Aproximación mediante polinomios de Bernstein deterministas y aleatorios

Autor: Miguel Abardía Serrano
Trabajo de fin de grado
Universidad de Zaragoza

Director del trabajo: José Antonio Adell Pascual
Fecha: 12 de Junio de 2024

Resumen

El objeto central del trabajo es el estudio de los polinomios de Bernstein tanto deterministas como aleatorios. Los operadores lineales positivos son uno de los métodos de aproximación de funciones más estudiados. Uno de los ejemplos más importantes son los polinomios de Bernstein. Estos polinomios se definen como

$$B_n f(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-k)^{n-k}, \quad x \in [0,1], n \geq 1,$$

donde $f : [0,1] \to \mathbb{R}$ es una función arbitraria. A principios del siglo XX, el matemático ruso S. Bernstein realizó trabajos que llevaron a la aproximación de funciones utilizando la teoría de la probabilidad. Introdujo los polinomios que llevan su nombre. Son muchos los usos de los polinomios de Bernstein, en economía, para modelizar mercados, en ingeniería forestal para replantar árboles, para el cálculo de órbitas de satélites, en medicina para modelizar el ritmo cardiaco... pero uno de los usos más trascendentes en matemáticas es su utilización para demostrar de forma constructiva el teorema de aproximación de Weierstrass, según el cual toda función continua en un compacto se puede aproximar uniformemente mediante polinomios.

El capítulo primero trata sobre los operadores lineales positivos y sus representaciones probabilísticas. Los polinomios de Bernstein sirvieron como base para aplicar métodos probabilísticos en la aproximación mediante operadores lineales positivos, por lo que están relacionados directamente con la probabilidad. Aunque el enfoque principal de este trabajo es el análisis de polinomios de Bernstein, tanto en sus formas deterministas como aleatorias, también se pretende demostrar que estos polinomios son un ejemplo paradigmático de operadores lineales de aproximación y extender la representación probabilística de los polinomios de Bernstein a través de procesos empíricos a otros operadores lineales usuales, unificando así las demostraciones de propiedades clave como la preservación de la monotonía, φ-variación, la suavidad global y las constantes de Lipschitz. Analizamos las representaciones probabilísticas de algunos operadores lineales positivos, en términos de varios procesos estocásticos a doble índice, entre ellos los procesos empíricos, los procesos de Poisson y los procesos Gamma. Los métodos probabilísticos suelen ser más eficaces si la familia Z contiene variables aleatorias con fuertes relaciones de dependencia.

En el segundo capítulo, estudiamos la velocidad de convergencia de los polinomios de Bernstein respecto a distintos módulos de continuidad. Para funciones f con distintos grados de derivabilidad, obtenemos la velocidad de convergencia de los polinomios de Bernstein respecto al primer módulo de continuidad. De hecho, se pueden dar cotas fácilmente en términos del primer módulo de continuidad, usando la fórmula de Taylor probabilística, obteniendo asimismo estimaciones precisas para los momentos centrales pares. Seguidamente, mediante las medias de Steklov, analizamos la velocidad de convergencia respecto al segundo módulo de continuidad. Para funciones continuas, la velocidad de convergencia de los polinomios de Bernstein se puede caracterizar en términos del segundo módulo de continuidad de Ditzian-Totik.

Finalmente, en el capítulo tercero, presentamos variantes estocásticas de los polinomios clásicos de Bernstein asociados a una función continua f, creadas a partir de un arreglo triangular de nodos aleatorios

en el intervalo $[0,1]$. Analizamos la convergencia uniforme en probabilidad del proceso de aproximación que representan, proporcionando además las velocidades de convergencia. Cuando los nodos aleatorios son los estadísticos ordenados de una muestra de variables aleatorias uniformes en $[0,1]$, se dará una respuesta positiva a una conjetura planteada por Wu y Zhou, sobre una tasa exponencial de convergencia en probabilidad. Este análisis no solo confirma la conjetura, sino también ilustra el potencial de los métodos estocásticos en la teoría de aproximación.

Este trabajo pretende así ofrecer una visión integral de los polinomios de Bernstein, mostrando tanto sus aplicaciones deterministas como sus versiones estocásticas, y proporcionando herramientas útiles para futuras investigaciones en el campo de la aproximación de funciones.

Abstract

The main subject of this work is the study of both deterministic and random Bernstein polynomials. Positive linear operators constitute one of the main tools for approximating functions. One of the most important examples are Bernstein polynomials, which are defined as

$$B_n f(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-k)^{n-k}, \quad x \in [0,1], n \geq 1,$$

where $f : [0,1] \to \mathbb{R}$ is an arbitrary function. In the early 20th century, Russian mathematician S. Bernstein introduced probabilistic methods to deal with functions approximation. He introduced the polynomials that bear his name. Bernstein polynomials have numerous applications, in economics, to model markets; in forestry engineering, to replant trees; for calculating satellite orbits; in medicine, to model heart rhythms. However, one of the most significant applications of Bernstein polynomials in mathematics is that they provide a construcitive proof of the Weierstrass approximation theorem, which states that any continuous function on a compact interval can be uniformly approximated by polynomials.

This work is organized as follows. The first chapter covers positive linear operators and their probabilistic representations. Bernstein polynomials constitute a paradigmatic example for applying probabilistic methods in approximation using positive linear operators, thus directly linking them to probability. Although the main subject of this work is the analysis of Bernstein polynomials, both deterministic and random forms, also it is going to demonstrate that these polynomials are a paradigmatic example of approximation linear operators and to extend the probabilistic representation of Bernstein polynomials through empirical processes to other usual linear operators. This unifies the demonstrations of key properties such as the preservation of monotonicity, φ-variation, global smoothness, and Lipschitz constants. We analyse the probabilistic representations of some positive linear operators in terms of various double-index stochastic processes, including empirical processes, Poisson processes, and Gamma processes. Probabilistic methods are often more effective if the family Z contains random variables with strong dependence relationships.

In the second chapter, we study the convergence rate of Bernstein polynomials concerning different modules of continuity. To f functions with various degrees of differentiability, we obtain the convergence rate of Bernstein polynomials with respect to the first modulus of continuity. In fact, bounds can be easily given in terms of the first modulus of continuity, using the probabilistic Taylor formula, obtaining in this way, estimates for the even central moments. Subsequently, using Steklov means, we discuss the rate of convergence related to the second modules of continuity. For continuous functions, the rate of convergence of Bernstein polynomials can be characterized in terms of the second modules of continuity of Ditzian-Totik.

Finally, in the third chapter, we introduce stochastic variants of the classic Bernstein polynomials associated with a continuous function f, built up from a general triangular array of random nodes in the interval $[0,1]$. We discuss the uniform convergence in probability of the approximation process that they represent, providing at the same time convergence rates. In the particular case in which random nodes are the order statistics of a sample of uniform random variables on $[0,1]$, we give a positive answer to a conjecture raised in Wu and Zhou about an exponential rate of convergence in probability. This analysis not only confirms this conjecture, but also illustrates the potential of stochastic methods in approximation theory.

This work thus aims to offer a comprehensive view of Bernstein polynomials, showing both their deterministic applications and stochastic versions and providing useful tools for future research in the field of function approximation.

Índice general

Capítulo 1

Operadores lineales positivos y sus representaciones probabilísticas

Aún cuando el objeto central del trabajo es el estudio de los polinomios de Bernstein, tanto deterministas como aleatorios, en este capítulo se persigue un doble objetivo: en primer lugar, mostrar que los polinomios de Bernstein son un ejemplo paradigmático de los operadores lineales de aproximación. En segundo lugar, probar que la representación probabilística de los polinomios de Bernstein en términos de procesos empíricos puede extenderse a otros operadores lineales usuales, dando lugar a demostraciones unificadas de sus propiedades de preservación, tales como la monotonía, φ-variación, suavidad global y constantes de Lipschitz.

1. Operadores lineales positivos

A lo largo del trabajo, denotamos $\mathbb{N}$ al conjunto de enteros positivos y suponemos que $n \in \mathbb{N}$. Denotamos mediante I a un intervalo de la recta real, habitualmente, I=[0,1], I=[0,∞) o I=$\mathbb{R}$. El conjunto de funciones reales, medibles y acotadas definidas en I, se denota como $M_b(I)$, mientras que $C_b(I) \subseteq M_b(I)$ denota la clase de funciones continuas y acotadas. Finalmente 1_A denota el indicador del conjunto A.

Definición 1.1. Un operador lineal positivo L es una aplicación lineal

$$L : M_b(I) \rightarrow \ M_b(I)$$

$$f \rightarrow \ Lf,$$

tal que $Lf(x) \geq 0$, si $f(x) \geq 0$, $x \in I$.

Definición 1.2. Una sucesión $(L_n)_{n\geq 1}$ de operadores lineales positivos se llama sucesión de aproximación si

$$L_n f(x) \underset{n\to\infty}{\longrightarrow} f(x), \quad x \in I, \ \ f \in C_b(I).$$

En muchos casos concretos, la condición de acotación en las definiciones anteriores puede suprimirse. Veamos algunos ejemplos.

Ejemplo 1. Operadores de sumación:

- *Polinomios de Bernstein*. Para cada $x \in [0,1]$, se define

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

- *Operadores de Bernstein-Kantorovich*. Para n≥1, m≥0 y $x \in [0,1]$, se define

$$K_{n,m} f(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-k)^{n-k} \int_0^1 \dots \int_0^1 f\left(\frac{k+u_1+\dots+u_m}{n+m}\right) du_1 \dots du_m.$$

- *Operadores de Szász-Mirakyan*. Para cada x≥0, se define

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

Ejemplo 2. Operadores integrales:

- *Operadores Gamma*. Para cada x≥0, se define

$$G_n f(x) := \frac{1}{\Gamma(n)} \int_0^{\infty} f\left(\frac{x\theta}{n}\right) \theta^{n-1} e^{-\theta} d\theta.$$

- *Operadores de Weierstrass*. Para cada $x \in \mathbb{R}$, se define

$$W_n f(x) := \frac{1}{\sqrt{2\pi n}} \int_{-\infty}^{\infty} f\left(x + \frac{\theta}{n}\right) e^{-\theta^2/(2n)} d\theta.$$

Como se ve en los ejemplos anteriores, $L_n f$ es una función suave, aún cuando f no lo sea y por ello, $L_n f$ es un aproximante suave de f. Por ejemplo, $B_n f$ es indefinidamente derivable para cualquier función f definida en [0,1]. Análogamente, $W_n f$ es indefinidamente derivable para cualquier $f \in M_b(\mathbb{R})$.

2. Representaciones probabilísticas

Muchas propiedades de las sucesiones $(L_n)_{n\geq 1}$, tanto de preservación como de aproximación, pueden probarse de manera unificada mediante representaciones probabilísticas adecuadas. Más concretamente, supongamos que $\mathbb{Z} = (Z_n(x), x \in I, n \in \mathbb{N})$ es una familia de variables aleatorias a doble índice con valores en I, definidas en el mismo espacio de probabilidad completo (Ω, F, P). Supongamos que

$$L_n f(x) = Ef(Z_n(x)), \quad x \in I, \ \ f \in M_b(I), \tag{2}$$

donde E denota la esperanza matemática. Observemos que la representación (2) implica que

$$L_n e_o(x) = Ee_o(Z_n(x)) = 1, \quad x \in I, \quad e_o(y) = 1, \ \ y \in I.$$

Esta condición de normalización se satisface en la mayoría de ejemplos habituales. Por tanto, las propiedades estocásticas de la familia $\mathbb{Z}$ permiten probar propiedades de la sucesión $(L_n)_{n\geq 1}$. Por ejemplo, supongamos que

$$Z_n(x) \underset{n\to\infty}{\overset{(\mathscr{L})}{\longrightarrow}} x,$$

donde $(\mathscr{L})$ denota convergencia en ley o en distribución. El Teorema de Helly-Bray permite así afirmar que

$$L_n f(x) = Ef(Z_n(x)) \underset{n\to\infty}{\longrightarrow} f(x), \quad x \in I, \quad f \in M_b(I).$$

Dicho de otra forma, $(L_n)_{n\geq 1}$ es una sucesión de aproximación.

Observemos que la representación dada en (2) no es única. En este sentido, escribimos $X \overset{(\mathscr{L})}{=} Y$ si las variables aleatorias X e Y (posiblemente definidas en espacios de probabilidad distintos) tienen la misma ley o función de distribución. Si $\widetilde{\mathbb{Z}} = (\widetilde{Z}_n(x),\, x \in I,\, n \in \mathbb{N})$ es una familia de variables aleatorias tal que $\widetilde{Z}_n(x) \overset{(\mathscr{L})}{=} Z_n(x)$, $x \in I$, $n \in \mathbb{N}$, se tiene obviamente:

$$L_n f(x) = Ef(Z_n(x)) = Ef(\widetilde{Z}_n(x)), \quad x \in I, \ \ n \in \mathbb{N}, \ \ f \in M_b(I).$$

Los métodos probabilísticos suelen ser más eficaces si la familia $\mathbb{Z}$ contiene variables aleatorias con fuertes relaciones de dependencia. Para representar los operadores vistos anteriormente, consideremos los siguientes procesos estocásticos a doble índice.

(a) **Proceso empírico**. $I=[0,1]$. Sea $(U_k)_{k\geq 1}$ una sucesión de copias independientes de una variable aleatoria U con distribución uniforme en [0,1] y sea $x\in[0,1]$. Definimos:

$$S_n(x)=\sum_{k=1}^{n} 1_{[0,x]}(U_k). \tag{3}$$

Claramente, $S_n(x)$ tiene la distribución binomial de parámetros n y x, es decir,

$$P(S_n(x)=k)=\binom{n}{k}x^k(1-x)^{n-k}. \tag{4}$$

Por tanto, se tiene

$$B_nf(x)=Ef\left(\frac{S_n(x)}{n}\right), \quad f\in M_b([0,1]), \tag{5}$$

que es una representación de los polinimios de Bernstein en términos del proceso empírico $S_n(x)/n$.
Por otra parte, sea $(V_k)_{k\geq 1}$ una sucesión de copias independientes de una variable aleatoria V con ley uniforme en [0,1]. Supongamos que $(U_k)_{k\geq 1}$ y $(V_k)_{k\geq 1}$ son independientes. Obtenemos así, la siguiente representación de los operadores de Bernstein-Kantorovich:

$$K_{n,m}(x)=Ef\left(\frac{S_n(x)+V_1+\ldots+V_m}{n+m}\right), \quad f\in M_b([0,1]). \tag{6}$$

(b) **Proceso de Poisson**. Un proceso $(N_t)_{t\geq 0}$ es un *proceso de Poisson*, si es un proceso con incrementos estacionarios independientes tal que, $N_0=0$, $c.s$ y para cualesquiera $0\leq s<t$, $N_t-N_s\equiv\mathscr{P}(t-s)$, denotando por $\mathscr{P}$ la distribución de Poisson. Es decir,

$$P(N_t-N_s=k)=P(N_{t-s}=k)=\frac{(t-s)^k}{k!}e^{-(t-s)}, \quad k=0,1,\ldots, \quad 0\leq s<t.$$

Una forma de construir dicho proceso es la siguiente. Supongamos que $(W_k)_{k\geq 1}$ es una sucesión de copias independientes de una variable aleatoria $W_1\equiv exp(1)$. Para cada $k\geq 1$, definimos $T_k=W_1+\ldots+W_k$ ($T_0=0$). Entonces, el proceso

$$N_t=\text{máx}\{n:T_n\leq t\}=\sum_{k=1}^{\infty}\mathbf{1}_{[0,t]}(T_k), \quad t\geq 0, \tag{7}$$

es un proceso de Poisson. Se tiene entonces que

$$S_nf(x)=Ef\left(\frac{N_{nx}}{n}\right), \quad x\geq 0, \;\; f\in M_b([0,\infty)), \tag{8}$$

que es una representación de los operadores de Szász-Mirakyan.

(c) **Proceso Gamma**. Un proceso $(V_t)_{t\geq 0}$ es un *proceso Gamma*, si es un proceso con incrementos estacionarios independientes tal que, $V_0=0$, $c.s$ y para cualesquiera $0\leq s<t$, $V_t-V_s\equiv Gamma(t-s,1)$, cuya función de densidad es

$$f_{t-s}(\theta)=\frac{\theta^{t-s-1}e^{-\theta}}{\Gamma(t-s)}, \quad \theta\geq 0,$$

siendo Γ la función Gamma de Euler.
Notar que $V_t\equiv(Gamma(t,1))$. Por tanto, obtenemos la siguiente representación de los operadores Gamma

$$G_nf(x)=Ef\left(\frac{xV_n}{n}\right), \quad x\geq 0, \;\; f\in M_b([0,\infty)). \tag{9}$$

Finalmente, podemos representar los operadores de Weierstarss de la siguiente manera

$$W_nf(x)=Ef\left(x+\frac{Z}{\sqrt{n}}\right), \quad x\in\mathbb{R}, \;\; f\in M_b(\mathbb{R}), \tag{10}$$

donde Z tiene la ley normal estándar, es decir, Z≡N(0,1).

3. Preservación de monotonía y φ-variación.

Salvo que se diga lo contrario, en esta sección supondremos que $n \geq 1$ es fijo, y denotaremos

$$Z(x) := Z_n(x), \quad Lf(x) := Ef(Z(x)), \quad x \in I.$$

Asimismo, denotamos mediante

$$\Phi = \{\varphi : [0,\infty) \to \mathbb{R} : \varphi \text{ es convexa, estrictamente creciente y } \varphi(0) = 0\}.$$

Definición 3.1. Sea $f \in M_b(I)$. Para cada $\varphi \in \Phi$, se llama φ-*variación* de f a la función

$$V_\varphi f(\varepsilon) := sup \sum_{i=1}^{n} \varphi(|f(x_i) - f(x_{i-1})|), \quad 0 < \varepsilon \leq \infty,$$

donde el supremo se toma sobre el conjunto de sucesiones finitas $x_0 \leq x_1 \leq \ldots \leq x_n$, tales que x_0, $x_n \in I$ y $x_i - x_{i-1} \leq \varepsilon$, $i = 1, \ldots, n$.
La cantidad (posiblemente infinita) $V_\varphi f := V_\varphi f(\infty)$ se llama φ-*variación total* de f.

Los ejemplos más comunes suelen ser la 1-variación o variación usual, con $\varphi(x) = x$, y la p-variación, cuando $\varphi(x) = x^p$, con $p \geq 1$.

La preservación de la monotonía tiene un significado evidente, mientras que la preservación de la φ-variación se puede interpretar diciendo que los aproximantes tienen menores fluctuaciones que la función original.

Teorema 3.2. *Supongamos que, para cualesquiera $x, y \in I$ con $x \leq y$, se satisface*

$$Z(x) \leq Z(y), \quad c.s. \tag{11}$$

Entonces Lf tiene la misma monotonía que f. Si además, $\varphi \in \Phi$ y $V_\varphi(f) < \infty$, entonces

$$V_\varphi(Lf) \leq V_\varphi f.$$

Demostración. Sean $x, y \in I$ con $x \leq y$, y supongamos que f no es decreciente. Entonces (11) implica que $f(Z(x)) \leq f(Z(y))$ c.s, y tomando esperanzas a ambos lados se tiene

$$Lf(x) = Ef(Z(x)) \leq Ef(Z(y)) = Lf(y).$$

Si f es no creciente, considerar $-f$, que es no decreciente, y aplicar lo anterior. Para la segunda parte, sean $x_0 \leq x_1 \leq \ldots \leq x_n$ puntos tales que $x_0, x_n \in I$. Por la hipótesis (11), se tiene

$$Z(x_0) \leq Z(x_1) \leq \ldots \leq Z(x_n), \quad c.s.$$

Utilizando sucesivamente que φ es creciente, la desigualdad de Jensen (por ser φ convexa), la linealidad de la esperanza, y la definición de variación total, se tiene

$$\begin{aligned}\sum_{i=1}^{n} \varphi(|Lf(x_i) - Lf(x_{i-1})|) &= \sum_{i=1}^{n} \varphi(|Ef(Z(x_i)) - Ef(Z(x_{i-1}))|) \leq \sum_{i=1}^{n} \varphi(E|f(Z(x_i)) - Ef(Z(x_{i-1}))|) \\ &\leq E\left(\sum_{i=1}^{n} \varphi(|f(Z(x_i)) - f(Z(x_{i-1}))|)\right) \leq EV_\varphi f = V_\varphi f.\end{aligned}$$

Tomando supremos sobre el conjunto de sucesiones finitas de puntos como los anteriores, se tiene el resultado. □

Corolario 3.3. *Los polinomios de Bernstein y los operadores de Bernstein-Kantorovich, Szász-Mirakyan, Gamma y Weierstrass preservan la monotonía y la φ-variación total.*

Demostración. Para los polinimoios de Bernstein, el resultado se sigue de la representación (5) y del Teorema 3.2, teniendo en cuenta que, por construcción, $S_n(x) \leq S_n(y)$, $0 \leq x \leq y \leq 1$. La argumentación para los operadores de Bernstein-Kantorovich es análoga, teniendo en cuenta la representación (6). En cuanto a los operadores de Szász-Mirakyan, el resultado se sigue análogamente, teniendo en cuenta (8) y el hecho de que $N_{nx} \leq N_{ny}$ si $0 \leq x \leq y$, tal como se sigue en la construcción (7). Finalmente, para los operadores Gamma y Weierstrass, basta tener en cuenta las representaciones (9) y (10), y las desigualdades

$$\frac{xV_n}{n} \leq \frac{yV_n}{n}, \quad 0 \leq x \leq y, \quad x + \frac{Z}{\sqrt{n}} \leq y + \frac{Z}{\sqrt{n}}, \quad x \leq y.$$

□

4. Preservación de módulos de continuidad y constantes de Lipschitz

Recordemos la definición del módulo de continuidad usual y sus principales propiedades.

Definición 4.1. Sea $I \subseteq \mathbb{R}$ un intervalo y $f : I \to \mathbb{R}$ una función. Se llama módulo de continuidad de f a la función $\omega(f;\cdot) : [0,\infty) \to [0,\infty)$ dada por

$$\omega(f;h) = sup\{|f(x) - f(y)| : x,y \in I, |x-y| \leq h\}, \quad h \geq 0.$$

Se dice que una función $g : [0,\infty) \to \mathbb{R}$ es subaditiva si

$$g(h+\delta) \leq g(h) + g(\delta), \quad h,\delta \geq 0.$$

Lema 4.2. *El módulo de continuidad satisface las siguientes propiedades:*

(a) $\omega(f;\cdot)$ *es no decreciente, subaditiva y tal que* $\omega(f;0) = 0$. *Además,*

$$|f(x) - f(y)| \leq \omega(f;|y-x|), \quad [x,y] \subseteq I.$$

(b) $\omega(f;ah) \leq (1+a)\omega(f;h), \quad a,h \geq 0.$

Demostración. (a) Sean $h,h' \geq 0$ y $x,y \in I$ con $|x-y| \leq h+h'$. Sea u un punto intermedio entre x e y con $|x-u| \leq h$ y $|u-y| \leq h'$. Entonces

$$|f(x) - f(y)| \leq |f(x) - f(u)| + |f(u) - f(y)| \leq \omega(f;h) + \omega(f;h').$$

Tomando supremos obtenemos que $\omega(f;h+h') \leq \omega(f;h) + \omega(f;h')$.

(b) Denotamos mediante:

$$\lceil a \rceil = \text{ínf}\{k \in \mathbb{Z} : k \geq a\}, \quad a \in \mathbb{R}.$$

Notar que $a \leq \lceil a \rceil \leq a+1$, para $a \in \mathbb{R}$. Ahora, tenemos que si g es aditiva, entonces $g(nh) \leq ng(h), n \in \mathbb{N}, h \geq 0$.Por tanto, obtenemos

$$\omega(f;ah) \leq \omega(f;\lceil a \rceil h) \leq \lceil a \rceil \omega(f;h) \leq (a+1)\omega(f;h),$$

donde la segunda desigualdad se debe a que si g es aditiva, entonces $g(nh) \leq ng(h)$, $n \in \mathbb{N}$, $h \geq 0$.

□

Definición 4.3. Sean $A > 0, \alpha \in (0,1]$. Diremos que f es de clase Lipschitz con constantes A, α si

$$|f(x) - f(y)| \leq A|x-y|^{\alpha}, \quad x,y \in I.$$

Usaremos la notación $f \in Lip(A,\alpha)$. Notar que si $f \in Lip(A,\alpha)$, entonces f es continua, y además $\omega(f;\delta) \leq A\delta^{\alpha}$.

Las propiedades del siguiente teorema se conocen habitualmente como *preservación de la suavidad global*. La preservación de las constantes de Lipschitz y del módulo de continuidad significan que, salvo un factor de proporcionalidad, los aproximantes no presentan variaciones grandes entre dos puntos cercanos si la función original se comporta así. Es decir, respetan la suavidad de la función original.

Teorema 4.4. *Supongamos que, para cualquier $x \in I$, se satisface*

$$EZ(x) = cx + d, \tag{12}$$

donde c y d son constantes, con $c > 0$. Entonces se verifica

(a) Si $f \in Lip(A, \alpha)$, entonces $Lf \in Lip(Ac^{\alpha}, \alpha)$.

(b) $\omega(Lf;\delta) \leq (1+c)\omega(f;\delta), \quad \delta \geq 0$.

(c) Si f es continua en I y $\omega(f;\cdot)$ es una función cóncava, entonces

$$\omega(Lf;\delta) \leq \omega(f;c\delta), \quad \delta \geq 0.$$

Demostración. Sean $x \leq y$ en I, y supongamos $f \in Lip(A.\alpha)$. Notar que la función x^{α}, $x \geq 0$ es cóncava, ya que $\alpha \in (0,1]$. Aplicando las hipótesis, la condición de Lipschitz y la desigualdad de Jensen, tenemos que

$$|Lf(y) - Lf(x)| \leq E|f(Z(y)) - f(Z(x))| \leq AE(Z(y) - Z(x))^{\alpha} \leq AE^{\alpha}(Z(y) - Z(x)) = Ac^{\alpha}(y-x)^{\alpha}.$$

Para probar (b) y (c), sean $\delta > 0$, $x, y \in I$ con $0 \leq y - x \leq \delta$. Entonces, se sigue del Lema 4.2(a) que

$$|Lf(y) - Lf(x)| \leq E|f(Z(y)) - f(Z(x))| \leq E\omega(f;Z(y) - Z(x)). \tag{13}$$

Si $\delta = 0$ la propiedad (d) es trivial. Supongamos pues que $\delta > 0$. Utilizando el Lema 4.2(b), se tiene

$$|Lf(y) - Lf(x)| \leq E\left(1 + \frac{Z(y) - Z(x)}{\delta}\right)\omega(f;\delta) = \left(1 + \frac{c(y-x)}{\delta}\right)\omega(f;\delta) \leq (1+c)\omega(f;\delta).$$

Tomando supremos en ambos miembros, se tiene (b). Finalmente, si $\omega(f,\cdot)$ es cóncavo, podemos utilizar la desigualdad de Jensen en (13) y obtenemos que

$$|Lf(y) - Lf(x)| \leq \omega(f;EZ(y) - E(Z(x)) \leq \omega(f, c(y-x)) \leq \omega(f;c\delta).$$

Tomando de nuevo supremos en ambos lados, se tiene (c). □

Corolario 4.5. *Los polinomios de Bernstein y los operadores de Bernstein-Kantorovich, Szász-Mirakyan, Gamma y Weierstrass preservan la suavidad global.*

Demostración. Para los polinomios de Bernstein (recuérdese (5)), se tiene

$$E\left(\frac{S_n(x)}{n}\right) = x.$$

De acuerdo con (6), para los operadores de Bernstein-Kantorovich, se tiene

$$E\left(\frac{S_n(x) + V_1 + \cdots + V_m}{n+m}\right) = \frac{1}{n+m}\left(nx + \frac{m}{2}\right) = \frac{n}{n+m}x + \frac{m}{2(n+m)}.$$

Recordando (8), (9) y (10) para los operadores de Szász-Minakyan, Gamma y Weierstrass se tiene, respectivamente, que

$$E\left(\frac{N_{nx}}{n}\right) = x = E\left(\frac{xV_n}{n}\right) = E\left(x + \frac{Z}{\sqrt{n}}\right).$$

Vemos pues que en todos los casos se satisface la condición del Teorema 4.4.

□

Capítulo 2

Velocidades de convergencia de los polinomios de Bernstein

En este capítulo, estudiaremos la velocidad de convergencia de los polinomios de Bernstein en el intervalo $[0,1]$ respecto a distintos módulos de continuidad. Veremos que se pueden dar cotas fácilmente en términos del primer módulo de continuidad, usando la fórmula de Taylor probabilística, y daremos estimaciones precisas para los momentos centrales pares. Demostraremos un teorema de velocidad de convergencia en términos del segundo módulo de continuidad usual, mediante una aproximación intermedia por medias de Steklov. Probaremos también que, para funciones de clase $\mathscr{C}^{(2)}[0,1]$, se puede medir la velocidad de convergencia respecto al segundo módulo de continuidad de Ditzian-Totik. Finalmente, mencionaremos, sin demostración, pues se utilizan técnicas muy avanzadas en ella, el mejor teorema conocido hasta ahora de velocidad de convergencia de los polinomios de Bernstein, en términos del segundo módulo de continuidad de Ditzian-Totik.

1. Velocidad en términos del primer módulo de continuidad

A lo largo del trabajo, denotaremos

$$\|f\| = sup\{|f(x)| : x \in [0,1]\}.$$

Para cada $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, denotaremos β_m a una variable aleatoria absolutamente continua con densidad $\rho(\theta) = m(1-\theta)^{m-1}, \theta \in [0,1]$, si $m \in \mathbb{N}$ ($\beta_0 = 1$). Denotamos asimismo

$$\mathscr{C}^{(m)}(I) = \{f : I \to \mathbb{R}; f^{(m)} \text{existe y es uniformemente continua en } I\},$$

siendo $I \subset \mathbb{R}$ un intervalo arbitrario. Denotaremos mediante $\mathring{I}$ al interior de I.

Teorema 1.1. ***(Fórmula de Taylor probabilística)*** *Sea $m \in \mathbb{N}_0, f \in \mathscr{C}^{(m)}(I)$ y X una variable aleatoria independiente de β_m que toma valores en I, con $E|X|^{m+2} < \infty$. Para cada $x \in \mathring{I}$ y $\delta > 0$, se tiene*

$$\left| Ef(X) - \sum_{j=0}^{m} \frac{f^{(j)}(x)}{j!} E(X-x)^j \right| \leq \frac{1}{m!} \left(E|X-x|^m + \frac{2E|X-x|^{m+2}}{\delta^2(m+1)(m+2)} \right) \omega(f^{(m)}; \delta).$$

Demostración. Podemos escribir la formula de Taylor de orden m de f con resto integral de la siguiente forma

$$f(y) - \sum_{j=0}^{m} \frac{f^{(j)}(x)}{j!}(y-x)^j = \frac{(y-x)^m}{m!} E(f^{(m)}(x+(y-x)\beta_m) - f^{(m)}(x)), \quad y \in I.$$

Gracias a la subdatividad del módulo de continuidad vista en el Lema 4.2 del Capítulo 1, tenemos que para cada $y \in I, \delta > 0$,

$$\left| f^{(m)}(x+(y-x)\beta_m) - f^{(m)}(x) \right| \leq \omega\big(f^{(m)}; |y-x|\beta_m\big) \leq \left(1 + \frac{|y-x|^2 {\beta_m}^2}{\delta^2}\right) \omega\big(f^{(m)}; \delta\big).$$

Ahora, sustituimos y por X, que es independiente de β_m y toma valores en I. Tomando esperanzas, se obtiene que

$$\left|Ef(X)-\sum_{j=0}^{m}\frac{f^{(j)}(x)}{j!}E(X-x)^j\right| \leq \frac{1}{m!}E\left|(X-x)^m\left(f^{(m)}(x+(X-x)\beta_m)-f^{(m)}(x)\right)\right|$$
$$\leq \frac{1}{m!}E\left[|X-x|^m\left(1+\frac{|X-x|^2\beta_m^2}{\delta^2}\right)\right]\omega\left(f^{(m)};\delta\right).$$

Puesto que β_m es independiente de X, se tiene

$$E\left(|X-x|^{m+2}\beta_m^2\right)=E|X-x|^{m+2}E\beta_m^2=\frac{2}{(m+1)(m+2)}E|X-x|^{m+2},$$

completando así la demostración. □

Para cada $n,j\in\mathbb{N}_0$, denotaremos

$$\mu_{n,j}(x)=E\left(\frac{S_n(x)}{n}-x\right)^j,\quad M_{n,j}(x)=E\left|\frac{S_n(x)}{n}-x\right|^j$$
$$\nu_{n,j}(x)=n^j\mu_{n,j}(x)=E(S_n(x)-nx)^j. \tag{14}$$

Como consecuencias del Teorema 1.1, damos el siguiente resultado de aproximación para los polinomios de Bernstein.

Corolario 1.2. *Sea $m\in\mathbb{N}_0$ y $f\in\mathscr{C}^{(m)}[0,1]$. Para cada $x\in(0,1)$, se tiene*

$$\left|B_nf(x)-\sum_{j=0}^{m}\frac{f^{(j)}(x)}{j!}\mu_{n,j}(x)\right|\leq\frac{1}{m!}\left(M_{n,m}(x)+\frac{2n}{(m+1)(m+2)}M_{n,m+2}(x)\right)\omega\left(f^{(m)};\frac{1}{\sqrt{n}}\right).$$

Demostración. Basta aplicar el Teorema 1.1 para $I=[0,1]$, $X=S_n(x)/n$ y $\delta=1/\sqrt{n}$. □

Sobre este resultado, hacemos las siguientes observaciones.

En primer lugar, se trata de una estimación puntual para cada $x\in(0,1)$. Una estimación uniforme se dará en el Teorema 1.8, una vez se hayan estimado uniformemente los momentos $M_{n,m}(x)$.

En segundo lugar, se conocen las expresiones explícitas de los primeros momentos $\mu_{n,j}(x)$ (veáse [1,(14)]). Denotando con $\sigma^2(x)=x(1-x)$, se tiene

$$\mu_{n,2}(x)=\frac{\sigma^2(x)}{n},\quad \mu_{n,3}(x)=\frac{(1-2x)\sigma^2(x)}{n^2} \tag{15}$$

$$\mu_{n,4}(x)=\frac{3\sigma^4(x)}{n^2}+\frac{\sigma^2(x)(1-6\sigma^2(x))}{n^3} \tag{16}$$

y

$$\mu_{n,5}(x)=(1-2x)\left(\frac{10\sigma^4(x)}{n^3}+\frac{\sigma^2(x)(1-12\sigma^2(x))}{n^4}\right). \tag{17}$$

De acuerdo con el Teorema Central del Límite, se tiene para cada $m\in\mathbb{N}$

$$\mu_{n,m}(x)=\left(\frac{x(1-x)}{n}\right)^{m/2}E\left(\frac{S_n(x)-nx}{\sqrt{nx(1-x)}}\right)^m\sim\left(\frac{x(1-x)}{n}\right)^{m/2}EZ^m,\quad n\to\infty, \tag{18}$$

así como

$$M_{n,m}(x)=\left(\frac{x(1-x)}{n}\right)^{m/2}E\left|\frac{S_n(x)-nx}{\sqrt{nx(1-x)}}\right|^m\sim\left(\frac{x(1-x)}{n}\right)^{m/2}E|Z|^m,\quad n\to\infty, \tag{19}$$

donde Z es una variable aleatoria con distribución normal estándar.

De (15) y (16) se sigue que el orden de magnitud de $\mu_{n,3}(x)$ y $\mu_{n,4}(x)$ es el mismo y vale $1/n^2$, cuando $n \to \infty$. Esto no es casual. Si m es impar, entonces $EZ^m = 0$ y puede probarase a partir de (18) y (19) lo siguiente: Si $m \geq 3$ es un impar, entonces

$$\mu_{n,m}(x) \sim M_{n,m+1}(x) \sim \frac{1}{n^{(m+1)/2}}, \quad n \to \infty.$$

Esto implica que el Corolario 1.2 no tiene sentido si $m \geq 3$ es un impar, puesto que en tal caso tendríamos

$$\mu_{n,m}(x) \sim \frac{1}{n^{(m+1)/2}}, \quad M_{n,m}(x) \sim \frac{1}{n^{m/2}}, \quad n \to \infty.$$

Es decir, la cota en el Corolario 1.2 tendría un orden de magnitud menor o igual que el orden de magnitud del término principal. De ahí que lo adecuado es escoger un m par, tal como se hará en el Teorema 1.7.

Corolario 1.3. *Sea $f \in \mathscr{C}[0,1]$. Entonces,*

$$\|B_n f - f\| \leq \frac{5}{4}\omega\left(f; \frac{1}{\sqrt{n}}\right).$$

Demostración. Aplicando el Corolario 1.2 para $m = 0$, obtenemos

$$\|B_n f(x) - f(x)\| \leq (1 + x(1-x))\omega\left(f; \frac{1}{\sqrt{n}}\right) \leq \frac{5}{4}\omega\left(f; \frac{1}{\sqrt{n}}\right),$$

ya que $M_{n,0}(x) = 1$ y $M_{n,2}(x) = \frac{x(1-x)}{n}$, como se sigue de (15). Basta entonces tomar supremos en $x \in [0,1]$. $\square$

Se puede ver una demostración analítica de este resultado en [6].

Corolario 1.4. ***(Fórmula de Voronovskaja)*** *Sea $f \in \mathscr{C}^{(2)}[0,1]$. Para cada $x \in (0,1)$, se tiene*

$$\left| B_n f(x) - f(x) - \frac{\sigma^2(x)}{2n} f''(x) \right| \leq \frac{\sigma^2(x)}{2n}\left(1 + \frac{\sigma^2(x)}{2} + \frac{1 - 6\sigma^2(x)}{6n}\right)\omega\left(f''; \frac{1}{\sqrt{n}}\right).$$

En consecuencia,

$$\|B_n f(x) - f(x) - \frac{\sigma^2(x)}{2n} f''(x)\| \leq \frac{1}{16n}\left(\frac{9}{4} + \frac{1}{3n}\right)\omega\left(f''; \frac{1}{\sqrt{n}}\right).$$

Demostración. La primera desigualdad se sigue del Corolario 1.2 con $m = 2$ y de las expresiones de $\mu_{n,2}(x)$ y $\mu_{n,4}(x)$ dadas en (15) y (16). La segunda desigualdad se sigue tomando supremos en $x \in [0,1]$, teniendo en cuenta que $\sigma^2(x) \leq 1/4$. $\square$

Nota. Se pueden dar cotas similares a los Corolarios 1.2, 1.3 y 1.4 para otros operadores tomando $X = Z_n(x)$, el proceso correspondiente al operador (veáse la sección 2 del capítulo 1). Sin embargo, cuando el intervalo I no es compacto, los momentos análogos a $\mu_{n,j}(x), M_{n,j}(x)$ no están acotados en I, en general, por lo que solo se pueden dar cotas uniformes en subintervalos compactos.

A continuación, introduciremos dos lemas en los que nos apoyaremos para obtener unas estimaciones precisas de los momentos centrales pares.

Lema 1.5. *Sea $n \in \mathbb{N}$ y $x \in [0,1]$. Para cualquier $j \in \mathbb{N}$, con $j \geq 2$, tenemos que*

$$v_{n+1,j}(x) = v_{n,j}(x) + x(1-x)\sum_{i=0}^{j-2}\binom{j}{i} v_{n,i}(x) p_{j-1-i}(x),$$

donde $v_{n,j}(x)$ está definido en (14) y $p_k(x) = (1-x)^k - (-x)^k, \quad k \in \mathbb{N}_0$.

Demostración. Se sigue de (3) que

$$S_{n+1}(x)-(n+1)x = S_n(x)-nx+Y_{n+1}(x), \quad Y_{n+1}(x) := 1_{[0,x]}(U_{n+1})-x.$$

También tenemos que

$$E(Y_{n+1}(x))^k = (1-x)^k x + (-x)^k(1-x) = x(1-x)p_{k-1}(x), \quad k \in \mathbb{N}.$$

Dado que las variables aleatorias $S_n(x)$ e $Y_{n+1}(x)$ son independientes, obtenemos que

$$\begin{aligned} v_{n+1,j}(x) &= E(S_n(x)-nx+Y_{n+1}(x))^j = \sum_{i=0}^{j}\binom{j}{i}E(S_n(x)-nx)^i E(Y_{n+1}(x))^{j-i} \\ &= v_{n,j}(x) + x(1-x)\sum_{i=0}^{j-2}\binom{j}{i}v_{n,i}(x)p_{j-1-i}(x), \end{aligned}$$

obteniendo de esta forma el resultado que buscábamos. □

Sea ahora $(V_k)_{k\geq 1}$ una copia independiente de $(U_k)_{k\geq 1}$ como se indica en (3). Denotamos

$$T_n(x) = \sum_{k=1}^{n} 1_{[0,x]}(V_k), \quad n \in \mathbb{N}, \quad T_0(x) = 0, \quad x \in [0,1].$$

Observamos que las familias de variables aleatorias $\big(S_n(x),\ 0 \leq x \leq 1,\ n \in \mathbb{N}_0\big)$ y $\big(T_n(x),\ 0 \leq x \leq 1,$ $n \in \mathbb{N}_0\big)$ son independientes y $S_n(x) \overset{(\mathscr{L})}{=} T_n(x)$, $n \in \mathbb{N}_0$, $x \in [0,1]$.

Lema 1.6. *Sean $x, y \in [0,1]$. Para cualquier $n \in \mathbb{N}_0$, tenemos*

$$S_{T_n(y)}(x) \overset{(\mathscr{L})}{=} S_n(xy).$$

Demostración. Utilizando la independencia entre las dos familias de variables aleatorias consideradas y recordando (4), tenemos

$$\begin{aligned} P(S_{T_n(y)}(x)=k) &= \sum_{m=k}^{n} P(S_m(x)=k, T_n(y)=m) = \sum_{m=k}^{n}\binom{m}{k}x^k(1-x)^{m-k}\binom{n}{m}y^m(1-y)^{n-m} \\ &= \binom{n}{k}(xy)^k \sum_{m=k}^{n}\binom{n-k}{m-k}((1-x)y)^{m-k}(1-y)^{n-m} \\ &= P(S_n(xy)=k), \end{aligned}$$

para cualquier $k = 0, 1, \ldots n$, lo que prueba el resultado. □

Denotemos mediante

$$\mu_{n,2j} = \sup_{0\leq x\leq 1}\mu_{n,2j}(x), \quad v_{n,2j} = \sup_{0\leq x\leq 1} v_{n,2j}(x), \quad j \in \mathbb{N}_0.$$

Teorema 1.7. *Para cualquier $n \in \mathbb{N}$ y $j \in \mathbb{N}_0$, tenemos que*

$$\mu_{n,2j}\left(\frac{1}{2}\right) \leq \frac{(2j)!}{8^j j! n^j}, \tag{20}$$

$$\mu_{n,2j} \leq \frac{(2j)!}{6^j j! n^j}. \tag{21}$$

Demostración. Sea $n \in \mathbb{N}$ y $x \in [0,1]$. Puesto que el resultado es obvio para $j = 0$, supondremos que $j \in \mathbb{N}$. Por (14), tenemos que la desigualdad (20) es equivalente a

$$v_{n,2j}\left(\frac{1}{2}\right) \leq \frac{(2j)!}{8^j j!} n^j. \tag{22}$$

Esta desigualdad, la cual es trivial para $n = 1$, se mostrará por inducción sobre n. Vemos en (4) que $S_n(x) \overset{(\mathscr{L})}{=} (n - S_n(1-x))$. Esto, junto con (14) implica que

$$v_{n,j}(x) = (-1)^j v_{n,j}(1-x),$$

lo que, a su vez, implica que $v_{n,2j+1}(1/2) = 0$. Por lo tanto, usando el Lema 1.5, y el supuesto de inducción, tenemos que

$$\begin{aligned} v_{n+1,2j}\left(\frac{1}{2}\right) &= v_{n,2j}\left(\frac{1}{2}\right) + \sum_{i=0}^{j-1} \binom{2j}{2i} v_{n,2i}\left(\frac{1}{2}\right) \frac{1}{4^{j-i}} = \sum_{i=0}^{j} \binom{2j}{2i} v_{n,2i}\left(\frac{1}{2}\right) \frac{1}{4^{j-i}} \\ &\leq \frac{(2j)!}{j!4^j} \sum_{i=0}^{j} \binom{j}{i} \frac{(j-i)!}{(2j-2i)!} \left(\frac{n}{2}\right)^i \\ &\leq \frac{(2j)!}{j!4^j} \left(\frac{n+1}{2}\right)^j, \end{aligned}$$

donde la última desigualdad se debe a que

$$\frac{s!}{(2s)!} \leq \frac{1}{2^s}, \quad s \in \mathbb{N}_0.$$

Por tanto, la desigualdad (22) se deriva fácilmente de lo anterior.
Para la desigualdad (21), bastará ver que

$$v_{n,2j}(x) \leq \frac{(2j)!}{j!6^j} n^j. \tag{23}$$

Por (14), sabemos que la función $v_{n,2j}(x)$ es continua en [0,1], simétrica respecto $1/2$ y satisface $v_{n,2j}(0) = v_{n,2j}(1) = 0$. Por lo tanto, $v_{n,2j} = v_{n,2j}(x_0)$ para algún $x_0 \in (0, 1/2]$. Si $x_0 = 1/2$, la desigualdad (23) se sigue de (20). De esta forma, supongamos que $v_{n,2j} = v_{n,2j}(y/2)$, para algún $y \in (0,1)$.
Por el Lema 1.6, podemos escribir

$$S_n\left(\frac{y}{2}\right) - \frac{ny}{2} \overset{(\mathscr{L})}{=} S_{T_n(y)}\left(\frac{1}{2}\right) - \frac{ny}{2} = S_{T_n(y)}\left(\frac{1}{2}\right) - \frac{T_n(y)}{2} + \frac{T_n(y) - ny}{2}.$$

Debido a la independencia entre las variables aleatorias involucradas y el hecho de que $v_{n,2j+1}(1/2) = 0$, obtenemos que

$$\begin{aligned} v_{n,2j} = v_{n,2j}\left(\frac{y}{2}\right) &= E\left(S_{T_n(y)}\left(\frac{1}{2}\right) - \frac{T_n(y)}{2} + \frac{T_n(y) - ny}{2}\right)^{2j} \\ &= \sum_{k=0}^{n} E\left(S_k\left(\frac{1}{2}\right) - \frac{k}{2} + \frac{k - ny}{2}\right)^{2j} P(T_n(y) = k) \\ &= \sum_{k=0}^{n} P(T_n(y) = k) \sum_{i=0}^{2j} \binom{2j}{i} E\left(S_k\left(\frac{1}{2}\right) - \frac{k}{2}\right)^i \left(\frac{k - ny}{2}\right)^{2j-i} \\ &= \sum_{k=0}^{n} P(T_n(y) = k) \sum_{p=0}^{j} \binom{2j}{2p} v_{k,2p}\left(\frac{1}{2}\right) \left(\frac{k - ny}{2}\right)^{2(j-p)}. \end{aligned}$$

Asi tenemos de (22), también válido para $n \in \mathbb{N}_0$, que

$$\begin{aligned} v_{n,2j} &\le \frac{1}{4^j}\sum_{p=0}^{j}\binom{2j}{2p}\frac{(2p)!}{p!}\sum_{k=0}^{n}\left(\frac{k}{2}\right)^p (k-ny)^{2(j-p)}P(T_n(y)=k) \le \frac{1}{4^j}\sum_{p=0}^{j}\binom{2j}{2p}\frac{(2p)!}{p!}\left(\frac{n}{2}\right)^p v_{n,2(j-p)}(y) \\ &\le \frac{1}{4^j}v_{n,2j}+\frac{1}{4^j}\sum_{p=1}^{j}\binom{2j}{2p}\frac{(2p)!}{p!}\left(\frac{n}{2}\right)^p v_{n,2(j-p)}. \end{aligned} \tag{24}$$

Esta última desigualdad nos permite utilizar inducción sobre j para mostrar (23). De hecho, se sigue de (24) que

$$\left(1-\frac{1}{4^j}\right)v_{n,2j} \le \frac{1}{4^j}\frac{(2j)!}{j!}\frac{n^j}{6^j}\sum_{p=1}^{j}\binom{j}{p}3^p = \frac{(2j)!}{j!6^j}n^j\frac{4^j-1}{4^j}.$$

Esto demuestra (23) y completa la prueba. $\square$

Si Z es una variable aleatoria con distribución normal estándar, se sabe que

$$EZ^{2j} = \frac{(2j)!}{j!2^j}, \quad j \in \mathbb{N}_0.$$

Por otra parte, sigue de (18) que

$$\mu_{n,2j}(1/2) \sim \left(\frac{1}{4n}\right)^j EZ^{2j} = \frac{(2j)!}{8^j j! n^j}, \quad j \in \mathbb{N}_0.$$

Esto demuestra que la estimación dada en (20) es asintóticamente óptima.

Finalizamos la sección dando una estimación uniforme del Corolario 1.2.

Teorema 1.8. *Sea $m \in \mathbb{N}_0$ y $f \in \mathscr{C}^{(m)}[0,1]$. Entonces,*

$$\|B_n f(x) - \sum_{j=0}^{2m}\frac{f^{(j)}(x)}{j!}\mu_{n,j}(x)\| \le \frac{1}{6^m m! n^m}\left(1+\frac{1}{9}\frac{2m+3}{(m+1)(2m+1)(2m+2)}\frac{1}{n}\right)\omega\left(f^{(m)};\frac{1}{\sqrt{n}}\right).$$

Demostración. Reemplazando m por $2m$ en el Corolario 1.2, aplicando la cota dada en (21) y haciendo los cálculos correspondientes, se llega al resultado. $\square$

Observemos finalmente que el orden de magnitud de la estimación en el Teorema 1.8 es

$$\frac{1}{n^m}\omega\left(f^{(m)};\frac{1}{\sqrt{n}}\right).$$

2. Medias de Steklov

A lo largo de esta sección, supondremos $I \subset \mathbb{R}$ un intervalo arbitrario y $f : I \to \mathbb{R}$.

Definición 2.1. Para cada $h \in \mathbb{R}$, definimos los operadores traslación T_h y diferencia $\triangle_h$ como

$$T_h f(x) := f(x+h), \quad \triangle_h f(x) := (T_h - J)f(x) = f(x+h) - f(x),$$

siendo J el operador identidad, siempre que x, $x+h \in I$.

Para cada $m \in \mathbb{N}$, definimos sus iteradas T_h^m y $\triangle_h^m$ como

$$T_h^m f(x) := f(x+mh), \quad \triangle_h^m f(x) := (T_h - J)^m f(x),$$

siempre que $x, x+mh \in I$, con $T_h^0 = \triangle_h^0 = J$. A la cantidad $\triangle_h^m f(x)$ la llamaremos *diferencia de orden m de f en x*. Por el Teorema del binomio de Newton, se tiene que

$$\triangle_h^m f(x) = \sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k}T_h^k f(x) = \sum_{k=0}^{m}\binom{m}{k}(-1)^{m-k}f(x+kh).$$

Sea $(U_j)_{j\geq 1}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con $U_1 \equiv U(0,1)$. Denotamos mediante

$$S_m = U_1 + \cdots + U_m, \quad m = 1,2,\ldots (S_0 = 0).$$

En el siguiente resultado damos una expresión probabilística para las diferencias de orden m de f cuando f es una función suave.

Lema 2.2. *Sea $m \in \mathbb{N}_0$ y $h \in \mathbb{R}$. Si $f \in \mathscr{C}^{(m)}(I)$, entonces*

$$\triangle_h^m f(x) = h^m E f^{(m)}(x + hS_m), \quad x,\ x + mh \in I.$$

Demostración. Para $m = 0$ la igualdad es trivial. Supongamos la igualdad cierta para $m \in \mathbb{N}_0$. Observemos que

$$\triangle_h f(x) = f(x+h) - f(x) = h\int_0^1 f'(x+hu)du = hEf'(x+hU_{m+1}). \tag{25}$$

Sea $f \in \mathscr{C}^{(m+1)}(I)$ y supongamos que el resultado es cierto para m. Entonces se sigue de (25) que

$$\begin{aligned}\triangle_h^{m+1} f(x) &= \triangle_h (\triangle_h^m f(x)) = \triangle_h (h^m E f^{(m)}(x+hS_m)) = h^m \triangle_h E f^{(m)}(x+hS_m)\\ &= h^{m+1} E f^{(m+1)}(x + hS_m + hU_{m+1}) = h^{m+1} E f^{(m+1)}(x+hS_{m+1}),\end{aligned}$$

lo que prueba el resultado. □

Esta igualdad se puede escribir también de esta forma

$$\triangle_h^m f(x) = h^m \int_0^1 \cdots \int_0^1 f^{(m)}(x + h(u_1 + \cdots + u_m))du_1 \cdots du_m.$$

A continuación, daremos dos lemas auxiliares para la siguiente sección. Supondremos ahora $I = [0,1]$ y $f : [0,1] \to \mathbb{R}$, pero valdría cualquier intervalo arbitrario.

Definición 2.3. Sea V una variable aleatoria que toma valores en [0,1]. Para cada $m \in \mathbb{N}$, $0 \leq h \leq 1/m$, definimos la función

$$P_{m,h}f(x) = (-1)^{m-1}\sum_{k=1}^{m}\binom{m}{k}(-1)^{m-k}Ef(x+kh(V-x)), \quad x \in [0,1]. \tag{26}$$

A los operadores de la forma $P_{m,h}$ se les llama *medias de Steklov*.

Notar que, para cada $k = 0,1,\ldots,m$, se tiene $0 \leq kh \leq 1$ y

$$x + kh(V-x) = (1-kh)x + khV \in [0,1],$$

por ser combinaciones convexas de puntos en $[0,1]$. También podemos observar que

$$P_{m,h}f(x) - f(x) = (-1)^{m-1}E \triangle_{h(V-x)}^m f(x), \quad x \in [0,1]. \tag{27}$$

Una forma típica de construir V, consiste en tomar $(U_j)_{j\geq 1}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, tales que $U_1 \equiv U(0,1)$ y definir

$$V = \frac{U_1 + \cdots + U_n}{n}, \quad n \in \mathbb{N}. \tag{28}$$

En tal caso, se tiene la expresión analítica

$$P_{m.h}f(x) = (-1)^{m-1}\sum_{k=1}^{m}\binom{m}{k}(-1)^{m-k}\int_0^1 \cdots \int_0^1 f\left(x + kh\left(\frac{u_1 + \cdots + u_n}{n} - x\right)\right)du_1 \cdots du_n.$$

Esta construcción se debe a Petrushev y Popov [9]. En lo que resta de sección, fijamos V como en (28), con $n = 2$, y denotaremos

$$P_h f(x) := P_{2,h} f(x), \quad x \in [0,1], \quad 0 \leq h \leq 1/2.$$

Llamaremos $f_{(m)}$ a una antiderivada de f de orden $m \in \mathbb{N}$, es decir, a una función $f_{(m)} : [0,1] \to \mathbb{R}$ derivable m veces en $[0,1]$ tal que $f_{(m)}^{(m)} = f$.

Lema 2.4. *Sea* $0 < h \leq 1/2$. *Entonces, para cada* $x \in [0,1]$, *se tiene*

$$\begin{aligned} P_h f(x) &= 2Ef\left(x + h\left(\frac{U_1+U_2}{2} - x\right)\right) - Ef\left(x + 2h\left(\frac{U_1+U_2}{2} - x\right)\right) \\ &= \frac{1}{h^2}\left(8 \triangle_{h/2}^2 f_{(2)}(x(1-h)) - \triangle_h^2 f_{(2)}(x(1-2h))\right). \end{aligned}$$

Demostración. La primera igualdad se sigue de (26), y la segunda se cumple gracias al Lema 2.2. □

Definición 2.5. Se define el *segundo módulo de continuidad de f como*

$$\omega_2(f;\delta) := sup\left\{| \triangle_h^2 f(x-h)| : x-h, x+h \in I, 0 \leq h \leq \delta\right\}, \quad \delta \geq 0.$$

Observemos que si $f \in \mathscr{C}^{(2)}(I)$ se tiene a partir del Lema 2.2 que

$$\triangle_h^2 f(x-h) = h^2 E f^{(2)}(x - h + hS_2).$$

Por tanto, si $f^{(2)}$ está acotada, entonces

$$| \triangle_h^2 f(x-h) \leq h^2 \|f^{(2)}\|$$

y en consecuencia

$$\omega_2(f;\delta) \leq \|f^{(2)}\|\delta^2, \quad \delta \geq 0.$$

Es decir, a diferencia del primer módulo de continuidad, $\omega_2(f;\delta)$ es del orden δ^2, cuando $\delta \to 0$, en el caso de funciones dos veces derivables con derivada segunda acotada. Este resultado se generalizará en el Lema 4.3.

Lema 2.6. *Sean* $0 < h \leq 1/2$ *y* $f \in \mathscr{C}[0,1]$. *Entonces, para cada* $x \in [0,1]$, *se tiene*

(a) $|P_h f(x) - f(x)| \leq \omega_2(f;h)$.

(b) $|(P_h f)''(x)| \leq \frac{1}{h^2}(8\omega_2(f;h/2) + \omega_2(f;h))$.

Demostración. La parte (a) se sigue de la igualdad (27). con $m = 2$,

$$|P_h f(x) - f(x)| = |E \triangle_{h(V-x)}^2 f(x)| \leq \omega_2(f;h), \quad x \in [0,1],$$

puesto que $|V - x| \leq 1$ para todo $x \in [0,1]$. La parte (b) se sigue de la segunda igualdad del Lema 2.4 y de la Definición 2.5. □

Este resultado indica que $P_h f$ es un aproximante suave de f, puesto que $P_h f \in \mathscr{C}^2[0,1]$.

3. Velocidad de convergencia en términos del segundo módulo de continuidad usual

En esta sección, estudiamos para funciones $f \in \mathscr{C}[0,1]$, la velocidad de convergencia de los polinomios de Bernstein en términos del segundo módulo de continuidad usual.

Teorema 3.1. *Sea $f \in \mathscr{C}[0,1]$. Entonces, para cada $n \geq 4$, se tiene que*

$$|B_n f(x) - f(x)| \leq \left(2 + \frac{1}{8}\right) \omega_2\left(f; \frac{2\sqrt{x(1-x)}}{\sqrt{n}}\right) + \omega_2\left(f; \frac{\sqrt{x(1-x)}}{\sqrt{n}}\right), \quad x \in [0,1].$$

Como consecuencia,

$$\|B_n f - f\| \leq \left(2 + \frac{1}{8}\right) \omega_2\left(f; \frac{1}{\sqrt{n}}\right) + \omega_2\left(f; \frac{1}{2\sqrt{n}}\right).$$

Demostración. Sean $n \geq 4$, $x \in [0,1]$ y $0 \leq h \leq 1/2$. Podemos escribir

$$B_n f(x) - f(x) = (B_n f(x) - B_n P_h f(x)) + (P_h f(x) - f(x)) + (B_n P_h f(x) - P_h f(x)).$$

Por el Lema 2.6(a), se tienen las siguientes desigualdes,

$$|B_n f(x) - B_n P_h f(x)| \leq E\left| f\left(\frac{S_n(x)}{n}\right) - P_h f\left(\frac{S_n(x)}{n}\right)\right| \leq \omega_2(f;h), \tag{29}$$

$$|P_h f(x) - f(x)| \leq \omega_2(f;h). \tag{30}$$

Ahora, utilizando la fórmula de Taylor probabilística (veáse Teorema 1.1), se tiene que

$$B_n P_h f(x) - P_h f(x) = E\left[P_h f\left(\frac{S_n(x)}{n}\right) - P_h f(x)\right] = \frac{1}{2} E\left[P_h'' f\left(x + \left(\frac{S_n(x)}{n} - x\right)\beta_2\right)\left(\frac{S_n(x)}{n} - x\right)^2\right].$$

Juntando esto con la parte (b) del Lema 2.6, se obtiene que

$$|B_n P_h f(x) - P_h f(x)| \leq \frac{x(1-x)}{2n} \frac{1}{h^2} (8\omega_2(f;h/2) + \omega_2(f;h)).$$

Todo lo anterior se cumple para cualquier $0 \leq h \leq 1/2$. Se cumplirá también si tomamos

$$h = \frac{2\sqrt{x(1-x)}}{\sqrt{n}},$$

por ser $n \geq 4$. Para este h, la última desigualdad queda de la forma

$$|B_n P_h f(x) - P_h f(x)| \leq \omega_2\left(f; \frac{\sqrt{x(1-x)}}{\sqrt{n}}\right) + \frac{1}{8} \omega_2\left(f; \frac{2\sqrt{x(1-x)}}{\sqrt{n}}\right).$$

Sumando esto a (29) y (30), se obtiene la cota puntual. La cota uniforme se obtiene tomando supremos en ambos miembros de la cota puntual. $\square$

4. Velocidad de convergencia en términos del segundo módulo de Ditzian-Totik

Esta sección se enfocará en ver que, para funciones $f \in \mathscr{C}^{(2)}[0,1]$, la velocidad de convergencia de los polinomios de Bernstein se puede medir en términos del segundo módulo de continuidad de Ditzian-Totik.

Definición 4.1. Llamaremos *peso* a cualquier función $\sigma : I \to [0,1]$ tal que $\sigma(x) > 0$ para todo $x \in \mathring{I}$.

Definición 4.2. Se define el *segundo módulo de continuidad de Ditzian-Totik* con peso σ de f como

$$\omega_\sigma^2(f;\delta) := \left\{ |\triangle_{h\sigma(x)}^2 f(x - h\sigma(x))| : x - h\sigma(x),\ x + h\sigma(x) \in I,\ 0 \leq h \leq \delta \right\}, \quad \delta \geq 0.$$

El módulo de Ditzian-Tokik es un segundo módulo de continuidad a distancia variable, que cambia con el punto. El peso σ está relacionado con la desviación típica del proceso del operador que vayamos a estudiar. En los polinomios de Bernstein, $Var(S_n(x)/n) = x(1-x)/n$, y se escoge $\sigma(x) = \sqrt{x(1-x)}$, $x \in [0,1]$.

Supondremos entonces $I = [0,1]$ y tomaremos $\sigma(x) = \sqrt{x(1-x)}$, $x \in [0,1]$. Nos restringiremos a funciones $f \in \mathscr{C}[0,1]$. Aunque $\|f\|$ es una cantidad asociada a la función, en ocasiones escribiremos $\|f(x)\|$ por cuestiones de claridad. Recordar además la desigualdad triangular inversa,

$$\Big| \|f\| - \|g\| \Big| \leq \|f - g\|, \quad f, g \in \mathscr{C}[0,1].$$

El siguiente resultado describe el módulo de Ditzian-Totik para funciones suaves.

Lema 4.3. *Sea $f \in \mathscr{C}^{(2)}[0,1]$ tal que $\|\sigma^2 f''\| < \infty$. Entonces se verifica*

$$\Big| \omega_\sigma^2(f;\delta) - \delta^2 \|\sigma^2 f''\| \Big| \leq \frac{\delta^2}{8} \omega_\sigma^2(f'';\delta), \quad \delta \geq 0.$$

Demostración. Sean $x \in [0,1], h \geq 0$ tales que $x - h, x + h \in [0,1]$. Por la fórmula de Taylor probabilistica se tiene

$$f(x-h) = f(x) - f'(x)h + \frac{f''(x)}{2}h^2 + \frac{h^2}{2}E(f''(x - h\beta_2) - f''(x)),$$

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2}h^2 + \frac{h^2}{2}E(f''(x + h\beta_2) - f''(x)).$$

Sumando ambas expresiones, se verifica que

$$\triangle_h^2 f(x-h) = f''(x)h^2 + \frac{h^2}{2}E(f''(x - h\beta_2) - 2f''(x) + f''(x + h\beta_2)).$$

Utilizando la desigualdad triangular inversa y la igualdad anterior, se tiene que, para aquellos $0 \leq h \leq \delta$ tales que $x - h\sigma(x), x + \sigma(x) \in [0,1]$,

$$\begin{aligned} \Big| \omega_\sigma^2(f;\delta) - \delta^2 \|\sigma^2 f''\| \Big| &\leq \| \triangle_{h\sigma(x)}^2 f(x - h\sigma(x)) - h^2\sigma^2(x) f''(x) \| \leq \frac{h^2}{2} \|\sigma^2(x)\| \omega_\sigma^2(f'';\delta) \\ &\leq \frac{\delta^2}{8} \omega_\sigma^2(f'';\delta), \end{aligned}$$

para cualquier $\delta > 0$, puesto que $\|\sigma^2\| = 1/4$. □

Observación. En las condiciones del Lema 4.3, tomando $\delta = 1/\sqrt{n}$, se sigue que el orden de magnitud de $\omega_\sigma^2(f;1/\sqrt{n})$ es $1/n$.

Teorema 4.4. *Sea $f \in \mathscr{C}^{(2)}[0,1]$. Entonces, para cada $n \in \mathbb{N}$, se verifica*

$$\left| \|B_n f - f\| - \frac{1}{2}\omega_\sigma^2\left(f;\frac{1}{\sqrt{n}}\right)\right| \le \frac{1}{n}\left(\frac{7}{48}\omega\left((f'';\frac{1}{\sqrt{n}}\right) + \frac{1}{16}\omega_\sigma^2\left(f'';\frac{1}{\sqrt{n}}\right)\right).$$

En consecuencia, si f no es lineal,

$$\operatorname*{lím}_{n\to\infty} \frac{\|B_n - f\|}{\omega_\sigma^2(f;\frac{1}{\sqrt{n}})} = \frac{1}{2}.$$

Demostración. Utilizando las desigualdades triangulares y aplicando el Lema 4.3 con $\delta = 1/\sqrt{n}$, obtenemos que

$$\begin{aligned}\left| \|B_n f - f\| - \frac{1}{2}\omega_\sigma^2\left(f;\frac{1}{\sqrt{n}}\right)\right| &\le \left| \|B_n f - f\| - \frac{1}{2n}\|\sigma^2 f''\|\right| + \left|\frac{1}{2n}\|\sigma^2 f''\| - \frac{1}{2}\omega_\sigma^2\left(f;\frac{1}{\sqrt{n}}\right)\right| \\ &\le \|B_n f(x) - f(x) - \frac{1}{2n}\sigma^2(x) f''(x)\| + \frac{1}{16n}\omega_\sigma^2\left(f'';\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Ahora, gracias al Corolario 1.4, se tiene

$$\begin{aligned}\left\|B_n f(x) - f(x) - \frac{1}{2n}\sigma^2(x) f''(x)\right\| &\le \frac{1}{2n}\left\|\sigma(x)\left(1+\frac{\sigma(x)}{2}+\frac{1-6\sigma(x)}{6n}\right)\right\|\omega\left(f'';\frac{1}{\sqrt{n}}\right) \\ &\le \frac{\|\sigma(x)\|}{2n}\left\|1+\frac{1}{6n}-\frac{\sigma(x)}{2}\right\|\omega\left(f'';\frac{1}{\sqrt{n}}\right) \\ &\le \frac{1}{8n}\left(1+\frac{1}{6n}\right)\omega\left(f'';\frac{1}{\sqrt{n}}\right) \\ &\le \frac{7}{48n}\omega\left(f'';\frac{1}{\sqrt{n}}\right),\end{aligned}$$

puesto que $0 \le \sigma(x) \le 1/4$ para todo $x \in [0,1]$.

Para la segunda parte del teorema, notar que si f no es lineal, por ser $f \in \mathscr{C}^2[0,1]$, existe $x_0 \in (0,1)$ tal que $f''(x_0) \neq 0$, por lo que, para cada $n \in \mathbb{N}$, podemos tomar $0 \le h \le 1/\sqrt{n}$ tal que

$$|\triangle_{h\sigma(x_0)}^2 f(x_0 - h\sigma(x_0))| > 0.$$

Por tanto, $\omega_\sigma^2(f;1/\sqrt{n}) > 0$ para todo $n \in \mathbb{N}$, así que tiene sentido dividir entre $\omega_\sigma^2(f;1/\sqrt{n})$. Por la observación anterior al teorema, $\omega_\sigma^2(f;1/\sqrt{n})$ tiene el mismo orden de magnitud que $1/n$. Si hacemos la división en la desigualdad del enunciado, el factor $1/n$ se compensa con $\omega_\sigma^2(f;1/\sqrt{n})$, y los dos módulos de continuidad que quedan tienden a 0 cuando $n \to \infty$, por ser f'' continua. $\square$

Este teorema se puede interpretar diciendo que, para funciones $f \in \mathscr{C}^{(2)}[0,1]$, la velocidad de convergencia se mide en términos del segundo módulo de continuidad de Ditzian-Totik, y que en una desigualdad de la forma

$$\|B_n f - f\| \le C\omega_\sigma^2(f;\frac{1}{\sqrt{n}}),$$

la mejor constante C que la verifica no puede ser inferior a $1/2$.

Actualmente, el mejor resultado en términos del segundo módulo de continuidad de Ditzian-Totik, que nombramos sin demostración, ya que se utilizan técnicas muy avanzadas, es el siguiente:

Teorema 4.5. *Si $f \in \mathscr{C}[0,1]$ y $\sigma(x) = \sqrt{x(1-x)}$, $x \in [0,1]$, entonces*

$$K\omega_\sigma^2\left(f;\frac{1}{\sqrt{n}}\right) \le \|B_n f - f\| \le \frac{5}{2}\omega_\sigma^2\left(f;\frac{1}{\sqrt{n}}\right), \quad n \in \mathbb{N}, \tag{31}$$

para alguna constante absoluta K.

La desigualdad inferior fue probada por Ditzian e Ivanov en 1993 (veáse [5] y [7]), sin dar un valor para la constante K, y aún no se conoce ninguno. La demostración se basa en los llamados *K-funcionales*, que fueron introducidos por Peetre en 1963. En [8], se estudia esta desigualdad para funciones suaves.

La desigualdad superior fue demostrada por Păltănea, después de varios refinamientos de otros autores. No se sabe si la constante $5/2$ es óptima, pero no puede ser inferior a $1/2$, como se sigue del Teorema 4.4.

Capítulo 3

Polinomios de Bernstein aleatorios

En este capítulo, introduciremos variantes estocásticas de los polinomios clásicos de Bernstein asociados a una función continua f, construidos a partir de un arreglo triangular de nodos aleatorios en el intervalo $[0,1]$. Discutiremos la convergenica uniforme en probabilidad del proceso de aproximación que representan, proporcionando al mismo tiempo velocidades de convergencia. En el caso particular en el que los nodos aleatorios son los estadísticos ordenados de una muestra de variables aleatorias uniformes en $[0,1]$, se da una respuesta positiva a una conjetura planteada por Wu y Zhou [14], sobre una tasa exponencial de convergencia en probabilidad.

1. Definición y planteamiento del problema

En la definición de los polinomios de Bernstein (recordar (1)), se asume implícitamente que es posible evaluar la función f en el conjunto de nodos equidistantes k/n, $k = 0,1,\ldots,n$. Sin embargo, en los problemas reales, no se suele disponer de datos en lugares igualmente espaciados, y a veces contienen errores aleatorios debido a diversos factores. Estas son las motivaciones que subyacen a la definición de los polinomios de Bernstein aleatorios, introducidos recientemente por Wu et al. [13], desarrollados en [11] y [14], y extendidos en un entorno multivariante en [4].

Más concretamente, sea $\mathbb{Y} = (Y_{n,k}, n \in \mathbb{N}, k = 0,1,\ldots,n)$ un arreglo triangular de variables aleatorias tales que

$$0 \leq Y_{n,0} \leq Y_{n,1} \leq \cdots \leq Y_{n,n} \leq 1, \quad n \in \mathbb{N}. \tag{32}$$

Wu et al. [13] definen el n-ésimo polinomio de Bernstein aleatorio como

$$B_n(f,\mathbb{Y};x) = \sum_{k=0}^{n} f(Y_{n,k}) p_{n,k}(x), \quad p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}. \tag{33}$$

Observar que la fórmula (33) define una función aleatoria continua en $[0,1]$, la cuál no necesariamente interpola f en los puntos extremos, o una función determinista contina, en el caso de que $Y_{n,k}$ sean puntos deterministas. Obviamente, si $Y_{n,k} = k/n$, $k = 0,1,\ldots,n$, entonces la definición 33 coincide con la definición de los polinomios de Bernstein clásicos

Recordemos que una sucesión de variables aleatorias $(X_n)_{n\geq 1}$ converge a 0 en probabilidad si

$$\lim_{n\to\infty} P(|X_n| > \varepsilon) = 0, \quad \varepsilon > 0,$$

y se denota de la forma $X_n \xrightarrow{(P)} 0$.

Las velocidades de convergencia uniforme de los polinomios clásicos de Bernstein se caracterizan por lo visto en (31). Por otro lado, Sikkema [10] demostró que

$$\|B_n f - f\| \leq c\omega\left(f; \frac{1}{\sqrt{n}}\right), \quad c = \frac{4306 + 837\sqrt{6}}{5832} = 1{,}089887\ldots, \tag{34}$$

donde c es la mejor constante que cumple la desigualdad.
En el entorno estocástico descrito anteriormente, vemos que los resultados análogos a (31) y (34) consisten en analizar la convergencia uniforme en probabilidad,

$$P\big(\|B_n(f,\mathbb{Y};x)-f(x)\|>\varepsilon\big), \quad \varepsilon>0,$$

intentando al mismo tiempo dar velocidades de convergencia. Se trata de construir una banda determinista $[f(x)-\varepsilon, f(x)+\varepsilon]$, $x\in[0,1]$ que contenga a la función aleatoria $B_n(f,\mathbb{Y};x)$, $x\in[0,1]$. O, en términos probabilísticos, construir una banda (aleatoria) de confianza $[B_n(f,\mathbb{Y};x)-\varepsilon, B_n(f,\mathbb{Y};x)+\varepsilon]$, $x\in[0,1]$ que contenga a la función determinista $f(x)$, $x\in[0,1]$.

Un caso particular de la definición (33) es el siguiente. Para cada $n\in\mathbb{N}$, sea $(V_j)_{j=1}^{n+1}$ una sucesión finita de variables aleatorias independientes idénticamente distribuidas, que siguen una distribución uniforme en $[0,1]$. Sean $V_{n+1:1}\le\cdots\le V_{n+1:n+1}$ los estadísticos de orden obtenidos al ordenar $(V_j)_{j=1}^{n+1}$ en orden creciente de magnitud. Denotaremos,

$$Y_{n,k}=V_{n+1:k+1}, \quad n\in\mathbb{N}, \quad k=0,1,\ldots,n. \tag{35}$$

Por construcción, el arreglo triangular definido en (35) satisface la condición (32). Además, la variable aleatoria $Y_{n,k}=V_{n+1:k+1}$ tiene una función de densidad beta (veáse, por ejemplo, Arnold et al. [3], Capítulo 2),

$$\rho_k(\theta)=(n+1)p_{n,k}(\theta), \quad \theta\in[0,1], \quad n\in\mathbb{N}, \quad k=0,1,\ldots,n, \tag{36}$$

donde $p_{n,k}(x)$ está definido en (33). Para el arreglo triangular en (35), Wu y Zhou [14] han obtenido la estimación

$$P\big(\|B_n(f,\mathbb{Y};x)-f(x)\|>\varepsilon\big)\le C(r)\frac{n\omega^{2r}(f;1/\sqrt{n})}{\varepsilon^{2r}}, \quad \varepsilon>0, \quad r\in\mathbb{N}, \tag{37}$$

siempre que $\omega(f;1/\sqrt{n})<\varepsilon/4.2$, donde $C(r)$ es una constante que depende solo de r. El caso $r=3$, donde $C(r)=C(3)=15/64$, ya había sido obtenido por Wu et al. [13]. Además, Sun y Wu [11] han demostrado la estimación

$$P\big(\|B_n(f,\mathbb{Y};x)-f(x)\|>\varepsilon\big)\le 40\frac{\omega^2(f;1/\sqrt{n})}{\varepsilon^2}, \quad \varepsilon>0,$$

siempre que $\omega(f;1/\sqrt{n})<\varepsilon/7.2$.

Nótese que la estimación (37) garantiza la convergencia uniforme en probabilidad cuando f pertenece a cualquier clase de Hölder. Por otro lado, Wu y Zhou [14] conjeturaron, para un tamaño n suficientemente grande, que

$$P\big(\|B_n(f,\mathbb{Y};x)-f(x)\|>\varepsilon\big)\le Cn\exp\left(-\frac{1}{4}\frac{\varepsilon^2}{\omega^2(f;1/\sqrt{n})}\right), \tag{38}$$

para alguna constante positiva C.

El objetivo en este capítulo es mostrar que, bajo la configuración general dada para $\mathbb{Y}$ en (32), y bajo el supuesto de que

$$M_n:=\max_{0\le k\le n}\left|Y_{n,k}-\frac{k}{n}\right|\xrightarrow{(P)}0, \tag{39}$$

para cualquier función $f\in\mathscr{C}[0,1]$, se tiene

$$\lim_{n\to\infty}P\big(\|B_n(f,\mathbb{Y};x)-f(x)\|>\varepsilon\big)=0, \quad \varepsilon>0,$$

proporcionando al mismo tiempo velocidades de convergencia. Asimismo, los resultados obtenidos se aplicarán al caso particular dado por los estadísticos de orden, y se dará una respuesta positiva a la conjetura escrita en (38).

La doble desigualdad en (31) para los polinomios de Bernstein clásicos, implica que las tasas de convergencia mejoran para funciones suaves f. Esto sugiere considerar la siguiente variante de los polinomios de Bernstein aleatorios definidos en (33). Supongamos que $f\in\mathscr{C}^2[0,1]$ y, lo más importante, que podemos observar los valores de la primera derivada f' en los puntos aleatorios $Y_{n,k}$. Bajo estas circunstancias, definimos la siguiente variante de polinomios de Bernstein aleatorios.

$$B_n^*(f,\mathbb{Y};x) = \sum_{k=0}^{n} \big(f(Y_{n,k}) - f'(Y_{n,k})(Y_{n,k}-x)\big) p_{n,k}(x). \tag{40}$$

Tales polinomios preservan funciones afines, es decir, si $f(x) = ax+b$, para algunas constantes reales a y b, entonces

$$B_n^*(f,\mathbb{Y};x) = f(x).$$

Posteriormente, veremos en el Teorema 2.2 y el Corolario 3.4, que tales operadores estocásticos aproximan funciones $f \in \mathscr{C}^2[0,1]$ con mejores tasas de convergencia que los definidos en (33).

2. Convergencia uniforme en probabilidad

Sean $I = [a,b]$ y $J = [A,B]$ dos intervalos finitos cerrados. Denotemos mediante $\overline{\mathscr{L}}(I,J)$ al conjunto de funciones continuas a la izquierda y no decrecientes $f : I \to J$ tal que $f(a) = A$ y $f(b) = B$. A su vez, denotemos por $\widetilde{\mathscr{L}}(J,I)$ al conjunto de funciones continuas a la derecha y no decrecientes $g : J \to I$ tal que $g(A) = a$ y $g(B) = b$. Si $f \in \overline{\mathscr{L}}(I,J)$, su inversa continua a la derecha se define como

$$\widetilde{f}(y) = \sup\{x \in I : f(x) \le y\}, \quad y \in J. \tag{41}$$

Será necesario el siguiente resultado auxiliar. Aunque es tedioso, su demostración es sencilla (para más detalles, veáse Winter [12]).

Lema 2.1. *Si $f \in \overline{\mathscr{L}}(I,J)$, entonces $\widetilde{f} \in \widetilde{\mathscr{L}}(J,I)$. Además $f(x) \le y$ si y solo si $x \le \widetilde{f}(y)$, $x \in I$, $y \in J$. En particular,*

$$x \le \widetilde{f}(f(x)), \quad x \in I.$$

Este resultado se aplicará como sigue. Sea $f \in \mathscr{C}[0,1]$, $I = [0,1]$, y $J = [0, \omega(f;1)]$. Obviamente, la función $\omega(f;\cdot)$ pertenece a $\overline{\mathscr{L}}(I,J)$. Denotemos mediante $\widetilde{\omega}(f;\cdot)$ su inversa continua a la derecha, tal como se define en (41). Entonces, el Lema 2.1 nos dice, en particular, que

$$\delta \le \widetilde{\omega}(f;\omega(f;\delta)), \quad 0 \le \delta \le 1. \tag{42}$$

A partir de ahora, $\mathbb{Y}$ es un arreglo triangular de variables aleatorias que satisfacen (32).

Teorema 2.2. *Sea $f \in \mathscr{C}[0,1]$ y sea c la constante vista en (34). Si $M_n \xrightarrow{(P)} 0$, entonces*

$$\|B_n(f,\mathbb{Y};x) - f(x)\| \xrightarrow{(P)} 0.$$

En tal caso, tenemos que para cualquier $0 < \delta \le 1$,

$$P\big(\|B_n(f,\mathbb{Y};x) - f(x)\| > (1+c)\omega(f;\delta)\big) \le P(M_n > \delta), \quad n \ge 1/\delta^2.$$

Demostración. Usando (33), (34) y la desigualdad triangular, vemos que

$$\begin{aligned}
|B_n(f,\mathbb{Y};x) - f(x)| &= \left| \sum_{k=0}^{n} \big(f(Y_{n,k}) - f(x)\big) p_{n,k}(x) \right| \\
&\le \|B_n f(x) - f(x)\| + \sum_{k=0}^{n} \omega\left(f; \left|Y_{n,k} - \frac{k}{n}\right|\right) p_{n,k}(x) \\
&\le c\omega\left(f; \frac{1}{\sqrt{n}}\right) + \omega(f;M_n). \qquad (43)
\end{aligned}$$

Sea $\varepsilon \in (0, \omega(f;1)]$. Siempre que

$$\omega\left(f; \frac{1}{\sqrt{n}}\right) \le \varepsilon, \tag{44}$$

se sigue de (43) que

$$P\big(\|B_n(f,\mathbb{Y};x) - f(x)\| > (1+c)\varepsilon\big) \le P(\omega(f;M_n) > \varepsilon) = P(M_n > \widetilde{\omega}(f;\varepsilon)),$$

donde la última desigualdad se deduce del Lema 2.1 y de los comentarios que le siguen. La segunda igualdad se sigue de escoger $\varepsilon = \omega(f;\delta), \delta \ge 1/\sqrt{n}$, teniendo en cuenta (42) y el hecho de que la desigualdad (44) obviamente se cumple. Queda así probado el resultado. □

Para los polinomios estocásticos definidos en (40), obtenemos mejores estimaciones que las del Teorema 2.2.

Teorema 2.3. *Sea $f \in \mathscr{C}^2[0,1]$. Si $M_n \xrightarrow{(P)} 0$, entonces*

$$\|B_n^*(f,\mathbb{Y};x) - f(x)\| \xrightarrow{(P)} 0.$$

En tal caso, tenemos que para cualquier $\varepsilon > 0$,

$$P\left(\|B_n^*(f,\mathbb{Y};x) - f(x)\| > \frac{5}{4}\|f''\|\varepsilon\right) \le P(M_n > \sqrt{\varepsilon}), \quad n \ge 1/\varepsilon.$$

Demostración. Recordemos que el segundo momento central para los polinomios de Bernstein viene dado por

$$\sum_{k=0}^{n} \left|\frac{k}{n} - x\right|^2 p_{n,k}(x) = \frac{x(1-x)}{n}. \tag{45}$$

Por tanto, utilizando la fórmula de Taylor y la desigualdad

$$(a+b)^2 \le 2(a^2+b^2), \quad a,b \in \mathbb{R},$$

obtenemos que

$$\begin{aligned} |B_n^*(f,\mathbb{Y};x) - f(x)| &\le \frac{\|f''\|}{2} \sum_{k=0}^{n} (Y_{n,k} - x)^2 p_{n,k}(x) \\ &\le \|f''\| \left(\sum_{k=0}^{n} \left|\frac{k}{n} - x\right|^2 p_{n,k}(x) + \sum_{k=0}^{n} \left|Y_{n,k} - \frac{k}{n}\right|^2 p_{n,k}(x) \right) \\ &\le \|f''\| \left(\frac{1}{4n} + M_n^2\right). \end{aligned} \tag{46}$$

Para $n \ge 1/\varepsilon$, esto implica el resultado. □

Para comparar los Teoremas 2.2 y 2.3, supongamos que $f \in \mathscr{C}^2[0,1]$. En tal caso, tenemos que

$$\omega(f;\varepsilon) \le \|f'\|\varepsilon, \quad \varepsilon > 0.$$

Por tanto, se sigue del Teorema 2.2 que

$$P\big(\|B_n(f,\mathbb{Y};x) - f(x)\| > (1+c)\|f'\|\varepsilon\big) \le P(M_n > \varepsilon), \quad n \ge 1/\varepsilon^2.$$

Por otro lado, el Teorema 2.3 nos da la cota superior $P(M_n > \sqrt{\varepsilon})$, la cual es obviamente menor o igual que $P(M_n > \varepsilon)$ para $0 < \varepsilon \le 1$.
Este hecho queda ilustrado por el Lema 3.1 que veremos a continuación, en el caso de los estadísticos de orden uniforme.

3. El caso de los estadísticos de orden uniforme

En esta sección, asumiremos que el arreglo triangular $\mathbb{Y}$ de variables aleatorias, viene dada en términos de estadísticos de orden uniforme, tal como se definen en (35) y (36).

En este caso, la estimación de las probabilidades de cola $P(M_n > \varepsilon)$ se basará en dos hechos. En primer lugar, la desigualdad generalizada de Chebyshev, que afirma que dada una variable aleatoria X y una función no decreciente $\varphi : (0,\infty) \to (0,\infty)$, se tiene

$$P(|X| > \varepsilon) \leq \frac{E\varphi(|X|)}{\varphi(\varepsilon)}, \quad \varepsilon > 0. \tag{47}$$

En segundo lugar, la estimación uniforme de los momentos centrales pares de los polinomios de Bernstein, vista en (21).

Lema 3.1. *Sea $\varepsilon > 0$ y $0 < r < 1$. Entonces,*

$$P(M_n > \varepsilon) \leq \frac{n+1}{\sqrt{1-r}} \exp\left(-\frac{3r}{2} n\varepsilon^2\right).$$

Demostración. Sea $\alpha > 0$, la cual elegiremos más tarde. Se sigue de (39) que

$$P(M_n > \varepsilon) = P\left(\bigcup_{k=0}^{n} \left\{\left|Y_{n,k} - \frac{k}{n}\right| > \varepsilon\right\} \leq \sum_{k=0}^{n} P\left(\left|Y_{n,k} - \frac{k}{n}\right|\right) > \varepsilon\right).$$

De esta forma, usando la desigualdad de Chebysehv (47) con $\varphi(x) = exp(\alpha x)$, tenemos que

$$P(M_n > \varepsilon) \leq \sum_{k=0}^{n} P\left(\left|V_{n,k} - \frac{k}{n}\right|^2 > \varepsilon^2\right) \leq exp(-\alpha\varepsilon^2) \sum_{k=0}^{n} Eexp\left(\alpha\left|V_{n,k} - \frac{k}{n}\right|^2\right). \tag{48}$$

Por otro lado, se sigue de (36), (45) y de la estimación (21) que

$$\begin{aligned}
\sum_{k=0}^{n} E\exp\left(\alpha\left|V_{n,k} - \frac{k}{n}\right|^2\right) &= (n+1)\int_0^1 \sum_{k=0}^{n} \exp\left(\alpha\left|\theta - \frac{k}{n}\right|^2\right) p_{n,k}(\theta) d\theta \\
&= (n+1)\int_0^1 E\exp\left(\alpha\left|\frac{S_n(\theta)}{n} - \theta\right|^2\right) d\theta \\
&= (n+1)\sum_{j=0}^{\infty} \frac{\alpha^j}{j!} \int_0^1 E\left|\frac{S_n(\theta)}{n} - \theta\right|^{2j}\Bigg) d\theta \\
&\leq (n+1)\sum_{j=0}^{\infty} \binom{2j}{j}\left(\frac{\alpha}{6n}\right)^j.
\end{aligned}$$

Ahora, aplicando las siguientes igualdades

$$\binom{2j}{j} = \binom{-1/2}{j}(-4)^j, \quad j \in \mathbb{N}_0,$$

$$(1+x)^{-1/2} = \sum_{j=0}^{\infty} \binom{-1/2}{j} x^{-1/2}, \quad |x| < 1,$$

obtenemos que

$$(n+1)\sum_{j=0}^{\infty} \binom{2j}{j}\left(\frac{\alpha}{6n}\right)^j = (n+1)\sum_{j=0}^{\infty} \binom{-1/2}{j}\left(\frac{-4\alpha}{6n}\right)^j = (n+1)\left(1 - \frac{2\alpha}{3n}\right)^{-1/2}. \tag{49}$$

Tomando $\alpha = 3rn/2$, el resultado se sigue de (48) y (49). $\square$

Ahora estamos en condiciones de aplicar el Teorema 2.2 y el Teorema 2.3 para el caso de arreglos triangulares basadas en estadísticos de orden uniforme. Para ello, sea $(\tau(n))_{n\geq 1}$ una sucesión de números reales que cumplen las condiciones

$$\lim_{n\to\infty} \tau(n) = \infty, \quad \lim_{n\to\infty} \frac{\tau(n)}{n} = 0, \quad \tau(n) \geq 1, \quad n \in \mathbb{N}. \tag{50}$$

Corolario 3.2. *Sea $0 < r < 1$, y sean c y $(\tau(n))_{n\geq 1}$ vistos en (34) y (50) respectivamente. Para cualquier $f \in \mathscr{C}[0,1]$ y $n \in \mathbb{N}$, tenemos que*

$$P\left(\|B_n(f,\mathbb{Y};x) - f(x)\| > (1+c)\omega\left(f;\sqrt{\frac{\tau(n)}{n}}\right)\right) \leq \frac{n+1}{\sqrt{1-r}}\exp\left(-\frac{3r}{2}\tau(n)\right). \tag{51}$$

Demostración. Basta tomar $\delta = \sqrt{\tau(n)/n}$ en el Teorema 2.2 y aplicar el Lema 3.1, observando que $n \geq 1/\delta^2$, ya que $\tau(n) \geq 1$, gracias a la suposición (50). □

Las condiciones dadas en (50) relativas al comportamiento límite de la sucesión $(\tau(n))_{n\geq 1}$ no se utilizan en la demostración del Corolario 3.2. Sin embargo, dichas condiciones son razonables a la vista de la desigualdad (51). De hecho, el Corolario 3.2 solo tiene sentido si la parte derecha de (51) tiende a 0, cuando $n \to \infty$. Podriamos elegir, por ejemplo, $\tau(n) \sim \log(n+1)$ o $\tau(n) \sim n^{\alpha}$, $0 < \alpha < 1$. Al hacer esta elección, tenemos que equilibrar la anchura de la banda de confianza, cuyo orden de magnitud es $\omega(f;\sqrt{\tau(n)/n})$, y la velocidad de convergencia a 0 de la parte derecha de (51).

Para reformular el Corolario 3.2 en términos de la conjetura (38), damos el siguiente resultado.

Corolario 3.3. *Sea $\varepsilon > 0$ y $a, r \in (0,1)$. Supongamos que n es suficientemente grande para que*

$$\omega\left(f;\frac{1}{\sqrt{n}}\right) \leq \frac{2a}{3}\varepsilon. \tag{52}$$

Entonces,

$$P\big(\|B_n(f,\mathbb{Y};x) - f(x)\| > \varepsilon\big) \leq \frac{n+1}{\sqrt{1-r}}\exp\left(-\frac{3r}{2}(1-a)^2\frac{\varepsilon^2}{\omega^2(f;1/\sqrt{n})}\right).$$

Demostración. A partir de (33) y de la propiedad de subaditividad del primer módulo de continuidad vista en el Lema 4.2 del Capítulo 1, se tiene

$$\begin{aligned}
|B_n(f,\mathbb{Y};x) - f(x)| &= \left|\sum_{k=0}^{n}\big(f(Y_{n,k}) - f(x)\big)p_{n,k}(x)\right| \\
&\leq \sum_{k=0}^{n}\omega(f;|Y_{n,k}-x|)p_{n,k}(x) \leq \omega\left(f;\frac{1}{\sqrt{n}}\right)\sum_{k=0}^{n}\big(1+\sqrt{n}|Y_{n,k}-x|\big)p_{n,k}(x) \\
&\leq \omega\left(f;\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\sum_{k=0}^{n}\left|\frac{k}{n}-x\right|p_{n,k}(x) + \sqrt{n}M_n\right).
\end{aligned} \tag{53}$$

Utilizando (45) y la desigualdad de Schwarz, se sigue de (52) y (53) que

$$|B_n(f,\mathbb{Y};x) - f(x)| \leq a\varepsilon + \sqrt{n}\omega(f;1/\sqrt{n})M_n.$$

Por tanto, gracias al Lema 3.1 tenemos que

$$\begin{aligned}
P\big(\|B_n(f,\mathbb{Y};x) - f(x)\| > \varepsilon\big) &\leq P\big(\sqrt{n}\omega(f;1/\sqrt{n})M_n > (1-a)\varepsilon\big) \\
&\leq \frac{n+1}{\sqrt{1-r}}\exp\left(-\frac{3r}{2}(1-a)^2\frac{\varepsilon^2}{\omega^2(f;1/\sqrt{n})}\right).
\end{aligned}$$

De esta forma, queda probado el resultado. □

Además de dar constantes explícitas, el Corolario 3.3 nos da mejores tasas de convergencia que las vistas en (34).

Para funciones suaves, damos el siguiente resultado para la variante de los polinomios de Bernstein aleatorios definida en (40).

Corolario 3.4. *Sea* $0 < r < 1$ *y sea* $(\tau(n))_{n\geq 1}$ *definida como en (50). Para cualquier* $f \in \mathscr{C}^2[0,1]$ *y* $n \in \mathbb{N}$, *tenemos que*

$$P\left(\|B_n^*(f,\mathbb{Y};x) - f(x)\| > \frac{5}{4}\|f''\|\frac{\tau(n)}{n} \right) \leq \frac{n+1}{\sqrt{1-r}} \exp\left(-\frac{3r}{2}\tau(n) \right).$$

Demostración. Basta tomar $\varepsilon = \tau(n)/n$ en el Teorema 2.3 y aplicar el Lema 3.1. □

Concluimos, comparando las afirmaciones del Corolario 3.2 y el Corolario 3.4 con el escenario determinista descrito en (34). La desigualdad (34) nos dice que, con probabilidad uno, el n-ésimo polinomio de Bernstein $B_n f(;x)$ está dentro del intervalo $f(x) \pm c\omega(f;1/\sqrt{n})$, $x \in [0,1]$.
De acuerdo con el Corolario 3.2, el n-ésimo polinomio de Bernstein estocástico $B_n(f,\mathbb{Y};x)$ está dentro del intervalo $f(x) \pm (1+c)\omega(f;\sqrt{\tau(n)/n})$ con alta probabilidad (asintóticamente uno).
Una afirmación análoga es válida para la n-ésima variante $B_n^*(f,\mathbb{Y};x)$ considerada en el Corolario 3.4. La principal diferencia en este último caso, es que la longitud del intervalo, es decir,

$$\frac{5}{4}\|f''\|\frac{\tau(n)}{n}$$

es, asintóticamente, mucho más corta que la del Corolario 3.2.

Bibliografía

[1] Adell, J.A., Cárdenas-Morales, D.: Quantitative generalized Voronovskaja's formulae for Bernstein polynomials. J. Aprrox. Theory, **231** (2018), 41-52.

[2] Adell, J.A., Cárdenas-Morales, D.: Stochastic Bernstein polynomials: Uniform convergence in probability with rates. Adv. Comput. Math. **46**(2) (2020), 16 pp.

[3] Arnold, B.C., Balakrishnan, N., Nagaranja, H.N.: *A First Course in Order Statistics*. SIAM (2008).

[4] Cao, F., Xia, S.: Random sampling scattered data with multivariate Bernstein polynomials, Chin. Ann. Math. **35B**(4), 607-618 (2014).

[5] Ditzian, Z., Ivanov, K.G.: Strong converse inequalities. J. Anal. Math. **61**, 61-111, (1993).

[6] Lorentz, G.G.: *Bernstein Polynomials*, Chelsea Publishing Company, (1986).

[7] Knoop, H., Zhou, X.: The lower estimate for linear positive operators (II), Birkhäuser, **25** (1994).

[8] Păltănea, R.: Asymptotic constant in approximation of twice differentiable functions by a class of positive linear operators, Results Math. (2018),73:64.

[9] Petrushev, P.P., Popov, V.A.: *Rational Approximation of Real Function*, Cambridge University Press, (2011).

[10] Sikkema, P.C.: Der Wert einiger Konstanten in der Theorie der Approximation mit Bernstein-Polynomen. Numer. Math. **3**, 107-116 (1961).

[11] Sun, X., Wu, Z.: Chebyshev type inequality for stochastic Bernstein polynomials. Proc. Amer. Math. Soc. **147**(2), 671-679 (2019).

[12] Winter, B.B.: Transformationes of Lebesque-Stieljies integrals. J. Math. Anal. Appl. **205**, 471-484 (1997).

[13] Wu, Z., Sun, X., Ma, L.: Sampling scattered data with Bernstein polynomials: stochastic and deterministic error estimates. Adv. Comput. Math. **38**, 187-205 (2013).

[14] Wu, Z., Zhou, X.: Polynomial convergence order of stochastic Bernstein approximation. Adv. Comput. Math. **46**, 8 (2020).