ScienceAsia 27 (2001) : **63-66**

Approximation of Lipschitz Functions on $\mathbb{R}^N$ by Bernstein Polynomials

K. Neammanee
Department of Mathematics, Faculty of Science, Chulalongkorn University Bangkok, Thailand.

Received 23 May 2000
Accepted 15 Dec 2000

ABSTRACT Let $f : \mathbb{R}^N \to \mathbb{R}^P$ be a Lipschitz function, i.e. $\| f(x)-f(y) \| \le C \| x - y \|$ for some positive constant C and all $x, y \in \mathbb{R}^N$. In this paper, we use the probabilistic tools to approximate f on any compact set by Bernstein polynomials. We prove that the rate of uniform convergence is at least of order $\sqrt{\frac{\ln n}{n}}$. Moreover, one can apply the same argument for Hölder functions.

KEYWORDS: Bernstein polynomials, Lipschitz functions, Hölder functions, Weierstrass approximation.

INTRODUCTION

Bernstein[1] first used his eponymous polynomials to prove the Weierstrass approximation theorem which states that the set of real polynomials over [0, 1] is dense in the space of all continuous real functions on [0, 1] in 1912. It is a classic application of probability theory to real analysis that finds its way into many textbooks[2, 3] and journals.[4] It took more than twenty years before results appeared concerning the rate of convergence of the approximation by Popoviciu[5] and by Kac[6, 7] became available. While Popoviciu established the speed of convergence in terms of the modulus of continuity, Kac originally proved the following theorem.

Theorem 1 If the function $f : [0, 1] \to \mathbb{R}$ is Hölder with exponent α ($0 < \alpha \le 1$) and constant L, i.e. $| f(x) - f(y)| \le L | x - y |^{\alpha}$ for all $x, y \in [0, 1]$, then

$$| f(x) - B_n(f, x) | \le L\left(\frac{x(1-x)}{n}\right)^{\frac{\alpha}{2}}$$

for all $n \in \mathbb{N}$, all $x \in [0, 1]$, where $B_n(f, x)$ is a Bernstein polynomial

$$B_n(f,x) = \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j} f\left(\frac{j}{n}\right).$$

More recently, Gzyl and Palacios[8], refined the original Bernstein argument to treat the specialization to Lipschitz functions on [0, 1] by Bernstein polynomials. The authors of [8] provided a rate of uniform convergence of $B_n(f, \cdot)$ to f using large deviation techniques. The theorem is the following.

Theorem 2 Let $f : [0, 1] \to \mathbb{R}$ be a Lipschitz function, i.e. there exists a constant C such that $| f(x) - f(y) | \le C | x - y |$ for all $x, y \in [0, 1]$. Then there exists a constant K such that for every $n \in \mathbb{N}$

$$| f(x) - B_n(f, x) | \le K\sqrt{\frac{\ln n}{n}}.$$

In this paper we extend Theorem 2 to Lipschitz functions on a compact set in $\mathbb{R}^N$. The theorem is the following.

Theorem 3 Let $f : \mathbb{R}^N \to \mathbb{R}^P$ be a Lipschitz function, i.e. $\| f(x) - f(y) \| \le C \| x - y \|$ for some positive constant C and all $x, y \in \mathbb{R}^N$. Then for every compact set D in $\mathbb{R}^N$ and $n \in \mathbb{N}$ there exists a polynomial $P_n : D \to \mathbb{R}^P$ and a constant K, depending on f, such that

$$\| f(x) - P_n(x) \| \le K\sqrt{\frac{\ln n}{n}}$$

for all $x \in D$.

PROOF OF THEOREM

Before we prove Theorem 3, we note that the important probabilistic tool in [2]-[4] is Chebyschev's inequality, and if the argument is applied to a function satisfying a Lipschitz condition, the rate of convergence of the Bernstein polynomials to the function can be shown to be at least of order $\frac{1}{n^{1/3}}$.

If instead of Chebyschev's inequality, we use another probabilistic tools (Lemma 1), we can prove that the rate of convergence is at least of order $\sqrt{\frac{\ln n}{n}}$.

Lemma 1 ([8], p650) Let X be a binomial random variable with parameters n and x. Then for a > 0, we have

$$P(|X-nx|\geq a)\leq 2e^{-\frac{2a^2}{n}}.$$

Proof of Theorem 3

In the (N+1)- faced die tossing experiment, in which each face has probability t_i $(0\leq t_i\leq 1, t_1+t_2+\ldots+t_{N+1}=1)$, let J_i be the random variable whose value is the number of the i^{th} -faced when we toss the die n times. Hence J_i is the binomial random variable with parameters n and t_i. So $E(J_i)=nt_i$, $Var(J_i)=nt_i(1-t_i)$ and

$$P(J_1=j_1, J_2=j_2, \ldots, J_{N+1}=j_{N+1}) = \binom{n}{j_1, j_2, \ldots, j_{N+1}} t_1^{j_1} t_2^{j_2} \ldots t_{N+1}^{j_{N+1}}.$$

Step 1 We first consider the case where $D=\prod_{i=1}^{N}\left[0,\frac{1}{N}\right]$.

Since $\prod_{i=1}^{N}[0,1]$ is compact and f satisfies the Lipschitz condition, f is bounded and uniformly continuous on $\prod_{i=1}^{N}[0,1]$. So there is an M > 0 such that $\|f(t)\| < M$ on $\prod_{i=1}^{N}[0,1]$.

For each $n\in \mathbb{N}$, let $P_n : \prod_{i=1}^{N}\left[0,\frac{1}{N}\right]\to \mathbb{R}^P$ be a polynomial function defined by $P_n(t_1,t_2,\ldots,t_N)=$

$$\sum_{j_1+j_2+\ldots+j_{N+1}=n} f\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\binom{n}{j_1, j_2, \ldots, j_{N+1}} t_1^{j_1} t_2^{j_2} \ldots t_{N+1}^{j_{N+1}}$$

where $t_{N+1}=1-(t_1+t_2+\ldots+t_N)$.

Since

$$\begin{aligned} 1 &= 1^n \\ &= [t_1+t_2+\ldots+t_{N+1}]^n \\ &= \sum_{j_1+j_2+\ldots+j_{N+1}=n}\binom{n}{j_1, j_2, \ldots, j_{N+1}} t_1^{j_1} t_2^{j_2} \ldots t_{N+1}^{j_{N+1}}, \end{aligned}$$

we have $f(t_1, t_2, \ldots, t_N)=$

$$\sum_{j_1+j_2+\ldots+j_{N+1}=n} f(t_1,t_2,\ldots,t_N)\binom{n}{j_1, j_2, \ldots, j_{N+1}} t_1^{j_1} t_2^{j_2} \ldots t_{N+1}^{j_{N+1}}.$$

So, one obtains

$$\begin{aligned} &\|f(t_1, t_2, \ldots, t_N) - P_n(t_1, t_2,\ldots, t_N)\| \\ &\leq \sum_{j_1+j_2+\ldots+j_{N+1}=n}\left\|f(t_1,t_2,\ldots,t_N)-f\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\right\| \binom{n}{j_1, j_2, \ldots, j_{N+1}} t_1^{j_1} t_2^{j_2} \ldots t_{N+1}^{j_{N+1}} \\ &= \sum_{j_1+j_2+\ldots+j_{N+1}=n}\left\|f(t_1,t_2,\ldots,t_N)-f\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\right\| P(j_1,j_2,\ldots,j_{N+1}) \end{aligned}$$

where $P(j_1, j_2, \ldots, j_{N+1}) = P(J_1=j_1, J_1=j_2, \ldots, j_{N+1}=j_{N+1})$.

For a > 0, let

$$S_1=\{(j_1, j_2,\ldots, j_N) : j_i\in\{0, 1, 2,\ldots, n\} \text{ and } \|(nt_1, nt_2,\ldots, nt_N)-(j_1, j_2,\ldots, j_N)\|\leq a\}$$

and

$$S_2=\{(j_1, j_2,\ldots, j_N) : j_i\in\{0, 1, 2, \ldots, n\} \text{ and } \|(nt_1, nt_2,\ldots, nt_N)-(j_1, j_2, \ldots, j_N)\| > a\}.$$

Then $\|f(t_1, t_2,\ldots, t_N) - P_n(t_1, t_2,\ldots, t_N)\|$

$$\begin{aligned} &\leq \sum_{\substack{j_1+j_2+\ldots+j_{N+1}=n\\ (j_1,j_2,\ldots,j_N)\in S1}}\left\|f(t_1,t_2,\ldots,t_N)-f\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\right\| P(j_1,j_2,\ldots,j_{N+1}) \\ &+ \sum_{\substack{j_1+j_2+\ldots+j_{N+1}=n\\ (j_1,j_2,\ldots,j_N)\in S2}}\left\|f(t_1,t_2,\ldots,t_N)-f\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\right\| P(j_1,j_2,\ldots,j_{N+1}). \quad \text{.....(1)} \end{aligned}$$

On S_1, we know that

$$\begin{aligned} &\left\|f(t_1,t_2,\ldots,t_N)-f\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\right\| \\ &\leq C\left\|f(t_1,t_2,\ldots,t_N)-\left(\frac{j_1}{n},\frac{j_2}{n},\ldots\frac{j_N}{n}\right)\right\| \\ &= \frac{C}{n}\|(nt_1, nt_2, \ldots, nt_N)-(j_1, j_2, \ldots, j_N)\| \\ &\leq \frac{Ca}{n}. \end{aligned}$$

So, one obtains

$$\sum_{\substack{j_1+j_2+\dots+j_{N+1}=n \\ (j_1,j_2,\dots,j_N)\in S_1}} \| f(t_1,t_2,\dots,t_N) - f(\frac{j_1}{n},\frac{j_2}{n},\dots\frac{j_N}{n}) \|$$

$$P(j_1,j_2,\dots,j_{N+1})$$

$$\le \frac{Ca}{n} \sum_{j_1+j_2+\dots+j_{N+1}=n} P(j_1,j_2,\dots,j_{N+1})$$

$$\le \frac{Ca}{n}. \qquad .....(2)$$

Since $\| f(t) \| \le M$ on $\prod_{i=1}^{N}[0,1]$, we have

$$\sum_{\substack{j_1+j_2+\dots+j_{N+1}=n \\ (j_1,j_2,\dots,j_N)\in S_2}} \| f(t_1,t_2,\dots,t_N) - f(\frac{j_1}{n},\frac{j_2}{n},\dots\frac{j_N}{n}) \|$$

$$P(j_1,j_2,\dots,j_{N+1})$$

$$\le \quad 2M \sum_{\substack{j_1+j_2+\dots+j_{N+1}=n \\ (j_1,j_2,\dots,j_N)\in S_2}} P(j_1,j_2,\dots,j_{N+1}). \qquad .....(3)$$

We now let

$A_1 = \{(j_1, j_2, \dots, j_{N+1}) \in S : (j_1, j_2, \dots, j_N) \in S_2$ and $|j_1 - nt_1| \ge \frac{a}{\sqrt{N}}\}$,

$A_2 = \{(j_1, j_2, \dots, j_{N+1}) \in S : (j_1, j_2, \dots, j_N) \in S_2$ and $|j_2 - nt_2| \ge \frac{a}{\sqrt{N}}\}$,

$\vdots$

$A_N = \{(j_1, j_2, \dots, j_{N+1}) \in S : (j_1, j_2, \dots, j_N) \in S_2$ and $|j_N - nt_N| \ge \frac{a}{\sqrt{N}}\}$,

where S is the sample space of (N+1)-faced die tossed n times, i.e.

$S = \{(j_1, j_2, \dots, j_{N+1}) : j_i \in \{0, 1, 2, \dots, n\}\}$.

By Lemma 1 we have

$$P(A_i) \le P(| J_i - nt_i | \ge \frac{a}{\sqrt{N}}) \;\; \le 2e^{\frac{-2a^2}{nN}}$$

for $i \in \{0, 1, 2, \dots, N\}$.

Since for each $(j_1, j_2, \dots, j_N) \in S_2$ there is j_0 such that $| j_0 - nt_0 | \ge \frac{a}{\sqrt{N}}$,

$$\bigcup_{i=1}^{N} A_i = \{(j_1, j_2, \dots, j_{N+1}) : (j_1, j_2, \dots, j_N) \in S_2\}.$$

Hence $\displaystyle\sum_{\substack{j_1+j_2+\dots+j_{N+1}=n \\ (j_1,j_2,\dots,j_N)\in S_2}} P(j_1,j_2,\dots,j_{N+1})$

$$= P(\bigcup_{i=1}^{N} A_i)$$

$$\le \sum_{i=1}^{N} P(A_i)$$

$$\le 2Ne^{\frac{-2a^2}{nN}}. \qquad .....(4)$$

From (1) - (4)

$$\| f(t_1, t_2, \dots, t_N) - P_n(t_1, t_2, \dots, t_N) \| \le \frac{Ca}{n} + 4MNe^{\frac{-2a^2}{nN}}.$$

If $a = \frac{1}{2}\sqrt{Nn\,(ln\, n)}$, then

$$\| f(t_1, t_2,\dots, t_N) - P_n(t_1, t_2, \dots, t_N) \| \le K \sqrt{\frac{ln n}{n}}$$

for some constant K.

In step 1 we proved Theorem 3 in case $D = \prod_{i=1}^{N}\left[0, \frac{1}{N}\right]$. In the following step, we extend our result to the case where $D = \prod_{i=1}^{N}[a_i, b_i]$.

Step 2 We now consider the case where $D = \prod_{i=1}^{N}[a_i, b_i]$.

If we let $H: \prod_{i=1}^{N} [a_i, N(b_i-a_i) + a_i] \to \prod_{i=1}^{N}[0,1]$ be a function defined by

$$H(x_1, x_2, \dots, x_N) = (H_1(x_1), H_2(x_2), \dots, H_N(x_N))$$

where $H_i: \prod_{i=1}^{N} [a_i, N(b_i-a_i) + a_i] \to [0, 1]$ is defined by

$$H_i(x_1, x_2, \dots, x_N) = \frac{x_i - a_i}{N(b_i - a_i)},$$

then H is a bijective continuous function and $H(\prod_{i=1}^{N}[a_i, b_i]) = \prod_{i=1}^{N}[0, \frac{1}{N}]$.

We now define a function $F : \prod_{i=1}^{N} [0,1] \to \mathbb{R}^P$ by $F(x) = f \circ H^{-1}(x)$.

Hence, for $x = (x_1, x_2, ..., x_N)$ and $y = (y_1, y_2, ..., y_N)$ we have

$$\| F(x)-F(y) \|$$
$$= \| f \circ H^{-1}(x_1, x_2, ..., x_N) - f \circ H^{-1}(y_1, y_2, ..., y_N) \|$$
$$= \| f (x_1N(b_1 - a_1) + a_1, ..., x_NN(b_N - a_N) + a_N) - f (y_1N(b_1 - a_1) + a_1, ..., y_N N(b_N - a_N) + a_N) \|$$
$$\le C \| (x_1 - y_1)N(b_1 - a_1), (x_2 - y_2) N(b_2 - a_2), ..., (x_N - y_N)N(b_N - a_N) \|$$
$$\le CN \max_{1\le i\le N} (b_i - a_i) \| x - y \|.$$

Therefore, F is a Lipschitz function. Obviously, F is continuous on $\prod_{i=1}^{N} [0,1]$ and $f = F \circ H$. By step 1, for $n \in N$ there exists a polynomial function $P_n : \prod_{i=1}^{N} [0, \frac{1}{N}] \to \mathbb{R}^P$ and a constant K such that

$$\| F(y_1, y_2, ..., y_N) - P_n(y_1, y_2, ..., y_N) \| \le K \sqrt{\frac{ln n}{n}}$$

for $(y_1, y_2, ..., y_N) \in \prod_{i=1}^{N} [0, \frac{1}{N}]$.

We next define a polynomial function

$\tilde{P}_n : \prod_{i=1}^{N} [a_i, b_i] \to \mathbb{R}^P$ by

$$\tilde{P}_n (x_1, x_2, ..., x_N) = P_n \circ H (x_1, x_2, ..., x_N).$$

Since $H(\prod_{i=1}^{N} [a_i, b_i]) = \prod_{i=1}^{N} [0, \frac{1}{N}]$, $\tilde{P}_n$ is well-defined and

$$\| f(x_1, x_2, ..., x_N) - \tilde{P}_n (x_1, x_2, ..., x_N) \|$$
$$= \| F \circ H (x_1, x_2, ..., x_N) - P_n \circ H (x_1, x_2, ..., x_N) \|$$
$$= \| F(H(x_1, x_2, ..., x_N)) - P_n (H(x_1, x_2, ..., x_N)) \|$$
$$\le K \sqrt{\frac{ln n}{n}}.$$

Step 3 Finally , we consider the case where D is a compact set in $\mathbb{R}^N$.

From the fact that every compact subset in $\mathbb{R}^N$ is bounded, we can find a set $\prod_{i=1}^{N} [a_i, b_i]$ such that $D \subseteq \prod_{i=1}^{N} [a_i, b_i]$. So we can apply step 2 to a compact set D.

Conclusion

In this paper, we have shown that a Lipschitz funtion may be approximated by the Bernstein polynomials and given the rate of uniform convergence. We can use the same argument to prove Theorem 3 in the case of f being a Hölder function.

References

1. Bernstein, S (1912) Démonstration du théoréme de weierstrass fondée sur le calcul des probabilities, *Comm. Soc. Math. Kharkov* 13, 1-2.
2. Alon, N and Spencer, J (1992) The Probabilistic Method, Wiley, New York.
3. Chung, K.L. (1974) A course in Probability Theory, 2nd ed., Academic Press, New York.
4. Levasseur, K.M. (1984) A probabilistic proof of the Weierstrass approximation theorem, *Amer. Math. Monthty* 91, 249-250.
5. Popoviciu, T (1934) Sur l'approximation des fonctions convexes d'ordre supérieur, *Mathematica* 10, 49-54.
6. Kac, M (1938) Une remarque sur les polynomes de M.S. Bernstein, *Studia math.* 7, 49-51.
7. Kac, M (1939) Reconnaissance de priorité relative à ma note, "Une remarque sur les polynomes de M.S. Bernstein", *Studia Math.* 8, 170.
8. Gzyl, H and Palacios, J.L. (1997) The Weierstrass approximation theorem and large deviations, *Amer. Math. Monthly* 104, 650-653.