Bernsteinpolynome

Vortrag zum Proseminar zur Analysis, 6. 10. 2010
Malte Milatz

In diesem Vortrag wird der bereits im Skript zur Analysis II zitierte Approximationssatz von Weierstraß mithilfe der Bernsteinpolynome bewiesen. Der ursprüngliche Beweis stammt vom Namensgeber dieser Polynome, Sergej N. Bernstein.[1]

§ 1. Der Satz von Weierstraß

Satz. *Es sei $f : [a, b] \to \mathbb{R}$ eine stetige reelle Funktion auf einem kompakten Intervall $K = [a, b] \subset \mathbb{R}$. Dann gibt es eine Folge von Polynomfunktionen $\phi_n : K \to \mathbb{R}$, die gleichmäßig gegen f konvergiert.*

Es sei daran erinnert, daß die Folge ϕ *gleichmäßig gegen f konvergiert*, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so daß für alle $n \geq n_0$ gilt:

$$|f(x) - \phi_n(x)| < \varepsilon \quad \text{für alle } x \in D. \tag{1}$$

In Kurzschreibweise bedeutet das

$$\lim_{n\to\infty} \|\phi_n - f\|_\infty = 0. \tag{2}$$

Hierbei bezeichnet $\|\cdot\|_\infty$ die Supremumsnorm des Raumes $C(K)$ aller stetigen Funktionen von K nach $\mathbb{R}$.[2]

Man kann (2) so lesen, daß die Funktionen ϕ_n die Funktion f beliebig gut approximieren (gemäß der Supremumsnorm). Der Satz beantwortet dann die Frage, unter welchen Voraussetzungen man zu gegebenem f eine Folge besonders „einfacher“ Funktionen finden kann, die f beliebig gut approximieren. Betrachtet man $(C(K), \|\cdot\|_\infty)$ als normierten Vektorraum und definiert $P(K)$ als die Menge aller Polynomfunktionen von K nach $\mathbb{R}$, d. h.

$$P(K) := \{g : K \to \mathbb{R} \; : \; \exists \psi \in \mathbb{R}[X] \; \forall x \in K : \; g(x) = \psi(x)\},$$

so liefert der Satz also beliebig gute Näherungen von $f \in C(K)$ im Teilraum $P(K)$.

[1] Bernstein, S. N., Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités, Commun. Soc. Math. Kharkow, Vol. 12, No. 2, pp. 1-2, 1912/1913.

[2] Auf $C(K)$ handelt es sich hierbei tatsächlich um eine Norm – nicht so auf $K^{\mathbb{R}}$, wo $\|\cdot\|_\infty$ den Wert ∞ annehmen kann.

§2. Bernsteinpolynome

Definition. Es sei $f : [0,1] \to \mathbb{R}$ eine Funktion. Das *n-te Bernsteinpolynom von f* ist das Polynom
$$\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} t^k (1-t)^{n-k} \in \mathbb{R}[t].$$
Die von diesem Polynom auf [0,1] induzierte Polynomfunktion wird mit $B_n(f)$ bezeichnet:
$$B_n(f)(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}, \quad x \in [0,1]. \qquad \diamond$$

Bei den Bernsteinpolynomen handelt es sich also um bestimmte Linearkombinationen von Polynomen der Form $b_{n,k}(t) := \binom{n}{k} t^k (1-t)^{n-k}$, die für den Fall $n = 3$ in Abbildung 1 dargestellt sind.

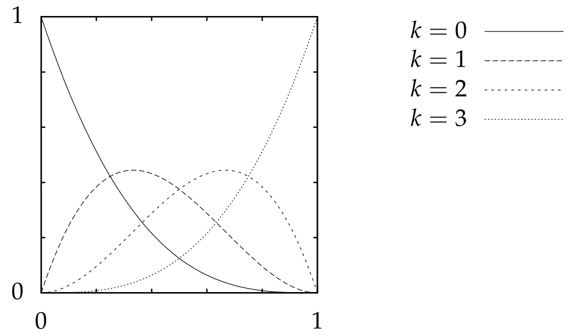

Abbildung 1: Die Polynome $b_{3,k}$, $k = 0, \ldots, 3$, auf dem Einheitsintervall.

Bernstein gibt folgende stochastische Motivation für seine Polynome: Er betrachtet ein Ereignis (etwa jenes, daß eine gegebene Münze nach einmaligem Wurf „Kopf" zeigt), dessen Wahrscheinlichkeit $x \in [0,1]$ betrage. Mit einem Glücksspieler wird vereinbart, das Experiment n-fach zu wiederholen und dem Spieler den Gewinn $f\left(\frac{k}{n}\right)$ auszuschütten, falls das Ereignis genau k-mal eingetreten ist. Dann beträgt der Erwartungswert für den Gewinn
$$\sum_{k=0}^{n} f\left(\frac{k}{n}\right) \cdot P(\{\text{genau } k\text{-mal Kopf}\}) = B_n(f)(x).$$
Ein negativer Gewinn ist als Verlust zu lesen; ein negativer Erwartungswert bedeutet, daß es sich nicht lohnt zu spielen. Während Bernstein den gesamten Beweis in der Sprache von Wahrscheinlichkeiten und Erwartungswerten verfaßt, ziehen wir hier eine rein analytische Formulierung vor.

§3. Beweis des Satzes

Lemma 3.1. Es sei $K \subset \mathbb{R}$ kompakt und $f : K \to \mathbb{R}$ eine stetige Funktion. Dann ist f nicht nur stetig, sondern sogar gleichmäßig stetig.

Dieses Lemma ist bereits aus der Analysis I bekannt und hier nur der Vollständigkeit halber bewiesen.

Beweis. Angenommen, die Konvergenz sei nicht gleichmäßig. Dann gibt es ein $\varepsilon > 0$ und zu jedem $n \in \mathbb{N}$ Punkte $x_n, y_n \in K$ mit $|x_n - y_n| < \frac{1}{n}$, aber $|f(x_n) - f(y_n)| > \varepsilon$. Nach dem Satz von Bolzano-Weierstraß haben die beschränkten Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ konvergente Teilfolgen, deren Grenzwerte wegen der Abgeschlossenheit von K wieder in K liegen und wegen $\lim_{n\to\infty} |x_n - y_n| = 0$ übereinstimmen. Der gemeinsame Grenzwert sei mit x_0 bezeichnet.

Da f stetig ist, gibt es ein $\delta > 0$, so daß aus $|x - x_0| < \delta$ stets $|f(x) - f(x_0)| < \frac{\varepsilon}{2}$ und aus $|y - x_0| < \delta$ stets $|f(y) - f(x_0)| < \frac{\varepsilon}{2}$ folgt. Diese Ungleichungen sind erfüllt, wenn man $x = x_n$ und $y = y_n$ mit einem genügend großen n wählt. Es folgt

$$\varepsilon < |f(x_n) - f(y_n)| \le |f(x_n) - f(x_0)| + |f(x_0) - f(y)| < \varepsilon,$$

ein Widerspruch. □

Wir beschränken uns nun zunächst auf den Fall $K = [0, 1]$, so daß wir mit den Bernsteinpolynomen arbeiten können, und führen anschließend den allgemeinen Fall darauf zurück. Die folgenden Spezialfälle helfen bei späteren Abschätzungen.

Lemma 3.2. Es sei $n \in \mathbb{N}$.

a) Für $f : [0, 1] \to \mathbb{R}, x \mapsto 1$ gilt $B_n(f) = f$.

b) Für $f : [0, 1] \to \mathbb{R}, x \mapsto x$ gilt ebenfalls $B_n(f) = f$.

c) Für $f : [0, 1] \to \mathbb{R}, x \mapsto x(1 - x)$ und $x \in [0, 1]$ gilt $B_n(f)(x) = \left(1 - \frac{1}{n}\right) f(x)$.

Beweis. Man rechnet nach:

a) Mit dem binomischen Lehrsatz hat man $B_n(f)(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} = (x + 1 - x)^n = 1$.

b) Mithilfe von (a) ergibt sich für alle $x \in [0, 1]$

$$\begin{aligned} B_n(f)(x) &= \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1 - x)^{n-k} = \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} x^k (1 - x)^{n-k} \\ &= \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} x^{k+1} (1 - x)^{n-1-k} \\ &= x \cdot \sum_{k=0}^{n-1} \binom{n-1}{k} x^k (1 - x)^{n-1-k} = x. \end{aligned}$$

c) Wiederum mit (a) hat man für $x \in [0,1]$

$$\begin{aligned}
B_n(f)(x) &= \sum_{k=0}^{n} \frac{k(n-k)}{n^2}\binom{n}{k}x^k(1-x)^{n-k} \\
&= \frac{n-1}{n}\sum_{k=1}^{n-1}\binom{n-2}{k-1}x^k(1-x)^{n-k} \\
&= \frac{n-1}{n}\sum_{k=0}^{n-2}\binom{n-2}{k}x^{k+1}(1-x)^{n-1-k} \\
&= \frac{n-1}{n}x(1-x)\sum_{k=0}^{n-2}\binom{n-2}{k}x^k(1-x)^{n-2-k} \\
&= \left(1-\frac{1}{n}\right)x(1-x).
\end{aligned}$$

□

Lemma 3.3. Für $x \in [0,1]$ und $n \in \mathbb{N}$ gilt

$$0 \leq \sum_{k=0}^{n}\left(x-\frac{k}{n}\right)^2 \cdot \binom{n}{k}x^k(1-x)^{n-k} = \frac{x(1-x)}{n} \leq \frac{1}{4n}.$$

Beweis. Schreibt man abkürzend $b_k(x) := \binom{n}{k}x^k(1-x)^{n-k}$, so ergibt sich mit (3.2)

$$\begin{aligned}
\sum_{k=0}^{n}\left(x-\frac{k}{n}\right)^2 b_k(x) &= x^2\sum_{k=0}^{n} b_k(x) + (1-2x)\sum_{k=0}^{n}\frac{k}{n}b_k(x) - \sum_{k=0}^{n}\frac{k(n-k)}{n^2}b_k(x) \\
&= x^2 + (1-2x)x - \left(1-\frac{1}{n}\right)x(1-x) \\
&= \frac{x(1-x)}{n}.
\end{aligned}$$

Aus $0 \leq (x-\frac{1}{2})^2 = x^2 - x + \frac{1}{4}$ folgt weiter $x(1-x) \leq \frac{1}{4}$, womit alles gezeigt ist. □

Lemma 3.4. Es sei $f : [0,1] \to \mathbb{R}$ stetig. Dann konvergiert die Funktionenfolge $(B_n(f))_{n\in\mathbb{N}}$ gleichmäßig gegen f.

Beweis. Zu gegebenem $\varepsilon > 0$ ist zu zeigen, daß für genügend große $n \in \mathbb{N}$ gilt:

$$|f(x) - B_n(f)(x)| < \varepsilon \quad \text{für alle } x \in [0,1].$$

Da f nach (3.1) gleichmäßig stetig ist, gibt es ein $\delta > 0$, so daß aus $|x_1 - x_2| < \delta$ stets

$$|f(x_1) - f(x_2)| < \frac{\varepsilon}{2} \tag{3}$$

folgt. Es sei nun $x \in [0,1]$ gegeben. Um $|f(x) - B_n(f)(x)|$ geeignet abschätzen zu können, teilen wir die Menge $\{0, \ldots, n\}$ in zwei Teile auf, und zwar

$$A_n := \left\{ k \in \{0, \ldots, n\} : \left| x - \frac{k}{n} \right| < \delta \right\} \text{ und } \overline{A_n} := \left\{ k \in \{0, \ldots, n\} : \left| x - \frac{k}{n} \right| \geq \delta \right\}.$$

Man erhält[3] für beliebig gegebenes $x \in [0,1]$:

$$\begin{aligned}
|f(x) - B_n(f)(x)| &= \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) \binom{n}{k} x^k (1-x)^{n-k} \right| \\
\text{(Dreiecksungl.)} &\leq \sum_{k=0}^{n} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \\
&= \sum_{k \in A_n} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \\
&\quad + \sum_{k \in \overline{A_n}} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \\
(3) &\leq \frac{\varepsilon}{2} \sum_{k \in A_n} \binom{n}{k} x^k (1-x)^{n-k} + 2\|f\|_\infty \sum_{k \in \overline{A_n}} \binom{n}{k} x^k (1-x)^{n-k} \\
\text{(Def. von } \overline{A_n}\text{)} &\leq \frac{\varepsilon}{2} \sum_{k \in A_n} \binom{n}{k} x^k (1-x)^{n-k} + \frac{2\|f\|_\infty}{\delta^2} \sum_{k \in \overline{A_n}} \left(x - \frac{k}{n} \right)^2 \binom{n}{k} x^k (1-x)^{n-k} \\
&\leq \frac{\varepsilon}{2} + \frac{\|f\|_\infty}{2\delta^2 n} \\
&< \varepsilon \quad \text{für alle } n > n_0 := \frac{\|f\|_\infty}{\varepsilon \delta^2}.
\end{aligned}$$

Die Existenz von $\|f\|_\infty$ folgt hierbei aus dem Satz vom Minimum und Maximum. □

Man beachte, daß die Wahl des n_0 unabhängig von x erfolgt, was für die gleichmäßige Konvergenz wesentlich ist. Jetzt folgt der Satz wie folgt.

Beweis. Es erfülle $f : K \to \mathbb{R}$ die Voraussetzungen des Satzes, d. h. $K = [a,b] \subset \mathbb{R}$ ist ein kompaktes Intervall und f ist stetig. Definiere die Konvexkombination

$$\tilde{f} : \quad [0,1] \to \mathbb{R}, \quad x \mapsto f(a + x(b-a)).$$

Mit f ist auch $\tilde{f}$ stetig, so daß (3.4) eine Folge $(\tilde{\phi}_n)_{n \in \mathbb{N}}$ liefert, die gleichmäßig gegen $\tilde{f}$ konvergiert. Für $n \in \mathbb{N}$ definiere

$$\phi_n : \quad K \to \mathbb{R}, \quad x \mapsto \tilde{\phi}_n\left(\frac{x-a}{b-a}\right).$$

[3] Bernstein benutzt an dieser Stelle ein Ergebnis aus der Wahrscheinlichkeitsrechnung, nach dem der Ausdruck $\sum_{k \in \overline{A_n}} \binom{n}{k} x^k (1-x)^{n-k}$ für $n \to \infty$ gegen 0 konvergiert. Wir benutzen stattdessen die elementarere Aussage in (3.3).

Dann konvergiert $(\phi_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f, denn

$$\|f-\phi_n\|_\infty = \|\tilde{f}-\tilde{\phi}_n\|_\infty \xrightarrow{n\to\infty} 0. \qquad \square$$

§4. Anmerkungen zur Aussage des Satzes

— *Über die Voraussetzung eines kompakten Definitionsbereichs.* —

Die Annahme des Satzes, daß der Definitionsbereich kompakt ist, kann nicht weggelassen werden, da es z. B. zu den Funktionen $g_1 : (0,1) \to \mathbb{R}, x \mapsto \frac{1}{x}$ und $g_2 : [0,\infty) \to \mathbb{R}, x \mapsto e^{-x}$ keine beliebig guten polynomialen Näherungen gibt. Das liegt am Grenzwertverhalten dieser Funktionen: Es gelten

$$\lim_{x\to 0} g_1(x) = \infty \qquad \text{sowie} \qquad \lim_{x\to\infty} g_2(x) = 0,$$

aber für eine nicht-konstante Polynomfunktion ψ gilt stets

$$\lim_{x\to 0} \psi(x) = \psi(0) \in \mathbb{R} \qquad \text{und} \qquad \lim_{x\to\infty} \psi(x) \in \{\pm\infty\},$$

mithin $\|g_1-\psi\|_\infty = \|g_2-\psi\|_\infty = \infty$.

— *Über die Annahme, daß die Funktion f stetig ist.* —

Auch die Annahme der Stetigkeit von f ist notwendig. Zu einer nicht stetigen Funktion findet man im allgemeinen nicht einmal beliebig gute stetige Approximationen. Das sieht man zum einen daran, daß ein gleichmäßiger Grenzwert stetiger Funktionen stets eine stetige Funktion ist, oder an einem Gegenbeispiel: Betrachte

$$g : [0,1] \to \mathbb{R}, x \mapsto \operatorname{sgn}(x).$$

Angenommen, $(\psi_n)_{n\in\mathbb{N}}$ sei eine Folge stetiger Funktionen, die gleichmäßig gegen g konvergiere. Dann gibt es nach der Definition der gleichmäßigen Konvergenz ein $n_0 \in \mathbb{N}$ mit $\|g-\psi_n\|_\infty < \frac{1}{2}$ für alle $n \geq n_0$. Fixiert man ein solches n, so hat man insbesondere $\psi_n(0) < \frac{1}{2}$. Wegen der Stetigkeit von ψ_n gibt es nun ein Intervall $[0,\delta)$, so daß sogar $\psi_n(x) < \frac{1}{2}$ für alle $x \in [0,\delta)$ gilt. Dann gilt aber für $x \in (0,\delta)$:

$$g(x) - \psi_n(x) = 1 - \psi_n(x) > 1 - \frac{1}{2} = \frac{1}{2}.$$

Dies ist ein Widerspruch zu $\|g-\psi_n\|_\infty < \frac{1}{2}$.

— *Über die Annahme, daß der Definitionsbereich ein Intervall ist.* —

Schließlich ist aber eine Bedingung des Satzes verzichtbar: Wenn wir den Definitionsbereich $[a, b]$ durch ein beliebiges Kompaktum $K \subset \mathbb{R}$ ersetzen, bleibt die Aussage gültig.

Korollar. Ist $K \subset \mathbb{R}$ kompakt und $f : K \to \mathbb{R}$ eine stetige Funktion, so gibt es eine Folge von Polynomfunktionen auf K, die gleichmäßig gegen f konvergiert.

Beweis. Die Idee ist, f auf dem kompakten Intervall $[\min K, \max K]$ stetig fortzusetzen. Es sei dazu $x \in [\min K, \max K] \setminus K$. Da x nicht in K liegt und K kompakt ist, ist x kein Häufungspunkt von K, d. h. es gibt eine zu K disjunkte Umgebung $(x - \delta, x + \delta)$ mit $\delta > 0$. Das Maximum des kompakten Intervalls $[\min K, x - \delta] \cap K$ ist der größte Punkt in K, der kleiner ist als x und sei mit $a(x)$ bezeichnet. Genauso gibt es auch kleinsten Punkt $b(x)$ in K, der größer ist als x. Verbindet man die Punkte $(a(x), f(a(x)))$ und $(b(x), f(b(x)))$ durch eine Gerade, definiert man also

$$\overline{f} : [\min K, \max K] \to \mathbb{R}, x \mapsto \begin{cases} f(x), & x \in K, \\ f(a(x)) + \frac{x-a(x)}{b(x)-a(x)} \left(f(b(x) - f(a(x))\right), & x \notin K, \end{cases}$$

so ist $\overline{f}$ stetig und genügt den Voraussetzungen des Satzes, so daß es eine Folge von Polynomen gibt, die gleichmäßig gegen $\overline{f}$ und damit auch gleichmäßig gegen f konvergiert. $\square$

— *Über die numerische Anwendung.* —

Im Beweis des Satzes haben wir zwar eine explizite Polynomfolge konstruiert, die gleichmäßig gegen die gegebene Funktion konvergiert. Zur numerischen Anwendung sind die Bernsteinpolynome aber etwas unhandlich, da speziell die Funktionswerte an den Stellen $\frac{k}{n}$, $k = 0, \ldots, n$, bekannt sein müssen (bzw. bei Skalierung auf ein allgemeines kompaktes Intervall $[a, b]$ an den Stellen $a + \frac{k}{n}(b - a)$). Außerdem ist nicht klar, welche Wahl von n die gewünschte Genauigkeit garantiert, da man den im Beweis verwendeten Stetigkeitsmodul nicht zahlenmäßig bestimmen möchte.