UNIVERSIDAD DE LA RIOJA

TRABAJO FIN DE ESTUDIOS

Título

Polinomios ortogonales y teorema de Favard

Autor/es

SATURIO CARBONELL URTUBIA

Director/es

JUAN LUIS VARONA MALUMBRES

Facultad

Facultad de Ciencia y Tecnología

Titulación

Grado en Matemáticas

Departamento

MATEMÁTICAS Y COMPUTACIÓN

Curso académico

2017-18

publicaciones.unirioja.es
E-mail: publicaciones@unirioja.es

UNIVERSIDAD
DE LA RIOJA

Facultad de Ciencia y Tecnología

TRABAJO FIN DE GRADO

Grado en Matemáticas

Polinomios ortogonales y teorema de Favard

Realizado por:

Saturio Carbonell Urtubia

Tutelado por:

Juan Luis Varona Malumbres

Logroño, junio de 2018

Resumen/Abstract

Resumen

En este Trabajo Fin de Grado se hace una breve introducción a la teoría de los polinomios ortogonales (construidos a partir de una medida) sobre la recta real, dando la definición de lo que es una sucesión de polinomios ortogonales, proporcionando ejemplos variados y demostrando una serie de propiedades.

En lo que a nosotros concierne, destaca que cualquier sucesión de polinomios ortogonales verifica una relación de recurrencia a tres términos, que permite obtener cualquier polinomio de la sucesión como combinación de los dos polinomios anteriores. Esta relación básica, sus propiedades, y el estudio de los coeficientes que aparecen en dicha relación, son fundamentales para el desarrollo del contenido posterior del trabajo.

Una vez expuesta la teoría de las sucesiones en recurrencias, y demostrado todos los resultados previos que se necesitan, se aborda el objetivo final del trabajo, la demostración del teorema de Favard, que es un recíproco a la existencia de la relación de recurrencia a tres términos: asegura que, para cualquier sucesión de polinomios que satisfaga una cierta relación de recurrencia a tres términos, existe una medida de forma que la sucesión es una sucesión de polinomios ortogonales respecto a esa medida.

Abstract

In this Final Degree Project, a brief introduction is made to the theory of orthogonal polynomials (constructed from a measure) on the real line, giving the definition of what is an orthogonal polynomial sequence, providing many examples and demonstrating some properties.

As far as we are concerned, it stands out that any succession of orthogonal polynomials verifies a recurrence relation to three terms, which allows obtaining any polynomial of the sequence as a combination of the two previous polynomials. This basic relation, its properties, and the study of the coefficients that appear in this relation, are fundamental for the development of the later content of the work.

Once that the theory of recurrence relations is exposed, and all previous results has been shown, the final objective of the work is addressed, the demonstration of Favard's theorem, which is a reciprocal to the existence of the three terms recurrence relation: it ensures that, for any sequence of polynomials that satisfies a certain three terms relation of recurrence, there is a measure such that the sequence is an orthogonal polynomial sequence for that measure.

Índice general

Capítulo 1

Definiciones y ejemplos

1.1. Polinomios ortogonales sobre la recta real

En esta memoria, el marco en el que vamos a considerar la ortogonalidad de funciones en general, y de polinomios en particular, es el espacio de Hilbert $L^2(\mathbb{R}, \mu)$ (o, simplemente, $L^2(\mu)$) donde μ es una medida positiva de Borel o de Lebesgue en la recta real $\mathbb{R}$; en ocasiones, según convenga, también usaremos $d\mu$ para denotar la medida.

Recordemos que $L^2(\mathbb{R}, \mu)$ es, por definición, el espacio de las funciones medible respecto a μ y de cuadrado integrable, es decir,

$$L^2(\mathbb{R}, \mu) = \left\{ f : \mathbb{R} \to \mathbb{R} : \int_{\mathbb{R}} |f(x)|^2 \, d\mu(x) < \infty \right\}.$$

Aunque para nosotros no será ningún problema, recordemos así mismo que, formalmente, los elementos de este espacio son clases de equivalencia de funciones (iguales en casi todo punto, es decir, salvo un conjunto de medida nula) y no funciones individuales.

Sin pérdida de generalidad, considerar siempre medidas sobre $\mathbb{R}$ incluye el caso de las medidas sobre un intervalo I (realmente, no hace falta que sea un intervalo, sino cualquier subconjunto $I \subseteq \mathbb{R}$), pues una medida μ sobre I se extiende a medida sobre $\mathbb{R}$ fácilmente sin más que considerar que $\mu(\mathbb{R} \setminus I) = 0$. Cuando eso ocurre, se suele decir que la mediada está soportada en I.

Aquí nos ocuparemos de sucesiones de polinomios ortogonales sobre la recta real (o sobre un intervalo, caso que, como acabamos de explicar, no es necesario explicitar al dar la teoría general). Antes de proseguir, queremos comentar que también existen sucesiones de polinomios ortogonales sobre la circunferencia unidad, o sobre otras curvas del plano complejo, pero esto queda totalmente fuera de nuestro estudio.

Comencemos dando algunas definiciones básicas.

Definición 1 (Ortogonalidad)**.** Dos funciones f y g en $L^2(\mathbb{R}, \mu)$ se dicen *ortogonales* si

$$\int_{\mathbb{R}} f(x)g(x) \, d\mu(x) = 0.$$

En este trabajo, la noción de ortogonalidad la aplicaremos exclusivamente a polinomios. Pero, por otra parte, nos interesa aplicarla a una sucesión en lugar de solo a dos funciones, lo que nos lleva a la siguiente definición.

Definición 2 (Sucesión de Polinomios Ortogonales, SPO)**.** Una sucesión de polinomios $\{P_n\}_{n=0}^{\infty}$ decimos que es una *sucesión de polinomios ortogonales* (con respecto a una medida μ) si:

- P_n es un polinomio de grado n.
- $\int_{\mathbb{R}} P_m(x)P_n(x)\,d\mu(x) = 0$, $m \neq n$, para todo $n, m = 0, 1, 2, \dots$.
- $\int_{\mathbb{R}} P_n(x)^2\,d\mu(x) \neq 0$, para todo $n = 0, 1, 2, \dots$.

Las dos últimas condiciones las podemos resumir como

$$\int_{\mathbb{R}} P_m(x)P_n(x)\,d\mu(x) = A_n\delta_{n,m},$$

donde $A_n \neq 0$ y $\delta_{n,m}$ es la delta de Kronecker. Si $A_n = 1$ para todo n, entonces decimos que los polinomios son *ortonormales*.

Nota 1 (Delta de Kronecker)*.* Recordemos que la delta de Kronecker $\delta_{i,j}$ se define como

$$\delta_{i,j} = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$$

$\lhd$

Por supuesto, para poder hablar de que los polinomios son ortogonales en $L^2(\mathbb{R}, \mu)$, necesitamos previamente que los polinomios estén en ese espacio, es decir, que

$$\int_{\mathbb{R}} |P(x)|\,d\mu(x) < \infty$$

para cualquier polinomio $P(x)$. Por linealidad, esto es equivalente a que

$$\int_{\mathbb{R}} |x|^n\,d\mu(x) < \infty \quad \text{para todo } n \geq 0. \tag{1.1}$$

En todas las medidas de este trabajo asumiremos esta condición sin necesidad de recordarla. Nótese además que, si μ es una medida soportada en un intervalo I acotado, para que se cumpla (1.1) basta con exigir

$$\mu(I) = \int_I d\mu < \infty$$

(pues cada $|x|^n$ estará acotado, en I, por una constante), pero eso no es cierto en general si I no es acotado.

Conviene aquí, aunque sea adelantarse al desarrollo de la memoria, exponer otra formulación equivalente de la propiedad de ortogonalidad para una familia de polinomios, aunque ahora desde un punto de vista totalmente algebraico —sin el concurso de medidas ni integrales—. Esta formulación mostrará que toda familia de polinomios ortogonales admite una representación simplísima, pues que los polinomios sean ortogonales con respecto a una medida positiva equivale a lo que sigue. Si escribimos el polinomio $xP_n(x)$ como combinación de $P_{n+1}(x), P_n(x), \dots, P_0(x)$, lo que claramente siempre es posible si el coeficiente director de los P_n es distinto de cero, resulta

$$xP_n(x) = \beta_{n,n+1}P_{n+1}(x) + \beta_{n,n}P_n(x) + \beta_{n,n-1}P_{n-1}(x) + \cdots + \beta_{n,0}P_0(x);$$

entonces, los $\{P_n(x)\}_{n=0}^{\infty}$ son ortogonales si —y solo si— en esta combinación solo aparecen los tres primeros sumandos, esto es:

$$xP_n(x) = \beta_{n,n+1}P_{n+1}(x) + \beta_{n,n}P_n(x) + \beta_{n,n-1}P_{n-1}(x), \tag{1.2}$$

con una condición de positividad $\beta_{n,n-1}/\beta_{n-1,n} > 0$. Este resultado es conocido como teorema de Favard, y a él dedicaremos el último capítulo de esta memoria. Antes de seguir, conviene apuntar que la relación (1.2), llamada fórmula de recurrencia a tres términos, permite generar familias de polinomios ortogonales a partir de las tres sucesiones de coeficientes que aparecen en la relación, siempre que prefijemos $P_0(x)$ y $P_1(x)$.

Las sucesiones de polinomios ortogonales son el tema central de este trabajo, por lo que, sin más preámbulos, vamos a ver unos ejemplos para familiarizarnos con este concepto.

1.2. Ejemplos

Vamos a distinguir varios tipos de medidas: continuas, discretas y mixtas. En cada apartado daremos la definición de ese tipo de medida, una breve introducción y unos ejemplos representativos.

1.2.1. Medidas continuas (pesos)

Supongamos que tenemos una función w que es positiva en un conjunto I de medida (Lebesgue o Borel) positiva. En este caso, se tiene que

$$\int_I w(x)\,dx > 0.$$

En efecto, para cada entero positivo n, sea

$$I_n = \{x \in I : w(x) \geq 1/n\},$$

con lo cual, obviamente, siempre $I_n \subseteq I_{n+1}$ y $m(I_n) \leq m(I_{n+1})$ (estamos usando m para denotar la medida de Lebesgue). Además, es claro que $I = \bigcup_{n=1}^{\infty} I_n$ y, por las propiedades de la medidas, $m(I) = \lim_{n\to\infty} m(I_n)$. Como, por hipótesis, $m(I) > 0$, eso implica que $m(I_n) > 0$ para algún n. En consecuencia,

$$\int_I w(x)\,dx \geq \int_{I_n} w(x)\,dx \geq \int_{I_n} \frac{1}{n}\,dx = m(I_n)\frac{1}{n} > 0,$$

tal como queríamos comprobar.

Eso nos da pie a la siguiente definición:

Definición 3 (Función peso)**.** Sea w un función no negativa e integrable en un intervalo (a, b), que cumple que $w(x) > 0$ en un subintervalo de (a, b) (o en un subconjunto con medida de Lebesgue positiva), con lo cual

$$\int_a^b w(x)\,dx > 0.$$

La función $w(x)$ se denomina *función peso* o, simplemente, *peso*.

Definición 4 (Medida continua)**.** Una medida $d\mu$ se dice que es *continua* si es de la forma

$$d\mu(x) = w(x)\,dx,$$

con $w(x)$ una función peso.

En este apartado todas las medidas son continuas, por lo que simplemente daremos el valor de la función peso. Daremos una serie de caracterizaciones, expresiones explícitas, propiedades y relaciones de los polinomios ortogonales más usuales: los *polinomios ortogonales clásicos*. Conviene recalcar que los polinomios que manejemos estarán sin normalizar. En primer lugar, he aquí un cuadro de los polinomios ortogonales clásicos:

Peso	Soporte	Denominación	Notación
$(1-x^2)^{-1/2}$	$[-1,1]$	Chebyshev (de 1.ª especie)	$T_n(x)$
$(1-x^2)^{1/2}$	$[-1,1]$	Chebyshev de 2.ª especie	$U_n(x)$
1	$[-1,1]$	Legendre	$P_n(x)$
$(1-x^2)^{\lambda-1/2}$, $\lambda > -1/2$	$[-1,1]$	Gegenbauer o ultraesféricos	$P_n^{(\lambda)}(x)$
$(1-x)^\alpha(1+x)^\beta$, $\alpha,\beta > -1$	$[-1,1]$	Jacobi	$P_n^{(\alpha,\beta)}(x)$
$e^{-x}x^\alpha$, $\alpha > -1$,	$[0,\infty)$	Laguerre	$L_n^{(\alpha)}(x)$
e^{-x^2}	$\mathbb{R}$	Hermite	$H_n(x)$

El que precisamente a estas familias de polinomios ortogonales se les llame clásicos es porque son las únicas, salvo cambios de variable, que verifican una serie de propiedades como son:

- Sus derivadas (y sus derivadas r-ésimas para cualquier $r \geq 1$) siguen formando un sistema de polinomios ortogonales.
- Verifican una ecuación diferencial de la forma $A(x)y'' + B(x)y' + \lambda_n y = 0$, donde $A(x)$ y $B(x)$ son polinomios independientes de n, y λ_n independiente de x.
- Pueden expresarse mediante una fórmula de tipo Rodrigues

  $$P_n(x) = \frac{1}{E_n w(x)}\frac{d^n}{dx^n}(w(x)\rho(x)),$$

  donde $w(x)$ es el peso, $\rho(x)$ un polinomio independiente de n y E_n un coeficiente real independiente de x.

Una explicación más detallada de estas caracterizaciones puede verse en Chihara [3, pág. 150].

Aunque los polinomios de Chebyshev de primera y segunda especie, los de Legendre y los ultraesféricos (o de Gegenbauer) son casos particulares de los de Jacobi, los citaremos de forma particular. También veremos los polinomios de Laguerre y Hermite. Para la demostración de las propiedades que aquí enunciemos, además de muchas otras, pueden consultarse los libros de Álvarez-Nodarse [2], Chihara [3], Freud [5] o Szegő [10], así como el reciente texto de Koekoek, Lesky y Swarttouw [6]. Sin demostraciones, muchas propiedades y las notaciones estándar se pueden consultar en [8].

Ejemplo 1 (Chebyshev de primera especie)**.** Vamos a construir los polinomios de Chebyshev de primera especie a partir de resultados básicos de matemáticas, concretamente de la trigonometría. Partimos de una conocida igualdad:

$$2\cos(m\alpha)\cos(n\alpha) = \cos((m+n)\alpha) + \cos((m-n)\alpha).$$

Si integramos esta igualdad ente 0 y π obtenemos fácilmente la fórmula

$$\int_0^\pi \cos(m\alpha)\cos(n\alpha)\,d\alpha = 0, \quad m \neq n, \quad m,n = 0,1,2,\dots \tag{1.3}$$

Como el valor de la integral se ha anulado, podemos decir que que las funciones $\cos m\alpha$ y $\cos n\alpha$ son ortogonales en el intervalo $[0,\pi]$ para $m \neq n$. Cuando $n = m$, el valor de la integral es π si son ambos 0, y $\pi/2$ para el resto de naturales.

Así, podemos formar la sucesión $\{1, \cos\alpha, \cos 2\alpha, \dots, \cos n\alpha, \dots\}$ y decir que es una sucesión ortogonal en $[0,\pi]$. Hasta aquí el ejemplo trata sobre la ortogonalidad, pero todavía no han aparecido polinomios. Para construir unos polinomios $T_n(x)$ y formar la sucesión $\{T_n(x)\}_{n=0}^\infty$ vamos a utilizar la fórmula de De Moivre

$$(\cos\alpha + i \operatorname{sen}\alpha)^n = \cos(n\alpha) + i \operatorname{sen}(n\alpha). \tag{1.4}$$

Quedándonos solo con la parte real de la fórmula de De Moivre y realizando el cambio de variable $\cos\alpha = x$ tenemos

$$\cos n\alpha = \operatorname{Re}\left(\sum_{k=0}^{n} \binom{n}{k} x^{n-k} \left(i\sqrt{1-x^2}\right)^k\right) = \sum_{k=0}^{\lfloor n/2\rfloor} \binom{n}{2k} x^{n-2k}(x^2-1)^k.$$

Entonces, basta definir $T_n(x) = \cos(n \arccos(x))$ y ya tendríamos los polinomios.

Nota 2 (Función suelo). En la expresión anterior, $\lfloor n/2\rfloor$ denota la función suelo definida como

$$\operatorname{suelo}(x) = \lfloor x\rfloor = \max\{k \in \mathbb{Z} : k \leq x\}. \quad \triangleleft$$

Los polinomios $\{T_n(x)\}_{n=0}^\infty$ cumplen la relación de ortogonalidad ya que, aplicando el mismo cambio ($x = \cos\alpha$) a la integral (1.3), obtenemos

$$\int_{-1}^{1} T_m(x)T_n(x)(1-x^2)^{-1/2}\,dx = 0, \quad m \neq n.$$

Cuando $n = m$, el valor de la integral coincide con el de (1.3).

Por lo tanto, como la sucesión de polinomios $\{T_n(x)\}_{n=0}^\infty$ cumple

- T_n es un polinomio de grado n,
- $\int_{-1}^{1} T_m(x)T_n(x)w(x)\,dx = 0$, $m \neq n$, para todo $n, m = 0, 1, 2, \ldots$,
- $\int_{-1}^{1} T_n(x)^2 w(x)\,dx \neq 0$, para todo $n = 0, 1, 2, \ldots$,

con $w(x) = (1-x^2)^{-1/2}$, podemos concluir que $\{T_n(x)\}_{n=0}^{\infty}$ es una sucesión de polinomios ortogonales con respecto a la función peso $(1-x^2)^{-1/2}$ en el intervalo $[-1, 1]$.

Estos polinomios se conocen como polinomios de Chebyshev de primera especie. Ya hemos visto que su fórmula general es

$$T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} x^{n-2k}(x^2-1)^k.$$

Los primeros términos de la sucesión son

$$T_0(x) = 1,$$

$$T_1(x) = x,$$

$$T_2(x) = 2x^2 - 1.$$

Además, satisfacen la siguiente fórmula en recurrencias:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad n \geq 1. \qquad \Diamond$$

Ejemplo 2 (Chebyshev de segunda especie)**.** Los polinomios Chebyshev de segunda especie $U_n(x)$, al igual que los de primera especie, también tienen una expresión trigonométrica explícita que puede reformularse en forma polinómica, quedando así:

$$U_n(x) = \frac{\operatorname{sen}((n+1)\alpha)}{\operatorname{sen}\alpha} = \sum_{m=0}^{\lfloor n/2 \rfloor} \binom{n+1}{2m+1} x^{n-2m}(x^2-1)^m.$$

con $x = \cos\alpha$, $n = 0, 1, 2, \ldots$. La expresión polinómica se obtiene de a partir de la fórmula de De Moivre (1.4), pero esta vez quedándonos con la parte imaginaria. La ortogonalidad esta garantizada de forma trivial por trigonometría, resultando

$$\int_0^{\pi} \frac{\operatorname{sen}((m+1)\alpha)}{\operatorname{sen}\alpha} \frac{\operatorname{sen}((n+1)\alpha)}{\operatorname{sen}\alpha}\,d\alpha = 0, \quad m \neq n, \quad m, n = 0, 1, 2, \ldots; \tag{1.5}$$

y, cuando $m = n$, el valor de la integral es $\pi/2$. Realizando el cambio de variable $x = \cos(\alpha)$ en (1.5) obtenemos

$$\int_{-1}^{1} U_m(x)U_n(x)(1-x^2)^{1/2}\,dx = 0, \quad m \neq n.$$

Así podemos decir que $\{U_n(x)\}_{n=0}^{\infty}$ es una SPO con respecto a la función peso $w(x) = (1-x^2)^{1/2}$ en el intervalo $[-1, 1]$.

Los primeros polinomios de esta sucesión son:

$$U_0(x) = 1,$$

$$U_1(x) = 2x,$$

$$U_2(x) = 4x^2 - 1.$$

Además, satisfacen la ecuación en recurrencias

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \qquad n \geq 1. \qquad \Diamond$$

Ejemplo 3 (Polinomios de Legendre)**.** Los polinomios de Legendre, $P_n(x)$, son unos polinomios ortogonales en el intervalo $[-1, 1]$ con respecto a la función peso $w(x) = 1$.

La fórmula general de estos polinomios es

$$P_n(x) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k}^2 (x+1)^{n-k}(x-1)^k,$$

y satisfacen la relación de ortogonalidad

$$\int_{-1}^{1} P_m(x)P_n(x)\,dx = \frac{2}{2n+1}\,\delta_{m,n}.$$

Los primeros polinomios de esta familia son

$$P_0(x) = 1,$$

$$P_1(x) = x,$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1).$$

Satisfacen la siguiente ecuación en recurrencias:

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0, \qquad n \geq 1. \qquad \Diamond$$

Ejemplo 4 (Polinomios de Gegenbauer)**.** Los polinomios de Gegenbauer $C_n^{(\lambda)}(x)$ (también denominados ultraesféricos) son unos polinomios ortogonales en el intervalo $[-1, 1]$ con respecto a la función peso $w(x) = (1-x^2)^{\lambda-1/2}$, con $\lambda > -1/2$.

Estos polinomios generalizan los polinomios de Chebyshev y de Legendre. Los primeros, cuando tomamos $\lambda = 0$ para los de primera especie y $\lambda = 1$ para los de segunda especie; los segundos, cuando $\lambda = 1/2$.

La fórmula general de los polinomios de Gegenbauer es:

$$C_n^{(\lambda)}(x) = \binom{n+2\lambda-1}{n} \sum_{k=0}^{n} \frac{\binom{n}{k}(2\lambda+n)_k}{(\lambda+1/2)_k} \left(\frac{x-1}{2}\right)^k,$$

donde $(a)_k$ es el denominado símbolo de Pochhammer.

Nota 3 (Símbolo de Pochhammer). Dado $a \neq 0$ perteneciente a los complejos y n un entero no negativo, se define el símbolo de Pochhammer $(a)_n$ como

$$(a)_n = \prod_{k=1}^{n} (a+k-1) = a(a+1)(a+2)\cdots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)},$$

con $(a)_0 = 1$, y donde $\Gamma(a)$ es la función Gamma. $\triangleleft$

Volviendo a los polinomios de Gegenbauer, satisfacen la relación de ortogonalidad

$$\int_{-1}^{1} C_m^{(\lambda)}(x) C_n^{(\lambda)}(x)(1-x^2)^{\lambda-1/2}\, dx = 2^{1-2\lambda}\pi \,\frac{\Gamma(n+2\lambda)}{(n+\lambda)\Gamma(\lambda)^2\Gamma(n+1)}\,\delta_{m,n},$$

para $\lambda > -1/2$. Los primeros términos de la sucesión son

$$C_0^{(\lambda)}(x) = 1,$$

$$C_1^{(\lambda)}(x) = 2\lambda x,$$

$$C_2^{(\lambda)}(x) = 2(\lambda^2+\lambda)x^2 - \lambda.$$

La fórmula de recurrencia de estos polinomios es

$$C_n^{(\lambda)}(x) = \frac{1}{n}\Big(2x(n+\lambda-1)C_{n-1}^{(\lambda)}(x) - (n+2\lambda-2)C_{n-2}^{(\lambda)}(x)\Big). \qquad \diamondsuit$$

Ejemplo 5 (Polinomios de Jacobi)**.** Los polinomios de Jacobi $P_n^{(\alpha,\beta)}(x)$ son ortogonales en el intervalo $[-1,1]$ respecto a la función peso $w(x) = (1-x)^\alpha(1+x)^\beta$, con $\alpha, \beta > -1$.

Estos polinomios generalizan los de Gegenbauer (y, por tanto, también los de Chebyshev y Legendre) con $\alpha = \beta = \lambda - 1/2$. Los polinomios de Jacobi constituyen una de las tres familias clásicas de polinomios ortogonales continuos.

La fórmula general de los polinomios de Jacobi es

$$P_n^{(\alpha,\beta)}(x) = \frac{\Gamma(\alpha+n+1)}{n!\,\Gamma(\alpha+\beta+n+1)} \sum_{m=0}^{n} \binom{n}{m} \frac{\Gamma(\alpha+\beta+n+m+1)}{\Gamma(\alpha+m+1)} \left(\frac{x-1}{2}\right)^m.$$

Su relación de ortogonalidad es

$$\int_{-1}^{1} P_m^{(\alpha,\beta)}(x) P_n^{(\alpha,\beta)}(x)(1-x)^\alpha(1+x)^\beta\, dx = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1}\,\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!\,\Gamma(n+\alpha+\beta+1)}\,\delta_{m,n}.$$

Los primeros polinomios de la sucesión son

$$P_0^{(\alpha,\beta)}(x) = 1,$$

$$P_1^{(\alpha,\beta)}(x) = \frac{1}{2}(2(\alpha+1) + (\alpha+\beta+2)(x-1)),$$

$$P_2^{(\alpha,\beta)}(x) = \frac{1}{8}(4(\alpha+1)(\alpha+2) + 4(\alpha+\beta+3)(\alpha+2)(x-1) + (\alpha+\beta+3)(\alpha+\beta+4)(x-1)^2).$$

Y su fórmula de recurrencia es

$$\begin{aligned}
&2n(n+\alpha+\beta)(2n+\alpha+\beta-2)P_n^{(\alpha,\beta)}(x)\\
&\quad=(2n+\alpha+\beta-1)\Big\{(2n+\alpha+\beta)(2n+\alpha+\beta-2)x+\alpha^2-\beta^2\Big\}P_{n-1}^{(\alpha,\beta)}(x)\\
&\qquad-2(n+\alpha-1)(n+\beta-1)(2n+\alpha+\beta)P_{n-2}^{(\alpha,\beta)}(x).
\end{aligned}\qquad\Diamond$$

Ejemplo 6 (Polinomios de Laguerre (generalizados))**.** Los polinomios de Laguerre generalizados, $L_n^{(\alpha)}(x)$, son unos polinomios ortogonales en el intervalo $[0,\infty)$ con respecto a la función peso $w(x)=e^{-x}x^{\alpha}$ con $\alpha>-1$ (el término "generalizados" se emplea a veces porque, originariamente, se consideró solo el caso $\alpha=0$).

Su fórmula general es

$$L_n^{(\alpha)}(x)=\sum_{k=0}^{n}(-1)^k\binom{n+\alpha}{n-k}\frac{x^k}{k!},$$

donde $\binom{n+\alpha}{n-k}$ es el binomial de un número real.

Nota 4 (Binomial de un número real)*.* Hasta ahora han aparecido binomiales de números enteros no negativos, pero también se puede definir, para números reales cualesquiera (e incluso complejos), de la siguiente forma:

$$\binom{r}{k}=\frac{\Gamma(r+1)}{\Gamma(k+1)\Gamma(r-k+1)}.$$

Esta definición extiende la definición habitual del binomial. $\lhd$

La relación de ortogonalidad de los polinomios de Laguerre es

$$\int_0^{\infty}L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)x^{\alpha}e^{-x}\,dx=\frac{\Gamma(n+\alpha+1)}{n!}\delta_{n,m}.$$

Los primeros términos de la sucesión son

$$L_0^{\alpha}(x)=1,$$

$$L_1^{\alpha}(x)=-x+\alpha+1,$$

$$L_2^{\alpha}(x)=\frac{1}{2}(x^2-2(\alpha+2)x+(\alpha+1)(\alpha+2)).$$

Y su relación de recurrencia es

$$L_{k+1}^{(\alpha)}(x)=\frac{(2k+1+\alpha-x)L_k^{(\alpha)}(x)-(k+\alpha)L_{k-1}^{(\alpha)}(x)}{k+1}.\qquad\Diamond$$

Ejemplo 7 (Polinomios de Hermite)**.** Los polinomios de Hermite, $H_n(x)$, son unos polinomios ortogonales en el intervalo $(-\infty,\infty)$ con respecto a la función peso $w(x)=e^{-x^2}$.

Su fórmula general es

$$H_n(x) = n! \sum_{m=0}^{\lfloor n/2 \rfloor} \frac{(-1)^m}{m!\,(n-2m)!}(2x)^{n-2m}.$$

Son ortogonales ya que

$$\int_{-\infty}^{\infty} H_m(x)H_n(x)e^{-x^2}\,dx = 2^n n!\,\sqrt{\pi}\,\delta_{m,n}.$$

Los primeros polinomios de Hermite son

$$H_0(x) = 1,$$

$$H_1(x) = 2x,$$

$$H_2(x) = 4x^2 - 2.$$

Y su fórmula de recurrencia es

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \qquad \diamondsuit$$

1.2.2. Medidas discretas

Hasta ahora hemos visto unos cuantos ejemplos de SPO cuando la medida sobre la que integramos proviene de una función peso. Veamos ahora otros ejemplos cuando la medida es discreta.

Definición 5 (Medida discreta)**.** Una medida $d\mu$ definida sobre los conjuntos medibles Lebesgue de la recta real se dice *discreta* si existe una sucesión (finita o infinita) de números

$$s_0, s_1, s_2, \ldots$$

tales que

$$\mu(\mathbb{R} \setminus \{s_0, s_1, s_2, \ldots\}) = 0.$$

Las medidas discretas las describiremos como combinación de deltas de Dirac. Dado un punto s en la recta real, la delta de Dirac en s, que se denota δ_s (o, también, $\delta(x-s)$) es la medida definida por

$$\delta_s(X) = \begin{cases} 1, & \text{si } s \in X, \\ 0, & \text{si } s \notin X, \end{cases}$$

para cualquier conjunto medible Lebesgue X. Entonces, si tenemos una sucesión de puntos $s_0, s_1, s_2, \ldots$, la medida

$$d\mu = \sum_k a_k \delta_{s_k}, \qquad a_k > 0,$$

es una medida discreta.

En la definición anterior, los a_k los podemos dar con una función a la que llamaremos función de salto j, sin más que tomar $j(k) = a_k$; además, en los ejemplos que vamos a mostrar, será siempre $s_k = k$. Así pues, las medidas discretas que veremos en esta sección serán de la forma

$$d\mu = \sum_{k\geq 0} j(k)\delta_k, \tag{1.6}$$

y en los ejemplos concretos especificaremos cuál es la función $j(k)$ en cada caso.

Del mismo modo que hay unas familias de polinomios respecto a medidas continuas que se denominan clásicas (Jacobi, Laguerre y Hermite), también hay cuatro familias de polinomios ortogonales respecto a medidas discretas que se denominan polinomios clásicos discretos (por ciertas razones que no detallaremos). Estas familias son los polinomios de Charlier, Krawtchouk, Meixner y Hahn. Se pueden ver en los textos [2] o [6], y una recopilación de propiedades en [8].

Ejemplo 8 (Polinomios de Charlier)**.** Los polinomios de Charlier $c_k(x; a)$ son unos polinomios ortogonales con respecto a la medida $d\mu$; donde $d\mu$ de la forma (1.6) con $j(k) = e^{-a}a^k(k!)^{-1}$, k entero no negativo y a real positivo.

La fórmula general de estos polinomios es

$$c_n(x; a) = \sum_{\nu=0}^{n} (-1)^{n-\nu} \binom{n}{\nu} \nu!\, a^{-\nu} \binom{x}{\nu}.$$

Son ortogonales ya que

$$\sum_{x=0}^{\infty} c_n(x; a)c_m(x; a)\frac{e^{-a}a^x}{x!} = a^{-n}n!\, \delta_{m,n}.$$

Los primeros polinomios son

$$c_1(x; a) = 1,$$
$$c_2(x; a) = -\frac{a - x}{a},$$
$$c_3(x; a) = \frac{a^2 - x - 2ax + x^2}{a^2}.$$

Y los siguientes se pueden calcular a partir de la fórmula de recurrencia (para $n \geq 2$)

$$xc_n(x; a) = c_{n+1}(x; a) + (n + a)c_n(x; a) + anc_{n-1}(x; a). \qquad \Diamond$$

Ejemplo 9 (Polinomios de Krawtchouk)**.** Los polinomios de Krawtchouk $k_n^p(x; N)$ son unos polinomios ortogonales con respecto a la medida $d\mu$ de la forma (1.6) con $j(x) = \binom{N}{x}p^x q^{N-x}$ en $x = 0, 1, \ldots, N$, donde $p > 0$, $q > 0$ y $p + q = 1$. Merece la pena resaltar que la familia de polinomios de Krawtchouk no es —al contrario de todas las demás que hemos visto hasta ahora— una familia infinita, sino que solo tiene $N + 1$ polinomios (de grados 0 hasta N).

Su fórmula general es

$$k_n(x;p,N) = \sum_{\nu=0}^{n} (-1)^{n-\nu} \binom{N-x}{n-\nu} \binom{x}{\nu} p^{n-\nu} q^{\nu}, \quad n = 0, 1, \ldots, N.$$

Los polinomios de Krawtchouk son ortogonales según la expresión

$$\sum_{x=0}^{N} k_m(x;p,N) k_n(x;p,N) \binom{N}{x} p^x q^{N-x} = \binom{N}{n} p^n q^{N-n} \delta_{m,n}.$$

Los primeros son

$$k_0(x;p,N) = 1,$$

$$k_1(x;p,N) = -Np + x,$$

$$k_2(x;p,N) = \frac{1}{2}\Big(N^2p^2 + x(2p + x - 1) - Np(p + 2x)\Big).$$

Y la fórmula de recurrencia es

$$(n+1)k_{n+1}(x;p,N) = (x - n - (N-2))k_n(x;p,N) - pq(N-n+1)k_{n-1}(x;p,N). \quad \diamondsuit$$

Ejemplo 10 (Polinomios de Meixner)**.** Los polinomios de Meixner $M_n(x;\beta,\gamma)$, con $\beta > 0$ y $0 < \gamma < 1$, son unos polinomios ortogonales con respecto a la medida $d\mu$ de la forma (1.6) con $j(k) = \gamma^k (\beta)_k / k!$.

Su fórmula general es

$$M_n(x;\beta,\gamma) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{x}{k} k!\, (x-\beta)_{n-k} \gamma^{-k}.$$

Son ortogonales ya que

$$\sum_{x=0}^{\infty} M_m(x;\beta,\gamma) M_n(x;\beta,\gamma) \frac{\gamma^x (\beta)_x}{x!} = \frac{\gamma^{-n} n!}{(\beta)_n (1-\gamma)^{\beta}}\, \delta_{m,n}.$$

Los primeros polinomios son

$$M_0(x;\beta,\gamma) = 1,$$

$$M_1(x;\beta,\gamma) = \beta + x\Big(1 - \frac{1}{\gamma}\Big),$$

$$M_2(x;\beta,\gamma) = \frac{\beta(\beta+1)\gamma^2 + (\gamma-1)(2\beta\gamma + \gamma + 1)x + (\gamma-1)^2 x^2}{\gamma^2},$$

y la fórmula de recurrencia es

$$\gamma(n+\beta)M_{n+1}(x;\beta,\gamma) = (n + (n+\beta)\gamma + (\gamma-1)x)M_n(x;\beta,\gamma) - nM_{n-1}(x;\beta,\gamma). \quad \diamondsuit$$

Ejemplo 11 (Polinomios de Hahn)**.** Los polinomios de Hahn $Q_n\,(x;\alpha,\beta,N)$, para $\alpha > -1$ y $\beta > -1$ o $\alpha < -N$ y $\beta < -N$, son unos polinomios ortogonales con respecto a la medida $d\mu$ de la forma (1.6) con

$$j(k) = \frac{(\alpha+1)_k(\beta+1)_{N-k}}{k!\,(N-k)!} = \binom{\alpha+k}{k}\binom{\beta+N-k}{N-k}, \quad k = 0, 1 \ldots, N.$$

Como ocurría con los polinomios de Krawtchouk, son una familia finita, con n desde 0 hasta N.

Su fórmula general es

$$Q_n(x;\alpha,\beta,N) = \sum_{k=0}^{n} \frac{(-n)_k(n+\alpha+\beta+1)_k(-x)_k}{(\alpha+1)_k(-N+1)_k k!}.$$

Son ortogonales con

$$\sum_{x=0}^{N} Q_n(x;\alpha,\beta,N)Q_m(x;\alpha,\beta,N)\binom{\alpha+x}{x}\binom{\beta+N-x}{N-x}$$
$$= \frac{(-1)^n(n+\alpha+\beta+1)_{N+1}(\beta+1)_n n!}{(2n+\alpha+\beta+1)(\alpha+1)_n(-N)_n N!}\,\delta_{m,n}.$$

Los primeros términos de la sucesión son

$$Q_0(x;\alpha,\beta,N) = 1,$$
$$Q_1(x;\alpha,\beta,N) = 1 - \frac{(\alpha+\beta+2)x}{(\alpha+1)N},$$
$$Q_2(x;\alpha,\beta,N) = \frac{(\alpha+\beta+3)(\alpha+\beta+4)(1-x)x}{(\alpha+1)(\alpha+2)(1-N)N} - \frac{2(\alpha+\beta+3)x}{(\alpha+1)N} + 1.$$

Y su fórmula en recurrencias es la siguiente:

$$xQ_n(x;\alpha,\beta,N) = -\frac{(n+\alpha+\beta+1)(n+\alpha+1)(N-n)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}Q_{n+1}(x;\alpha,\beta,N)$$
$$+\left(\frac{(n+\alpha+\beta+1)(n+\alpha+1)(N-n)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} + \frac{n(n+\alpha+\beta+N+1)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)}\right)Q_n(x;\alpha,\beta,N)$$
$$-\frac{n(n+\alpha+\beta+N+1)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)}Q_{n-1}(x;\alpha,\beta,N). \qquad \Diamond$$

1.2.3. Medidas mixtas

Aunque son son mucho menos usadas, también hay familias de polinomios ortogonales sobre la recta real que se definen a partir de una medida que tiene parte continua y parte discreta, lo que denominamos una *medida mixta*.

Sin detenernos en detalles, veamos un par de ejemplos.

Posiblemente los conocidos son los llamados polinomios de Koornwinder, que fueron introducidos en el artículo [7]:

Ejemplo 12 (Polinomios de Koornwinder)**.** Son ortogonales en el intervalo $[-1,1]$ respecto a la medida

$$(1-x)^{\alpha}(1+x)^{\beta}\,dx + M\delta_1 + N\delta_{-1}$$

con $\alpha, \beta > -1$ y $M, N \geq 0$. $\Diamond$

Otros polinomios ortogonales respecto a una medida mixta son los de Krall, que a su vez son de de tres tipos, y que se definen como sigue. Se pueden ver, por ejemplo, en [2].

Ejemplo 13 (Polinomios de Krall, de tres tipos)**.** Los polinomios de Krall de tipo Laguerre, están soportados en $[0,\infty)$ y son ortogonales respecto a la medida

$$e^{-x}\,dx + M\delta_0, \qquad M > 0.$$

Los de tipo Legendre están soportados $[-1,1]$ y son ortogonales respecto a

$$\frac{\alpha}{2}\,dx + \frac{1}{2}\,\delta_{-1} + \frac{1}{2}\,\delta_1, \qquad \alpha > 0.$$

Finalmente, los de tipo Jacobi están soportados en $[0,1]$ y son ortogonales respecto a

$$(1-x)^{\alpha}\,dx + M\delta_0 \qquad M > 0, \quad \alpha > -1. \qquad \Diamond$$

Capítulo 2

Propiedades básicas de los polinomios ortogonales

En este capítulo vamos a profundizar en el concepto de polinomios ortogonales. Las SPO tienen gran cantidad de caracterizaciones y propiedades que han sido descubiertas y probadas a lo largo de los años, pero para el propósito de este Trabajo de Fin de Grado solo veremos unas pocas.

Muchas otras propiedades se pueden encontrar en los libros [2, 3, 5, 10].

2.1. Caracterización, coeficientes de Fourier y momentos

Comencemos con una caracterización básica de las SPO que, aunque es muy sencilla a partir de la definición 2, al comprobarla nos aparecerán, en el camino, algunos conceptos muy importantes.

Teorema 1 (Caracterización). *Sea $\{P_n(x)\}_{n=0}^{\infty}$ una sucesión de polinomios, con P_n de grado n, y sea μ una medida. Entonces se dan las siguientes equivalencias:*

(a) $\{P_n(x)\}_{n=0}^{\infty}$ *es una SPO con respecto a* μ.

(b) $\int_{\mathbb{R}} \pi(x)P_n(x)\,d\mu(x) = 0$ *para todo polinomio* $\pi(x)$ *de grado* $m < n$*, pero* $\int_{\mathbb{R}} \pi(x)P_n(x)\,d\mu(x) \neq 0$ *si* $m = n$.

(c) $\int_{\mathbb{R}} \pi(x)P_n(x)\,d\mu(x) = K_n\delta_{m,n}$ *para todo polinomio* $\pi(x)$ *de grado* $m \leq n$*, donde* $K_n \neq 0$*,* $m = 0, 1, \ldots, n$.

Antes de demostrar este teorema vamos dar un pequeño lema que nos ayudará en la demostración.

Lema 1. *Dada una SPO, los* $n+1$ *polinomios* $\{P_j\}_{j=0}^{n}$ *forman una base del espacio vectorial de los polinomios con coeficientes reales de grado a lo sumo* n.

Demostración. Es evidente ya que el grado del polinomio $P_k(x)$ es k. □

Con este pequeño lema ya podemos completar la demostración del teorema.

Demostración del teorema 1. Sea $\{P_n(x)\}_{n=0}^{\infty}$ una sucesión de polinomios ortogonales; dado que P_n tiene grado n, tenemos que $\{P_0(x), P_1(x), \ldots, P_m(x)\}$ forman una base para el espacio vectorial de los polinomios de grado a lo sumo m. Como $\pi(x)$ es de grado m, por el lema lo podemos expresar como

$$\pi(x) = \sum_{k=0}^{m} c_k P_k(x), \qquad c_m \neq 0,$$

donde los c_k son constantes que serán llamadas coeficientes de Fourier.

Si $\pi(x)$ lo multiplicamos escalarmente por los polinomios ortogonales, tenemos

$$\begin{aligned}\int_{\mathbb{R}} \pi(x) P_n(x)\, d\mu(x) &= \int_{\mathbb{R}} \sum_{k=0}^{m} c_k P_k(x) P_n(x)\, d\mu(x) \\ &= \sum_{k=0}^{m} c_k \Big(\int_{\mathbb{R}} P_k(x) P_n(x)\, d\mu(x) \Big).\end{aligned}$$

Las integrales del sumatorio se anulan cuando $k \neq n$; en consecuencia, si $m < n$ la k nunca llega a valer n y la expresión anterior es 0. Sin embargo, cuando $m = n$, el último sumando del sumatorio no se anula y la integral anterior es $c_n \int_{\mathbb{R}} P_n(x) P_n(x)\, d\mu(x)$.

Así hemos demostrado que (a) implica (b); como (c) es un caso particular de (b), (b) implica (c); y trivialmente, (c) implica (a), con lo que daríamos por concluida la demostración. □

Durante la demostración del teorema anterior han aparecido unos coeficientes que hemos llamado coeficientes de Fourier. Vamos a profundizar en ese concepto con el siguiente teorema.

Teorema 2 (Sobre los coeficientes de Fourier)**.** *Sea $\{P_n(x)\}_{n=0}^{\infty}$ una sucesión de polinomios ortogonales respecto a la medida μ. Entonces, para cualquier polinomio $\pi(x)$ de grado n se tiene que*

$$\pi(x) = \sum_{k=0}^{n} c_k P_k(x),$$

con

$$c_k = \frac{\int_{\mathbb{R}} \pi(x) P_k(x)\, d\mu(x)}{\int_{\mathbb{R}} P_k(x)^2\, d\mu(x)}, \qquad k = 0, 1, \ldots, n,$$

y estos coeficientes c_k se denominan coeficientes de Fourier.

Además, cada polinomio P_n de la SPO está determinado de manera única, salvo una constante multiplicativo.

Demostración. Como hemos visto en la demostración del teorema 1, si $\pi(x)$ en un polinomio de grado n entonces existen constantes c_k que satisfacen la siguiente igualdad:

$$\pi(x) = \sum_{k=0}^{n} c_k P_k(x).$$

Operando como allí, para $m \leq n$ obtenemos

$$\begin{aligned}\int_{\mathbb{R}} \pi(x) P_m(x)\, d\mu(x) &= \int_{\mathbb{R}} \sum_{k=0}^{n} c_k P_k(x) P_m(x)\, d\mu(x) \\ &= \sum_{k=0}^{n} c_k \left(\int_{\mathbb{R}} P_k(x) P_m(x)\, d\mu(x) \right) = c_m \int_{\mathbb{R}} P_m(x) P_m(x)\, d\mu(x).\end{aligned}$$

Como la última integral no es nula, la podemos pasar dividiendo y tenemos la expresión buscada para los coeficientes:

$$c_k = \frac{\int_{\mathbb{R}} \pi(x) P_k(x)\, d\mu(x)}{\int_{\mathbb{R}} P_k(x) P_k(x)\, d\mu(x)}. \qquad \square$$

Así, hemos visto que los coeficientes de Fourier son los coeficientes del desarrollo del polinomio $\pi(x)$ en la base $\{P_n(x)\}_{n=0}^{\infty}$ del espacio vectorial de los polinomios.

A continuación, vamos a ver otro concepto que resulta clave para nuestro trabajo, los momentos.

Definición 6 (Momentos de una medida)**.** Dada una medida no negativa μ, llamamos *momentos* de la medida a los valores

$$\int_{\mathbb{R}} x^n\, d\mu(x) = \mu_n, \quad n = 0, 1, 2, \ldots.$$

Por lo tanto, dada una SPO respecto a una medida tenemos una sucesión de momentos asociada a dicha medida.

Dichos momentos van a resultar cruciales para en el último capítulo demostrar el teorema de Favard. Para sacarles el máximo partido posible tenemos que dar una cuantas definiciones.

2.1.1. Sucesiones definidas positivas y matrices de Hankel

Como decíamos, veamos ahora unas definiciones que nos van a ser necesarias para este trabajo.

Definición 7 (Sucesión definida positiva)**.** Una sucesión de números reales $\{r_k\}_{k=0}^{\infty}$ se dice *definida positiva* si

$$\sum_{i,k=0}^{m} r_{i+k} a_i a_k > 0 \tag{2.1}$$

para cualquier vector no nulo $(a_0, \ldots, a_m)$, $m = 0, 1, 2 \ldots$.

Otra definición importante es la que presentamos a continuación.

Definición 8 (Determinante de Hankel)**.** Sea $\{r_i\}_{i=0}^{n}$ una sucesión de números reales. Llamamos *matriz de Hankel* H_n a la que tiene por entradas los valores $h_{i,j} = r_{i+j}$ (la primera entrada de la matriz es la $(0,0)$). Su determinante es

$$\Delta_n = \det(H_n) = \det(r_{i+j})_{i,j=0}^n = \begin{vmatrix} r_0 & r_1 & \cdots & r_n \\ r_1 & r_2 & \cdots & r_{n+1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ r_n & r_{n+1} & \cdots & r_{2n} \end{vmatrix}.$$

Estas dos definiciones están relacionadas mediante el siguiente teorema.

Teorema 3. *Una sucesión $\{r_k\}_{k=0}^{\infty}$ es definida positiva si y solo si sus determinantes de Hankel Δ_n son todos positivos.*

Previo a la demostración de este teorema vamos a ver un lema muy conocido que es el criterio de Sylvester. La demostración se ve en cualquier curso de álgebra lineal, así que no la vamos a incluir aquí. Para enunciar el lema, recordemos que una matriz cuadrada A, con entradas reales, se dice *definida positiva* si

$$a^T A a > 0$$

para cualquier vector columna a no nulo (y donde a^T denota el vector traspuesto); en esta definición a menudo se incluye el requisito de que la matriz A sea simétrica, pero esto a nosotros no nos va a afectar.

Lema 2 (Criterio de Sylvester)**.** *Una condición necesaria y suficiente para determinar si una matriz simétrica es definida positiva es que los menores principales de dicha matriz sean todos mayores que cero.*

Una vez visto el lema, pasemos ahora a demostrar el teorema.

Demostración del teorema 3. La noción de sucesión definida positiva (2.1) se puede escribir, en forma matricial, como

$$\begin{bmatrix} a_0 & a_1 & \cdot & \cdot & a_{m-1} & a_m \end{bmatrix} \begin{bmatrix} r_0 & r_1 & \cdots & r_m \\ r_1 & r_2 & \cdots & r_{m+1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ r_{m-1} & r_m & \cdots & r_{2m-1} \\ r_m & r_{m+1} & \cdots & r_{2m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_{m-1} \\ a_m \end{bmatrix} > 0.$$

Por lo tanto, que la sucesión sea definida positiva equivale a que la matriz de Hankel sea definida positiva. Teniendo en cuenta el criterio de Sylvester (ya que la matriz es simétrica) la matriz será definida positiva si y solo si los menores principales son todos positivos. Dichos menores corresponden a los determinantes de las matrices H_n, es decir, a Δ_n para $n = 0, 1, 2 \dots$. Así queda probada la equivalencia. $\square$

Continuemos, con —como hemos dicho antes— las condiciones que tienen que cumplir las medidas.

2.2. Condiciones sobre las medidas

Previamente hemos dicho que las SPO eran ortogonales respecto a una medida, pero nos podemos preguntar: ¿Todas las medidas tienen una SPO asociada? Si la respuesta es negativa, ¿qué condiciones tienen que cumplir las medidas para tener una SPO?

Antes de resolver estas dudas, veamos el siguiente ejemplo donde muestra que, aunque exista una medida, no tiene por qué tener una sucesión infinita de polinomios ortogonales asociados.

Ejemplo 14. Sea la medida $d\mu = \delta_0 + \delta_1$, y sean $P_0(x)$ y $P_1(x)$ los dos primeros polinomios ortogonales respecto a μ. En estas condiciones, no existe ningún polinomio $P_2(x)$ no nulo en $L^2(\mu)$ (es decir, no nulo en un conjunto de medida μ positiva) ortogonal a ambos. $\Diamond$

Comprobación. Notemos que, en el caso de que exista una SPO infinita para dicha medida también existirá una SPO mónica (ya que podemos dividir el polinomio original por su coeficiente director y así obtener un polinomio mónico). Así, en la demostración tomaremos que los polinomios son mónicos.

Sea $P_0(x) = 1$, $P_1(x) = x + a$ y $P_2(x) = x^2 + bx + c$. Vamos a aplicar la definición de SPO a dichos términos y durante el proceso encontraremos una contradicción que demuestra que no existe dicha SPO.

Comencemos con la condición $\int_{\mathbb{R}} P_k(x)^2\,d\mu \neq 0$:

$$\int_{\mathbb{R}} P_0(x)^2\,d\mu = 2,$$

$$\int_{\mathbb{R}} P_1(x)^2\,d\mu = 2a^2 + 2a + 1 \neq 0, \tag{2.2}$$

$$\int_{\mathbb{R}} P_2(x)^2\,d\mu = 2c^2 + b^2 + 1 + 2bc + 2b + 2c \neq 0; \tag{2.3}$$

y sigamos con la condición $\int P_k(x)P_t(x)\,d\mu = 0$:

$$\int_{\mathbb{R}} P_0(x)P_1(x)\,d\mu = 2a + 1 = 0, \tag{2.4}$$

$$\int_{\mathbb{R}} P_0(x)P_2(x)\,d\mu = 2c + b + 1 = 0, \tag{2.5}$$

$$\int_{\mathbb{R}} P_1(x)P_2(x)\,d\mu = ca + 1 + b + c + a + ab + ac = 0. \tag{2.6}$$

Aplicando (2.4) a (2.2), y como

$$0 \neq 2a^2 + 2a + 1 = 2a^2 + (2a + 1) = (2a + 1)a + a + 1,$$

obtenemos

$$a \neq 0 \qquad \text{y} \qquad a \neq -1.$$

Aplicando (2.5) a (2.6), y como

$$0 = ca + 1 + b + c + a + ab + ac = (2c + b + 1)a + (2c + b + 1) - c,$$

obtenemos

$$c = 0, \qquad \text{que implica} \qquad b = -1.$$

Aplicando estas últimas condiciones a (2.3) llegamos a una contradicción:

$$0 \neq 2c^2 + b^2 + 1 + 2bc + 2b + 2c = 2 \cdot 0^2 + (-1)^2 + 1 + 2 \cdot (-1) \cdot 0 + 2 \cdot (-1) + 2 \cdot 0 = 0.$$

Queda así probado que no existe una SPO infinita para la medida $d\mu = \delta_0 + \delta_1$.

Como aclaración, conviene mencionar que el polinomio $P_2(x) = x^2 - x = (x-1)x$ sí que es ortogonal a $P_0(x)$ y a $P_1(x)$ respecto a la medida $d\mu = \delta_0 + \delta_1$; pero, en esa medida, $P_2(x)$ es una función nula en $L^2(\mu)$ (y por eso $\int_{\mathbb{R}} P_2(x)^2\, d\mu = 0$), así que no puede formar parte de la SPO. $\square$

Tras ver este ejemplo, retomemos nuestra pregunta: ¿cuáles son las condiciones que tiene que cumplir una medida para que exista una SPO infinita asociada a ella? El siguiente teorema nos da una caracterización.

Teorema 4. *Sea μ una medida de Borel o Lebesgue y sea $\{\mu_n\}_{n=0}^{\infty}$ su sucesión de momentos asociada. Una condición necesaria y suficiente para la existencia de una SPO infinita para μ es que*

$$\Delta_n \neq 0, \ \ \textit{para } n = 0, 1, 2, \ldots$$

Demostración. Escribamos

$$P_n(x) = \sum_{k=0}^{n} c_{n,k} x^k.$$

Utilizando la caracterización del teorema 1, que esos polinomios sean ortogonales equivale a que

$$\int_{\mathbb{R}} x^m P_n(x)\, d\mu(x) = \sum_{k=0}^{n} c_{n,k} \mu_{k+m} K_n \delta_{m,n} \text{ donde } K_n \neq 0, \ m < n,$$

que si lo ponemos en forma de sistema matricial es

$$\begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{bmatrix} \begin{bmatrix} c_{n,0} \\ c_{n,1} \\ \cdot \\ \cdot \\ c_{n,n-1} \\ c_{n,n} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ K_n \end{bmatrix}.$$

Por lo tanto, si existe una SPO asociada a la sucesión de momentos estará únicamente determinada por las constantes K_n de las ecuaciones anteriores. Como esas constantes tienen solución única se tiene que cumplir que $\Delta_n \neq 0$ para $n \geq 0$.

Recíprocamente, si $\Delta_n \neq 0$, entonces, para un $K_n \neq 0$ arbitrario, la ecuación matricial anterior tiene solución única; y por lo tanto $P_n(x)$, cumpliendo la relación de ortogonalidad, existe. Además,

$$c_{n,n} = \frac{K_n \Delta_{n-1}}{\Delta_n} \neq 0, \text{ para } n \geq 1, \tag{2.7}$$

que sería válido también para $n = 0$ si definimos $\Delta_{-1} = 1$. Como $P_n(x)$ es de grado n, se tiene que, efectivamente, $\{P_n\}_{n=0}^{\infty}$ es una SPO para la sucesión de momentos $\{\mu_n\}_{n=0}^{\infty}$. □

Veamos otras relaciones que se cumplen usando la notación antes descrita.

Teorema 5. *Sea $\{P_n\}_{n=0}^{\infty}$ una SPO infinita. Entonces, para cualquier polinomio $\pi_n(x)$ de grado n,*

$$\int_{\mathbb{R}} \pi_n(x) P_n(x)\, d\mu(x) = a_n \int_{\mathbb{R}} x^n P_n(x)\, d\mu(x) = \frac{a_n k_n \Delta_n}{\Delta_{n-1}}, \qquad \Delta_{-1} = 1,$$

donde a_n y k_n corresponden a los coeficientes directores de $\pi_n(x)$ y $P_n(x)$, respectivamente.

Demostración. Sea

$$\pi_n(x) = a_n x^n + \pi_{n-1}(x),$$

con $\pi_{n-1}(x)$ el polinomio de grado $n-1$ que satisface la igualdad anterior. Desarrollando la integral pedida en el enunciado,

$$\begin{aligned}\int_{\mathbb{R}} \pi_n(x) P_n(x)\, d\mu(x) &= a_n \int_{\mathbb{R}} x^n P_n(x)\, d\mu(x) + \int_{\mathbb{R}} \pi_{n-1}(x) P_n(x)\, d\mu(x) \\ &= a_n \int_{\mathbb{R}} x^n P_n(x)\, d\mu(x) = a_n K_n,\end{aligned}$$

ya que el segundo sumando se anula por las propiedades de las SPO. Además, por la expresión vista en (2.7), teniendo en cuenta que en nuestro caso denotamos al coeficiente director k_n, obtenemos

$$a_n K_n = a_n \frac{k_n \Delta_n}{\Delta_{n-1}},$$

con lo quedaría demostrado el teorema. □

Vamos a introducir una condición sobre la medida que nos va a permitir demostrar teoremas importantes. Más adelante refinaremos esa condición para adecuarla al estudio de los polinomios ortogonales que queremos realizar en este trabajo.

Definición 9 (Medida definida positiva)**.** Sea $d\mu$ una medida (de Borel o Lebesgue). Diremos que $d\mu$ es una medida *definida positiva* si $\int_{\mathbb{R}} \pi(x)\, d\mu(x) > 0$ para todo polinomio $\pi(x)$ que no sea idénticamente nulo y tal que $\pi(x) \geq 0$ para todo x perteneciente a los reales.

Esta medida definida positiva permite construir una sucesión de polinomios ortogonales mediante el método de Gram-Schmidt.

Teorema 6 (Existencia de la SPO)**.** *Dada una medida $d\mu$ definida positiva, existe una SPO $\{p_n\}_{n=0}^{\infty}$ formada por polinomios con coeficientes directores positivos y que cumplen la relación de ortonormalidad.*

Demostración. Vamos a comenzar construyendo sucesivamente los polinomios. Para $k = 0$ tomamos

$$p_0(x) = \frac{1}{\sqrt{\int_{\mathbb{R}} d\mu(x)}},$$

que es de grado 0 y positivo.

Supongamos que tenemos construidos los polinomios $p_k(x)$ hasta $k = n$, de modo que

$$\int_{\mathbb{R}} p_k(x)p_j(x)\, d\mu(x) = 0, \text{ para } j = 0, 1, \ldots, k-1,$$

y

$$\int_{\mathbb{R}} p_k(x)^2\, d\mu(x) = 1.$$

Vamos a construir ahora $p_{n+1}(x)$ que cumpla lo mismo. Sea

$$\hat{p}_{n+1}(x) = x^{n+1} - \sum_{i=0}^{n} a_i p_i(x), \text{ donde } a_i = \int_{\mathbb{R}} x^{n+1} p_i(x)\, d\mu(x).$$

El polinomio $\hat{p}_{n+1}(x)$ tiene grado $n + 1$ y su coeficiente director es positivo. Además, es ortogonal a los de grado menor:

$$\begin{aligned}
&\int_{\mathbb{R}} \hat{p}_{n+1}(x)p_i(x)\, d\mu(x) \\
&\quad = \int_{\mathbb{R}} \Big(x^{n+1} - \sum_{j=0}^{n} a_j p_j(x)\Big) p_i(x)\, d\mu(x) \\
&\quad = \int_{\mathbb{R}} x^{n+1} p_i(x)\, d\mu(x) - \sum_{j=0}^{n} a_j \int_{\mathbb{R}} p_j(x)p_i(x)\, d\mu(x) \\
&\quad = a_i - a_i = 0, \quad \text{para } i \in \{0, \ldots, n\}.
\end{aligned}$$

Una vez visto que el polinomio $\hat{p}_{n+1}(x)$ que hemos construido es ortogonal con los anteriores, ya solo quedaría normalizarlo.

Sea

$$\int_{\mathbb{R}} \hat{p}_{n+1}(x)^2\, d\mu(x) = K > 0,$$

que es positivo por la definición de medida definida positiva (el polinomio $\hat{p}_{n+1}(x)^2$ solo se anula en sus ceros). Por lo que, tomando

$$p_{n+1}(x) = \frac{\hat{p}_{n+1}(x)}{\sqrt{K}},$$

dicho polinomio cumple las condiciones anteriores, terminamos la inducción y ya hemos probado la existencia. □

El teorema anterior está enunciado para medidas definidas positivas, pero a partir de ahora vamos a considerar un tipo más específico de ellas, las medidas que tienen un número infinito de puntos de crecimiento efectivo.

Definición 10 (Puntos de crecimiento efectivo)**.** Sea μ una medida positiva de Lebesgue o de Borel en los reales. Decimos que $t \in \mathbb{R}$ un *punto de crecimiento efectivo* para μ si

$$\int_{t-\varepsilon}^{t+\varepsilon} d\mu(x) > 0 \text{ para todo } \varepsilon > 0.$$

Y decimos que μ *tiene infinitos puntos de crecimiento efectivo* si el conjunto

$$\mathcal{E}(\mu) = \Big\{ t \in \mathbb{R} : t \text{ es de crecimiento efectivo para } \mu \Big\}$$

es infinito.

Habitualmente, que una medida tenga o no infinitos puntos de crecimiento efectivo es fácil de comprobar y, además, es un concepto bastante intuitivo. Al estudiar sucesiones de polinomios ortogonales respecto de medidas, a menudo se usa como condición previa que la medida tenga infinitos puntos de crecimiento efectivo, pues este requisito garantiza que dicha medida sea definida positiva (lo que, como hemos demostrado en el teorema 6, implica la existencia de la correspondiente SPO). Veámoslo:

Teorema 7. *Una medida tiene infinitos puntos de crecimiento efectivo si y solo si es definida positiva.*

Demostración. Sea μ una medida con infinitos puntos de crecimiento efectivo, y sea $\pi(x)$ un polinomio no idénticamente nulo y tal que $\pi(x) \geq 0$ para todo $x \in \mathbb{R}$. El polinomio π tendrá un número finito de raíces reales (quizás ninguna). Así pues, si una medida μ tiene infinitos puntos de crecimiento efectivo, el conjunto

$$\mathcal{E}(\mu) \setminus \{\text{raíces reales de } \pi\}$$

también es infinito; de hecho, nos basta con saber que tiene al menos un elemento, que denotamos como t. En un entorno $[t-\varepsilon, t+\varepsilon]$ de t, el polinomio $\pi(x)$ no tiene más ceros, y está acotado inferiormente por una constante positiva c. Entonces,

$$\int_{\mathbb{R}} \pi(x)\, d\mu(x) \geq \int_{t-\varepsilon}^{t+\varepsilon} \pi(x)\, d\mu(x) \geq c \int_{t-\varepsilon}^{t+\varepsilon} d\mu(x) > 0.$$

Recíprocamente, sea μ una medida definida positiva, y supongamos que solo tiene un número finito N de puntos de crecimiento efectivo, que denotamos $x_1, x_2, \ldots, x_N$. Sea ahora el polinomio

$$\pi(x) = (x - x_1)^2 (x - x_2)^2 \cdots (x - x_N)^2.$$

Entonces,

$$\int_{-\infty}^{\infty} (x - x_1)^2 (x - x_2)^2 \cdots (x - x_N)^2 \, d\mu(x) = 0,$$

que es absurdo porque $\pi(x) \geq 0$ y la medida es definida positiva. □

Merece la pena aquí citar el ejemplo 14. Allí tomábamos la medida $d\mu = \delta_0 + \delta_1$ cuyos únicos puntos de crecimiento efectivo son 0 y 1. Y lo que veíamos es que solo existían dos polinomios ortogonales, P_0 y P_1. De hecho, lo mismo ocurre en las SPO discretos de Krawtchouk y Hahn (apartado 1.2.2): solo hay $N + 1$ puntos de crecimiento efectivo y $N + 1$ polinomios ortogonales en la familia.

2.2.1. Relaciones entre la medida, los momentos y los polinomios

Un resultado muy interesante es el siguiente (recordemos que con Δ_n estamos denotando los determinantes de la matriz de Hankel asociada a la medida):

Teorema 8. *Una medida es definida positiva si y solo si $\Delta_n > 0$ para todo $n \geq 0$.*

Para probarlo, usaremos un lema previo:

Lema 3. *Sea $\pi(x)$ un polinomio que es no negativo para todo x perteneciente a los reales. Entonces existen polinomios reales $p(x), q(x)$ tales que*

$$\pi(x) = p^2(x) + q^2(x).$$

Demostración. Como $\pi(x) \geq 0$ para cualquier real, entonces $\pi(x)$ tiene sus raíces reales de multiplicidad par y las raíces complejas aparecen en pares conjugados. Por tanto podemos escribir

$$\pi(x) = r^2(x) \prod_{k=1}^{m} (x - \alpha_k - i\beta_k)(x - \alpha_k + i\beta_k),$$

donde los coeficientes de $r(x)$, α_k y β_k son todos números reales. Además multiplicando solamente los factores en los que β_k aparece restando obtenemos

$$\prod_{k=1}^{m} (x - \alpha_k - i\beta_k) = A(x) + iB(X),$$

con $A(x)$ y $B(x)$ polinomios reales. De igual forma tenemos que

$$\prod_{k=1}^{m} (x - \alpha_k + i\beta_k) = A(x) - iB(X),$$

por lo que conjuntamente

$$\pi(x) = r^2(x)(A(x) + iB(X))(A(x) - iB(X)) = r^2(x)(A^2(x) + B^2(X)),$$

y tenemos los polinomios buscados. □

Y ya podemos demostrar el teorema:

Demostración del teorema 8. Comencemos suponiendo que tenemos una sucesión μ_n con $\Delta_n > 0$, $n \geq 0$. Según el teorema 4, existe una SPO $\{P_n\}_{n=0}^{\infty}$ para dicha sucesión de momentos. Además, sin pérdida de generalidad, podemos asumir que los P_n son mónicos. Por la fórmula vista en el teorema 5 sabemos que

$$\int_{\mathbb{R}} P_n(x)^2 \, d\mu = \frac{\Delta_n}{\Delta_{n-1}} > 0.$$

Por lo tanto, si $p(x)$ es un polinomio de grado m, entonces

$$p(x) = \sum_{k=0}^{m} a_k P_k(x),$$

donde los a_k son todos reales y $a_m \neq 0$. De esa forma

$$\int_{\mathbb{R}} p(x)^2 \, d\mu = \sum_{j,k=0}^{m} a_j a_k \int_{\mathbb{R}} P_j(x) P_k(x) \, d\mu = \sum_{k=0}^{m} a_k^2 \int_{\mathbb{R}} P_k(x)^2 \, d\mu > 0.$$

Por el lema previo sabemos que todo polinomio no negativo lo ponemos poner como suma de dos polinomios al cuadrado, si calculamos la integral de dicho polinomio no negativo respecto a μ podemos concluir que μ es una medida definida positiva, sin más que viendo que su integral es positiva por ser el resultado de las integrales de los polinomios al cuadrado en los que lo hemos partido, es decir,

$$\pi(x) = p(x)^2 + q(x)^2,$$

y ya hemos visto que la integral de $p(x)^2$ es positiva y lo mismo podemos concluir con $q(x)^2$.

Veamos ahora la otra implicación (que si μ es definida positiva entonces los momentos asociados a dicha medida existen y $\Delta_n > 0$ para todo $n \geq 0$).

Si μ es definida positiva, por el teorema 6, sus momentos son reales y existe una SPO $\{P_n\}_{n=0}^{\infty}$ con respecto a dicha medida. Supongamos (sin pérdida de generalidad) que dicha SPO está formada por polinomios mónicos. Como antes, tenemos

$$0 < \int_{\mathbb{R}} P_n(x)^2 \, d\mu(x) = \frac{\Delta_n}{\Delta_{n-1}},$$

que se cumple para $n \geq 0$. Como $\Delta_{-1} = 1$, entonces $\Delta_n > 0$ para todo $n > 0$ y queda probado el teorema. □

Además, con los momentos también tenemos una forma de construir la SPO. Esta forma viene dada a partir de la matriz de Hankel y es la siguiente. Dado que este resultado no lo vamos a utilizar en el memoria, no incluimos su demostración que, no obstante, es sencilla y se puede encontrar en la mayoría de los textos de polinomios ortogonales.

Teorema 9. *Sea $\{P_n\}_{n=0}^{\infty}$ una SPO mónica y sea $\{\mu_n\}_{n=0}^{\infty}$ la sucesión de momentos asociada a dicha SPO. Entonces podemos expresar los polinomios de la sucesión en función de los momentos con la siguiente fórmula:*

$$P_n(x) = \frac{1}{\Delta_{n-1}} \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \mu_{n-1} & \mu_n & \cdots & \mu_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix},$$

donde Δ_{n-1} es el determinante de Hankel.

2.3. La fórmula de recurrencia a tres términos

En esta sección vamos a demostrar que los polinomios ortogonales satisfacen fórmulas recursivas que ligan términos de distintos grados, de manera análoga a lo que ocurre con la sucesión de Fibonacci $F_{n+1} = F_n + F_{n-1}$. Como hemos dicho al comienzo del trabajo, la fórmula de recurrencia a tres términos es una propiedad importantísima de las SPOs. De hecho, como veremos más adelante con la demostración del teorema de Favard, es una caracterización de las SPOs. Sin más dilación, vamos a probar la existencia de dicha fórmula en recurrencias.

Teorema 10 (Existencia de la fórmula de recurrencia a tres términos)**.** *Sea μ una medida definida positiva y sea $\{P_n\}_{n=0}^{\infty}$ la correspondiente SPO asociada que además vamos a tomar sin pérdida de generalidad que sea mónica. Entonces existen constantes c_n y $\lambda_n > 0$ tales que*

$$P_n(x) = (x - c_n)P_{n-1}(x) - \lambda_n P_{n-2}(x), \qquad n = 1, 2, \ldots, \tag{2.8}$$

definiendo $P_{-1}(x) = 0$.

Demostración. Como $xP_n(x)$ es un polinomio de grado $n + 1$, se puede escribir como

$$xP_n(x) = \sum_{k=0}^{n+1} a_{n,k} P_k(x), \qquad a_{n,k} = \frac{\int_{\mathbb{R}} xP_n(x)P_k(x)\, d\mu(x)}{\int_{\mathbb{R}} P_k(x)^2\, d\mu(x)}.$$

Si observamos los valores que puede tomar $a_{n,k}$ podemos decir que, como xP_k es un polinomio de grado $k + 1$, $a_{n,k} = 0$ cuando $0 \leq k < n - 1$. Además sabemos que $P_n(x)$ es mónico, luego el coeficiente director de $xP_n(x)$, $a_{n,n+1} = 1$. Por lo tanto tenemos que

$$xP_n(x) = P_{n+1}(x) + a_{n,n}P_n(x) + a_{n,n-1}P_{n-1}(x), \qquad n \geq 1.$$

Por comodidad en la demostración vamos a realizar los cambios de variable n por $n - 1$, $a_{n,n}$ por c_n y $a_{n,n-1}$ por λ_n, con lo que la expresión anterior nos queda

$$xP_{n-1}(x) = P_n(x) + c_nP_{n-1}(x) + \lambda_nP_{n-2}(x), \qquad n \geq 2,$$

y esta expresión es similar a la que nos pedían probar (2.8) excepto porque no tenemos asegurado que se cumpla para $n = 1$ y que $\lambda_n > 0$. En el caso $n = 1$, la expresión obtenida en la demostración se cumple cuando definimos $P_{-1}(x) = 0$ y elegimos $c_1 = -P_1(0)$.

Si además calculamos

$$\int_{\mathbb{R}} x^{n-2} P_n(x)\, d\mu(x) = \int_{\mathbb{R}} x^{n-1} P_{n-1}(x)\, d\mu(x) \\ - c_n \int_{\mathbb{R}} x^{n-2} P_{n-1}(x)\, d\mu(x) - \lambda_n \int_{\mathbb{R}} x^{n-2} P_{n-2}(x)\, d\mu(x)$$

obtenemos que

$$0 = \int_{\mathbb{R}} x^{n-1} P_{n-1}(x)\, d\mu(x) - \lambda_n \int_{\mathbb{R}} x^{n-2} P_{n-2}(x)\, d\mu(x).$$

Por lo tanto, según el teorema 5 obtenemos que

$$\lambda_{n+1} = \frac{\int_{\mathbb{R}} x^n P_n(x)\, d\mu(x)}{\int_{\mathbb{R}} x^{n-1} P_{n-1}(x)\, d\mu(x)} = \frac{\Delta_{n-2}\Delta_n}{\Delta_{n-1}^2},$$

y como sabemos que $\Delta_n > 0$ para $n \geq -1$, $\lambda_{n+1} > 0$ y ya quedan probadas todas las condiciones. $\square$

A partir de la ecuación de recurrencia probada en el teorema anterior se pueden obtener diversas propiedades

Teorema 11 (Propiedades de la fórmula de recurrencia)**.** *Dada una SPO (cuyos polinomios son mónicos) $\{P_n\}_{n=0}^{\infty}$ con respecto a una medida definida positiva μ, la fórmula de recurrencia* (2.8) *cumple las siguientes propiedades:*

(a)

$$\lambda_{n+1} = \frac{\int_{\mathbb{R}} P_n(x)^2\, d\mu(x)}{\int_{\mathbb{R}} P_{n-1}(x)^2\, d\mu(x)} = \frac{\Delta_{n-2}\Delta_n}{\Delta_{n-1}^2}.$$

(b) $\int_{\mathbb{R}} P_n(x)^2\, d\mu(x) = \lambda_1\lambda_2 \cdots \lambda_{n+1}$ *si definimos* $\lambda_1 = \mu_0 = \Delta_0$.

(c)

$$c_n = \frac{\int_{\mathbb{R}} x P_{n-1}(x)^2\, d\mu(x)}{\int_{\mathbb{R}} P_{n-1}(x)^2\, d\mu(x)}.$$

(d) El coeficiente de x^{n-1} *en* $P_n(x)$ *es* $-(c_1 + c_2 + \cdots + c_n)$.

Demostración. (a) En la demostración del teorema anterior llegamos a la siguiente fórmula:

$$\lambda_{n+1} = \frac{\int_{\mathbb{R}} x^n P_n(x)\, d\mu(x)}{\int_{\mathbb{R}} x^{n-1} P_{n-1}(x)\, d\mu(x)} = \frac{\Delta_{n-2}\Delta_n}{\Delta_{n-1}^2}.$$

Como $\int_{\mathbb{R}} x^n P_n(x)\, d\mu(x) = \int_{\mathbb{R}} P_n(x) P_n(x)\, d\mu(x)$, por las propiedades de ortogonalidad cuando los polinomios son mónicos, podemos expresar λ_{n+1} directamente en función de los elementos de la SPO.

(b) Trivialmente podemos ver que

$$\prod_{i=2}^{n+1} \lambda_i = \frac{\int_{\mathbb{R}} P_n(x)^2\, d\mu(x)}{\int_{\mathbb{R}} P_0(x)^2\, d\mu(x)},$$

por la simplificación telescópica de los factores en el numerador y denominador. Además, por el teorema 5 tenemos que

$$\int_{\mathbb{R}} P_0(x)^2\, d\mu(x) = \frac{\Delta_0}{\Delta_{-1}} \qquad (\Delta_{-1} = 1),$$

por lo tanto, si como dice el enunciado definimos $\lambda_1 = \mu_0 = \Delta_0$, se cumple que la expresión pedida.

(c) Si la identidad obtenida en (2.8) la multiplicamos escalarmente por P_{n-1} tenemos:

$$\begin{aligned}\int_{\mathbb{R}} & P_n(x) P_{n-1}(x)\, d\mu(x) \\ &= \int_{\mathbb{R}} (x - c_n) P_{n-1}(x)^2\, d\mu(x) - \int_{\mathbb{R}} \lambda_n P_{n-2} P_{n-1}(x)\, d\mu(x),\end{aligned}$$

que, por las propiedades de las SPO,

$$0 = \int_{\mathbb{R}} x P_{n-1}(x)^2\, d\mu(x) - c_n \int_{\mathbb{R}} P_{n-1}(x)^2\, d\mu(x) - 0,$$

por lo que finalmente nos queda

$$c_n = \frac{\int_{\mathbb{R}} x P_{n-1}(x)^2\, d\mu(x)}{\int_{\mathbb{R}} P_{n-1}(x)^2\, d\mu(x)}.$$

(d) Sea d_n el coeficiente del término x^{n-1} del polinomio $P_n(x)$. Si nos fijamos en los coeficientes de x^{n-1} en la expresión (2.8) tenemos que $d_n = d_{n-1} - c_n$. Para los primeros casos tenemos que $d_2 = d_1 - c_2$ y $d_1 = -c_1$, luego efectivamente se cumple la relación buscada. □

El teorema de la recurrencia del comienzo de esta sección lo hemos visto para SPO mónicos, ya que enunciarlo y probarlo resulta más sencillo y se puede hacer sin pérdida de generalidad. Veamos ahora como quedaría en la forma más general.

Teorema 12. *Dada una SPO $\{P_n\}_{n=0}^{\infty}$ con respecto a una medida definida positiva μ, se cumple la siguiente relación:*

$$P_{n+1}(x) = (A_n x + B_n) P_n(x) - C_n P_{n-1}(x), \qquad n \geq 0.$$

Demostración. Vamos a escribir $P_n(x) = k_n \hat{P}_n(x)$ con $\hat{P}_n(x)$ mónico. Entonces si obtenemos los valores de las sucesiones $\{c_n\}_{n=1}^{\infty}$ y $\{\lambda_n\}_{n=1}^{\infty}$ asociados a la SPO mónica $\{\hat{P}_n\}_{n=0}^{\infty}$, podemos expresar las sucesiones A_n, B_n y C_n de la siguiente forma:

$$A_n = k_n^{-1} k_{n+1}, \qquad B_n = -c_{n+1} k_n^{-1} k_{n+1}, \qquad C_n = \lambda_{n+1} k_{n-1}^{-1} k_{n+1},$$

para $n \geq 0$ y definiendo $k_{-1} = 1$. Además, como μ es definida positiva se tiene que $C_n A_n A_{n-1} > 0$, cuando $n \geq 1$. □

2.3.1. Fórmula de Christoffel-Darboux

Como ya hemos visto, para una SPO $\{P_n(x)\}_{n=0}^{\infty}$ con respecto a una medida definida positiva μ hay una fórmula de recurrencia para los polinomios (los suponemos mónicos) que es la siguiente:

$$P_n(x) = (x - c_n)P_{n-1}(x) - \lambda_n P_{n-2}(x), \qquad n = 1, 2, \dots,$$

y, por lo tanto, existen asociadas a dicha SPO dos sucesiones de números reales

$$\{c_n\}_{n=1}^{\infty} \quad \text{y} \quad \{\lambda_n\}_{n=1}^{\infty}, \qquad \lambda_n > 0,$$

que satisfacen la ecuación de recurrencia.

A partir de dicha ecuación vamos a probar una propiedad muy conocida, la identidad de Christoffel-Darboux.

Teorema 13 (Identidad de Christoffel-Darboux). *Sea una SPO $\{P_n(x)\}_{n=0}^{\infty}$ que cumple la relación de recurrencia*

$$P_{n+1}(x) = (x - c_{n+1})P_n(x) - \lambda_{n+1}P_{n-1}(x), \qquad n = 0, 1, \dots,$$

definiendo $P_{-1}(x) = 0$ y $P_0(x) = 1$, con las dos sucesiones $\{c_n\}_{n=1}^{\infty}$ y $\{\lambda_n\}_{n=1}^{\infty}$ ($\lambda_n > 0$) asociadas a la fórmula anterior. Entonces se cumple

$$\sum_{k=0}^{n} \frac{P_k(x)P_k(u)}{\lambda_1\lambda_2\cdots\lambda_{k+1}} = (\lambda_1\lambda_2\cdots\lambda_{k+1})^{-1}\frac{P_{n+1}(x)P_n(u) - P_n(x)P_{n+1}(u)}{x-u}. \tag{2.9}$$

Demostración. Por la fórmula de recurrencia, tenemos

$$xP_n(x)P_n(u) = P_{n+1}(x)P_n(u) + c_{n+1}P_n(x)P_n(u) + \lambda_{n+1}P_{n-1}(x)P_n(u),$$

$$uP_n(u)P_n(x) = P_{n+1}(u)P_n(x) + c_{n+1}P_n(u)P_n(x) + \lambda_{n+1}P_{n-1}(u)P_n(x),$$

Restando ambas expresiones nos queda

$$\begin{aligned}(x-u)P_n(x)P_n(u) &= P_{n+1}(x)P_n(u) \\ &- P_{n+1}(u)P_n(x) - \lambda_{n+1}(P_{n-1}(u)P_n(x) - P_{n-1}(x)P_n(u)).\end{aligned} \tag{2.10}$$

Si ahora denotamos la parte derecha de (2.9) como $F_n(x,u)$, podemos escribir (2.10) como

$$\frac{P_m(x)P_m(u)}{\lambda_1\lambda_2\cdots\lambda_{m+1}} = F_m(x,u) - F_{m-1}(x,u), \qquad m \geq 0.$$

Sumando esta expresión desde $m = 0$ hasta n, obtenemos

$$\sum_{m=0}^{n} \frac{P_m(x)P_m(u)}{\lambda_1\lambda_2\cdots\lambda_{m+1}} = F_n(x,u) - F_{-1}(x,u),$$

y como $F_{-1}(x,u) = 0$ tenemos probada la identidad. □

Estas son algunas de las propiedades de los polinomios ortogonales construidos a partir de una medida. En el siguiente capítulo veremos las propiedades que cumplen los polinomios que satisfacen una ecuación en recurrencias a tres términos.

Capítulo 3

Propiedades de sucesiones en recurrencias de polinomios

En este capítulo queremos estudiar una serie de propiedades de los polinomios definidos mediante una ecuación en recurrencia a tres términos. Queremos recalcar que, de momento, estos polinomios no son ortogonales respecto a ninguna medida. De hecho, lo que perseguimos —y lo veremos en el último capítulo de esta memoria— es que a partir de la ecuación de recurrencia surge la ortogonalidad.

Vamos a partir de una sucesión de polinomios $\{y_k(x)\}_{k=0}^{\infty}$ que satisfacen la ecuación en recurrencias

$$b_k y_{k+1}(x) + a_k y_k(x) + b_{k-1} y_{k-1}(x) = x y_k(x), \qquad k = 1, 2, 3, \dots, \tag{3.1}$$

y, cuando convenga, impondremos las condiciones iniciales fijando los primeros polinomios y_0 e y_1 (de grados 0 y 1 respectivamente); por supuesto, esto requiere que los coeficientes b_k sean no nulos. Por lo tanto, a partir de y_0 e y_1 podemos construir los siguientes polinomios y_k para $k > 1$. Otras veces tomaremos $y_{-1} = 0$ y, así, para desarrollar la recurrencia solo hará falta explicitar y_0.

Más en concreto, tomando

$$y_0(x) = P_0(x) = 1, \quad y_1(x) = P_1(x) = \frac{x - a_0}{b_0} \quad \text{e} \quad y_k(x) = P_k(x),$$

la relación (3.1) permite construir una sucesión de polinomios $\{P_n(x)\}_{n=0}^{\infty}$ tal que el grado de $P_k(x)$ es k. Esta sucesión de polinomios es base del espacio vectorial de los polinomios en una variable con coeficientes reales, y cualquier polinomio puede ser expresado como combinación lineal de términos de dicha sucesión.

3.1. Funcional asociado a una sucesión en recurrencia

Como hemos comentado previamente, queremos que los polinomios de la relación de recurrencia sean ortogonales respecto a una medida. De momento, y como paso previo, esa ortogonalidad la vamos a dar respecto a un funcional (en lugar de respecto a una medida como teníamos en los dos capítulos anteriores); es decir, el funcional va a jugar el papel de la integral con respecto a una medida. De hecho, es habitual en otros contextos definir la ortogonalidad de polinomios respecto a un funcional (por ejemplo en el libro de Chihara [3]). Veamos cómo se define el funcional.

Partimos de que tenemos una sucesión $\{P_n(x)\}_{n=0}^{\infty}$ que satisface la relación de recurrencia a tres términos (3.1).

Deseamos construir un funcional lineal $\mathcal{S}$ cuyo dominio es todo el espacio de los polinomios. Queremos que cumpla

$$\mathcal{S}\{P_i(x)P_k(x)\} = \delta_{i,k},$$

es decir, de momento lo definimos así para ciertos polinomios de forma especial. A partir de esa expresión vamos a intentar definir el funcional para un polinomio cualquiera $R(x) = A(x)B(x)$ con

$$A(x) = \sum_{i=0}^{m} A_i P_i(x), \qquad B(x) = \sum_{k=0}^{n} B_k P_k(x).$$

Aplicando que $\mathcal{S}$ debe ser un funcional lineal, tenemos

$$\mathcal{S}\{R(x)\} = \sum_{i=0}^{m}\sum_{k=0}^{n} A_i B_k \mathcal{S}\{P_i(x)P_k(x)\} = \sum_{i=0}^{\text{mín}\{m,n\}} A_i B_i.$$

Para que esta definición sea válida hay que comprobar que, independientemente de la forma que elijamos los factores del polinomio $R(x)$, el valor es el mismo.

Realmente, basta probar que para polinomios arbitrarios $M(x)$, $N(x)$ (con grados m, n respectivamente), se cumple

$$\mathcal{S}\{A_1(x)B_1(x)\} = \mathcal{S}\{A_2(x)B_2(x)\}, \tag{3.2}$$

cuando

$$A_1(x) = xM(x), \quad B_1(x) = N(x), \quad A_2(x) = M(x), \quad B_2(x) = xN(x).$$

En efecto, si

$$\mathcal{S}\{(M(x)x)N(x)\} = \mathcal{S}\{M(x)(xN(x))\},$$

entonces, por tener los polinomios factorizaciones finitas, basta probar que, cambiando un factor genérico $f(x)$ de lado, el resultado es el mismo, es decir, que

$$\mathcal{S}\{(M(x)f(x))N(x)\} = \mathcal{S}\{M(x)(f(x)N(x))\}.$$

Sea $f(x) = \sum_{i\geq 0} f_i x^i$ (obviamente, nos estamos refiriendo a sumas finitas aunque no lo indiquemos expresamente); por la linealidad del funcional,

$$\mathcal{S}\Big\{\Big(M(x)\sum_{i\geq 0} f_i x^i\Big)N(x)\Big\} = \sum_{i\geq 0} f_i \mathcal{S}\{(M(x)x^i)N(x)\},$$

y, según nuestra hipótesis, podemos cambiar las x de lado, luego

$$\sum_{i\geq 0} f_i \mathcal{S}\{(M(x)x^i)N(x)\} = \sum_{i\geq 0} f_i \mathcal{S}\{M(x)(x^i N(x))\},$$

y, de forma análoga,

$$\sum_{i\geq 0} f_i \mathcal{S}\{M(x)(x^i N(x))\} = \mathcal{S}\Big\{M(x)\Big(\sum_{i\geq 0} f_i x^i N(x)\Big)\Big\} = \mathcal{S}\{M(x)(f(x)N(x))\}.$$

Por otra parte, de nuevo por la linealidad del funcional, es suficiente comprobar (3.2) para $M(x) = P_m(x)$ y $N(x) = P_n(x)$. En efecto, si lo tenemos probado con $P_m(x)$ y $P_n(x)$, para $M(x)$ y $N(x)$ arbitrarios de la forma

$$M(x) = \sum_{i=0}^{m} m_i P_i(x) \quad \text{y} \quad N(x) = \sum_{j=0}^{n} n_j P_j(x)$$

tenemos (donde m, n son los grados de los polinomios)

$$\mathcal{S}\{(xM(x))N(x)\} = \mathcal{S}\Big\{\Big(x\sum_{i=0}^{m} m_i P_i(x)\Big)\Big(\sum_{j=0}^{n} n_j P_j(x)\Big)\Big\} = \sum_{i=0}^{m} m_i \sum_{j=0}^{n} n_j \mathcal{S}\{(xP_i(x))P_j(x)\},$$

que por la hipótesis $\mathcal{S}\{(xP_i(x))P_j(x)\} = \mathcal{S}\{P_i(x)(xP_j(x))\}$, con lo que finalmente obtenemos

$$\sum_{i=0}^{m} m_i \sum_{j=0}^{n} n_j \mathcal{S}\{P_i(x)(xP_j(x))\} = \mathcal{S}\Big\{\Big(\sum_{i=0}^{m} m_i P_i(x)\Big)\Big(x\sum_{j=0}^{n} n_j P_j(x)\Big)\Big\} = \mathcal{S}\{M(x)(xN(x))\},$$

luego bastaría probarlo solo para ese caso.

Así pues, como hemos dicho, analicemos el caso $M(x) = P_m(x)$ y $N(x) = P_n(x)$. En esta situación tenemos

$$A_1(x) = xP_m(x) = b_m P_{m+1}(x) + a_m P_m(x) + b_{m-1} P_{m-1}(x),$$

$$B_2(x) = xP_n(x) = b_n P_{n+1}(x) + a_n P_n(x) + b_{n-1} P_{n-1}(x),$$

y, por lo tanto,

$$\begin{aligned}\mathcal{S}\{A_1(x)B_1(x)\} &= \mathcal{S}\{(b_m P_{m+1}(x) + a_m P_m(x) + b_{m-1} P_{m-1}(x))P_n(x)\} \\ &= b_m \mathcal{S}\{P_{m+1}(x)P_n(x)\} + a_m \mathcal{S}\{P_m(x)P_n(x)\} + b_{m-1}\mathcal{S}\{P_{m-1}(x)P_n(x)\};\end{aligned}$$

análogamente,

$$\begin{aligned}\mathcal{S}\{A_2(x)B_2(x)\} &= \mathcal{S}\{P_m(x)(b_n P_{n+1}(x) + a_n P_n(x) + b_{n-1} P_{n-1}(x))\} \\ &= b_n \mathcal{S}\{P_{n+1}(x)P_m(x)\} + a_n \mathcal{S}\{P_n(x)P_m(x)\} + b_{n-1}\mathcal{S}\{P_{n-1}(x)P_m(x)\},\end{aligned}$$

y los valores de $\mathcal{S}\{A_1(x)B_1(x)\}$ y $\mathcal{S}\{A_2(x)B_2(x)\}$ son 0 salvo en los casos $n = m + 1$, $n = m$ y $n = m - 1$ (o, respectivamente, $m = n - 1$, $m = n$ y $m = n + 1$). En los casos descritos es muy sencillo ver que el valor obtenido en ambos funcionales es el mismo.

Por lo tanto, ya tenemos la definición del funcional completa.

Recapitulando, tenemos un funcional lineal $\mathcal{S}$, cuyo dominio es todo el espacio de los polinomios, y que cumple (para los polinomios de la sucesión $\{P_n(x)\}_{n=0}^{\infty}$) la siguiente propiedad:

$$\mathcal{S}\{P_i(x)P_k(x)\} = \delta_{i,k}.$$

Es importante destacar que el comportamiento del funcional es similar al comportamiento de una integral —con respecto a una medida y su sucesión de polinomios ortogonales— visto en los capítulos anteriores. El funcional $\mathcal{S}$ que hemos definido es clave para las propiedades que veremos a continuación. Su correspondencia intuitiva con una medida resultará fundamental, finalmente, en el próximo capítulo, con la demostración del teorema de Favard.

3.1.1. Sucesión de momentos del funcional

Una vez definido el funcional lineal $\mathcal{S}$ a partir de los polinomios $\{P_n(x)\}_{n=0}^{\infty}$ de la relación de recurrencia, podemos asociarle una sucesión de momentos (al igual que había momentos respecto a una medida). En lo que sigue, ya no recordaremos a cada paso la linealidad ni que el funcional procede de la relación de recurrencia de los polinomios.

Definición 11 (Momentos de un funcional)**.** Dado el funcional $\mathcal{S}$, podemos definir la *sucesión de momentos* $\{s_k\}_{k=0}^{\infty}$ de la siguiente forma:

$$s_k = \mathcal{S}\{x^k\}, \qquad k = 0, 1, 2, \dots. \tag{3.3}$$

Tras dar la definición, vamos a probar una propiedad que tienen dichos momentos.

Teorema 14 (Positividad de la sucesión de momentos)**.** *La sucesión de momentos* $\{s_k\}_{k=0}^{\infty}$ *del funcional* $\mathcal{S}$ *siempre es definida positiva (según la definición 7).*

Demostración. Si tomamos

$$A(x) = \sum_{i=0}^{m} a_i x_i,$$

es claro que

$$\sum_{i,k=0}^{m} s_{i+k} a_i a_k = \mathcal{S}\{A(x)^2\}.$$

Por otra parte, el polinomio $A(x)$ puede expresarse en función de los polinomios ortogonales $P_k(x)$ en la forma

$$A(x) = \sum_{k=0}^{m} c_k P_k(x),$$

por lo que

$$\mathcal{S}\{A(x)^2\} = \sum_{k=0}^{m} c_k^2.$$

En consecuencia,

$$\sum_{i,k=0}^{m} s_{i+k} a_i a_k = \sum_{k=0}^{m} c_k^2. \qquad \square$$

Implícitamente, hemos probado que

$$\mathcal{S}\{A(x)^2\} \geq 0$$

y solo es $= 0$ cuando $A(x)$ es el polinomio nulo. En muchos textos, esto se expresa diciendo que el funcional $\mathcal{S}$ es *definido positivo.*

3.1.2. Polinomios de segunda clase

Recordemos la sucesión en recurrencias con la que hemos empezado el capítulo:

$$b_k y_{k+1} + a_k y_k + b_{k-1} y_{k-1} = x y_k, \qquad k = 1, 2, 3, \ldots.$$

Cuando fijábamos las condiciones iniciales

$$P_0(x) = 1, \qquad P_1(x) = \frac{x - a_0}{b_0},$$

obteníamos la sucesión de polinomios ortogonales $\{P_n(x)\}_{n=0}^{\infty}$.

Por otra parte, fijando las condiciones iniciales

$$Q_0(x) = 0, \qquad Q_1(x) = \frac{1}{b_0},$$

obtenemos otra sucesión de polinomios $\{Q_n(x)\}_{n=0}^{\infty}$ que cumple la relación en recurrencias, y que en lo que sigue nos va a resultar útil. En dicha sucesión, $Q_k(x)$ es un polinomio de grado $k-1$. Estos polinomios $Q_k(x)$ se llaman polinomios de segunda clase, mientras que los descritos inicialmente $P_k(x)$ se llaman polinomios de primera clase.

Ambos tipos de polinomios están relacionados entre ellos a través del funcional, como vemos a continuación:

Teorema 15. *Si $\{P_n\}_{n=0}^{\infty}$ y $\{Q_n\}_{n=0}^{\infty}$ son las anteriores sucesiones de polinomios, se cumple*

$$Q_k(x) = \mathcal{S}_z\left\{\frac{P_k(x) - P_k(z)}{x - z}\right\}, \tag{3.4}$$

donde $\mathcal{S}_z$ denota que el funcional se calcula sobre la variable z.

Demostración. Veamos por inducción que, efectivamente, se cumple (3.4). Los casos $k = 0$ y $k = 1$ quedan

$$\mathcal{S}_z\left\{\frac{P_0(x) - P_0(z)}{x - z}\right\} = \mathcal{S}_z\left\{\frac{1 - 1}{x - z}\right\} = \mathcal{S}_z\{0\} = 0 = Q_0,$$

$$\mathcal{S}_z\left\{\frac{P_1(x) - P_1(z)}{x - z}\right\} = \mathcal{S}_z\left\{\frac{\frac{x-a_0}{b_0} - \frac{z-a_0}{b_0}}{x - z}\right\} = \mathcal{S}_z\left\{\frac{1}{b_0}\right\} = \frac{1}{b_0} = Q_1.$$

Lo suponemos cierto para $k = t$ y lo vamos a probar para $k = t + 1$. Por la fórmula de recurrencia $b_k y_{k+1}(x) = (x - a_k)y_k(x) - b_{k-1}y_{k-1}(x)$,

$$\begin{aligned}
&\mathcal{S}_z\left\{\frac{P_{t+1}(x) - P_{t+1}(z)}{x - z}\right\} = \mathcal{S}_z\left\{\frac{\frac{(x-a_t)P_t(x)-b_{t-1}P_{t-1}(x)}{b_t} - \frac{(z-a_t)P_t(z)-b_{t-1}P_{t-1}(z)}{b_t}}{x - z}\right\}\\
&= \frac{1}{b_t}\left(\mathcal{S}_z\left\{\frac{xP_t(x) - zP_t(z)}{x - z}\right\} - a_t\mathcal{S}_z\left\{\frac{P_t(x) - P_t(z)}{x - z}\right\} - b_{t-1}\mathcal{S}_z\left\{\frac{P_{t-1}(x) - P_{t-1}(z)}{x - z}\right\}\right)\\
&= \frac{1}{b_t}\left(\mathcal{S}_z\left\{\frac{xP_t(x) - xP_t(z)}{x - z}\right\} - \mathcal{S}_z\left\{\frac{xP_t(z) - zP_t(z)}{x - z}\right\} - a_kQ_t(x) - b_{t-1}Q_{t-1}(x)\right)\\
&= \frac{1}{b_t}(xQ_t(x) - 0 - a_kQ_t(x) - b_{t-1}Q_{t-1}(x)) = Q_{t+1}(x),
\end{aligned}$$

con lo que queda probada la relación. □

Pero la relación (3.4) no es la única interesante entre dichos polinomios. Veamos ahora otras propiedades y fórmulas que nos van a ayudar a sacarle todo el partido posible a la ecuación en recurrencias.

3.2. Propiedades de las sucesiones de polinomios

Vamos a mostrar una serie de interesantes fórmulas que son análogas a las que aparecen en la teoría de ecuaciones diferenciales lineales, y por ello le daremos el mismo nombre que allí tienen.

Comencemos con la fórmula de Liouville-Ostrogradski, que relaciona las distintas soluciones que se obtienen, en una ecuación en recurrencias, cuando se parte de valores iniciales distintos.

Teorema 16 (Fórmula de Liouville-Ostrogradski)**.** *Dados P_k y Q_k con las definiciones de la sección anterior, se tiene*

$$P_{k-1}(x)Q_k(x) - P_k(x)Q_{k-1}(x) = \frac{1}{b_{k-1}} \qquad k = 1, 2, 3, \ldots.$$

Demostración. Vamos a probarlo por inducción. Para el caso $k = 1$ queremos probar que

$$P_0(x)Q_1(x) - P_1(x)Q_0(x) = \frac{1}{b_0}.$$

Por definición,

$$P_0(x) = 1, \quad P_1(x) = \frac{x - a_0}{b_0}, \quad Q_0(x) = 0, \quad Q_1(x) = \frac{1}{b_0},$$

así que

$$P_0(x)Q_1(x) - P_1(x)Q_0(x) = 1 \cdot \frac{1}{b_0} - \frac{x - a_0}{b_0} \cdot 0 = \frac{1}{b_0},$$

y concuerda con la fórmula a probar.

Suponemos cierto el resultado hasta $k = t$, es decir,

$$P_{t-1}(x)Q_t(x) - P_t(x)Q_{t-1}(x) = \frac{1}{b_{t-1}}. \tag{3.5}$$

Veamos que se cumple para $k = t + 1$. Partimos de la ecuación en recurrencias inicial

$$b_k y_{k+1} + a_k y_k + b_{k-1} y_{k-1} = x y_k, \qquad k = 1, 2, 3, \ldots.$$

Tomando $y_k = P_k$ y multiplicando por Q_k obtenemos la expresión

$$b_k P_{k+1} Q_k + a_k P_k Q_k + b_{k-1} P_{k-1} Q_k = x P_k Q_k.$$

De forma equivalente con $y_k = Q_k$ y multiplicando por P_k,

$$b_k Q_{k+1} P_k + a_k Q_k P_k + b_{k-1} Q_{k-1} P_k = x Q_k P_k.$$

Restando ahora estas expresiones obtenemos

$$b_{k-1}(P_{k-1}Q_k - Q_{k-1}P_k) - b_k(P_k Q_{k+1} - P_{k+1} Q_k) = 0.$$

Con $k = t$ y aplicando la hipótesis de inducción (3.5) nos queda

$$b_{t-1} \frac{1}{b_{t-1}} = b_t(P_t Q_{t+1} - P_{t+1} Q_t),$$

que simplificando resulta la fórmula que queríamos probar para $k = t + 1$,

$$P_t Q_{t+1} - P_{t+1} Q_t = \frac{1}{b_t},$$

con lo que hemos finalizado la inducción y probado el teorema. □

Otra relación entre las distintas soluciones de la ecuación (3.1) es la que viene dada por la fórmula de Green, que relaciona los productos cruzados de dos soluciones distintas con una suma de productos de soluciones del mismo orden.

Teorema 17 (Fórmula de Green)**.** *Sea y_k una solución a la ecuación en recurrencias con variable y, y z_k otra solución con variable z. Entonces,*

$$b_{n-1}(y_{n-1}z_n - y_n z_{n-1}) - b_{m-1}(y_{m-1}z_m - y_m z_{m-1}) = (z - y) \sum_{k=m}^{n-1} y_k z_k.$$

Demostración. Partimos de la sucesión en recurrencias

$$b_k y_{k+1}(y) = (y - a_k) y_k(y) - b_{k-1} y_{k-1}(y).$$

Multiplicando por $z_k(z)$ obtenemos

$$b_k y_{k+1}(y) z_k(z) = (y - a_k) y_k(y) z_k(z) - b_{k-1} y_{k-1}(y) z_k(z).$$

Si tomamos la fórmula en recurrencias con solución $z_k(z)$ y multiplicamos esta vez por $y_k(y)$ nos queda

$$b_k z_{k+1}(z) y_k(y) = (z - a_k) z_k(z) y_k(y) - b_{k-1} z_{k-1}(z) y_k(y).$$

Restando las dos expresiones anteriores y reorganizando, obtenemos

$$b_k(y_k(y) z_{k+1}(z) - y_{k+1}(y) z_k(z)) = (z - y) y_k(y) z_k(z) + b_{k-1}(y_{k-1}(y) z_k(z) - y_k(y) z_{k-1}(z)).$$

Denotando $B_k = b_k(y_k(y) z_{k+1}(z) - y_{k+1}(y) z_k(z))$, podemos expresar la ecuación anterior como

$$B_k = (z - y) y_k(y) z_k(z) + B_{k-1}.$$

Usando esta fórmula recurrente para los B_k, la diferencia entre dos términos la podemos expresar como

$$B_{n-1} - B_{m-1} = (z - y) \sum_{k=m}^{n-1} y_k(y) z_k(z),$$

es decir, la expresión que teníamos que probar. $\square$

La última relación que vamos a ver en esta sección ya la habíamos visto antes en el capítulo anterior, y es la fórmula de Christoffel-Darboux. Sin embargo, en este caso, la obtenemos a partir del funcional y no a partir de la medida.

Teorema 18 (Fórmula de Christoffel-Darboux)**.** *Sea $\{P_n\}_{n=0}^{\infty}$ una familia de polinomios obtenida a partir de la relación de recurrencia a tres términos. Entonces,*

$$(z - y) \sum_{k=0}^{n-1} P_k(y) P_k(z) = b_{n-1}(P_{n-1}(y) P_n(z) - P_n(y) P_{n-1}(z).)$$

Demostración. Partiendo de la fórmula de Green y tomando $y_k = P_k(y)$, $z_k = P_k(z)$ y $m = 1$, obtenemos

$$\begin{aligned}
(z - y) &\sum_{k=1}^{n-1} P_k(y) P_k(z) \\
&= b_{n-1}(P_{n-1}(y) P_n(z) - P_n(y) P_{n-1}(z)) - b_0(P_0(y) P_1(z) - P_1(y) P_0(z)) \\
&= b_{n-1}(P_{n-1}(y) P_n(z) - P_n(y) P_{n-1}(z)) - b_0 \left(\frac{y - a_0}{b_0} - \frac{z - a_0}{b_0} \right) \\
&= b_{n-1}(P_{n-1}(y) P_n(z) - P_n(y) P_{n-1}(z)) - (z - y),
\end{aligned}$$

que introduciendo el $(z - y)$ final en el sumatorio de la izquierda queda

$$(z - y) \sum_{k=0}^{n-1} P_k(y) P_k(z) = b_{n-1}(P_{n-1}(y) P_n(z) - P_n(y) P_{n-1}(z)).$$ $\square$

3.3. Polinomios cuasiortogonales

Vamos a introducir a continuación unos polinomios que no tienen mucho interés en sí mismos, pero que nos van a resultar útiles en las demostraciones que seguirán.

Definición 12 (Polinomio cuasiortogonal)**.** Sea $\{P_n(x)\}_{n=0}^{\infty}$ una sucesión de polinomios ortogonales para el funcional $\mathcal{S}$. Para cada n, sea el polinomio

$$P_n(x;\alpha,\beta) = \alpha P_n(x) - \beta P_{n-1}(x),$$

donde α y β son parámetros que no valen cero a la vez (y estamos tomando $P_{-1}(x) \equiv 0$). Dicho polinomio $P_n(x;\alpha,\beta)$ se dice que es *cuasiortogonal* para el funcional $\mathcal{S}$ porque cumple

$$\mathcal{S}\{P_n(x;\alpha,\beta)x^k\} = 0, \qquad k = 1,2,3,\dots,n-2,$$

es decir, es ortogonal a todos los polinomios cuyo grado sea al menos dos unidades menor (en lugar de solo una unidad menor que es la propiedad de los polinomios ortogonales).

En la definición anterior, el caso $\alpha = 0$ es trivial; y, cuando $\alpha \neq 0$, podemos simplificar la notación de la siguiente forma:

$$P_n(x;\alpha,\beta) = \alpha P_n\left(x, \frac{\beta}{\alpha}\right)$$

con

$$P_n(x,t) = P_n(x) - tP_{n-1}(x). \tag{3.6}$$

Muchas veces será la notación dada en (3.6) la que usemos cuando queramos referirnos a polinomios cuasiortogonales, ya que $P_n(x,t)$ también es un polinomio cuasiortogonal.

Estos polinomios cuasiortogonales tienen un propiedad destacable que es la que vamos a ver a continuación.

Teorema 19. *Todos los ceros de un polinomio cuasiortogonal son simples y reales.*

Demostración. Tenemos el polinomio cuasiortogonal $P_n(x,t)$ y asumimos que solo cambia de signo en los valores

$$x_1 < x_2 < \cdots < x_m.$$

Entonces el polinomio

$$R(x) = P_n(x,t)M(x), \qquad \text{con} \quad M(x) = (x-x_1)(x-x_2)\dots(x-x_m),$$

es no negativo (el coeficiente director de cada $P_n(x)$ es positivo) y distinto de 0. Por lo tanto, aplicando el funcional a $R(x)$ tenemos

$$\mathcal{S}\{R(x)\} \geq 0.$$

Sin embargo, por ser $P_n(x,t)$ un polinomio cuasiortogonal, para que el valor del funcional $\mathcal{S}\{P_n(x,t)M(x)\}$ sea distinto de cero se tiene que cumplir que $m \geq n-1$. Lo que significa que $P_n(x,t)$ tiene al menos $n-1$ ceros simples. Como el grado de $P_n(x,t)$ es n y tiene al menos $n-1$ ceros simples reales podemos concluir que $P_n(x,t)$ tiene n ceros simples reales. $\square$

Nota 5. Conviene destacar que para el caso $t = 0$ estamos probando que los ceros de los polinomios ortogonales son simples y reales. Esta es una propiedad muy importante y que tiene grandes repercusiones en la aplicación de los polinomios ortogonales al campo de la física. En este trabajo, esta propiedad no tiene gran relevancia, pero sí que merecía la pena reseñarla. $\triangleleft$

3.4. Fórmulas de cuadratura

Las fórmulas que vamos a obtener en esta sección juegan un papel fundamental en la demostración del teorema de Favard, ya que sugieren como puede ser la medida que finalmente se construye. En lugar de presentar la fórmula como "teorema-demostración" vamos a obtenerla de forma justificada a partir de lo probado anteriormente.

Queremos encontrar una fórmula explícita que nos dé el valor de aplicar el funcional $\mathcal{S}$ a un polinomio cualquiera.

Antes de proseguir, vamos a recordar la fórmula de interpolación de Lagrange:

Nota 6 (Fórmula de interpolación de Lagrange). Sean n números reales distintos x_1, $x_2, \ldots, x_n$, y sus correspondientes imágenes $y_1, y_2, \ldots, y_n$ respecto a la función $f(x)$, es decir, $y_i = f(x_i)$. Entonces, existe un único polinomio $\pi_{n-1}(x)$ de grado menor que n tal que $\pi_{n-1}(x_i) = y_i$. Además ese polinomio se puede calcular de la siguiente forma:

$$\pi_{n-1}(x) = \sum_{k=1}^{n} y_k L_k(x),$$

con

$$L_k(x) = \frac{F(x)}{F'(x)(x - x_k)}, \qquad F(x) = \prod_{k=1}^{n}(x - x_k). \qquad \triangleleft$$

Partimos del polinomio cuasiortogonal de grado n,

$$P_n(x,t) = P_n(x) - tP_{n-1}(x),$$

donde, como hemos explicado antes, t es un número real. Además, denotamos por

$$x_1 < x_2 < \cdots < x_n$$

a los ceros del polinomio $P_n(x,t)$.

Tomamos ahora un polinomio arbitrario $R_{2n-2}(x)$ de grado $2n - 2$ y lo descomponemos de la siguiente forma:

$$R_{2n-2}(x) = P_n(x,t)R_{n-2}(x) + R_{n-1}(x), \tag{3.7}$$

donde $R_{n-2}(x)$ y $R_{n-1}(x)$ son polinomios, de grados $\leq n-2$ y $\leq n-1$ respectivamente, que satisfacen la igualdad. Notemos que, por la propiedad de cuasiortogonalidad, en dicha descomposición tenemos

$$\mathcal{S}\{R_{2n-2}(x)\} = \mathcal{S}\{R_{n-1}(x)\}. \tag{3.8}$$

Por otro lado, apliquemos la fórmula de interpolación de Lagrange a la función $f(x) = R_{n-1}(x)$ en los puntos $x_1, x_2, \dots, x_n$ (raíces de $P_n(x,t)$) y con $y_k = R_{n-1}(x_k)$. Así, obtenemos

$$R_{n-1}(x) = P_n(x,t) \sum_{k=1}^{n} \frac{R_{n-1}(x_k)}{P_n'(x_k,t)(x-x_k)},$$

donde $F(x) = P_n(x,t) = \prod_{k=1}^{n}(x-x_k)$ y P_n' denota la derivada respecto a la primera variable (el polinomio P_n lo podemos considerar mónico pues, si no lo fuera, su coeficiente director se simplificaría con el de P_n' en la fracción anterior).

Como $R_{2n-2}(x_k) = R_{n-1}(x_k)$ (porque $P_n(x_k,t) = 0$), podemos continuar la igualdad anterior con

$$R_{n-1}(x) = P_n(x,t) \sum_{k=1}^{n} \frac{R_{n-1}(x_k)}{P_n'(x_k,t)(x-x_k)} = P_n(x,t) \sum_{k=1}^{n} \frac{R_{2n-2}(x_k)}{P_n'(x_k,t)(x-x_k)}. \tag{3.9}$$

Combinando las ecuaciones (3.8) y (3.9) con la linealidad del funcional obtenemos

$$\mathcal{S}\{R_{2n-2}(x)\} = \sum_{k=1}^{n} \frac{R_{2n-2}(x_k)}{P_n'(x_k,t)} \mathcal{S}_x \left\{ \frac{P_n(x,t)}{x-x_k} \right\}.$$

Teniendo en cuenta que $P_n(x_k,t) = 0$, los $\mathcal{S}_x$ de esta expresión podemos escribirlos como

$$\mathcal{S}_x \left\{ \frac{P_n(x,t)}{x-x_k} \right\} = \mathcal{S}_x \left\{ \frac{P_n(x,t) - P_n(x_k,t)}{x-x_k} \right\} = Q_n(x_k,t).$$

Además, merece la pena estudiar en detenimiento el caso $t = 0$, ya que se tiene $P_n(x,t) = P_n(x)$ y el polinomio es ortogonal (en lugar de solo cuasiortogonal). Así, la descomposición (3.7) queda $R_{2n-1}(x) = P_n(x)R_{n-1}(x) + R_n(x)$ cumpliendo $\mathcal{S}\{R_{2n-1}(x)\} = \mathcal{S}\{R_n(x)\}$.

En consecuencia, hemos probado lo siguiente:

Teorema 20. *Para un polinomio arbitrario $R(x)$ de grado $\leq 2n-2$ se cumple la siguiente fórmula de cuadratura:*

$$\mathcal{S}\{R(x)\} = \sum_{k=0}^{n} q_k R(x_k) \qquad \text{con} \qquad q_k = q_k^{(n)}(t) = \frac{Q_n(x_k,t)}{P_n'(x_k,t)}. \tag{3.10}$$

Además, para $t = 0$, la fórmula es válida para cualquier polinomio $R(x)$ de grado $\leq 2n-1$.

Los coeficientes que hemos denotado por q_k tienen otras dos expresiones importantes que vamos a necesitar.

Partimos de los coeficientes de la fórmula (3.10), donde, además, podemos poner

$$q_k = \frac{Q_n(x_k) - tQ_{n-1}(x_k)}{P_n'(x_k) - tP_{n-1}'(x_k)}. \tag{3.11}$$

También sabemos que $P_n(x_k,t) = P_n(x_k) - tP_{n-1}(x_k) = 0$, es decir,

$$P_n(x_k) = tP_{n-1}(x_k).$$

Multiplicando numerador y denominador de (3.11) por $P_{n-1}(x_k)$ y simplificando, obtenemos

$$q_k = \frac{P_{n-1}(x_k)Q_n(x_k) - P_n(x_k)Q_{n-1}(x_k)}{P_{n-1}(x_k)P_n'(x_k) - P_n(x_k)P_{n-1}'(x_k)}.$$

El numerador de la fracción es, por la fórmula de Liouville-Ostrogradskii, $1/b_{n-1}$. Para simplificar el denominador partimos de la fórmula de Christoffel-Darboux

$$(z-y)\sum_{k=0}^{n-1} P_k(y)P_k(z) = b_{n-1}(P_{n-1}(y)P_n(z) - P_n(y)P_{n-1}(z)),$$

donde, si hacemos tender $z \to y$, introducimos el término $P_{n-1}(y)P_n(y)$ y usamos la definición de derivada, obtenemos

$$\sum_{k=0}^{n-1} P_k(y)P_k(z) = b_{n-1}(P_{n-1}(y)P_n'(y) - P_n(y)P_{n-1}'(y)).$$

Así, nuestro denominador quedaría

$$P_{n-1}(x_k)P_n'(x_k) - P_n(x_k)P_{n-1}'(x_k) = \frac{\sum_{k=0}^{n-1} P_k(x_k)^2}{b_{n-1}}.$$

Por lo tanto, la expresión final resultante es

$$q_k = \frac{1}{\sum_{k=0}^{n-1} P_k(x_k)^2}.$$

Además, nótese que esta fórmula demuestra que los coeficientes q_k son positivos.

La segunda expresión para los q_k la obtenemos a partir del polinomio

$$\pi(x) = \left(\frac{P_n(x,t)}{P_n'(x_k,t)(x-x_k)}\right)^2.$$

Este polinomio es de grado $2n-2$, es igual a 1 en $x = x_k$, y vale 0 en el resto de raíces de $P_n(x,t)$, que denotamos como $x = x_i$. Si introducimos ahora $\pi(x)$ en la fórmula de cuadratura (3.10) obtenemos

$$\mathcal{S}\{\pi(x)\} = \sum_{i=0}^{n} q_i \pi(x_i) = q_k \pi(x_k) = q_k,$$

es decir,

$$q_k = \mathcal{S}\left\{\left(\frac{P_n(x,t)}{P_n'(x_k,t)(x-x_k)}\right)^2\right\}.$$

Con estos coeficientes q_k, se puede encontrar una expresión que resulta de gran utilidad para calcular los momentos del funcional asociado a una sucesión en recurrencias. En concreto, a partir de la fórmula de cuadratura (3.10), tomando $R(x) = x^m$ tenemos lo siguiente:

Teorema 21. *Con la notación que estamos usando, y recordando que $q_k = q_k^{(n)}(t)$, se cumple*

$$\mathcal{S}\{x^m\} = \sum_{k=1}^{n} q_k x_k^m = s_m, \qquad m = 0, 1, \ldots, 2n-2.$$

Además, para $t = 0$, se cumple

$$\mathcal{S}\{x^m\} = \sum_{k=1}^{n} q_k x_k^m = s_m, \qquad m = 0, 1, \ldots, 2n-1. \tag{3.12}$$

Con este teorema ya tenemos todas las herramientas necesarias (que las sucesiones en recurrencias nos proporcionan) para poder abordar el teorema de Favard.

Capítulo 4

Teorema de Favard

El teorema de Favard es importante desde el punto de vista teórico porque establece una relación entre medidas sobre la recta real y sucesiones de polinomios ortogonales (esta relación no es biyectiva, pues una sucesión de polinomios puede ser ortogonal respecto a varias medidas distintas, pero esto queda fuera de los objetivos de este trabajo, a nosotros nos basta con que exista una medida).

Como curiosidad, aquí encontramos una mala atribución de nombres al teorema, pues antes de que Favard en 1935 probara su resultado (véase [4]), este ya había sido demostrado, entre otros, por Perron (1929), Wintner (1929) y Stone (1932). Además, anterior a estos, Stieltjes había demostrado una primera versión del teorema (con funcionales) y Chebyschev lo había probado para medidas con soporte finito.

Existen distintas formas de probar este teorema. Nosotros vamos a desarrollar una que calcula una medida de ortogonalidad para una familia de polinomios que verifican una fórmula de recurrencia a tres términos según el siguiente esquema: de la fórmula de recurrencia se deducen las propiedades de los ceros y la fórmula de cuadratura (ya lo hemos visto en el capítulo anterior); la fórmula de cuadratura genera entonces una sucesión de medidas discretas; esa sucesión acumula en, por lo menos, una medida (teoremas de Helly), y se ve que esa medida ortogonaliza los polinomios de partida. Loa textos que más se aproximan a nuestro desarrollo son los de [1, capítulo 2] y [5, capítulo 2].

Así pues, antes de comenzar con el teorema de Favard necesitamos introducir los teoremas de Helly.

Todas las medidas que vamos a manejar serán medidas de Stieltjes, que es un caso particular de medidas de Lebesgue (esto no es ningún problema para nuestros objetivos de demostrar la existencia de una medida a partir de la relación de recurrencia a tres términos). En consecuencia, a menudo utilizaremos la notación típica de medidas de Stieltjes df, donde f sería una función no decreciente (a la que a menudo se alude con el nombre de función de distribución).

4.1. Teoremas de Helly

Veamos los teoremas de Helly que nos van a permitir hacer converger una sucesión de medidas hacia otra medida. En lo que a nosotros respecta, la idea subyacente es hacer converger una sucesión de medidas, cada una de ellas con un número finito de puntos de crecimiento efectivo, hacia otra que ya tendrá infinitos puntos de crecimiento efectivo. Para profundizar más en los consecuencias de estos teoremas y conceptos relacionados puede consultarse el libro de Shohat y Tamarkin [9] o el de Freud [5].

Teorema 22 (Teorema de selección de Helly)**.** *Sea una sucesión de funciones $\{f_n(x)\}$ monótonas no decrecientes sobre $\mathbb{R}$ con cotas superior e inferior comunes para cada x. Entonces existe un función f y una subsucesión $\{f_{n_k}(x)\}$ que converge a $f(x)$ para todo x.*

Demostración. Elegimos un conjunto denso numerable $\mathbf{M}$ (por ejemplo, los racionales $\mathbb{Q}$) en los reales y lo ordenamos como $\eta_1, \eta_2, \dots$. Elegimos de nuestra sucesión de funciones $\{f_n\}$ una subsucesión $\{f_{n_1}\}$ tal que $\lim_{n_1\to\infty} f_{n_1}(\eta_1)$ exista. Esto es posible, ya que la sucesión $\{f_{n_1}\}$ esta acotada. Con $\{f_{n_1}\}$ seleccionamos una subsucesión $\{f_{n_2}\}$ de forma análoga, y así sucesivamente.

La "sucesión diagonal" $\{f_{n_n}\}$ converge para todo $x \in \mathbf{M}$ y $\lim_{n\to\infty} f_{n_n}(x) = f(x)$ es una función no decreciente, definida para todo x perteneciente a $\mathbf{M}$. Como $\mathbf{M}$ es denso en los reales, la función $f(x)$ puede ser extendida de forma única a una función no decreciente en $\mathbb{R}\backslash\mathbf{M}^*$, donde $\mathbf{M}^*$ es el conjunto numerable de los puntos de discontinuidad de $f(x)$. Así, se puede ver que

$$\lim_{n\to\infty} f_{n_n}(x) = f(x)$$

para los x pertenecientes a $\mathbb{R} \setminus \mathbf{M}^*$.

Ahora realizamos la misma construcción diagonal para el conjunto $\mathbf{M}^*$ con la sucesión $\{f_{n_n}\}$, es decir, creamos una nueva sucesión diagonal a partir de la sucesión $\{f_{n_n}\}$. De esta forma obtenemos una subsucesión convergente para todo punto y el teorema queda demostrado. □

Teorema 23 (Teorema de convergencia de Helly)**.** *Sea $\{f_n\}$ una sucesión de funciones monótonas no decrecientes en $[a, b]$ (intervalo finito) uniformemente acotadas, y convergentes para todo x en $[a, b]$ a una función $f(x)$. Sea además $g(x)$ una función continua en $[a, b]$. Entonces*

$$\lim_{n\to\infty} \int_a^b g(x)\, df_n(x) = \int_a^b g(x)\, df(x).$$

Demostración. Como $[a, b]$ es finito, $f(x)$ es uniformemente continua en $[a, b]$. Para un $\varepsilon > 0$ arbitrario, elegimos un δ tal que para $|h| < \delta$ se cumpla la desigualdad $|f(x+h) - f(x)| < \varepsilon$.

Elegimos una subdivisión $a = \eta_0 < \eta_1 < \cdots < \eta_n = b$ arbitraria del intervalo cuyas partes no sean mayores que δ. Las sumas de Riemann

$$\sigma_n = \sum_{k=1}^{n} g(\eta_k)\big(f_n(\eta_k) - f_n(\eta_{k-1})\big)$$

y

$$\sigma = \sum_{k=1}^{n} g(\eta_k)(f(\eta_k) - f(\eta_{k-1}))$$

difieren, respectivamente, en menos de $2M\varepsilon$ (para alguna constante M) de las siguientes integrales:

$$\int_a^b g(x)\,df_n(x), \qquad \int_a^b g(x)\,df(x).$$

Debido a que $f_n(\eta_k) \to f(\eta_k)$ tenemos que $\sigma_n \to \sigma$, y para valores suficientemente grandes de n, esto implica

$$\left| \int_a^b g(x)\,df_n(x) - \int_a^b g(x)\,df(x) \right| < 3M\varepsilon.$$

Como $\varepsilon > 0$ era arbitrario, esto demuestra el teorema. □

4.2. Teorema de Favard

Como hemos ido anunciando previamente, la demostración del teorema de Favard es el objetivo último de este trabajo.

Teorema 24 (Teorema de Favard)**.** *Sea una sucesión de polinomios cualesquiera $\{P_n\}_{n=0}^{\infty}$ (P_n de grado n), y dos sucesiones de números reales $\{\alpha_n\}$ y $\{\beta_n\}$ (con $\beta_n > 0$) de modo que*

$$\beta_{k-1}P_{k-1}(x) + \alpha_k P_k(x) + \beta_k P_{k+1}(x) = xP_k(x), \qquad k = 1, 2, 3, \dots,$$

con $P_{-1}(x) = 0$, $P_0(x) = 1$. Entonces existe una medida μ tal que

$$\int_{\mathbb{R}} P_m(x)P_n(x)\,d\mu(x) = 0, \tag{4.1}$$

para $n \neq m$, y

$$\int_{\mathbb{R}} P_n(x)^2\,d\mu(x) > 0; \tag{4.2}$$

es decir, la sucesión $\{P_n(x)\}_{n=0}^{\infty}$ es ortogonal con respecto a la medida μ.

Demostración. Como paso previo, vamos a construir una sucesión $\{s_n\}_{n=0}^{\infty}$ que deberán ser los momentos de la medida. Como aún no tenemos medida, esos s_n los definimos como los momentos del funcional, tal como hemos hecho en (3.3):

$$s_k = \mathcal{S}\{x^k\}.$$

Una vez que ya tenemos los $\{s_n\}_{n=0}^{\infty}$, vamos a hacer la demostración del teorema en dos partes. Primero, para cada n, vamos a intentar encontrar un medida que cumpla

$$\int_{-\infty}^{\infty} x^k\,d\mu(x) = s_k, \qquad k = 0, 1, \dots, 2n-1.$$

Como los polinomios cumplen la sucesión en recurrencias podemos aplicar los teoremas y fórmulas que hemos visto en el capítulo anterior.

Concretamente, por la fórmula (3.12) se tiene

$$s_k = \sum_{i=1}^{n} q_i x_i^k, \qquad k = 0, 1, \dots, 2n-1.$$

Con esta información podemos construir una medida μ_n a partir de una función de distribución $M_n(x)$ constante a trozos que comienza valiendo 0 y incrementa su valor en q_i en los puntos x_i, es decir,

$$q_i = M_n(x_i^+) - M_n(x_i^-), \qquad i = 1, \dots, n+1;$$

con notación de deltas de Dirac lo podemos escribir como

$$\mu_n = \sum_{i=1}^{n+1} q_i \delta_{x_i}.$$

De esta forma se cumple

$$s_k = \int_{-\infty}^{\infty} x^k \, d\mu_n(x), \qquad k = 0, 1, \dots, 2n-1.$$

Así pues, para cada n finito hemos encontrado una medida $d\mu_n$ cuyos momentos s_k coinciden con el valor de $\mathcal{S}\{x^k\}$ para $k = 0, 1, \dots, 2n-1$. Como la medida y el funcional coinciden sobre las funciones x^k, también coinciden sobre los polinomios. Entonces, sin más que tomar los polinomios ortogonales $\{P_k(x)\}_{k=0}^n$ que salían en la definición del funcional, dichos polinomios son también ortogonales respecto a la medida μ_n.

En segundo lugar, vamos a intentar construir una medida μ a partir de la sucesión de medidas anteriores $\{\mu_n\}_{n=1}^\infty$. En realidad, vamos a construir una función de distribución $M(x)$ a partir de la sucesión de funciones de distribución anteriores $\{M_n(x)\}_{n=0}^\infty$. Para ello vamos a aplicar el teorema de selección de Helly.

Podemos aplicar dicho teorema, ya que, para cualquier n, la función $M_n(x) = \int_{-\infty}^x d\mu_n$ es no negativa y monótona no decreciente, y su variación total es

$$\int_{-\infty}^{\infty} d\mu_n(x) = s_0;$$

esto implica que $0 \le M_n(x) \le s_0$ para todo $x \in \mathbb{R}$. Así, existe una función $M(x)$ no decreciente y una sucesión $\{M_{n_i}(x)\}_{i=1}^\infty$ tales que, en todos los puntos de continuidad de $M(x)$, se cumple la igualdad

$$M(x) = \lim_{i \to \infty} M_{n_i}(x).$$

Vamos ahora a probar que, en efecto, la medida μ correspondiente a la distribución $M(x)$ (y que, para cualquier conjunto medible X, es $\mu(X) = \int_X dM$) es la medida que ortogonaliza los polinomios tal como pide (4.1). Para ello, simplemente hay que ver que

$$s_k = \int_{-\infty}^{\infty} x^k \, d\mu(x), \qquad k = 0, 1, \dots. \tag{4.3}$$

Comprobemos (4.3). Asumimos que $-A \leq -1$ y $B \geq 1$ son dos puntos de continuidad de $M(x)$. Entonces, por el teorema de convergencia de Helly,

$$\int_{-A}^{B} x^k \, d\mu(x) = \lim_{i \to \infty} \int_{-A}^{B} x^k \, d\mu_{n_i}(x). \tag{4.4}$$

Vamos ahora a intentar acotar superiormente la siguiente expresión:

$$\left| \int_{-A}^{B} x^k \, d\mu(x) - s_k \right|. \tag{4.5}$$

Para k fijo y n_i suficientemente grande los momentos s_k satisfacen

$$s_k = \int_{-\infty}^{\infty} x^k \, d\mu_{n_i}(x),$$

por lo que (4.5) es igual a

$$\left| \int_{-A}^{B} x^k \, d\mu(x) - \int_{-\infty}^{\infty} x^k \, d\mu_{n_i}(x) \right|. \tag{4.6}$$

Además, descompongamos

$$\int_{-\infty}^{\infty} x^k \, d\mu_{n_i}(x) = \int_{-A}^{B} x^k \, d\mu_{n_i}(x) + \int_{-\infty}^{-A} x^k \, d\mu_{n_i}(x) + \int_{B}^{\infty} x^k \, d\mu_{n_i}(x),$$

con lo cual (4.6) es igual a

$$\left| \left(\int_{-A}^{B} x^k \, d\mu(x) - \int_{-A}^{B} x^k \, d\mu_{n_i}(x) \right) - \left(\int_{-\infty}^{-A} x^k \, d\mu_{n_i}(x) + \int_{B}^{\infty} x^k \, d\mu_{n_i}(x) \right) \right|$$

Aplicando la desigualdad triangular ($|a - b| \leq |a| + |b|$) nos queda

$$\begin{aligned} &\left| \int_{-A}^{B} x^k \, d\mu(x) - s_k \right| \\ &\quad \leq \left| \int_{-A}^{B} x^k \, d\mu(x) - \int_{-A}^{B} x^k \, d\mu_{n_i}(x) \right| + \left| \int_{-\infty}^{-A} x^k \, d\mu_{n_i}(x) + \int_{B}^{\infty} x^k \, d\mu_{n_i}(x) \right|. \end{aligned} \tag{4.7}$$

Si $2r$ es un número par mayor que k entonces, para $n_i > 2r$,

$$\begin{aligned} &\left| \int_{-\infty}^{-A} x^k \, d\mu_{n_i}(x) + \int_{B}^{\infty} x^k \, d\mu_{n_i}(x) \right| = \left| \int_{-\infty}^{-A} \frac{x^{2r}}{x^{2r-k}} \, d\mu_{n_i}(x) + \int_{B}^{\infty} \frac{x^{2r}}{x^{2r-k}} \, d\mu_{n_i}(x) \right| \\ &\leq \frac{1}{A^{2r-k}} \int_{-\infty}^{-A} x^{2r} \, d\mu_{n_i}(x) + \frac{1}{B^{2r-k}} \int_{B}^{\infty} x^{2r} \, d\mu_{n_i}(x) \leq \left(\frac{1}{A^{2r-k}} + \frac{1}{B^{2r-k}} \right) s_{2r}. \end{aligned} \tag{4.8}$$

Combinando (4.7) con (4.8),

$$\left| \int_{-A}^{B} x^k \, d\mu(x) - s_k \right| \leq \left| \int_{-A}^{B} x^k \, d\mu(x) - \int_{-A}^{B} x^k \, d\mu_{n_i}(x) \right| + \left(\frac{1}{A^{2r-k}} + \frac{1}{B^{2r-k}} \right) s_{2r}.$$

Tomando límites en i y usando (4.4), queda

$$\left| \int_{-A}^{B} x^k \, d\mu(x) - s_k \right| \leq \left(\frac{1}{A^{2r-k}} + \frac{1}{B^{2r-k}} \right) s_{2r}.$$

La expresión de la derecha es independiente de k y tiende a cero cuando A y B tienden a infinito. Por lo tanto,

$$\int_{-\infty}^{\infty} x^k \, d\mu(x) = \lim_{A,B\to\infty} \int_{-A}^{B} x^k \, d\mu(x) = s_k,$$

que prueba (4.3).

Además, teniendo en cuenta que, por el teorema 14, la sucesión de momentos del funcional es definida positiva, la sucesión momentos asociados a la medida que hemos construido también es definida positiva; esto equivale, según el teorema 3, a que los determinantes de Hankel sean positivos. Así, aplicando el teorema 8 (en el capítulo 2), confirmamos que la medida que hemos construido es definida positiva y, por tanto, tiene asociada una SPO infinita (lo cual garantiza (4.2)).

De hecho, también se puede probar que la medida μ que hemos construido (a partir de la función de distribución $M(x)$) tiene infinitos puntos de crecimiento efectivo (lo que también garantizaría (4.2)). En efecto, si μ tuviera un número finito N de puntos de crecimiento, sean $x_1, x_2, \ldots, x_N$ dichos puntos, podríamos construir el polinomio

$$\pi(x) = (x - x_1)^2 (x - x_2)^2 \cdots (x - x_N)^2.$$

Aplicando el funcional a dicho polinomio (o de forma equivalente la integral) obtendríamos

$$\mathcal{S}\{\pi(x)\} = \int_{-\infty}^{\infty} (x - x_1)^2 (x - x_2)^2 \cdots (x - x_N)^2 \, d\mu(x) = 0,$$

que es absurdo porque, como se ve en la demostración del teorema 14, el resultado de aplicar el funcional a un polinomio (no nulo) al cuadrado siempre es mayor que cero. Por lo tanto, la medida μ tiene infinitos puntos de crecimiento efectivo. $\square$

Capítulo 5

Conclusiones

La realización de este Trabajo Fin de Grado ha supuesto un gran reto debido a la necesidad de desarrollar, de forma fundamentada y ordenada, una teoría matemática medianamente larga. Habitualmente en matemáticas, los estudiantes estamos acostumbrados a estudiar la teoría y luego aplicarla. Así, el proceso de construir la teoría y hacerlo adecuadamente no se aprecia y se da por supuesto.

Tras esta reflexión, quiero destacar que en el contenido de esta memoria se han desarrollado dos teorías: una introducción a los polinomios ortogonales (respecto a una medida) y otra a las sucesiones en recurrencias. Además, ambas teorías se han interrelacionado mediante la demostración del teorema de Favard. Esa fusión, junto con los desarrollos ordenados de cada una de ellas, ha sido una tarea de elevada complejidad.

Así, en la parte de polinomios ortogonales, hemos visto su definición formal y los ejemplos más representativos (primer capítulo), seguido de un gran número de resultados y conceptos relacionados, donde destaca la relación de recurrencia (con la que cerramos el segundo capítulo).

Sobre las sucesiones en recurrencias de polinomios, en el tercer capítulo hemos desarrollado un estudio básico de sus propiedades (sin asumir que dichos polinomios sean ortogonales), obteniendo todas las herramientas necesarias para abordar —ya en el último capítulo— el resultado que permitirá ligar las sucesiones en recurrencias con los polinomios ortogonales.

En concreto, en el último capítulo —previa demostración de los teoremas de Helly— damos una del Teorema de Favard, que prueba la correspondencia entre ortogonalidad y recurrencia, creando así un círculo perfecto —¿o una banda de Möbius?— entre ambas teorías.

Realmente, este ha sido un trabajo apasionante, motivador, desafiante e incluso desesperante en algunos casos, pero el resultado ha merecido la pena y me ha ayudado a mi desarrollo personal como matemático.

Por último, quiero comentar que el tema de este trabajo podría ser continuado en futuras investigaciones. No en vano, la Universidad de La Rioja es puntera en el ámbito de los polinomios ortogonales y la teoría de aproximación, donde se encuentra inmerso este trabajo.

Bibliografía

[1] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Oliver & Boyd, 1965.

[2] R. Álvarez-Nodarse, *Polinomios hipergemétricos y q-polinomios*, Monografías del Seminario Matemático "García de Galdeano", número 26, Prensas Universitarias de Zaragoza, 2003.

[3] T. S. Chihara, *An introduction to orthogonal polynomials*, Gordon & Breach, New York, 1978. Reimpresión: Dover, 2011.

[4] J. Favard, Sur les polynômes de Tchebicheff, *C. R. Acad. Sci. Paris* **200** (1935), 2052–2053.

[5] G. Freud, *Orthogonal polynomials*, Pergamon Press, New York, 1971.

[6] R. Koekoek, P. A. Lesky y R. F. Swarttouw, *Hypergeometric orthogonal oolynomials and their q-Analogues*, Springer, Berlin, 2010,

[7] T. H. Koornwinder, Orthogonal polynomials with weight function $(1-x)^\alpha(1+x)^\beta + M\delta(x+1) + N\delta(x-1)$, *Canad. Math. Bull.* **27** (1984), 205–214.

[8] T. H. Koornwinder, R. Wong, R. Koekoek y R. F. Swarttouw, Orthogonal Polynomials, *NIST handbook of mathematical functions* (editado por F. W. F. Olver, D. W. Lozier, R. F. Boisvert y C. W. Clark), 435–484, National Institute of Standards and Technology, Washington, DC, y Cambridge University Press, Cambridge, 2010. Disponible en `https://dlmf.nist.gov/18`.

[9] J. A. Shohat y J. D. Tamarkin, *The problem of moments*, Mathematical Surveys and Monographs, vol. 1, Amer. Math. Soc., Providence, R. I., 1943.

[10] G. Szegő, *Orthogonal polynomials*, 3.ª ed., American Mathematical Society Colloquium Publications, vol. 23, Amer. Math. Soc., Providence, R. I., 1967.