(k\Theta)\ \sin\Theta\, d\Theta \ = \ 0, \qquad k = 1, \ldots, n.$$

Für $k = 0$ folgt dies direkt aus $I_1 = 0$. Damit ist (10.12) gezeigt. $\square$

Folgerung (10.16)

Aus dem obigen Beweis halten wir fest: Setzt man für ein vorgegebenes $n \in \mathbb{N}$

$$\tau_{j,n} \ := \ \frac{j\,\pi}{n+1}, \qquad j = 0, \ldots, n+1,$$

sowie

$$\sigma_n(t) \ := \ \begin{cases} (-1)^j & : \ \tau_{j,n} < t < \tau_{j+1,n} \\ 0 & : \ t \ = \ \tau_{j,n}, \end{cases}$$

so gilt $\quad I_{m,n} \ := \displaystyle\int_0^{\pi} \sigma_n(t)\ \sin(m\,t)\, dt \ = \ 0, \quad$ für $m = 0, \ldots, n.$

Beispiel (10.17)

Gesucht sei die L_1–Bestapproximation von $f(t) := t$ auf $[0, \pi]$ bezüglich des folgenden linearen Teilraumes von $C[0, \pi]$:

$$V_n \ = \ \mathrm{Spann}\{\sin(kt) : \ k = 1, 2, \ldots, n\}.$$

Da V_n auf dem offenen Intervall $]0, \pi[$ ein Haarscher Raum der Dimension n ist, suchen wir Punkte $0 < \tau_1 < \ldots < \tau_n < \pi$, so dass mit der zugehörigen Vorzeichenfunktion σ_n gilt

$$\forall\, q \in V_n : \quad \int_0^{\pi} \sigma_n(t)\, q(t)\, dt \ = \ 0.$$

Nach Folgerung (10.16) sind diese Punkte durch $\tau_j = j\pi/(n+1)$ gegeben. Die Bestapproximation erhält man also durch Lösung der *trigonometrischen Interpolationsaufgabe*

$$p(t) = \sum_{k=1}^{n} a_k \sin(kt); \quad \text{mit} \quad p(\tau_j) = \tau_j, \; j = 1, \ldots, n.$$

Wir berechnen die *Minimalabweichung*

$$\begin{aligned}
\int_0^\pi |t - p(t)|\, dt &= \left| \int_0^\pi \sigma_n(t)\, (t - p(t))\, dt \right| = \left| \int_0^\pi \sigma_n(t)\, t\, dt \right| \\
&= \left| \sum_{j=0}^{n} (-1)^j \int_{\tau_j}^{\tau_{j+1}} t\, dt \right| = \left| \sum_{j=0}^{n} (-1)^j\, (\tau_{j+1}^2 - \tau_j^2)/2 \right| \\
&= \frac{\pi^2}{2\,(n+1)^2} \left| \sum_{j=0}^{n} (-1)^j\, (2j+1) \right| = \frac{\pi^2}{2\,(n+1)}\,.
\end{aligned}$$

Man vergleiche hierzu auch den Beweis des ersten Jackson-Satzes (7.12).