JOURNAL OF APPROXIMATION THEORY 59, 307–315 (1989)

Lipschitz Constants for Some Approximation Operators of a Lipschitz Continuous Function

MOHAMMAD K. KHAN AND MICHAEL A. PETERS

Department of Mathematical Sciences, Kent State University, Kent, Ohio 44242, U.S.A.

Communicated by Oved Shisha

Received June 22, 1987

1. INTRODUCTION

For a function f, continuous on $[0, 1]$, the nth ($n \geqslant 1$) Bernstein polynomial of f is the polynomial of degree $\leqslant n$ defined by

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

$B_n(f)$ is a polynomial, so there exists for every $\mu \in (0, 1]$ a constant A such that $B_n(f) \in \mathrm{Lip}_A\, \mu$ (where A may depend on f and μ). Brown *et al.* [2] gave an elementary proof that the Lipschitz constant is preserved, i.e.:

THEOREM 1. *If $f \in \mathrm{Lip}_A\, \mu$, then for all $n \geqslant 1$, $B_n(f) \in \mathrm{Lip}_A\, \mu$.*

For a brief history of the problem see [2]. By using some simple results from probability theory, we shall prove that the identical result holds for a general class of univariate and multivariate approximation operators. This class includes the Szász, Bernstein, Gamma, and translation operators, the tensor product operators formed from them, and the Bernstein operators over simplices. The central idea is to exploit the splitting property of the probability distributions involved in these approximation operators.

In this paper a function from a convex subset $I_k \subseteq \mathbb{R}^k$ into $\mathbb{R}$ is said to be Lipschitz continuous of order μ, $\mu \in (0, 1]$, if there exists a constant $A \geqslant 0$ such that for every pair of points $x, y \in I_k$ we have

$$|f(x) - f(y)| \leqslant A\, \|x - y\|^{\mu}$$

($\|\cdot\|$ is the l_1 norm in $\mathbb{R}^k$). We write this as $f \in \mathrm{Lip}_A\, \mu(I_k)$ or simply $f \in \mathrm{Lip}_A\, \mu$, suppressing the I_k.

0021-9045/89 $3.00

We shall use the multivariate Feller operator which generalizes the above-mentioned approximation operators. For the properties of the Feller operator see [4, 6, 9]. Let $f \in C(I_k)$. Suppose that $S_{n,x}$, $n \geq 1$, is a sequence of k-dimensional random vectors over I_k with $E(S_{n,x}) = nx$ and with $\mathrm{Cov}(S_{n,x}, S_{n,x}) = \sigma_n^2(x)$ a non-singular variance–covariance matrix, where x is a k-dimensional vector of parameters. Stancu [9] defined the multivariate version of the Feller operator:

$$L_n(f, x) := E\left\{f\left(\frac{S_{n,x}}{n}\right)\right\}$$
$$:= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f\left(\frac{t_1}{n}, ..., \frac{t_k}{n}\right) dF_{n,x}(t_1, ..., t_k),$$

where $F_{n,x}$ is the joint distribution function (df) of $S_{n,x}$ and $E\,|f(S_{n,x}/n)| < \infty$. In the next section the principal result is proved. A few special cases are outlined in the last section.

2. Principal Result

The tool used to prove Theorem 1 is to split a binomial random variable into a multinomial random vector. Consider the experiment of rolling n times a three-sided die that has sides one and two colored blue and side three colored white. Let S_n be the number of times blue comes up, U be the number of times side one comes up, and V be the number of times side two comes up. S_n has the same probability distribution as $U+V$, i.e., $P(S_n = k) = P(U+V=k)$, $k = 0, 1, ..., n$, which we write as $S_n \sim U+V$. This shows that a binomial random variable S_n can be split into a trinomial random vector (U, V) such that $S_n \sim U+V$. If the probability of a blue face coming up on one roll is x and the probability of side one coming up is y (clearly $x \geq y$) then the expected value of V, $E(V)$, is $n(x-y)$. Also note that $B_n(f, x) = Ef(S_n/n) = Ef((U+V)/n)$ and $B_n(f, y) = Ef(U/n)$. We formalize this concept of the splitting of a random variable in the following definition:

Definition. The random vector $S_{n,x}$ is said to have the splitting property if for every x, $y \in I_k$ there exist random vectors U, R_x, and R_y such that U, R_x, and R_y are defined over the same probability space with

$$S_{n,x} \sim U + R_x \qquad \text{and} \qquad S_{n,y} \sim U + R_y$$

and $E\,\|R_x - R_y\| \leq n\,\|x - y\|$.

The random vectors R_x and R_y need not be independent of U. For $k=1$ and $x>y$, the splitting property would hold if $S_{n,x} \sim U+R$, and $S_{n,y} \sim U$, where $E|R| \leqslant n(x-y)$ and U and R are defined over the same probability space. The main result of this paper is the following:

THEOREM 2. *Assume that $S_{n,x}$ has the splitting property. If $f \in \mathrm{Lip}_A \mu(I_k)$ then $L_n(f) \in \mathrm{Lip}_A \mu(I_k)$ $(0<\mu\leqslant 1)$.*

Proof. Split $S_{n,x}$ and $S_{n,y}$. Then

$$\begin{aligned}
|L_n(f,x)-L_n(f,y)| &= \left|Ef\left(\frac{S_{n,x}}{n}\right)-Ef\left(\frac{S_{n,y}}{n}\right)\right| \\
&= \left|Ef\left(\frac{U+R_x}{n}\right)-Ef\left(\frac{U+R_y}{n}\right)\right| \\
&\leqslant E\left|f\left(\frac{U+R_x}{n}\right)-f\left(\frac{U+R_y}{n}\right)\right|.
\end{aligned}$$

The equality holds by the splitting property and the last inequality follows since the underlying probability space is the same. Consequently,

$$\begin{aligned}
|L_n(f,x)-L_n(f,y)| &\leqslant AE\left(\left\|\frac{R_x-R_y}{n}\right\|^{\mu}\right) \\
&\leqslant A\left(E\left\|\frac{R_x-R_y}{n}\right\|\right)^{\mu} \\
&\leqslant A\,\|x-y\|^{\mu}.
\end{aligned}$$

The first inequality is due to the Lipschitz hypothesis, the second inequality is by Jensen's inequality, and the last inequality follows by the splitting property. This completes the proof.

Remark. The converse of Theorem 2 is trivially true if $L_n(f,x) \to f(x)$ as $n \to \infty$.

Remark. Since the l_p norms have the property

$$C_p\|\cdot\| \leqslant \|\cdot\|_p \leqslant \|\cdot\|,$$

where $C_p = k^{1/p-1}$, it follows that if $f \in \mathrm{Lip}_A(\mu,p)$ (i.e., $\mathrm{Lip}_A\,\mu$ with respect to the l_p norm, $1<p$) and $S_{n,x}$ has the splitting property then $L_n(f) \in \mathrm{Lip}_{A^*}(\mu,p)$, where $A^* = C_p^{-\mu}A$.

3. Special Cases

1. *Bernstein Operator*

Let

$$P(S_{n,x}=k)=\binom{n}{k}x^k(1-x)^{n-k}; \qquad k=0, 1, .., n.$$

For this distribution $L_n(f, x)$ reduces to $B_n(f, x)$. Let (U, V) have a trinomial distribution [5] with parameters (n, x_1, x_2-x_1), where $0<x_2<1$, $0<x_1<x_2$, i.e.,

$$P(U=j, V=l)=\frac{n!}{j!\,l!\,(n-j-l)!}x_1^j(x_2-x_1)^l(1-x_2)^{n-j-l},$$

$j=0,\ 1,\ 2, ..., n;\ l=0,\ 1, ..., n;\ j+l\leqslant n$. The joint moment generating function (mgf) of (U, V) is

$$\phi_{U,V}(t_1, t_2)=E(e^{t_1U+t_2V})=(1-x_2+x_1e^{t_1}+(x_2-x_1)e^{t_2})^n,$$

Hence, by letting $t_1=t_2$ we get $S_{n,x_2}\sim U+V$. Also note that by letting $t_2=0$ we get $U\sim S_{n,x_1}$ and similarly $V\sim S_{n,x_2-x_1}$. Furthermore $E|V|=E(V)=n(x_2-x_1)$. Therefore by taking $V=R$, we have that $S_{n,x}$ has the splitting property and Theorem 1 follows from Theorem 2.

2. *Szász Operator*

Let $P(S_{n,x}=k)=\exp(-nx)(nx)^k/k!$; $k=0, 1, 2, ...,$ and $x>0$. $L_n(f, x)$ reduces to the Szász operator [6]

$$S_n(f, x)=e^{-nx}\sum_{k=0}^{\infty}f\left(\frac{k}{n}\right)\frac{(nx)^k}{k!}.$$

Define the joint probability distribution of (U, V) as

$$P(U=j, V=l)=e^{-nx_1}\frac{(nx_1)^j}{j!}e^{-n(x_2-x_1)}\frac{(n(x_2-x_1))^l}{l!},$$

$j=0,\ 1,\ 2, ...;\ l=0,\ 1, ...;\ 0<x_1\leqslant x_2$. Again, it is easily verified that

$$P(U+V=k)=\sum_{0\leqslant j+l=k}P(U=j, V=l)=P(S_{n,x_2}=k)$$

for all $k=0, 1, ...$ Also note that U, V are independent Poisson random variables and $E(V)=n(x_2-x_1)$. Taking $U=S_{n,x_1}$, and $V=R$, we have by Theorem 2 that if $f\in \mathrm{Lip}_A\,\mu$ then $S_n(f)\in \mathrm{Lip}_A\,\mu$.

3. *Baskakov Operator*

Let $S_{n,x}$ have a negative binomial distribution with parameters (n, x), $x > 0$, i.e.,

$$P(S_{n,x} = k) = (1+x)^{-n} \binom{n+k-1}{k} \left(\frac{x}{1+x}\right)^k, \qquad k = 0, 1, 2, \ldots.$$

The Feller operator reduces to the Baskakov operator

$$B_n^*(f, x) = \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right) P(S_{n,x} = k).$$

Let (U, V) have negative multinomial distribution [5] with parameters (n, p_1, p_2), where $0 < p_1, p_2$; and let $q = p_1 + p_2 + 1$.

$$P(U = j, V = l) = \frac{(n+j+l-1)!}{j!\, l!\, (n-1)!} q^{-n} \left(\frac{p_1}{q}\right)^j \left(\frac{p_2}{q}\right)^l,$$

$j = 0, 1, \ldots; l = 0, 1, \ldots$. The mgf of (U, V) is

$$\phi_{U,V}(t_1, t_2) = (q - p_1 e^{t_1} - p_2 e^{t_2})^{-n}.$$

By letting $t_1 = t_2$ we have that $U + V$ has a negative binomial distribution with parameters $(n, p_1 + p_2)$. By letting $t_2 = 0$ we get that $U \sim S_{n, p_1}$, and similarly $V \sim S_{n, p_2}$. Therefore, for $0 < x_1 < x_2$, by taking $p_1 = x_1$ and $p_2 = x_2 - x_1$ we have that

$$S_{n, x_2} \sim U + V.$$

Also, $E(V) = n(x_2 - x_1)$. Hence, by Theorem 2, if $f \in \mathrm{Lip}_A\, \mu$ then $B_n^*(f) \in \mathrm{Lip}_A\, \mu$.

4 *Gamma Operator*

For $x > 0$, the Gamma operator is defined as

$$G_n(f, x) = \frac{x^{-n}}{(n-1)!} \int_0^{\infty} f\left(\frac{y}{n}\right) y^{n-1} e^{-y/x}\, dy.$$

Let $S_{n,x}$ have the probability density function

$$g_{S_{n,x}}(y) = \begin{cases} \dfrac{x^{-n}}{(n-1)!} y^{n-1} e^{-y/x} & \text{if } \ y > 0 \\ 0 & \text{if } \ y \leqslant 0. \end{cases}$$

Then $L_n(f, x)$ reduces to $G_n(f, x)$. It is easy to check that $xS_{n,1} \sim S_{n,x}$.

Therefore, if $0 < x_1 < x_2$ then

$$\begin{aligned} S_{n,x_2} &\sim x_2 S_{n,1} \\ &\sim x_1 S_{n,1} + (x_2 - x_1) S_{n,1} \\ &\sim S_{n,x_1} + R. \end{aligned}$$

Note that S_{n,x_1} and R are dependent random variables and $R \sim S_{n,x_2-x_1}$. Also, $E(R) = n(x_2 - x_1)$, which implies, by Theorem 2, that $G_n(f) \in \mathrm{Lip}_A\, \mu$ if $f \in \mathrm{Lip}_A\, \mu$.

5. *Translation Operators*

To consider positive translation operators (also known as convolution operators [10]), let $H_n(t-x)$ be the kernel for the translation operator

$$\begin{aligned} T_n(f, x) &= \int_{-\infty}^{\infty} f(t)\, d_t H_n(t-x) \\ &= \int_{-\infty}^{\infty} f(u+x)\, dH_n(u). \end{aligned}$$

Let Z_n be a random variable having df $H_n(z)$ and define $S_{n,x} := n(x + Z_n)$. For this case $L_n(f, x)$ reduces to the translation operator $T_n(f, x)$. For $x_2 > x_1$

$$S_{n,x_2} \sim n(x_1 + Z_n) + n(x_2 - x_1) \sim S_{n,x_1} + R,$$

where R is a degenerate random variable taking the value $n(x_2 - x_1)$ with probability one. Hence, by Theorem 2, if $f \in \mathrm{Lip}_A\, \mu$ then $T_n(f) \in \mathrm{Lip}_A\, \mu$. In the following some of the typical translation operators are provided.

(i) Fejér Operator [3, p. 34]:

$$F_n(f, x) = \frac{1}{n} \int_{-1/2}^{1/2} f(u+x) \left[\frac{\sin n\pi u}{\sin \pi u} \right]^2 du.$$

(ii) Korovkin [1, 7]:

Let ϕ be a non-negative, even, and continuous function on $[-r, r]$, decreasing on $[0, r]$ and such that $\phi(0) = 1$, and $0 \leqslant \phi(t) < 1$ for $0 < t \leqslant r$. For a continuous f on $I = [a, b]$ with $b - a \leqslant r$,

$$K_n(f, x) = \rho_n \int_a^b f(t)\, \phi^n(t-x)\, dt, \qquad n = 1, 2, ...,$$

where

$$1 = 2\rho_n \int_0^r \phi^n(t)\,dt.$$

Since $K_n(f, x)$ is a special case of $T_n(f, x)$, if $f \in \mathrm{Lip}_A\,\mu$ then $K_n(f) \in \mathrm{Lip}_A\,\mu$. Several classical smoothing operators [8] such as the Weierstrass and Picard operators are special cases of $K_n(f, x)$.

6. *Bernstein Operator over a Simplex*

Let $S_{n,x} = (S_{n,x_1}, ..., S_{n,x_k})$ have a multinomial distribution with parameters $(n,\ x_1, ..., x_k,\ 1 - x_1 - \cdots - x_k)$, where $x = (x_1, ..., x_k) \in \Delta_k =$ the standard simplex. That is, $\Delta_k = \{(x_1, ..., x_k)\colon 0 \leqslant x_i \leqslant 1;\ i = 1, 2, ..., k;$ and $x_1 + \cdots + x_k \leqslant 1\}$. For a continuous f defined over Δ_k, the Bernstein operator is defined as

$$B_{n,k}(f, x) = \sum_j f(j/n)\, P(S_{n,x} = j),$$

where $j = (j_1, ..., j_k)$; $j_i \geqslant 0$; $i = 1, 2, ..., k$; $j_1 + \cdots + j_k \leqslant n$ and

$$P(S_{n,x} = j) = \frac{n!}{j_1!\, j_2! \cdots j_k!\, (n - j_1 - \cdots - j_k)!} \times (1 - x_1 - x_2 - \cdots - x_k)^{n - j_1 - \cdots - j_k} \prod_{i=1}^{k} x_i^{j_i}$$

is the multinomial distribution [5] with parameters $(n,\ x_1, ..., x_k,\ 1 - \sum_{i=1}^k x_i)$. Let $f \in \mathrm{Lip}_A\,\mu$. For $x \neq y$ define a $2k$-dimensional random vector (U, V) having multinomial distribution [5] with parameters

$$\left(n, x_1^*, ..., x_k^*, |x_1 - y_1|, ..., |x_k - y_k|, 1 - \sum_{i=1}^k x_i^* - \sum_{i=1}^k |x_i - y_i|\right),$$

where U is a k-dimensional multinomial with parameters $(n, x^*, 1 - \sum_{i=1}^k x_i^*)$ and V is a k-dimensional multinomial with parameters $(n, |x_1 - y_1|, ..., |x_k - y_k|,\ 1 - \sum_{i=1}^k |x_i - y_i|)$ and $x^* = (\min(x_1, y_1), ..., \min(x_k, y_k))$. Let $c_i = 1$ if $x_i > y_i$ and $c_i = 0$ otherwise, $i = 1, 2, ..., k$. Define $R_x = (V_1 c_1, ..., V_k c_k)$ and $R_y = (V_1(1 - c_1), ..., V_k(1 - c_k))$. Note that

$$S_{n,x} \sim U + R_x \qquad \text{and} \qquad S_{n,y} \sim U + R_y,$$

where $U \sim S_{n,x^*}$. Also since $2c_i - 1 = 1$ or -1, $\|R_x - R_y\| = V_1 + V_2 + \cdots + V_k$. Therefore, $\|R_x - R_y\|$ has a binomial distribution with parameters $(n, \|x-y\|)$. Hence, $S_{n,x}$ has the splitting property and therefore by Theorem 2 if $f \in \mathrm{Lip}_A\, \mu$ then $B_{n,k}(f) \in \mathrm{Lip}_A\, \mu$.

7. *Tensor Product Operators*

For the tensor product operators formed by the Bernstein, Szász, Baskakov, Gamma, and Weierstrass operators the same result holds. For this more general setup, we need to modify $L_n(f, x)$ to incorporate the tensor product operators with different n_i, $i=1, 2, ..., k$. Let S_{n_i, x_i}, $i=1, 2, ..., k$, be mutually independent univariate random variables defined over $I \subseteq \mathbb{R}$ having the univariate splitting property. Define $S_{n,x} = (S_{n_1,x_1}, ..., S_{n_k,x_k})$, where $x=(x_1, ..., x_k)$ and $n=(n_1, ..., n_k)$. The tensor product operator for $f \in C(I_k)$, $I_k = I \times I \times \cdots \times I$, is defined as

$$L_n(f, x) = Ef\left(\frac{S_{n_1,x_1}}{n_1}, \frac{S_{n_2,x_2}}{n_2}, ..., \frac{S_{n_k,x_k}}{n_k}\right).$$

If S_{n_i,x_i} has df $F_{n_i,x_i}(t_i)$, $i=1, 2, ..., k$, then by the mutual independence of the random variables the joint df of $S_{n,x}$ is

$$F_{n,x}(t) = \prod_{i=1}^{k} F_{n_i,x_i}(t_i),$$

where $t=(t_1, ..., t_k)$. Let $x, y \in I_k$ be fixed vectors. By the univariate splitting property there exist random variables U_i, $R^i_{x_i}$, and $R^i_{y_i}$ defined over the same probability space such that $S_{n_i,x_i} \sim U_i + R^i_{x_i}$, $S_{n_i,y_i} \sim U_i + R^i_{y_i}$, and $E\,|R^i_{x_i} - R^i_{y_i}| \leqslant n_i\,|x_i - y_i|$, $i=1, 2, ..., k$. Let $U=(U_1, ..., U_k)$, $R_x = (R^1_{x_1}, ..., R^k_{x_k})$, and $R_y = (R^1_{y_1}, ..., R^k_{y_k})$. Define the joint df of (U, R_x, R_y) to be the product measure of the measures of $(U_i, R^i_{x_i}, R^i_{y_i})$, $i=1, 2, ..., k$, so that the random vectors $(U_i, R^i_{x_i}, R^i_{y_i})$, $i=1, 2, ..., k$, are mutually independent. By the univariate splitting property and independence,

$$S_{n,x} \sim U + R_x \qquad \text{and} \qquad S_{n,y} \sim U + R_y.$$

Also, by the univariate splitting property, $E\,|R^i_{x_i} - R^i_{y_i}| \leqslant n_i\,|x_i - y_i|$. We have that

$$\begin{aligned} |L_n(f, x) - L_n(f, y)| &\leqslant AE\left(\sum_{i=1}^{k} \frac{|R^i_{x_i} - R^i_{y_i}|}{n_i}\right)^{\mu} \\ &\leqslant A\left(\sum_{i=1}^{k} E\frac{|R^i_{x_i} - R^i_{y_i}|}{n_i}\right)^{\mu} \\ &\leqslant A\,\|x-y\|^{\mu}. \end{aligned}$$

Hence, the tensor product operators are $\mathrm{Lip}_A\, \mu$ for $f \in \mathrm{Lip}_A\, \mu$ whenever the univariate operators have the splitting property.

ACKNOWLEDGMENT

We thank Professor A. S. Cavaretta, Jr., for his valuable suggestions.

REFERENCES

1. R. BOJANIC AND O. SHISHA, On the precision of uniform approximation of continuous functions by certain linear positive operators of convolution type, *J. Approx. Theory* **8** (1973), 101–113.
2. B. M. BROWN, D. ELLIOTT, AND D. F. PAGET, Lipschitz constants for the Bernstein polynomials of a Lipschitz continuous function, *J. Approx. Theory* **49** (1987), 196–199.
3. H. DYM AND H. P. MCKEAN, "Fourier Series and Integrals," Academic Press, New York, 1972.
4. W. FELLER, "An Introduction to Probability Theory and Its Applications II," Wiley, New York, 1966.
5. N. L. JOHNSON AND S. KOTZ, "Discrete Distributions," Houghton–Mifflin, Boston, 1969.
6. R. A. KHAN, Some probabilistic methods in the theory of approximation operators, *Acta. Math. Sci. Hungar.* **39** (1980), 193–203.
7. P. P. KOROVKIN, "Linear Operators and Approximation Theory," Hindustan Pub., Delhi, 1960 (translated from the Russian edition of 1959).
8. H. S. SHAPIRO, "Smoothing and Approximation of Functions," Van Nostrand–Reinhold, New York, 1969.
9. D. D. STANCU, Probabilistic methods in the theory of approximation of functions of several variables by linear positive operators, *in* "Approximation Theory" (A. Talbot, Ed.), pp. 329–342, Academic Press, New York, 1970.
10. R. DEVORE, "The Approximation of Continuous Functions by Positive Linear Operators," Lecture Notes in Mathematics, Vol. 293, Springer-Verlag, New York, 1972.