# Métodos probabilísticos en el estudio de operadores lineales positivos

Facultad de Ciencias Universidad Zaragoza

Universidad Zaragoza 1542

**Mario Pérez Madre**

Trabajo de fin de grado en Matemáticas

Universidad de Zaragoza

Director del trabajo: José Antonio Adell Pascual

Junio de 2019

# Prólogo

Los operadores lineales positivos son uno de los métodos de aproximación de funciones más estudiados. Dado $I \subset \mathbb{R}$ un intervalo y $\mathscr{S}(I)$ un conjunto de funciones $f : I \to \mathbb{R}$, se dice que un operador lineal $L : \mathscr{S}(I) \to \mathscr{S}(I)$ es positivo si $Lf(x) \geq 0, x \in I$, para cualquier función $f \in \mathscr{S}(I)$ tal que $f(x) \geq 0$, $x \in I$. Uno de los ejemplos más importantes son los polinomios de Bernstein,

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

donde $f : [0,1] \to \mathbb{R}$ es una función arbitraria, siempre que esté definido $B_n f(x)$. Fueron introducidos por S. Bernstein a principios del siglo XX para dar una demostración sencilla del teorema de aproximación de Weierstrass y, actualmente, tienen una gran importancia en el diseño asistido por ordenador, ya que respetan muchas propiedades de la función original, como veremos más adelante. Los polinomios de Bernstein fueron muy estudiados por G. G. Lorentz, quien demostró algunas de sus propiedades, como la preservación de la monotonía, la convexidad o la variación usual.

Los polinomios de Bernstein están muy unidos a la probabilidad, y fueron el punto de partida de un gran interés en aplicar métodos probabilísticos en la aproximación mediante operadores lineales positivos. Muchos de los operadores estudiados habitualmente en la literatura poseen representaciones probabilísticas de la forma

$$L_n f(x) = Ef(Z_n(x)), \qquad x \in I, n \geq 1,$$

donde $(L_n)_{n\geq 1}$ es la familia de operadores en cuestión, $(Z_n(x), n \geq 1, x \in I)$ es un proceso estocástico a doble índice definido en un espacio de probabilidad completo $(\Omega, \mathscr{F}, P)$, y $f$ es una función que pertenece al conjunto

$$\mathscr{L} := \mathscr{L}(I) = \{f : I \to \mathbb{R} \text{ medible}; E|f(Z_n(x))| < \infty \text{ para cualesquiera } x \in I, n \geq 1\}.$$

A las representaciones como esta, en la que los procesos estocásticos se construyan a partir de unos procesos concretos, las denominaremos *representaciones completas,* y permiten demostrar varias propiedades de los operadores lineales positivos de forma unificada y sencilla. Una herramienta fundamental en algunas de las demostraciones será la *esperanza condicionada a $\sigma$-álgebras,* una generalización de la esperanza condicionada a variables aleatorias.

En el primer capítulo, definiremos la esperanza de una variable aleatoria condicionada a una $\sigma$-álgebra y daremos algunas de sus propiedades más importantes, junto con una de sus principales aplicaciones, las *martingalas y martingalas inversas,* cuyo origen se debe a los juegos de azar, y que son útiles, por ejemplo, al aplicar los teoremas de convergencia monótona bajo convexidad.

Después, veremos las representaciones completas de algunos operadores lineales positivos, en términos de varios procesos estocásticos, entre ellos los procesos empíricos, los procesos de Poisson y los procesos gamma, aunque no demostraremos algunas de las propiedades de estos procesos. La característica fundamental de estos será la fuerte dependencia entre las variables aleatorias.

En el tercer capítulo, demostraremos varias propiedades de preservación y aproximación de los operadores lineales positivos, aunque nos restringiremos en la práctica a los polinomios de Bernstein y los operadores de Szász y de Weierstrass. Veremos la preservación de la monotonía, $\varphi$-*variación,* convexidad, constantes de Lipschitz y del módulo de continuidad. En cuanto a propiedades de aproximación, probaremos la convergencia monótona bajo convexidad y una cierta propiedad de tipo Lipschitz. A la hora de aplicar los teoremas, nos centraremos en los polinomios de Bernstein y los operadores de Szász y de Weierstrass, para dar una muestra de cómo se utilizan los diferentes teoremas.

Finalmente, estudiaremos la velocidad de convergencia de los polinomios de Bernstein respecto al primer módulo de continuidad, para funciones $f$ con distintos grados de derivabilidad, y después respecto al segundo módulo de continuidad, mediante las llamadas *medias de Steklov,* para funciones $f$ continuas en $[0,1]$. Veremos, asimismo, el segundo módulo de continuidad de Ditzian-Totik. Probaremos un resultado para funciones suaves que permite medir la velocidad de convergencia en términos de este módulo y da una cota inferior de $1/2$ para las constantes que satisfacen desigualdades del tipo

$$\|B_n f - f\| \leq C\omega_\sigma^2(f, 1/\sqrt{n}),$$

y, finalmente, mencionaremos el mejor resultado obtenido hasta ahora sobre la velocidad de convergencia de los polinomios de Bernstein, en términos del segundo módulo de continuidad de Ditzian-Totik, que no probaremos, ya que su demostración requiere técnicas muy avanzadas.

# Resumen

Positive linear operators are one of the most popular methods to approximate functions. Let $I \subset \mathbb{R}$ be any real interval. If $\mathscr{S}(I)$ denotes a set of functions $f : I \to \mathbb{R}$ of a certain type, a linear operator $L : \mathscr{S}(I) \to \mathscr{S}(I)$ is said to be positive if $Lf(x) \geq 0, x \in I$, for any function $f \in \mathscr{S}(I)$ such that $f(x) \geq 0$, $x \in I$. Bernstein polynomials were one of the first positive linear operators considered. They are defined as

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

where $f : [0,1] \to \mathbb{R}$ is any arbitrary function. They were firstly introduced by S. Bernstein a hundred years ago to give an easy proof of Weierstrass' approximation theorem, and nowadays they are still very useful in Computer Aided Design, as they preserve many of the shape properties of the original function. Bernstein polynomials are closely related to Probability, and they were the first of many other positive linear operators to be studied using probabilistic methods.

Many positive linear operators can be represented in terms of some stochastic processes. Such representations allow us to prove their properties in a unified and simple way, instead of using the analytic techniques usually considered in the literature, which are strongly dependent both on the operator and the property. The drawback is that some advanced probabilistic tools are needed, namely, the conditional expectation with respect to $\sigma$-algebras, a generalization of the conditional expectation with respect to a random variable.

In the first chapter, we define the conditional expectation of a random variable with respect to a $\sigma$-algebra $\mathscr{G}$. First of all, we prove its existence using the Radon-Nikodym theorem, and we relate it to the usual conditional expectation with respect to another random variable. Then, we prove some of its main properties, such as linearity, monotonicity, a monotone convergence theorem, which we need to prove a property for products of random variables where one of them is measurable with respect to $\mathscr{G}$; a *tower* property, useful when iterated conditional expectations are considered; Jensen's inequality, which we will need later in many proofs, and another property for independent random variables. Finally, we consider one of the most important applications of conditional expectations, known as *martingales and reverse martingales.* These notions are very useful for some results, such as monotone convergence under convexity.

In the second chapter, we define the concept of *stochastic processes,* although we do not prove some of their properties. Basically, we deal with empirical processes, Poisson processes and gamma processes. We give a constructive definition of the first two, and the usual definition of the third one that can be found in textbooks. Next, we show that many positive linear operators, such as Bernstein polynomials or Szász operators, can be represented by means of stochastic processes as

$$L_n f(x) = Ef(Z_n(x)), \qquad x \in I, n \geq 1,$$

where $(L_n)_{n\geq 1}$ is the sequence of positive linear operators under consideration, $(Z_n(x), n \geq 1, x \in I)$ is a double-indexed stochastic process defined on a complete probability space $(\Omega, \mathscr{F}, P)$, constructed using the previously defined stochastic processes, and $f$ is a function in the set

$$\mathscr{L} := \mathscr{L}(I) = \{f : I \to \mathbb{R} \text{ measurable; } E|f(Z_n(x))| < \infty \text{ for all } x \in I, n \geq 1\}.$$

Despite the fact that only marginal distributions are needed to describe probabilistically these positive linear operator, we emphasize that the internal structure of the stochastic process is more important than what one would think at first sight. For instance, monotonicity will be used in many proofs, and martingale-type properties are closely related to convexity preservation and monotone convergence under convexity.

In the third chapter, we deal with some preserving and approximation properties of positive linear operators. Attention is focused on Bernstein polynomials, Szász operators and Weierstrass operators. By preserving properties, we mean that, if $I \subset \mathbb{R}$ is an interval on the real line, $(L_n)_{n\geq 1}$ is a family of positive linear operators and $\mathscr{R}(I) \subset \mathscr{S}(I)$ is a set of functions $f : I \to \mathbb{R}$ satisfying a certain property (for instance, monotonic functions or convex functions), then $L_n(\mathscr{R}(I)) \subset \mathscr{R}(I)$. In this sense, we start proving monotonicity and *$\varphi$-variation* preservation. To prove these results, we only need to assume that the stochastic process defining the operator has monotone paths. Next, we prove the preservation of Lipschitz constants and modulus of continuity. Here, we additionally need some linearity assumptions on the expectations of the random variables of the process. Finally, we prove a result on convexity preservation, whenever the underlying stochastic process has monotone paths and satisfies a martingale-type property. As we will see, the fact that the underlying process has stationary and independent increments is important when applying this theorem, along with some of the independence properties of the first chapter.

Regarding approximation properties, this is, properties concerning the *paths* $n \to L_nf(x)$, for a fixed $x \in I$, we prove that, if the function $f$ is convex and the processes under consideration are reverse martingales with respect to a decreasing sequence of $\sigma$-algebras, then the approximation paths converge monotonically to the function $f$. Along with this property, we also see that, if $f$ is a Lipschitz function, then the approximation paths satisfy a Lipschitz-type property.

In the last chapter, we consider the rate of convergence of Bernstein polynomials of functions with different assumptions on their continuity or differentiability. Firstly, we write the well-known Taylor's formula from a probabilistic viewpoint to obtain an upper bound in terms of the first modulus of continuity, and we derive from it some upper bounds for Bernstein polynomials. The formula obtained from the Taylor's polynomial of second order, known as Voronovskaja's formula, is particularly important, historically. The upper bound obtained from the expansion of order 0 was proved in other way by G. G. Lorentz.

Giving a step further, we define the Ditzian-Totik second modulus of continuity, which is a second modulus of continuity associated to the standard deviation of the random variables involved. Previously, we define the usual differences of order $m$ of $f$, giving some useful formulae that we will need later. Then, we prove a theorem for twice-differentiable functions that allows us to measure the rate of convergence of Bernstein polynomials in terms of the Ditzian-Totik second modulus of continuity, to conclude that an upper bound of the form $\|B_nf - f\| \leq C\omega_\sigma^2(f, 1/\sqrt{n})$ is possible with a constant $C \geq 1/2$, where $\omega_\sigma^2$ denotes the Ditzian-Totik second modulus. We also define the so-called *Steklov averages* or *Steklov functions,* which we will use as an intermediate step in the approximation of continuous functions by Bernstein polynomials. Next, we prove the main result of the fourth chapter, a theorem showing the rate of convergence of Bernstein polynomials of continuous functions $f$ with respect to the usual second modulus of continuity. Finally, we state without proof the best known result up to date concerning the rate of convergence of Bernstein polynomials of continuous functions, in terms of the Ditzian-Totik second modulus of continuity.

# Índice general

**Prólogo** III

**Resumen** V

**1. Esperanzas condicionales y martingalas** 1
- 1.1. Definición y existencia de la esperanza condicional . . . 1
- 1.2. Relación con la esperanza condicionada a otra variable aleatoria . . . 2
- 1.3. Propiedades de la esperanza condicional . . . 3
- 1.4. Martingalas, submartingalas y supermartingalas . . . 6

**2. Representaciones completas de operadores lineales positivos** 8
- 2.1. Procesos estocásticos y representaciones completas . . . 8
  - 2.1.1. Proceso empírico . . . 9
  - 2.1.2. Proceso de Poisson . . . 9
  - 2.1.3. Proceso gamma . . . 9
- 2.2. Ejemplos de representaciones completas . . . 10

**3. Propiedades de preservación y aproximación** 11
- 3.1. Propiedades de preservación . . . 11
  - 3.1.1. Monotonía y $\varphi$-variación . . . 11
  - 3.1.2. Constantes de Lipschitz y módulo de continuidad . . . 12
  - 3.1.3. Convexidad . . . 14
- 3.2. Propiedades de los caminos de aproximación . . . 16
  - 3.2.1. Convergencia monótona bajo convexidad . . . 16
  - 3.2.2. Propiedades de tipo Lipschitz . . . 17

**4. Velocidades de convergencia de los polinomios de Bernstein** 18
- 4.1. Velocidad en términos del primer módulo de continuidad . . . 18
- 4.2. Segundo módulo de continuidad de Ditzian-Totik . . . 20
- 4.3. Medias de Steklov . . . 22
- 4.4. Velocidades en términos de segundos módulos de continuidad . . . 24
  - 4.4.1. Velocidad en términos del segundo módulo usual . . . 24
  - 4.4.2. Velocidad en términos del segundo módulo de Ditzian-Totik . . . 24

**Bibliografía** 25

# Capítulo 1

# Esperanzas condicionales y martingalas

En la asignatura de Teoría de la probabilidad, hemos dado la esperanza condicionada a variables aleatorias. En este capítulo, trataremos la esperanza condicional respecto a $\sigma$-álgebras, que generaliza la anterior, y veremos el concepto de martingala.

Durante todo el capítulo, supondremos que todas las variables aleatorias toman valores en $(\mathbb{R}, \mathscr{B})$, siendo $\mathscr{B}$ la $\sigma$-álgebra de Borel de $\mathbb{R}$, y están definidas en un espacio de probabilidad $(\Omega, \mathscr{F}, P)$ completo, salvo que se diga lo contrario. Diremos que una condición se cumple *casi seguramente*, y lo denotaremos por *c.s.,* si se verifica salvo en un conjunto de probabilidad $P$ nula. Cuando manejemos varias probabilidades, escribiremos $P-$c.s. para indicar la probabilidad en consideración. Si una cantidad numerable de condiciones se satisfacen, cada una, casi seguramente, todas ellas juntas también, ya que, si $(A_i)_{i\geq 1} \subset \mathscr{F}$ y $P(A_i) = 1$, para cada $i = 1, 2, \ldots$, entonces $P\left(\cap_{i=1}^{\infty} A_i\right) = 1$.

**Definición.** Una sucesión $(Z_n)_{n\geq 1}$ de variables aleatorias *converge casi seguramente* a una variable aleatoria $Z$ si $P\Big(\{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\}\Big) = 1$.

## 1.1. Definición y existencia de la esperanza condicional

**Definición.** Sea $Z$ una variable aleatoria integrable y $\mathscr{G} \subset \mathscr{F}$ una $\sigma$-álgebra. Se llama ***esperanza condicional*** de $Z$ respecto a $\mathscr{G}$ a cualquier variable aleatoria $\mathscr{G}$-medible e integrable $Y$ tal que

$$\int_C Y\,dP = \int_C Z\,dP, \qquad C \in \mathscr{G}. \tag{1.1}$$

Notar que la esperanza usual es un caso particular de la esperanza condicional, tomando $\mathscr{G} = \{\emptyset, \Omega\}$. Para probar la existencia de la esperanza condicional, necesitamos el teorema de Radon-Nikodym, del que ya se ha dado una demostración en el grado.

**Definición.** Dado un espacio medible $(\Omega, \mathscr{F})$ y dos medidas finitas $P, Q$ definidas en este espacio, se dice que $Q$ es *absolutamente continua* respecto a $P$, y se escribe $Q \ll P$, si

$$P(C) = 0 \implies Q(C) = 0, \qquad C \in \mathscr{F}.$$

**Teorema 1.1.** ***(Radon-Nikodym)*** *Sea $Q$ una medida finita definida en $(\Omega, \mathscr{F})$. Entonces $Q \ll P$ si y solo si existe una variable aleatoria $\mathscr{F}$-medible, no negativa e integrable $Y$ tal que*

$$Q(C) = \int_C Y\,dP, \qquad C \in \mathscr{F}.$$

*Además, $Y$ es única $P-$c.s.*

**Teorema 1.2.** *Sea $Z$ una variable aleatoria integrable y $\mathscr{G} \subset \mathscr{F}$ una $\sigma$-álgebra. Entonces existe la esperanza condicional de $Z$ respecto a $\mathscr{G}$ y es única $P-$c.s.*

*Demostración.* Sea $P|_{\mathscr{G}}$ la medida $P$ restringida a la $\sigma$-álgebra $\mathscr{G}$, y supongamos $Z \geq 0$. Definimos

$$Q(C) = \int_C Z\,dP, \qquad C \in \mathscr{G}.$$

Notar que $Q$ es una medida finita definida en $(\Omega, \mathscr{G})$, por ser $Z$ integrable, y además $Q \ll P|_{\mathscr{G}}$. Aplicando el teorema de Radon-Nikodym, existe una variable aleatoria $\mathscr{G}$-medible, no negativa e integrable $Y$ tal que

$$\int_C Z\,dP = Q(C) = \int_C Y\,dP, \qquad C \in \mathscr{G}.$$

Por tanto, $Y$ satisface la definición de la esperanza condicional de $Z$ respecto a $\mathscr{G}$, y se tiene la unicidad $P-$c.s. de la esperanza condicional por la del teorema de Radon-Nikodym.
Si $Z$ es arbitraria, considerar $Z = Z^+ - Z^-$, siendo $Z^+, Z^-$ la parte positiva y negativa, respectivamente, y aplicar lo anterior. Por la misma razón, se tiene la unicidad $P-$c.s. de la esperanza condicional. □

**Notación.** Usaremos la notación $E(Z|\mathscr{G})$ para referirnos a una versión de la esperanza condicional. Se entenderá que cualquier propiedad relativa a la esperanza condicional es cierta $P-$c.s., sin decirlo expresamente.

## 1.2. Relación con la esperanza condicionada a otra variable aleatoria

**Teorema 1.3.** *Sea $Z$ una variable aleatoria integrable y $X = (X_1, \ldots, X_n)$ un vector aleatorio. Si $P_X$ denota la ley de $X$, entonces existe una función $g : \mathbb{R}^n \to \mathbb{R}$ medible e integrable tal que*

$$\int_{\{X \in B\}} Z\,dP = \int_B g(x_1, \ldots, x_n)\,dP_X(x_1, \ldots, x_n), \qquad B \in \mathscr{B}^n, \tag{1.2}$$

*siendo $\mathscr{B}^n$ la $\sigma$-álgebra de Borel en $\mathbb{R}^n$. Además, $g$ es única $P_X-$c.s.*

*Demostración.* Supongamos $Z \geq 0$ (si $Z$ es arbitraria, separar $Z = Z^+ - Z^-$). Sea

$$Q(B) = \int_{\{X \in B\}} Z\,dP, \qquad B \in \mathscr{B}^n.$$

Notar que $Q$ es una medida finita definida en $(\mathbb{R}^n, \mathscr{B}^n)$, por ser $Z$ integrable, y además $Q \ll P_X$, ya que $P_X(B) = P(X \in B), B \in \mathscr{B}^n$. Aplicando el teorema de Radon-Nikodym, existe una función medible, no negativa e integrable $g : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\int_{\{X \in B\}} Z\,dP = Q(B) = \int_B g(x_1, \ldots, x_n)\,dP_X(x_1, \ldots, x_n), \qquad B \in \mathscr{B}^n,$$

y $g$ es única $P_X-$c.s. □

Usaremos la notación $E(Z|X_1 = x_1, \ldots, X_n = x_n) := g(x_1, \ldots, x_n)$, entendiendo la igualdad c.s., ya que coincide con la esperanza de $Z$ condicionada a $X_1 = x_1, \ldots, X_n = x_n$, vista en cursos anteriores.

**Corolario 1.4.** *Si $Z$ y $X$ son variables aleatorias discretas que toman valores $z_i$, $x_i$, respectivamente, con $P(Z = z_i), P(X = x_i) > 0$, $i = 1, 2, \ldots$, entonces*

$$E(Z|X = x_i) = \sum_{j=1}^{\infty} z_j P(Z = z_j | X = x_i), \qquad i = 1, 2, \ldots$$

*Demostración.* Sea $g(x_i) = \sum_{j=1}^{\infty} z_j P(Z = z_j | X = x_i)$, $i = 1, 2, \ldots$ Para cada $B \in \mathscr{B}$, se tiene

$$\int_{\{X \in B\}} Z\, dP = \sum_{x_i \in B} \int_{\{X = x_i\}} Z\, dP = \sum_{x_i \in B} P(X = x_i) \sum_{j=1}^{\infty} z_j \frac{P(Z = z_j, X = x_i)}{P(X = x_i)} = \int_B g\, dP_X.$$

□

**Definición.** Si $X = (X_1, \ldots, X_n)$ es un vector aleatorio, definimos la *$\sigma$-álgebra generada por $X$* como

$$\sigma\{X\} = \{X^{-1}(B) : B \in \mathscr{B}^n\}.$$

Notar que es la menor $\sigma$-álgebra que hace medible a $X$. Denotaremos $E(Z|X_1, \ldots, X_n) := E(Z|\sigma(X))$.

**Teorema 1.5.** *Con la notación del teorema 1.3, se tiene que $E(Z|X_1, \ldots, X_n) = g(X_1, \ldots, X_n)$.*

*Demostración.* Para cada $B \in \mathscr{B}^n$, utilizando el teorema 1.3, se tiene

$$\int_{\{X \in B\}} g(X_1, \ldots, X_n)\, dP = \int_B g(x_1, \ldots, x_n)\, dP_X(x_1, \ldots, x_n) = \int_{\{X \in B\}} Z\, dP.$$

□

## 1.3. Propiedades de la esperanza condicional

**Definición.** Diremos que una variable $Z$ es *independiente de una $\sigma$-álgebra $\mathscr{G}$* si, para cualesquiera $B \in \mathscr{B}$, $C \in \mathscr{G}$, se tiene $P(\{Z \in B\} \cap C) = P(Z \in B)P(C)$.

**Teorema 1.6.** *Sea $\mathscr{G} \subset \mathscr{F}$ una $\sigma$-álgebra, $Y, Z$ variables aleatorias integrables, $a, b, c \in \mathbb{R}$. Se verifica:*

*(a)* $E(E(Z|\mathscr{G})) = EZ$.

*(b) Si $Z$ es $\mathscr{G}$-medible, entonces* $E(Z|\mathscr{G}) = Z$.

*(c)* $E(aY + bZ|\mathscr{G}) = aE(Y|\mathscr{G}) + bE(Z|\mathscr{G})$.

*(d) Si $Y \leq Z$* c.s. *entonces* $E(Y|\mathscr{G}) \leq E(Z|\mathscr{G})$.

*(e)* $\left|E(Z|\mathscr{G})\right| \leq E\left(|Z|\,|\mathscr{G}\right)$.

*(f) Si $Z$ es independiente de $\mathscr{G}$, entonces* $E(Z|\mathscr{G}) = EZ$.

*Demostración.* (a) Basta tomar $C = \Omega$ en (1.1).

(b) Es inmediata por la definición.

(c) Solo hay que usar la linealidad de la integral, ya que $Y, Z$ son integrables y $aY + bZ$ también.

(d) Si $Y \leq Z$ c.s., se tiene

$$\int_C E(Z - Y|\mathscr{G})\, dP = \int_C (Z - Y)\, dP \geq 0, \qquad C \in \mathscr{G}.$$

La conclusión se sigue de que esta fórmula es cierta para $C = \{E(Z - Y|\mathscr{G}) < 0\} \in \mathscr{G}$.

(e) Escribiendo $Z = Z^+ - Z^-$, $|Z| = Z^+ + Z^-$, y utilizando (c) y (d), se tiene

$$\left|E(Z|\mathscr{G})\right| = \left|E(Z^+|\mathscr{G}) - E(Z^-|\mathscr{G})\right| \leq E(Z^+|\mathscr{G}) + E(Z^-|\mathscr{G}) = E\left(|Z|\,|\mathscr{G}\right).$$

(f) Notar que, para cada $C \in \mathscr{G}$, $Z$ es independiente de $\mathbf{1}_C$, luego

$$\int_C EZ\, dP = EZP(C) = E(Z\mathbf{1}_C) = \int_\Omega Z\mathbf{1}_C\, dP = \int_C Z\, dP.$$

$\square$

**Notación.** Si $(Y_n)_{n\geq 1}, Y$ son variables aleatorias tales que $\lim_{n\to\infty} Y_n = Y$, con $Y_1 \leq \cdots \leq Y_n \leq \cdots \leq Y$, escribiremos $Y_n \uparrow Y$. Si satisfacen las condiciones casi seguramente, pondremos $Y_n \uparrow Y$ c.s.

**Teorema 1.7.** *Sea $\mathscr{G} \subset \mathscr{F}$ una $\sigma$-álgebra. Si $(Z_n)_{n\geq 1}, Z$ son variables aleatorias no negativas e integrables tales que $Z_n \uparrow Z$* c.s., *entonces*

$$E(Z_n|\mathscr{G}) \uparrow E(Z|\mathscr{G}).$$

*Demostración.* Debido a la monotonía de la sucesión $(Z_n)_{n\geq 1}$, aplicando la propiedad 1.6 (d), se tiene

$$E(Z_1|\mathscr{G}) \leq \cdots \leq E(Z_n|\mathscr{G}) \leq \cdots \leq E(Z|\mathscr{G}).$$

Esto implica, además, que existe $\lim_{n\to\infty} E(Z_n|\mathscr{G})$. Utilizando dos veces el teorema de la convergencia monótona y la definición de esperanza condicional, se tiene que, para cualquier $C \in \mathscr{G}$,

$$\int_C \lim_{n\to\infty} E(Z_n|\mathscr{G})\, dP = \lim_{n\to\infty} \int_C E(Z_n|\mathscr{G})\, dP = \lim_{n\to\infty} \int_C Z_n\, dP = \int_C \lim_{n\to\infty} Z_n\, dP = \int_C Z\, dP,$$

lo que implica que

$$\lim_{n\to\infty} E(Z_n|\mathscr{G}) = E(Z|\mathscr{G}).$$

$\square$

**Teorema 1.8.** *Sea $\mathscr{G} \subset \mathscr{F}$ una $\sigma$-álgebra. Sean $Y, Z$ variables aleatorias tales que $Z$, $YZ$ son integrables y $Y$ es $\mathscr{G}$-medible. Entonces*

$$E(YZ|\mathscr{G}) = YE(Z|\mathscr{G}).$$

*Demostración.* Seguiremos el método estándar utilizado a menudo en probabilidad. Supongamos $Z \geq 0$. Si $Y = \mathbf{1}_A$ es un indicador, con $A \in \mathscr{G}$, entonces $A \cap C \in \mathscr{G}$ para cualquier $C \in \mathscr{G}$, luego

$$\int_C YE(Z|\mathscr{G})\, dP = \int_{A\cap C} E(Z|\mathscr{G})\, dP = \int_{A\cap C} Z\, dP = \int_C YZ\, dP, \qquad C \in \mathscr{G}.$$

Si $Y = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, es una variable aleatoria simple, con $n \geq 1$, $a_i > 0$, $A_i \in \mathscr{G}$, para cada $i = 1, \ldots, n$, utilizando la linealidad de la esperanza condicional, se tiene

$$E(YZ|\mathscr{G}) = \sum_{i=1}^n a_i E(\mathbf{1}_{A_i} Z|\mathscr{G}) = \sum_{i=1}^n a_i \mathbf{1}_{A_i} E(Z|\mathscr{G}) = YE(Z|\mathscr{G}).$$

Supongamos ahora $Y$ no negativa, por teoría de la medida, sabemos que existe una sucesión $(Y_n)_{n\geq 1}$ de variables aleatorias simples, medibles y no negativas tales que $Y_n \uparrow Y$. Por tanto, se tiene que $Y_n Z \uparrow YZ$, y por el teorema 1.7,

$$E(Y_n Z|\mathscr{G}) \uparrow E(YZ|\mathscr{G}).$$

Por otro lado, como $Y_n$ son simples y no negativas,

$$E(Y_n Z|\mathscr{G}) = Y_n E(Z|\mathscr{G}) \uparrow YE(Z|\mathscr{G}).$$

Por la unicidad del límite, se tiene $E(YZ|\mathscr{G}) = YE(Z|\mathscr{G})$. Si $Y = Y^+ - Y^-$ es una variable aleatoria real arbitraria, utilizando la linealidad de la esperanza condicional y lo que acabamos de probar, se tiene

$$E(YZ|\mathscr{G}) = E(Y^+ Z|\mathscr{G}) - E(Y^- Z|\mathscr{G}) = Y^+ E(Z|\mathscr{G}) - Y^- E(Z|\mathscr{G}) = YE(Z|\mathscr{G}).$$

Finalmente, si $Y, Z$ son variables aleatorias reales cualesquiera, razonando de forma análoga, se tiene

$$E(YZ|\mathscr{G}) = E(YZ^+|\mathscr{G}) - E(YZ^-|\mathscr{G}) = YE(Z^+|\mathscr{G}) - YE(Z^-|\mathscr{G}) = YE(Z|\mathscr{G}).$$

$\square$

El siguiente teorema es muy útil cuando se toman esperanzas condicionales de forma reiterada.

**Teorema 1.9.** *Sean $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$ dos $\sigma$-álgebras y $Z$ una variable aleatoria. Entonces:*

$$E(E(Z|\mathscr{G}_1)|\mathscr{G}_2) = E(Z|\mathscr{G}_1) = E(E(Z|\mathscr{G}_2)|\mathscr{G}_1).$$

*Demostración.* La primera igualdad es inmediata, puesto que $E(Z|\mathscr{G}_1)$ es $\mathscr{G}_1$-medible, luego también es $\mathscr{G}_2$-medible, y basta aplicar la propiedad 1.6 (b). Para la segunda igualdad, notar que

$$\int_C E(E(Z|\mathscr{G}_2)|\mathscr{G}_1)\, dP = \int_C E(Z|\mathscr{G}_2)\, dP = \int_C Z\, dP, \qquad C \in \mathscr{G}_1,$$

donde hemos usado la definición de esperanza condicional y, en la segunda igualdad, que $\mathscr{G}_1 \subset \mathscr{G}_2$. Notar que $E(E(Z|\mathscr{G}_2)|\mathscr{G}_1)$ es $\mathscr{G}_1$-medible. □

Ahora demostraremos la desigualdad de Jensen. Necesitamos el siguiente resultado:

**Lema 1.10.** *Sea $I \subset \mathbb{R}$ un intervalo abierto y $f : I \to \mathbb{R}$ una función convexa. Entonces existen dos familias numerables $(a_n)$ y $(b_n)$ de números reales tales que $f(x) = \sup_n\{a_n x + b_n\}, x \in I$.*

No damos la demostración porque es técnica, pero elemental, y no aporta ideas relevantes al trabajo. Se puede consultar en ([4], p. 253). La idea es probar que se puede tomar el supremo sobre un número contable de rectas llamadas *rectas de soporte.*

**Teorema 1.11.** ***(Desigualdad de Jensen)*** *Sea $\mathscr{G} \subset \mathscr{F}$ una $\sigma$-álgebra, $I \subset \mathbb{R}$ un intervalo abierto, $f : I \to \mathbb{R}$ convexa y $Z$ una variable aleatoria integrable con valores en $I$. Si $f(Z)$ es integrable, se tiene*

$$E(f(Z)|\mathscr{G}) \geq f(E(Z|\mathscr{G})).$$

*Demostración.* Supongamos $I = (a,b)$, con $-\infty < a < b < \infty$. Veamos que $E(Z|\mathscr{G})$ toma valores en $(a,b)$ c.s. Notar que $\{E(Z-a|\mathscr{G}) \leq 0\}, \{E(b-Z|\mathscr{G}) \leq 0\} \in \mathscr{G}$. Por tanto,

$$0 \geq \int_{\{E(Z-a|\mathscr{G})\leq 0\}} E(Z-a|\mathscr{G})dP = \int_{\{E(Z-a|\mathscr{G})\leq 0\}} (Z-a)dP \geq 0,$$
$$0 \geq \int_{\{E(b-Z|\mathscr{G})\leq 0\}} E(b-Z|\mathscr{G})dP = \int_{\{E(b-Z|\mathscr{G})\leq 0\}} (b-Z)dP \geq 0,$$

lo que implica que, en el suceso $\{E(Z-a|\mathscr{G}) \leq 0\}$, $Z = a$ c.s., y en el suceso $\{E(b-Z|\mathscr{G}) \leq 0\}$, $Z = b$ c.s. Pero, por hipótesis, $a < Z < b$, por lo que ambos sucesos deben tener probabilidad nula. Por tanto, $E(Z|\mathscr{G})$ toma valores en $(a,b)$ c.s., con lo que $f(E(Z|\mathscr{G}))$ está bien definida. Si $I$ solo está acotado por un lado, aplicar el razonamiento anterior, y si $I = \mathbb{R}$, $f(E(Z|\mathscr{G}))$ está bien definida automáticamente.

Utilizando el lema 1.10, se tiene que, para cada $n$, $f(x) \geq a_n x + b_n$, $x \in (a,b)$. Cambiando $x$ por $Z$, se tiene que $f(Z) \geq a_n Z + b_n$, y aplicando la monotonía y la linealidad de la esperanza condicional, se obtiene que, para cada $n$,

$$E(f(Z)|\mathscr{G}) \geq a_n E(Z|\mathscr{G}) + b_n.$$

Como las familias $(a_n), (b_n)$ son numerables, esta desigualdad se cumple c.s. para todo $n$ a la vez. Tomando supremos sobre $n$ en ambos miembros, y aplicando de nuevo el lema 1.10, se obtiene

$$E(f(Z)|\mathscr{G}) \geq \sup_n\{a_n E(Z|\mathscr{G}) + b_n\} = f(E(Z|\mathscr{G})).$$

□

**Teorema 1.12.** *Sea $Z$ una variable aleatoria integrable y $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_m)$ vectores aleatorios tales que $(Z, X)$ y $Y$ son independientes. Entonces*

$$E(Z|X,Y) = E(Z|X).$$

*Demostración.* Si $A \times B \in \mathscr{B}^n \times \mathscr{B}^m$, utilizando la independencia, el teorema 1.9 y el teorema 1.5, se tiene

$$\int_{\{X\in A, Y\in B\}} ZdP = E\left(Z\mathbf{1}_{\{X\in A\}}\mathbf{1}_{\{Y\in B\}}\right) = E\left(Z\mathbf{1}_{\{X\in A\}}\right)E\mathbf{1}_{\{Y\in B\}} = E\left(E\left(Z\mathbf{1}_{\{X\in A\}}|X\right)\right)E\mathbf{1}_{\{Y\in B\}}$$
$$= E\left(\mathbf{1}_{\{X\in A\}}E\left(Z|X\right)\right)E\mathbf{1}_{\{Y\in B\}} = E\left(\mathbf{1}_{\{X\in A\}}E\left(Z|X\right)\mathbf{1}_{\{Y\in B\}}\right) = \int_{\{X\in A, Y\in B\}} E\left(Z|X\right)dP.$$

Notar que también se cumple para uniones finitas disjuntas de rectángulos. Si $(C_n)_{n\geq 1}$ es una sucesión no decreciente de conjuntos en $\sigma\{X,Y\}$ que verifican la igualdad y $C = \cup_{n=1}^{\infty} C_n$, entonces

$$\int_C ZdP = \lim_{n\to\infty}\int_{C_n} ZdP = \lim_{n\to\infty}\int_{C_n} E\left(Z|X\right)dP = \int_C E\left(Z|X\right)dP,$$

y si $(C_n)_{n\geq 1}$ es una sucesión no creciente, también se verifica para $C = \cap_{n=1}^{\infty} C_n$. Utilizando el teorema de la clase monótona, se tiene el resultado. □

## 1.4. Martingalas, submartingalas y supermartingalas

En esta primera parte, supondremos que $(\mathscr{G}_n)_{n\geq 1}$ es una sucesión **creciente** de $\sigma$-álgebras en $\mathscr{F}$.

**Definición.** Una sucesión de variables aleatorias $(Z_n)_{n\geq 1}$ se dice adaptada si $Z_n$ es $\mathscr{G}_n$-medible, $n \geq 1$. En ese caso, usaremos la notación $(Z_n, \mathscr{G}_n)_{n\geq 1}$.

**Definición.** Una sucesión $(Z_n, \mathscr{G}_n)_{n\geq 1}$ de variables aleatorias integrables es una *martingala* (respectivamente, *submartingala, supermartingala*) si

$$E(Z_{n+1}|\mathscr{G}_n) = Z_n, \quad (\text{resp. } \geq Z_n,\ \leq Z_n), \quad n \geq 1.$$

**Observación.** Una martingala es también una submartingala y una supermartingala.

**Teorema 1.13.** *Una sucesión $(Z_n, \mathscr{G}_n)_{n\geq 1}$ de variables aleatorias integrables es una martingala (*resp. *submartingala, supermartingala) si y solo si*

$$E(Z_n|\mathscr{G}_m) = Z_m, \quad (\text{resp. } \geq Z_m,\ \leq Z_m), \quad 1 \leq m \leq n.$$

*Demostración.* Es evidente que es una condición suficiente, tomando $n = m+1$. Para ver que es necesaria, supongamos $1 \leq m < n$ (si $m = n$, se verifica la igualdad por ser $(Z_n)_{n\geq 1}$ adaptada). Utilizando el teorema 1.9 y que $(\mathscr{G}_n)_{n\geq 1}$ es creciente, se tiene

$$E(Z_n|\mathscr{G}_m) = E(E(Z_n|\mathscr{G}_{n-1})|\mathscr{G}_m) = E(Z_{n-1}|\mathscr{G}_m) = \cdots = E(Z_{m+1}|\mathscr{G}_m) = Z_m.$$

Si es submartingala, utilizando, además, la monotonía de la esperanza condicional, se tiene

$$E(Z_n|\mathscr{G}_m) = E(E(Z_n|\mathscr{G}_{n-1})|\mathscr{G}_m) \geq E(Z_{n-1}|\mathscr{G}_m) \geq \cdots \geq E(Z_{m+1}|\mathscr{G}_m) \geq Z_m.$$

y si es supermartingala, se obtiene esto mismo, cambiando los signos de las desigualdades. □

**Observación.** Si $(Y_n, \mathscr{G}_n)_{n\geq 1}$, $(Z_n, \mathscr{G}_n)_{n\geq 1}$ son martingalas y $a, b \in \mathbb{R}$, entonces $(aY_n + bZ_n, \mathscr{G}_n)_{n\geq 1}$ es una martingala. El resultado también es cierto cambiando las martingalas por submartingalas o supermartingalas, siempre que $a, b \geq 0$.

Veamos un ejemplo interesante de martingalas.

**Proposición 1.14.** *Sean $(Z_n)_{n\geq 1}$ variables aleatorias integrables, independientes e idénticamente distribuidas con media $EZ_1 = \mu$. Para cada $n \geq 1$, definamos $S_n = Z_1 + \cdots + Z_n - n\mu$, y supongamos $\mathscr{G}_n = \sigma\{Z_1, \ldots, Z_n\}$. Entonces $(S_n, \mathscr{G}_n)_{n\geq 1}$ es una martingala.*

*Demostración.* Sea $n \geq 1$. Notar que $Z_{n+1}$ es independiente de $Z_1,\dots,Z_n$, y que $Z_1,\dots,Z_n$ son $\mathscr{G}_n$-medibles. Utilizando la linealidad de la esperanza condicional y la propiedad 1.6 (f), se tiene

$$E(S_{n+1}|\mathscr{G}_n) = \sum_{k=1}^{n} E(Z_k|\mathscr{G}_n) + E(Z_{n+1}|\mathscr{G}_n) - (n+1)\mu = \sum_{k=1}^{n} Z_k + EZ_{n+1} - (n+1)\mu = S_n.$$

□

De aquí en adelante, supondremos que $(\mathscr{G}_n)_{n\geq 1}$ es una sucesión **decreciente** de $\sigma$-álgebras en $\mathscr{F}$.

**Definición.** Una sucesión $(Z_n,\mathscr{G}_n)_{n\geq 1}$ de variables aleatorias integrables es una *martingala inversa* (respectivamente, *submartingala inversa, supermartingala inversa*) si

$$E(Z_n|\mathscr{G}_{n+1}) = Z_{n+1}, \quad (\text{resp. } \geq Z_{n+1}, \leq Z_{n+1}), \qquad n \geq 1.$$

El teorema 1.13 se cumple para martingalas inversas si $1 \leq n \leq m$, usando el mismo razonamiento.

**Observación.** Si $(Y_n,\mathscr{G}_n)_{n\geq 1}$, $(Z_n,\mathscr{G}_n)_{n\geq 1}$ son martingalas inversas y $a,b \in \mathbb{R}$, entonces $(aY_n+bZ_n,\mathscr{G}_n)_{n\geq 1}$ es una martingala inversa. El resultado también es cierto si consideramos submartingalas o supermartingalas inversas, siempre que $a,b \geq 0$.

**Definición.** Sean $(Z_n)_{n\geq 1}$ variables aleatorias. Definimos la *$\sigma$-álgebra generada por* $(Z_n)$ como

$$\sigma\{(Z_n)\} := \sigma\{\cup_{i=1}^{\infty}\sigma\{Z_n\}\} = \sigma\{Z_1,Z_2,\dots\}.$$

Notar que es la menor $\sigma$-álgebra que hace medible a todas las $Z_n$, $n = 1,2,\dots$

**Proposición 1.15.** *Si* $S_n = Z_1 + \cdots + Z_n$, $n = 1,\dots,m$, *se tiene*

$$\sigma\{Z_1,\dots,Z_m\} = \sigma\{S_1,\dots,S_m\}.$$

*Demostración.* La $\sigma$-álgebra generada por un vector aleatorio es la menor que lo hace medible, y un vector aleatorio es medible si y solo si sus componentes lo son. Si las $Z_n$ son medibles, también lo son las $S_n$, luego $\sigma\{S_1,\dots,S_m\} \subset \sigma\{Z_1,\dots,Z_m\}$. Y recíprocamente, ya que $Z_n = S_n - S_{n-1}$, $n = 2,\dots,m$. □

Nombramos, sin demostración, que el teorema 1.12 y la proposición 1.15 se pueden probar para sucesiones de variables aleatorias. Veamos ahora un ejemplo importante de martingalas inversas.

**Proposición 1.16.** *Sean* $(Z_n)_{n\geq 1}$ *variables aleatorias integrables, independientes e idénticamente distribuidas. Para cada* $n \geq 1$, *definamos* $S_n = Z_1 + \cdots + Z_n$, *y supongamos* $\mathscr{G}_n = \sigma(S_n,S_{n+1},\dots)$. *Entonces* $(S_n/n,\mathscr{G}_n)_{n\geq 1}$ *es una martingala inversa.*

*Demostración.* Por la proposición 1.15, $\mathscr{G}_{n+1} = \sigma(S_{n+1},Z_{n+2},\dots)$, y por ser $Z_n$ independientes, usando el teorema 1.12, se tiene

$$E\left(\frac{S_n}{n}\middle|\mathscr{G}_{n+1}\right) = \frac{1}{n}\sum_{k=1}^{n} E\left(Z_k|\mathscr{G}_{n+1}\right) = \frac{1}{n}\sum_{k=1}^{n} E(Z_k|S_{n+1}) = \frac{1}{n}\sum_{k=1}^{n}\frac{S_{n+1}}{n+1} = \frac{S_{n+1}}{n+1},$$

ya que, al ser $Z_k$ idénticamente distribuidas, $E(Z_1|S_{n+1}) = E(Z_k|S_{n+1})$ para cada $k = 1,\dots,n$, luego

$$(n+1)E(Z_1|S_{n+1}) = \sum_{k=1}^{n+1} E(Z_k|S_{n+1}) = E(S_{n+1}|S_{n+1}) = S_{n+1}.$$

□

En [4] y [8], se pueden encontrar más propiedades de martingalas y martingalas inversas, como teoremas de convergencia, descomposiciones o desigualdades, entre otras.

# Capítulo 2

# Representaciones completas de operadores lineales positivos

Sea $I \subset \mathbb{R}$ un intervalo y sea $\mathscr{S}(I)$ una familia de funciones $f : I \to \mathbb{R}$ suficientemente rica. De aquí en adelante, consideraremos sucesiones $(L_n)_{n\geq 1}$ de operadores $L_n : \mathscr{S}(I) \to \mathscr{S}(I)$ lineales y positivos, es decir, tales que $L_n f(x) \geq 0$, si $f(x) \geq 0$, para cualesquiera $n \geq 1$, $x \in I$. En los siguientes capítulos, estudiaremos esencialmente tres tipos de propiedades. En primer lugar, propiedades de preservación, de la forma $L_n(\mathscr{R}(I)) \subset \mathscr{R}(I)$, donde $\mathscr{R}(I) \subset \mathscr{S}(I)$ es un conjunto de funciones en $\mathscr{S}(I)$ que cumplen una cierta propiedad, como el conjunto de funciones monótonas, convexas o de clase Lipschitz, entre otros. En segundo lugar, propiedades de aproximación, es decir, de las funciones $n \to L_n f(x)$, $x \in I$, siendo $f \in \mathscr{R}(I) \subset \mathscr{S}(I)$ una función con una determinada propiedad. Finalmente propiedades de convergencia, es decir,

$$L_n f(x) \to f(x), \qquad x \in I, f \in \mathscr{R}(I),$$

siendo $\mathscr{R}(I) \subset \mathscr{S}(I)$ un conjunto de funciones con ciertas propiedades, como la continuidad o la derivabilidad hasta un cierto orden. Las demostraciones analíticas que se encuentran habitualmente en la literatura dependen fuertemente del tipo de operador y de la propiedad considerada. Sin embargo, podemos conseguir demostraciones unificadas utilizando representaciones probabilísticas de la forma

$$L_n f(x) = E f(Z_n(x)), \qquad x \in I, n \geq 1, \tag{2.1}$$

siendo $(Z_n(x), n \geq 1, x \in I)$ un proceso estocástico a doble índice definido en un espacio de probabilidad completo $(\Omega, \mathscr{F}, P)$ y $f$ una función en el conjunto

$$\mathscr{L} := \mathscr{L}(I) = \{f : I \to \mathbb{R} \text{ medible; } E|f(Z_n(x))| < \infty \text{ para cualesquiera } x \in I, n \geq 1\}.$$

Como se ve en (2.1), en la definición de $L_n f(x)$ solo interviene la ley marginal de la variable aleatoria $Z_n(x)$. Sin embargo, si somos capaces de construir un proceso estocástico $(Z_n(x), n \geq 1, x \in I)$ con fuertes relaciones de dependencia entre las variables, muchas propiedades de preservación y convergencia se pueden estudiar de forma unificada y relativamente simple. Por ejemplo, si una construcción permite determinar fácilmente que $Z_n(x) \to x$ c.s. cuando $n \to \infty$, para cada $x \in I$, entonces los teoremas del cálculo de probabilidades garantizan que $L_n f(x) \to f(x)$, $n \to \infty$, $x \in I$, si la función $f$ tiene propiedades de continuidad adecuadas. Llamaremos *representaciones completas* a aquellas como (2.1) que se construyan con los procesos que veremos en este capítulo. Aunque daremos varios ejemplos de operadores con representaciones completas, en los capítulos siguientes nos centraremos en los polinomios de Bernstein y los operadores de Szász y Weierstrass, para dar una muestra de cómo se aplican los teoremas al considerar operadores sobre intervalos compactos, no acotados por un lado, o sobre $\mathbb{R}$.

## 2.1. Procesos estocásticos y representaciones completas

**Definición.** Un *proceso estocástico* es una familia de variables aleatorias $(Z_t)_{t\in T}$ definidas en un mismo espacio de probabilidad $(\Omega, \mathscr{F}, P)$, siendo $T$ un conjunto de índices arbitrario.

**Definición.** Fijado $\omega \in \Omega$, la aplicación $t \to Z_t(\omega)$ se llama *trayectoria.*

Una martingala es un proceso estocástico con $T = \mathbb{N}$. Otro conjunto típico es $T = [0,\infty)$. Los procesos de Poisson y gamma los definiremos para $T = [0,\infty)$, pero no demostraremos su existencia ni sus propiedades (por ejemplo, que se puede considerar que sus trayectorias son no decrecientes y continuas a derecha), ya que son temas muy avanzados y, además, ocuparían todo un trabajo entero.

**Definición.** Sea $T$ un conjunto ordenado. Se dice que un proceso estocástico $(Z_t)_{t\in T}$ tiene incrementos estacionarios e independientes si, para cualesquiera $t_1 < t_2 < \cdots < t_n$ en $T$, los incrementos $Z_{t_2} - Z_{t_1}, \ldots, Z_{t_n} - Z_{t_{n-1}}$ son independientes, y para cualesquiera $s < t$ en $T$, la ley de $Z_t - Z_s$ depende solo de $t - s$.

Usaremos la notación $Z_t \equiv \mu_t$ para expresar que una variable tiene la ley $\mu_t$.

### 2.1.1. Proceso empírico

Supongamos $U_1, U_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas definidas en $(\Omega, \mathscr{F}, P)$, tales que $U_1 \equiv U(0,1)$, la distribución uniforme en $[0,1]$. Definimos

$$S_n(x) = \sum_{k=1}^{n} \mathbf{1}_{[0,x]}(U_k), \qquad x \in [0,1], n \geq 1,$$

siendo $\mathbf{1}$ la variable *indicador.* El proceso $(S_n(x), n \geq 1, x \in [0,1])$ se llama *proceso empírico.* Por ser $(U_k)_{k\geq 1}$ independientes, se tiene que $S_n(x) \equiv \mathrm{Bin}(n,x)$, ya que

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

### 2.1.2. Proceso de Poisson

Habitualmente, en la literatura se suele definir el proceso de Poisson de esta forma:

**Definición.** Un proceso $(N_t)_{t\geq 0}$ es un *proceso de Poisson* si $N_0 = 0$ c.s. y es un proceso con incrementos estacionarios e independientes tal que, 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,

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

Se puede dar una definición constructiva del proceso de Poisson como sigue: Supongamos $W_1, W_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas con $W_1 \equiv \exp(1)$, siendo esta la distribución exponencial. Para cada $n \geq 1$, definimos $T_n = W_1 + \cdots + W_n$. Entonces, el proceso

$$N_t = \max\{n : T_n \leq t\} = \sum_{n=1}^{\infty} \mathbf{1}_{[0,t]}(T_n), \qquad t \geq 0,$$

es un proceso de Poisson. Podemos pensar que estamos observando un evento que ocurre muchas veces: $T_n$ es el tiempo total hasta que ocurre por $n$-ésima vez, y $W_n$ es el tiempo desde que ocurrió por $(n-1)$-ésima vez hasta la $n$-ésima. No daremos la demostración de que este proceso es un proceso de Poisson, por falta de espacio y porque no es el objetivo principal de este trabajo. Se puede ver en [5].

### 2.1.3. Proceso gamma

Se suele definir en la literatura de la siguiente forma:

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

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

siendo $\Gamma$ la función Gamma de Euler.

Tampoco demostraremos la existencia del proceso gamma, y en este caso no daremos una definición constructiva.

## 2.2. Ejemplos de representaciones completas

Para aplicar algunos teoremas del siguiente capítulo, **necesitaremos que los procesos de la sección anterior tengan trayectorias no decrecientes y continuas a derecha.** En los procesos empíricos y de Poisson, se sigue fácilmente de la definición constructiva. Para el proceso gamma, nombramos sin demostración, pues es muy avanzada, que es posible obtener una *versión* que satisfaga las dos condiciones.

Recordamos que, en los siguientes capítulos, nos centraremos en los polinomios de Bernstein, y los operadores de Szász y Weierstrass. Las representaciones de estos últimos y de los polinomios de Bernstein-Kantorovich, aunque no se construyan con los procesos vistos, también se considerarán completas. Las demás familias de operadores las exponemos para hacer ver la versatilidad de este tipo de representaciones. Además, aunque algunos operadores se pueden definir para $t \geq 0$, nos restringiremos a tiempos discretos. Para cada $n = 1, 2, \ldots$, se definen las siguientes familias de operadores:

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

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

- *Szász.* Para cada $x \geq 0$, se define

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

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

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

  donde $Z \equiv \mathcal{N}(0,1)$ tiene la ley normal estándar.

- *Gamma.* Para cada $x \geq 0$, se define

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

- *Bernstein-Kantorovich.* Para cada $m = 0, 1, \ldots$, $x \in [0,1]$, se define

$$\begin{aligned} K_{n,m} f(x) &:= \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \int_0^1 \cdots \int_0^1 f\left(\frac{k + u_1 + \cdots + u_m}{n+m}\right) du_1 \cdots du_m \\ &= Ef\left(\frac{S_n(x) + V_1 + \cdots + V_m}{n+m}\right), \end{aligned}$$

  siendo $(V_k)_{k\geq 1}$ una sucesión de variables independientes e idénticamente distribuidas con distribución uniforme en $[0,1]$, independientes de las $(U_k)_{k\geq 1}$ del proceso empírico.

Se pueden encontrar más ejemplos de operadores, como los operadores de Baskakov o Stancu, en [3], y los operadores beta y beta a doble índice, entre otros, en [2] y [1].

# Capítulo 3

# Propiedades de preservación y aproximación

En este capítulo estudiaremos varias propiedades de preservación y aproximación de los operadores lineales positivos, aunque nos centraremos en los polinomios de Bernstein y los operadores de Szász y Weierstrass. Es conveniente recordar que una de las aplicaciones de los polinomios de Bernstein es el diseño asistido por ordenador, ya que estos respetan muchas propiedades de forma de la función original, como veremos en las aplicaciones de los teoremas.

Como dijimos en el capítulo anterior, las representaciones probabilísticas completas permiten dar demostraciones unificadas. Veremos su eficacia al aplicarlas a los operadores concretos. La mayor parte de las propiedades que probaremos se pueden encontrar en [2] y [1].

Durante todo el capítulo, supondremos que $I \subset \mathbb{R}$ es un intervalo arbitrario, y que $(L_n)_{n\geq 1}$ es una familia de operadores lineales positivos con representación completa, es decir, de la forma

$$L_n f(x) = Ef(Z_n(x)), \qquad x \in I, n \geq 1,$$

siendo $(Z_n(x), n \geq 1, x \in I)$ un proceso estocástico a doble índice definido en un espacio de probabilidad completo $(\Omega, \mathscr{F}, P)$, construido como en el capítulo anterior, y $f : I \to \mathbb{R}$ una función tal que $f \in \mathscr{L}$, es decir, una función para la que estén definidos los aproximantes.

## 3.1. Propiedades de preservación

En toda esta sección, salvo que se diga lo contrario, supondremos $n \geq 1$ fijo, y denotaremos

$$Z(x) := Z_n(x), \qquad Lf(x) := L_n f(x) = Ef(Z(x)), \qquad x \in I.$$

### 3.1.1. Monotonía y $\varphi$-variación

**Definición.** Sea $\Phi = \{\varphi : [0,\infty) \to \mathbb{R}$ convexa y estrictamente creciente : $\varphi(0) = 0\}$. Para cada $\varphi \in \Phi$, definimos la *$\varphi$-variación* de $f$ como la función

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

donde el supremo se toma sobre el conjunto de sucesiones finitas de puntos $x_0 \leq x_1 \leq \cdots \leq x_n$, $n \in \mathbb{N}$, tales que $x_0, x_n \in I$ y $x_i - x_{i-1} \leq \varepsilon$, $i = 1, \ldots, n$.

**Definición.** Llamaremos *$\varphi$-variación total* de $f$ a la cantidad $V_\varphi f := V_\varphi f(\infty)$, posiblemente infinita. Notar que si $I$ es un intervalo compacto y $V_\varphi f < \infty$, entonces $f \in \mathscr{L}$.

Los ejemplos más habituales son la $1-$variación o variación usual, con $\varphi(x) = x$, y la $p-$variación, cuando $\varphi(x) = x^p$, con $p \geq 1$. En [6] y [10], se prueba la preservación de la monotonía y de la $1-$variación para los polinomios de Bernstein, calculando $(B_n f)'$, mientras que nosotros probaremos un resultado válido para muchos operadores. La preservación de la monotonía tiene un significado evidente, mientras que la preservación de la $\varphi-$variación se puede interpretar diciendo que los aproximantes tienen menores fluctuaciones que la función original.

**Teorema 3.1.** *Supongamos que, para cualesquiera $x, y \in I$ con $x \leq y$, se satisface*

$$Z(x) \leq Z(y), \qquad \text{c.s.} \tag{3.1}$$

*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 $f$ es no decreciente. Entonces (3.1) 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 $x_0 \leq x_1 \leq \cdots \leq x_n$ puntos tales que $x_0, x_n \in I$. Por la hipótesis (3.1), se tiene

$$Z(x_0) \leq Z(x_1) \leq \cdots \leq Z(x_n), \qquad \text{c.s.}$$

Utilizando, en este orden, que $\varphi$ es creciente, la desigualdad de Jensen (por ser $\varphi$ convexa), la linealidad de la esperanza, y la definición de variación total, se tiene

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

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

**Corolario 3.2.** *Los polinomios de Bernstein y los operadores de Szász y Weierstrass preservan la monotonía y la $\varphi$-variación total.*

*Demostración.* Por la construcción de los procesos empíricos, fijado $n \geq 1$, se tiene $S_n(x) \leq S_n(y)$, si $0 \leq x \leq y \leq 1$. Lo mismo ocurre con el proceso de Poisson, ya que, por la construcción que dimos del proceso, $N_{nx} \leq N_{ny}$, si $0 \leq x \leq y < \infty$. Y para los operadores de Weierstrass, se cumple trivialmente, puesto que $x + \frac{Z}{\sqrt{n}} \leq y + \frac{Z}{\sqrt{n}}$, para cualesquiera $-\infty < x < y < \infty$. □

### 3.1.2. Constantes de Lipschitz y módulo de continuidad

Recordamos la definición del módulo de continuidad y sus principales propiedades, enunciadas en la proposición 3.3, cuyas demostraciones se han dado ya durante el grado.

**Definición.** Se define el *módulo de continuidad* de $f$ como la función

$$\omega(f, \delta) := \sup\left\{\left|f(x+h) - f(x)\right| : x, x+h \in I,\ 0 \leq h \leq \delta\right\}, \qquad \delta \geq 0.$$

**Proposición 3.3.** *El módulo de continuidad es no decreciente y subaditivo, y dados $\alpha, \delta \geq 0$, verifica :*

*(a)* $\omega(f, \alpha\delta) \leq (1 + \alpha)\omega(f, \delta)$.

*(b)* $\omega(f,\alpha\delta) \leq (1+\alpha^2)\omega(f,\delta)$.

*(c)* *$f$ es uniformemente continua si y solo si* $\lim_{\delta\to 0}\omega(f,\delta) = \omega(f,0) = 0$.

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

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

Usaremos la notación $f \in \mathrm{Lip}(A,\alpha)$. Notar que si $f \in \mathrm{Lip}(A,\alpha)$, entonces $f$ es continua y además $\omega(f,\delta) \leq A\delta^{\alpha}$, $\delta \geq 0$.

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 no las presenta la función original. Es decir, respetan la suavidad de la función original.

**Teorema 3.4.** *Supongamos que, para cualesquiera $x,y \in I$, con $x \leq y$, se satisface* (3.1) *y, además,*

$$EZ(x) = cx + d, \tag{3.2}$$

*siendo $c > 0$ y $d$ constantes que pueden depender de $n$. Entonces se verifica:*

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

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

*(c)* *Si $\omega(f,\cdot)$ es una función cóncava, entonces*

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

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

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

Para probar (b) y (c), sean $\delta > 0$, $x,y \in I$ con $0 \leq y-x \leq \delta$. Se tiene

$$|Lf(y) - Lf(x)| \leq E\,|f(Z(y)) - f(Z(x))| \leq E\omega(f,Z(y)-Z(x)). \tag{3.3}$$

Utilizando la propiedad 3.3 (a), 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 (3.3), y se tiene

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

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

**Corolario 3.5.** *Los polinomios de Bernstein y los operadores de Szász y Weierstrass preservan la suavidad global.*

*Demostración.* La monotonía de sus procesos se ha probado en el corolario 3.2. Para la condición (3.2), notar que, para cada $x \in I$, $ES_n(x) = nx = EN_{nx}$, $EZ = 0$, por lo que se tiene

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

Por tanto, se puede aplicar el teorema 3.4 en los tres casos. □

### 3.1.3. Convexidad

La preservación de la convexidad es una propiedad importante en el diseño asistido por ordenador que satisfacen, entre otros, los polinomios de Bernstein. Es otra propiedad de forma que se debería respetar si se quiere obtener una aproximación similar a la función de partida. En [6] y [10] se puede ver una prueba para los polinomios de Bernstein, calculando $(B_n f)''$.

**Teorema 3.6.** *Supongamos que, para cualesquiera $x, y \in I$ con $x \le y$, se satisface* (3.1) *y que, para cualesquiera $x < y$ en $I$, existe una $\sigma$-álgebra $\mathscr{G}(x,y) \subset \mathscr{F}$ tal que $Z(x)$ y $Z(y)$ son $\mathscr{G}(x,y)$-medibles y se verifica*

$$E\left(Z(v) - Z(x) | \mathscr{G}(x,y)\right) = \frac{v-x}{y-x}(Z(y) - Z(x)), \qquad x < v < y. \tag{3.4}$$

*Entonces $Lf$ tiene la misma convexidad que $f$.*

*Demostración.* Sean $x, v, y \in I$ tales que $x < v < y$, y supongamos que $f$ es convexa. Veamos que

$$E\left(f(Z(v)) | \mathscr{G}(x,y)\right) \le \frac{y-v}{y-x} f(Z(x)) + \frac{v-x}{y-x} f(Z(y)). \tag{3.5}$$

En efecto, notar que $\{Z(x) = Z(y)\} \in \mathscr{G}(x,y)$ por ser ambas variables aleatorias $\mathscr{G}(x,y)$-medibles, y que en este suceso se verifica la igualdad, pues $Z(x) = Z(v) = Z(y)$ c.s., y $Z(x)$ es $\mathscr{G}(x,y)$-medible, luego, utilizando la propiedad 1.6 (b) de esperanzas condicionales, en este suceso se tiene

$$E\left(f(Z(v)) | \mathscr{G}(x,y)\right) = E\left(f(Z(x)) | \mathscr{G}(x,y)\right) = f(Z(x)) = f(Z(v)).$$

Recordar que la esperanza condicional es única c.s. Por otro lado, en el suceso $\{Z(x) < Z(y)\}$, que está en $\mathscr{G}(x,y)$ por la misma razón que antes, podemos poner

$$Z(v) = \frac{Z(y) - Z(v)}{Z(y) - Z(x)} Z(x) + \frac{Z(v) - Z(x)}{Z(y) - Z(x)} Z(y), \tag{3.6}$$

ya que los denominadores no se anulan; basta multiplicar por $Z(y) - Z(x)$ en ambos miembros para comprobar la igualdad. Además, observar que, debido a la monotonía de las variables aleatorias,

$$0 \le Z(y) - Z(v), Z(v) - Z(x) \le Z(y) - Z(x), \qquad \text{c.s.,}$$

por lo que $Z(v)$ está escrito en (3.6) como una combinación convexa de $Z(x)$ y $Z(y)$. Por ser $f$ convexa, podemos utilizar la desigualdad de Jensen, y se obtiene:

$$f(Z(v)) \le \frac{Z(y) - Z(v)}{Z(y) - Z(x)} f(Z(x)) + \frac{Z(v) - Z(x)}{Z(y) - Z(x)} f(Z(y)). \tag{3.7}$$

Notar que (3.4) se satisface trivialmente cuando $v = y$ por la propiedad 1.6 (b), y por linealidad, se tiene

$$E\left(Z(y) - Z(v) | \mathscr{G}(x,y)\right) = E\left(Z(y) - Z(x) | \mathscr{G}(x,y)\right) - E\left(Z(v) - Z(x) | \mathscr{G}(x,y)\right) = \frac{y-v}{y-x}(Z(y) - Z(x)).$$

Tomando esperanzas condicionales en (3.7) y aplicando la monotonía, la linealidad, el teorema 1.8 (notar que $f(Z(x)), f(Z(y))$ son $\mathscr{G}(x,y)$-medibles) y la hipótesis (3.4), se obtiene

$$\begin{aligned} E\left(f(Z(v)) | \mathscr{G}(x,y)\right) &\le E\left( \frac{Z(y) - Z(v)}{Z(y) - Z(x)} f(Z(x)) + \frac{Z(v) - Z(x)}{Z(y) - Z(x)} f(Z(y)) \middle| \mathscr{G}(x,y) \right) \\ &= E\left( \frac{Z(y) - Z(v)}{Z(y) - Z(x)} \middle| \mathscr{G}(x,y) \right) f(Z(x)) + E\left( \frac{Z(v) - Z(x)}{Z(y) - Z(x)} \middle| \mathscr{G}(x,y) \right) f(Z(y)) \\ &= \frac{y-v}{y-x} f(Z(x)) + \frac{v-x}{y-x} f(Z(y)). \end{aligned}$$

Con esto, hemos visto que se cumple (3.5). Finalmente, tomando en ella esperanzas usuales a ambos lados y aplicando el teorema 1.9, se tiene

$$Lf\left(\frac{y-v}{y-x}x+\frac{v-x}{y-x}y\right)=Lf(v)=Ef(Z(v))\leq\frac{y-v}{y-x}Lf(x)+\frac{v-x}{y-x}Lf(y),$$

lo que prueba la convexidad de $Lf$. Si $f$ es cóncava, considerar $-f$, que es convexa. □

**Corolario 3.7.** *Los polinomios de Bernstein preservan la convexidad.*

*Demostración.* Por la construcción de los procesos empíricos, si $0\leq x\leq y\leq 1$, entonces $S_n(x)\leq S_n(y)$. Sean $0\leq x<v<y\leq 1$. Llamemos

$$X:=S_n(v)-S_n(x),\qquad Y:=S_n(y)-S_n(v).$$

Notar que $(X,Y)\equiv\mathscr{M}(n,v-x,y-v,1-(y-x))$, siendo esta una distribución trinomial, es decir,

$$P(X=i,Y=j)=\frac{n!}{i!j!(n-(i+j))!}(v-x)^i(y-v)^j(1-(y-x))^{n-(i+j)},\qquad i,j\geq 0,\ i+j\leq n.$$

Por tanto, se tiene $X+Y\equiv\mathrm{Bin}(n,y-x)$, puesto que

$$P(X+Y=k)=\sum_{i=0}^{k}P(X=i,Y=k-i)=\binom{n}{k}(y-x)^k(1-(y-x))^{n-k},\qquad k=0,1,\ldots,n.$$

Además, la ley de $X$ condicionada a $X+Y=k$ es Bin$\left(k,\frac{v-x}{y-x}\right)$, $k=0,1,\ldots,n$, ya que

$$P(X=i|X+Y=k)=\frac{P(X=i,Y=k-i)}{P(X+Y=k)}=\binom{k}{i}\left(\frac{v-x}{y-x}\right)^i\left(\frac{y-v}{y-x}\right)^{k-i},\qquad i=0,1,\ldots,k.$$

De aquí, por el corolario 1.4, se sigue $E(X|X+Y=k)=k\frac{v-x}{y-x}$. Utilizando el teorema 1.5, se tiene

$$E(X|X+Y)=\frac{v-x}{y-x}(X+Y).$$

Tomando en ambos lados esperanzas condicionales respecto a $\mathscr{G}(x,y)=\sigma\{S_n(x),S_n(y)\}$ y utilizando el teorema 1.9 y la propiedad 1.6 (b), se tiene

$$E(S_n(v)-S_n(x)|S_n(y)-S_n(x))=\frac{v-x}{y-x}(S_n(y)-S_n(x)).$$

□

**Corolario 3.8.** *Los operadores de Szász y Weierstrass preservan la convexidad.*

*Demostración.* La monotonía de los procesos se ha visto en el corolario 3.2. Veamos que se satisface la condición (3.4). En el caso de los operadores de Weierstrass, sean $-\infty<x<v<y<\infty$. Notar que $Z(x)=x+\frac{Z}{\sqrt{n}}$, donde $Z$ tiene la ley normal estándar, luego $Z(v)-Z(x)=v-x$, por lo que (3.4) se verifica trivialmente.
Para los operadores de Szász, sean $0\leq x<v<y<\infty$, y sea $\mathscr{G}(x,y)=\sigma\{\frac{N_{nx}}{n},\frac{N_{ny}}{n}\}$. Por la proposición 1.15, se tiene

$$\mathscr{G}(x,y)=\sigma\{N_{nx},N_{ny}\}=\sigma\{N_{nx},N_{ny}-N_{nx}\},$$

y además, como los incrementos del proceso de Poisson son independientes y $N_0=0$ c.s., el vector aleatorio $(N_{nv}-N_{nx},N_{ny}-N_{nx})$ es independiente de $N_{nx}=N_{nx}-N_0$. Aplicando el teorema 1.12, se tiene

$$E(N_{nv}-N_{nx}|N_{nx},N_{ny}-N_{nx})=E(N_{nv}-N_{nx}|N_{ny}-N_{nx}).$$

Notar que $N_{nv} - N_{nx} \le N_{ny} - N_{nx}$ y que ambas toman valores en $\{0,1,2,\dots\}$. Llamemos

$$X = N_{nv} - N_{nx}, \qquad Y = N_{ny} - N_{nv}.$$

Observar que $X$ e $Y$ son independientes, y recordar que los incrementos del proceso de Poisson tienen la ley de Poisson con media la diferencia de los subíndices. Además, fijado $k \ge 0$, se tiene

$$\begin{aligned} P(X = i|X+Y = k) &= \frac{P(X = i, X+Y = k)}{P(X+Y = k)} = \frac{P(X = i, Y = k-i)}{P(X+Y = k)} = \frac{P(X = i)P(Y = k-i)}{P(X+Y = k)} \\ &= \frac{\left(e^{-n(v-x)}(n(v-x))^i/i!\right)\left(e^{-n(y-v)}(n(y-v))^{k-i}/(k-i)!\right)}{e^{-n(y-x)}(n(y-x))^k/k!} \\ &= \binom{k}{i}\left(\frac{v-x}{y-x}\right)^i\left(\frac{y-v}{y-x}\right)^{k-i}, \qquad i = 0,1,\dots k, \end{aligned}$$

luego $E(X|X+Y = k) = k\frac{v-x}{y-x}$. es la esperanza de una variable aleatoria con ley Bin $\left(k, \frac{v-x}{y-x}\right)$. Aplicando el teorema 1.5, se tiene

$$E(N_{nv} - N_{nx}|N_{ny} - N_{nx}) = E(X|X+Y) = (X+Y)\frac{v-x}{y-x} = (N_{ny} - N_{nx})\frac{v-x}{y-x}.$$

□

**Observación.** El argumento anterior se puede generalizar para procesos con incrementos estacionarios e independientes con $Z_0 = 0$ c.s.

## 3.2. Propiedades de los caminos de aproximación

Durante toda la sección, salvo que se diga lo contrario, supondremos $x \in I$ fijo, y denotaremos

$$Z_n := Z_n(x), \qquad p(n) := L_n f(x) = Ef(Z_n), \qquad n \ge 1.$$

Llamaremos a la función $p(\cdot)$ *camino de aproximación.*

### 3.2.1. Convergencia monótona bajo convexidad

La convergencia monótona es otra propiedad importante en el diseño asistido por ordenador que verifican los polinomios de Bernstein. Se puede ver una prueba analítica para estos en [6]. Si la función que queremos aproximar está dentro de la pantalla, el aproximante también debería estar. La convergencia monótona prueba que, si la función es convexa, la aproximación se hace por encima de la función.

**Teorema 3.9.** *Sea $(\mathscr{G}_n)_{n\ge1}$ una sucesión decreciente de $\sigma$-álgebras en $\mathscr{F}$ y $f$ una función convexa. Supongamos que se verifica una de las siguientes condiciones:*

*a)* $(Z_n, \mathscr{G}_n)_{n\ge1}$ *es una martingala inversa.*

*b)* $f$ *es no decreciente y* $(Z_n, \mathscr{G}_n)_{n\ge1}$ *es una submartingala inversa.*

*c)* $f$ *es no creciente y* $(Z_n, \mathscr{G}_n)_{n\ge1}$ *es una supermartingala inversa.*

*Entonces la función $p(\cdot)$ es no creciente.*

*Demostración.* Sean $1 \le n \le m$. Para probar (a), basta usar la condición de martingala inversa, la desigualdad de Jensen y el teorema 1.9:

$$p(m) = Ef(Z_m) = Ef\left(E(Z_n|\mathscr{G}_m)\right) \le E(E(f(Z_n)|\mathscr{G}_m)) = Ef(Z_n) = p(n).$$

De la misma forma, se tienen (b) y (c), cambiando la segunda igualdad por una desigualdad $\le$ . □

**Observación.** El teorema también se cumple sin exigir que $(\mathcal{G}_n)_{n\geq 1}$ sea una sucesión decreciente. Basta con que $Z_n$ sea $\mathcal{G}_n$-medible y se cumplan las igualdades o desigualdades de esperanzas condicionales correspondientes. Por otro lado, observar que si $f$ es cóncava, entonces $p(\cdot)$ es no decreciente.

**Corolario 3.10.** *Los polinomios de Bernstein y los operadores de Szász y de Weierstrass tienen caminos de aproximación no crecientes, si $f$ es convexa.*

*Demostración.* Para los polinomios de Bernstein, sea $\mathcal{G}_n = \sigma\{S_n(x), S_{n+1}(x), \ldots\}$, $n \geq 1$. Recordar que

$$S_n(x) = \mathbf{1}_{[0,1]}(U_1) + \cdots + \mathbf{1}_{[0,1]}(U_n), \qquad n \geq 1,$$

siendo $(U_n)_{n\geq 1}$ variables aleatorias independientes e idénticamente distribuidas con ley uniforme en $[0,1]$, por lo que $\left(\mathbf{1}_{[0,x]}(U_n)\right)_{n\geq 1}$ es una sucesión de variables independientes e idénticamente distribuidas, y podemos aplicar la proposición 1.16 para concluir que $(S_n(x)/n, \mathcal{G}_n)_{n\geq 1}$ es una martingala inversa.
Para los operadores de Szász, sea $\mathcal{G}_n = \sigma\{N_{nx}, N_{(n+1)x}, \ldots\}$, $n \geq 1$. En este caso, podemos poner

$$N_{nx} = (N_{nx} - N_{(n-1)x}) + (N_{(n-1)x} - N_{(n-2)x}) + \cdots + (N_x - N_0), \qquad n \geq 1,$$

y por las características del proceso de Poisson, se tiene que $\left(N_{nx} - N_{(n-1)x}\right)_{n\geq 1}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas. Aplicando la proposición 1.16, se tiene que $(N_{nx}/n, \mathcal{G}_n)_{n\geq 1}$ es una martingala inversa.
En el caso de los operadores de Weierstrass, notar que si $(X_j)_{j\geq 1}$ son variables aleatorias independientes e idénticamente distribuidas con ley $\mathcal{N}(0,1)$, entonces las variables $Z/\sqrt{n}$ y $(X_1 + \cdots + X_n)/n$ tienen la misma ley normal de media 0 y varianza $1/n$, por lo que

$$W_n f(x) = Ef\left(x + \frac{Z}{\sqrt{n}}\right) = Ef\left(x + \frac{X_1 + \cdots + X_n}{n}\right).$$

Por la proposición 1.16 $((X_1 + \cdots + X_n)/n, \mathcal{G}_n)_{n\geq 1}$ es una martingala inversa, y la conclusión se sigue por la linealidad de la esperanza condicional. □

**Observación.** El argumento anterior se puede extender a cualquier proceso con incrementos estacionarios e independientes.

### 3.2.2. Propiedades de tipo Lipschitz

Si $f \in \mathrm{Lip}(A, \alpha)$, los caminos de aproximación también verifican propiedades de tipo Lipschitz.

**Proposición 3.11.** *Sea $f \in \mathrm{Lip}(A, \alpha)$. Los polinomios de Bernstein y los operadores de Szász y de Weierstrass verifican:*

$$|p(m) - p(n)| \leq A\left(\frac{2x(m-n)}{m}\right)^{\alpha}, \qquad 1 \leq n \leq m.$$

*Demostración.* Llamemos $C = A\left(\frac{m-n}{m}\right)^{\alpha}$. Como el proceso de Poisson tiene incrementos estacionarios y $N_0 = 0$ c.s., las variables aleatorias $N_{mx} - N_{nx}$ y $N_{(m-n)x}$ tienen la misma distribución. Aplicando la condición de Lipschitz y la desigualdad de Jensen, se tiene

$$\begin{aligned}|p(m) - p(n)| &\leq E\left|f\left(\frac{N_{mx}}{m}\right) - f\left(\frac{N_{nx}}{n}\right)\right| \leq AE^{\alpha}\left|\frac{N_{mx}}{m} - \frac{N_{nx}}{n}\right| = CE^{\alpha}\left|\frac{N_{mx}}{m-n} - \frac{mN_{nx}}{n(m-n)}\right| \\ &= CE^{\alpha}\left|\frac{N_{mx} - N_{nx}}{m-n} - \frac{N_{nx}}{n}\right| = CE^{\alpha}\left|\frac{N_{(m-n)x}}{m-n} - \frac{N_{nx}}{n}\right| \leq CE^{\alpha}\left(\frac{N_{(m-n)x}}{m-n} + \frac{N_{nx}}{n}\right) = C(2x)^{\alpha}.\end{aligned}$$

Para los polinomios de Bernstein, se puede razonar igual, cambiando $N_{nx}$ por $S_n(x)$, ya que $S_m(x) - S_n(x) = \sum_{k=n+1}^{m} \mathbf{1}_{[0,x]}(U_k)$ tiene la misma distribución que $S_{m-n}(x)$, por ser $(U_k)_{k\geq 1}$ independientes. Y para los operadores de Weierstrass también, puesto que, con la notación del corolario 3.10, se tiene que $X_{n+1} + \cdots + X_m$ y $X_1 + \cdots + X_{m-n}$ tienen la misma ley, ya que $(X_k)_{k\geq 1}$ son variables aleatorias independientes e idénticamente distribuidas con ley normal estándar. □

**Observación.** Este teorema se puede extender a cualquier operador cuyo proceso tenga incrementos estacionarios, con $Z_0 = 0$ c.s.

# Capítulo 4

# 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. Probaremos 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. Demostraremos un teorema de velocidad de convergencia en términos del segundo módulo de continuidad usual, mediante una aproximación intermedia por *medias de Steklov*. 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.

## 4.1. Velocidad en términos del primer módulo de continuidad

Para cada $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, denotaremos $\beta_m$ a una variable aleatoria absolutamente continua con densidad $\rho(\theta) = m(1-\theta)^{m-1}, \theta \in [0,1]$, y llamaremos

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

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

**Teorema 4.1.** ***(Fórmula de Taylor probabilística)*** *Sea $m \in \mathbb{N}_0$, $f \in \mathscr{C}^{(m)}(I)$ y $X$ una variable aleatoria independiente de $\beta_m$ 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.* La fórmula de Taylor de orden $m$ de $f$ con resto integral se puede escribir así:

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

Por la subaditividad del módulo de continuidad (recordar 3.3), para cada $y \in I$, $\delta > 0$, se tiene

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

Sustituyendo $y$ por $X$, que toma valores en $I$ y es independiente de $\beta_m$, y tomando esperanzas y módulos, se tiene

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

Por ser $\beta_m$ independiente de $X$, se tiene

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

lo que completa la demostración. □

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

$$\mu_{n,j}(x)=E\left(\frac{S_n(x)}{n}-x\right)^j,\qquad M_{n,j}(x)=E\left|\frac{S_n(x)}{n}-x\right|^j.$$

**Corolario 4.2.** *Sea $m\in\mathbb{N}_0$ y $f\in\mathscr{C}^{(m)}[0,1]$. Para cada $x\in(0,1)$, se tiene*

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

*Demostración.* Basta aplicar el teorema 4.1 con $I=[0,1]$, $X=S_n(x)/n$, $\delta=1/\sqrt{n}$. □

**Nota.** Utilizando el teorema central del límite, se tiene

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

donde $Z$ es una variable aleatoria con distribución normal estándar. Es decir, el orden de magnitud de $M_{n,j}(x)$ es $n^{-j/2}$ cuando $n\to\infty$. Por tanto, el resto del corolario 4.2 tiene orden de magnitud

$$\frac{1}{n^{m/2}}\omega\left(f^{(m)},\frac{1}{\sqrt{n}}\right).$$

**Corolario 4.3.** *Sea $f\in\mathscr{C}[0,1]$. Para cada $x\in(0,1)$, se tiene*

$$|B_nf(x)-f(x)|\le\frac{5}{4}\omega\left(f,\frac{1}{\sqrt{n}}\right).$$

*Demostración.* Utilizando el corolario 4.2 con $m=0$, como $M_{n,0}(x)=1$, $M_{n,2}(x)=\frac{x(1-x)}{n}$, se tiene

$$|B_nf(x)-f(x)|\le(1+x(1-x))\,\omega\left(f,\frac{1}{\sqrt{n}}\right)\le\frac{5}{4}\omega\left(f,\frac{1}{\sqrt{n}}\right).$$

□

Se puede ver una demostración analítica de este resultado en [10].

**Corolario 4.4.** ***(Fórmula de Voronovskaja)*** *Sea $f\in\mathscr{C}^{(2)}[0,1]$. Para cada $x\in(0,1)$, se tiene*

$$\left|B_nf(x)-f(x)-\frac{\sigma^2(x)}{2n}f''(x)\right|\le\frac{\sigma^2(x)}{2n}\left(1+\frac{\sigma^2(x)}{2}+\frac{1-6\sigma^2(x)}{6n}\right)\omega\left(f'',\frac{1}{\sqrt{n}}\right),\tag{4.1}$$

*siendo $\sigma(x)=\sqrt{x(1-x)}$, $x\in(0,1)$.*

*Demostración.* Utilizando el corolario 4.2 con $m = 2$, habiendo calculado previamente los momentos

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

se tiene el resultado. □

**Nota.** Se pueden dar cotas similares a las de los corolarios 4.2, 4.3 y 4.4 para otros operadores, tomando $X = Z_n(x)$, el proceso correspondiente al operador (véase la sección 2.2). 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.

## 4.2. Segundo módulo de continuidad de Ditzian-Totik

En esta sección veremos 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. Por ahora, supondremos $I \subset \mathbb{R}$ un intervalo arbitrario y $f : I \to \mathbb{R}$.

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

$$T_h f(x) := f(x+h), \qquad \Delta_h f(x) := (T_h - J)f(x) = f(x+h) - f(x),$$

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

Para cada $m \in \mathbb{N}_0$, definimos sus iteradas $T_h^m$ y $\Delta_h^m$ como

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

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

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

**Lema 4.5.** *Sea $m \in \mathbb{N}_0$ y $h \in \mathbb{R}$. Si $f \in \mathscr{C}^{(m)}(I)$, entonces*

$$\Delta_h^m f(x) = h^m E f^{(m)}(x + hS_m),$$

*siempre que $x, x+mh \in I$, siendo $(U_j)_{j\geq 1}$ variables aleatorias independientes e idénticamente distribuidas tales que $U_1 \equiv U(0,1)$, y $S_m = U_1 + \cdots + U_m$, con $S_0 = 0$.*

*Demostración.* Para $m = 0$, la igualdad es trivial. Supongamos que es cierta para $m \in \mathbb{N}_0$. Si $f \in \mathscr{C}^{(m+1)}(I)$, se tiene que

$$\begin{aligned}\Delta_h^{m+1} f(x) &= \Delta_h (\Delta_h^m f(x)) = \Delta_h \left( h^m E f^{(m)}(x+hS_m) \right) = h^m E \left( \Delta_h f^{(m)}(x+hS_m) \right) \\ &= h^m E \left( \int_0^h f^{(m+1)}(x+hS_m+t)dt \right) = h^{m+1} E f^{(m+1)}(x+hS_{m+1}).\end{aligned}$$

□

Esta igualdad se puede escribir analíticamente de esta forma:

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

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

$$\omega_2(f,\delta) := \sup\left\{ \left|\Delta_h^2 f(x-h)\right| : x-h, x+h \in I,\ 0 \le h \le \delta \right\}, \qquad \delta \ge 0.$$

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

**Definición.** Se define el *segundo módulo de continuidad de Ditzian-Totik con peso* $\sigma$ de $f$ como

$$\omega_\sigma^2(f,\delta) := \sup\left\{ \left|\Delta_{h\sigma(x)}^2 f(x-h\sigma(x))\right| : x-h\sigma(x), x+h\sigma(x) \in I,\ 0 \le h \le \delta \right\}, \qquad \delta \ge 0.$$

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

De aquí en adelante, nos centraremos en los polinomios de Bernstein. Supondremos $I = [0,1]$ y tomaremos $\sigma(x) = \sqrt{x(1-x)}$, $x \in [0,1]$. Nos restringiremos a funciones $f \in \mathscr{C}[0,1]$, y denotaremos

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

Aunque $\|f\|$ es una cantidad asociada a la función, en ocasiones escribiremos $\|f(x)\|$ por cuestiones de claridad. Recordar además la desigualdad triangular inversa,

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

**Lema 4.6.** *Sea $f \in \mathscr{C}^{(2)}[0,1]$. Entonces, se verifica:*

$$\left| \omega_\sigma^2(f,\delta) - \delta^2 \left\|\sigma^2 f''\right\| \right| \le \frac{\delta^2}{8}\omega_\sigma^2(f'',\delta), \qquad \delta \ge 0.$$

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

$$\begin{aligned} f(x-h) &= f(x) - f'(x)h + \frac{f''(x)}{2}h^2 + \frac{h^2}{2}E(f''(x-h\beta_2) - f''(x)), \\ f(x+h) &= f(x) + f'(x)h + \frac{f''(x)}{2}h^2 + \frac{h^2}{2}E(f''(x+h\beta_2) - f''(x)). \end{aligned}$$

Sumando ambas expresiones, se verifica que

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

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

$$\begin{aligned} \left| \omega_\sigma^2(f,\delta) - \delta^2 \left\|\sigma^2 f''\right\| \right| &\le \left\| \Delta_{h\sigma(x)}^2 f(x-h\sigma(x)) - h^2\sigma^2(x)f''(x) \right\| \\ &\le \frac{h^2}{2}\left\|\sigma^2(x)\right\| \omega_\sigma^2(f'',\delta) \\ &\le \frac{\delta^2}{8}\omega_\sigma^2(f'',\delta), \end{aligned}$$

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

**Observación.** Tomando $\delta = 1/\sqrt{n}$, se sigue que el orden de magnitud de $\omega_\sigma^2(f, 1/\sqrt{n})$ es $1/n$.

**Teorema 4.7.** *Sea $f \in \mathscr{C}^{(2)}[0,1]$. Entonces, para cada $n \in \mathbb{N}$, se verifica:*

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

*Además, si $f$ no es lineal, se tiene*

$$\lim_{n\to\infty} \frac{\|B_n f - f\|}{\omega_\sigma^2(f, 1/\sqrt{n})} = \frac{1}{2}.$$

*Demostración.* Utilizando las desigualdades triangulares y el lema 4.6, con $\delta = 1/\sqrt{n}$, se tiene

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

Utilizando la desigualdad (4.1), se tiene

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

puesto que $0 \leq \sigma(x) \leq 1/4$ 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 $0 \leq h \leq 1/\sqrt{n}$ tal que

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

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. □

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

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

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

## 4.3. Medias de Steklov

En esta sección, daremos dos lemas auxiliares para la siguiente sección. Supondremos $I = [0,1]$ y $f : [0,1] \to \mathbb{R}$, pero se podría considerar cualquier intervalo arbitrario.

**Definición.** 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

$$P_{m,h} f(x) = (-1)^{m-1} \sum_{k=1}^{m} \binom{m}{k} (-1)^{m-k} E f\left(x + kh(V-x)\right), \qquad x \in [0,1]. \tag{4.2}$$

A los operadores de la forma de $P_{m,h}$ 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 $[0,1]$. Observar además que

$$P_{m,h} f(x) - f(x) = (-1)^{m-1} E \Delta^m_{h(V-x)} f(x), \qquad x \in [0,1]. \tag{4.3}$$

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

$$V = \frac{U_1 + \cdots + U_n}{n}, \qquad n \in \mathbb{N}. \tag{4.4}$$

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 generaliza la de otros autores como Petrushev y Popov en [12]. En lo que resta de sección, fijaremos $V$ como en (4.4), con $n = 2$, y denotaremos

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

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

$$\begin{aligned} P_h f(x) &= 2E f\left(x + h\left(\frac{U_1 + U_2}{2} - x\right)\right) - E f\left(x + 2h\left(\frac{U_1 + U_2}{2} - x\right)\right) \\ &= \frac{1}{h^2}\left(8\Delta^2_{h/2} f_{(2)}(x(1-h)) - \Delta^2_h f_{(2)}(x(1-2h))\right) \end{aligned}$$

*Demostración.* La primera igualdad se tiene directamente por (4.2), y la segunda por el lema 4.5. □

**Lema 4.9.** *Sean $0 < h \leq 1/2$, $f \in \mathscr{C}[0,1]$ y $\omega_2(f,\cdot)$ el segundo módulo de continuidad usual de $f$. Entonces, para cada $x \in [0,1]$, se tiene*

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

*b)* $|(P_h f)''(x)| \leq \dfrac{1}{h^2}\left(8\omega_2(f,h/2) + \omega_2(f,h)\right)$.

*Demostración.* La parte (a) se sigue de la igualdad (4.3), con $m = 2$, ya que

$$|P_h f(x) - f(x)| = \left|E \Delta^2_{h(V-x)} f(x)\right| \leq \omega_2(f,h), \qquad x \in [0,1],$$

puesto que $|V - x| \leq 1$ para todo $x \in [0,1]$. La parte (b) se sigue de la segunda igualdad del lema 4.8. □

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

## 4.4. Velocidades en términos de segundos módulos de continuidad

### 4.4.1. Velocidad en términos del segundo módulo usual

**Teorema 4.10.** *Sea $f \in \mathscr{C}[0,1]$. Entonces, para cada $n \geq 4$, se tiene:*

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

*Como consecuencia, se tiene*

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

*Demostración.* Sean $n \geq 4$, $x \in [0,1]$, $0 \leq h \leq 1/2$. 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 4.9 (a), se tienen las desigualdades

$$|B_n f(x) - B_n P_h f(x)| \leq E \left| f\left(\frac{S_n(x)}{n}\right) - P_h f\left(\frac{S_n(x)}{n}\right) \right| \leq \omega_2(f,h), \tag{4.5}$$

$$|P_h f(x) - f(x)| \leq \omega_2(f,h). \tag{4.6}$$

Utilizando la fórmula de Taylor probabilística, 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 4.9, se obtiene:

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

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

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

Notar que $0 \leq h \leq 1/2$, por ser $n \geq 4$. Para este $h$, la última desigualdad queda:

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

Sumando a esto (4.5) y (4.6), se obtiene la cota puntual. La cota uniforme se obtiene tomando supremos en ambos miembros de la cota puntual. □

### 4.4.2. Velocidad en términos del segundo módulo de Ditzian-Totik

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

**Teorema 4.11.** *Si $f \in \mathscr{C}[0,1]$ y $\sigma(x) = \sqrt{x(1-x)}, x \in [0,1]$, entonces*

$$K\omega_\sigma^2 \left(f, \frac{1}{\sqrt{n}}\right) \leq \|B_n f - f\| \leq \frac{5}{2} \omega_\sigma^2 \left(f, \frac{1}{\sqrt{n}}\right), \qquad n \in \mathbb{N},$$

*para alguna constante absoluta $K$.*

La desigualdad inferior fue probada por Ditzian e Ivanov en 1993 (véase [7] y [9]), sin dar un valor de $K$, y aún no se conoce ninguno. La demostración se basa en los *K-funcionales*, introducidos por Peetre en 1963 (véase [6], capítulo 6). En [11], 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 4.7.

# Bibliografía


[1] J. A. ADELL, J. DE LA CAL: Bernstein-type operators diminish the $\varphi$-variation, Constructive Approximation, (1996), 12: 489-507.

[2] J. A. ADELL, J. DE LA CAL: Using stochastic processes for studying Bernstein-type operators, Supplemento ai Rendiconti del Circolo Matematico di Palermo, **11** (33) (1993), 124-141.

[3] F. ALTOMARE, M. CAMPITI: *Korovkin-type Approximation Theory and its Applications*, Walter de Gruyter, (1994).

[4] R. B. ASH, C. A. DOLÉANS-DADE: *Probability & Measure Theory*, Harcourt Academic Press, (2000).

[5] P. BILLINGSLEY: *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics, (1995).

[6] R. A. DEVORE, G. G. LORENTZ: *Constructive Approximation*, Springer-Verlag, (1993).

[7] Z. DITZIAN, K. G. IVANOV: Strong converse inequalities, Journal D'Analyse Mathématique, **61** (1993).

[8] A. GUT: *Probability: A Graduate Course*, Springer Texts in Statistics, (2005).

[9] H. KNOOP, X. ZHOU: The lower estimate for linear positive operators (II), Birkhäuser, **25** (1994).

[10] G. G. LORENTZ:, *Bernstein Polynomials*, Chelsea Publishing Company, (1986).

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

[12] P. P. PETRUSHEV, V. A. POPOV: *Rational Approximation of Real Functions*, Cambridge University Press, (2011).