Sección de Matemáticas
Universidad de La Laguna

Pedro Manuel de San Gil González

Espacios de Hilbert con núcleo reproductor

Reproducing kernel Hilbert spaces

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, julio de 2023

Dirigido por
María Isabel Marrero Rodríguez

María Isabel Marrero Rodríguez
Departamento de Análisis
Matemático
Universidad de La Laguna
Apartado 456
38200 La Laguna, Tenerife

Agradecimientos

Quiero agradecer a mi tutora María Isabel Marrero Rodríguez por su ayuda, paciencia y, sobre todo, por los ánimos. También a mis padres y mi hermana, por su apoyo estos años. A mis amigos por estar ahí, en especial a Pablo, por enseñarme a usar LaTeX.

Pedro Manuel de San Gil González
La Laguna, julio de 2023

Resumen · Abstract

Resumen

Un espacio de Hilbert con núcleo reproductor (EHNR) es un espacio de Hilbert de funciones sobre el que los funcionales evaluación son continuos. Vía el teorema de representación de Fréchet-Riesz, todo EHNR tiene asociado un núcleo que permite reproducir cualquier función del espacio, en el sentido de que el valor de la función en un punto de su dominio se obtiene multiplicándola escalarmente por otra función del espacio, determinada por el núcleo. El teorema de Moore-Aronszajn establece una correspondencia biyectiva entre los núcleos reproductores de espacios de Hilbert y las denominadas funciones núcleo, o núcleos (semi)definidos positivos, y se puede decir que la teoría de EHNRs consiste en el estudio de las consecuencias que se derivan de esta biyección.

Los EHNRs encuentran aplicación en una amplísima variedad de áreas. En este trabajo se ilustra su relevancia en la teoría de la aproximación, probando que toda función núcleo tiene asociado un EHNR de modo que la interpolación equivale a la interpolación de norma mínima en ese espacio. Además, se aborda una variante de este esquema interpolatorio que, a diferencia del original, permite reproducir polinomios de un grado determinado. En este nuevo escenario, el papel de las funciones núcleo y de los EHNRs es desempeñado, respectivamente, por los núcleos condicionalmente definidos positivos y por los llamados espacios de Pontryagin con núcleo reproductor.

Palabras clave: *Aproximación – Interpolación – Datos dispersos – Espacio de Pontryagin – Función núcleo – Núcleo reproductor – Núcleo definido positivo – Núcleo semidefinido positivo – Núcleo condicionalmente definido positivo – Teorema de Moore-Aronszajn.*

Abstract

A reproducing kernel Hilbert space (RKHS, for short) is a Hilbert space of functions on which all point functionals are continuous. Via the Fréchet-Riesz representation theorem, every RKHS is associated with a kernel that allows to evaluate any function in the space by performing its inner product against another function in the space, determined by the kernel. The Moore-Aronszajn theorem establishes a bijective correspondence between the reproducing kernels of Hilbert spaces and the so-called kernel functions, or positive (semi)definite kernels, and it can be said that the theory of RKHSs amounts to studying the consequences derived from such a bijection.

RKHSs find applications in a wide variety of fields. Here, its relevance in approximation theory is illustrated by proving that for every kernel function there exists an RKHS so that interpolation equals minimal norm interpolation within this space. Furthermore, a variant of such an interpolation scheme is addressed that, unlike the original, allows to reproduce polynomials of a given degree. In this new setting, the role of the kernel functions and of the RKHSs are respectively played by the so-called conditionally positive definite kernels and reproducing kernel Pontryagin spaces.

Keywords: *Approximation – Interpolation – Scattered data – Pontryagin space – Kernel function – Reproducing kernel – Positive definite kernel – Positive semidefinite kernel – Conditionally positive definite kernel – Moore-Aronszajn theorem.*

Contenido

Introducción

En análisis funcional, un espacio de Hilbert con núcleo reproductor (EHNR) es un espacio de Hilbert $\mathcal{H}$ de funciones definidas sobre un conjunto X tal que los funcionales evaluación en puntos de X son continuos sobre $\mathcal{H}$. Así pues, si dos funciones $f, g \in \mathcal{H}$ están próximas en norma, también están próximas puntualmente (el recíproco no es necesariamente cierto).

Por ejemplo, los espacios L^2 no son propiamente espacios de Hilbert de funciones sino, como es bien conocido, de clases de equivalencia de funciones, lo que impide que sean EHNRs. Sin embargo, existen espacios de Hilbert cuya norma es de tipo L^2 que sí son EHNRs. De hecho, importantes espacios de funciones analíticas (concretamente, los de Hardy y Bergman) lo son.

Todo EHNR $\mathcal{H}$ tiene asociado un núcleo que permite reproducir cualquier función $f \in \mathcal{H}$, en el sentido de que, para cada $x \in X$, el valor $f(x)$ se obtiene efectuando el producto interior de f contra otra función determinada por el núcleo. En virtud del teorema de representación de Fréchet-Riesz, la existencia de este núcleo reproductor está garantizada si, y solo si, todo funcional evaluación es continuo.

Los núcleos reproductores fueron utilizados por primera vez a principios del siglo XX por S. Zaremba. En su trabajo de 1907 [21] sobre problemas de valores de frontera para funciones armónicas y biarmónicas, Zaremba enunció la propiedad de reproducción en espacios de Hilbert apropiados e introdujo los núcleos correspondientes a estas clases de funciones, aunque sin distinguirlos con una denominación específica ni abundar en una teoría al respecto. Paralelamente, J. Mercer [11] examinó las funciones que satisfacen la propiedad de reproducción en el marco de la teoría de ecuaciones integrales desarrollada por Hilbert. Mercer denominó núcleos definidos positivos a estas funciones y mostró sus buenas propiedades respecto a los otros núcleos continuos que comparecen en dicha teoría. La idea no reaparecería hasta casi veinte años después, en las disertaciones de G. Szegö (1921), S. Bergman (1922) y S. Bochner (1922) (cf. [4]). Bergman definió núcleos reproductores en una y varias variables para la clase de las funciones armónicas y la clase de las funciones analíticas, y los denominó funciones núcleo. En 1935, E.H. Moore [12] examinó los núcleos definidos positivos, que asoció con las matrices hermitianas positivas, y descubrió el teorema fundamental que vincula estos núcleos con los EHNRs. Alrededor de 1950, la teoría de núcleos reproductores fue sistematizada por N. Aronszajn [3] y el propio Bergman [5]. En 1964, L. Schwartz [19] introdujo la noción de subespacio de Hilbert de un espacio vectorial topológico y probó la correspondencia entre dichos subespacios y los núcleos reproductores, generalizando aún más la teoría.

Durante los últimos 50 años hemos asistido a un espectacular incremento de las aplicaciones de los EHNRs en una amplia variedad de campos, incluyendo el análisis complejo, el análisis armónico, la teoría de la aproximación, los operadores integrales, los procesos estocásticos o la mecánica cuántica. Son particularmente importantes en la teoría del aprendizaje estadístico [15] debido al célebre teorema del representante, un resultado muy útil en la práctica porque simplifica el problema de minimizar el riesgo empírico[1] trasladándolo de un problema en dimensión infinita a otro en dimensión finita.

La presente memoria está estructurada en cuatro capítulos. Para el desarrollo de los tres primeros hemos seguido, fundamentalmente, las referencias [13, 14], mientras que el cuarto se basa en [6].

En el Capítulo 1 se formaliza la definición de espacio de Hilbert con núcleo reproductor y se presentan las propiedades más básicas de estos espacios, junto con algunos ejemplos que pretenden ilustrar mínimamente la diversidad de las áreas en que comparecen.

En el Capítulo 2 se establecen los fundamentos de la teoría correspondiente. En particular, se prueba el importante teorema de Moore-Aronszajn, que caracteriza aquellas funciones que son núcleos de algún EHNR. Como quedó dicho más arriba, tales funciones suelen ser denominadas en la literatura funciones núcleo o funciones (semi)definidas positivas.

El Capítulo 3 está dedicado a la resolución de problemas de interpolación y aproximación, que constituye una de las aplicaciones principales de la teoría. Se probará que toda función núcleo tiene asociado un EHNR de modo que la interpolación equivale a la interpolación de norma mínima en ese espacio.

Por último, en el Capítulo 4 se aborda una variante de este esquema interpolatorio que, a diferencia del original, permite reproducir polinomios de un grado determinado. En este nuevo contexto, el papel de las funciones núcleo y de los EHNRs es desempeñado, respectivamente, por los núcleos condicionalmente definidos positivos y por los llamados espacios de Pontryagin con núcleo reproductor.

[1] La minimización del riesgo empírico es un principio de la teoría del aprendizaje estadístico que define una familia de algoritmos de aprendizaje y se utiliza para establecer límites teóricos sobre su rendimiento. La idea central es que no podemos saber exactamente lo bien que funcionará un algoritmo en la práctica («riesgo verdadero») porque desconocemos la distribución real de los datos a los que se aplicará el algoritmo, pero podemos medir su desempeño sobre un conjunto conocido de datos de entrenamiento («riesgo empírico»).

1

Definición, primeras propiedades y ejemplos

Los espacios de Hilbert con núcleo reproductor surgen en áreas muy diversas de la matemática pura y aplicada, entre ellas la teoría de la aproximación, la estadística, la teoría del aprendizaje automático, la teoría de la representación de grupos y varias ramas del análisis complejo. Sin embargo, el estudio de su estructura y fundamentos tan solo requiere un conocimiento previo de la teoría de espacios de Hilbert y el análisis real.

En este capítulo se presenta la definición formal de espacio de Hilbert con núcleo reproductor y sus propiedades más básicas, así como algunos ejemplos que pretenden ilustrar mínimamente la variedad de áreas en las que comparecen estos espacios.

1.1. Definición de espacio de Hilbert con núcleo reproductor

Consideraremos espacios de Hilbert sobre el cuerpo de los números reales, $\mathbb{R}$, o de los complejos, $\mathbb{C}$. Usaremos $\mathbb{K}$ para representar cualquiera de los dos, $\mathbb{R}$ ó $\mathbb{C}$, de modo que cuando queramos enunciar una definición o resultado que es cierto para ambos cuerpos, escribiremos $\mathbb{K}$.

Sea X un conjunto. Denotaremos por $\mathcal{F}(X, \mathbb{K}) = \{f : X \to \mathbb{K}\}$ la clase de todas las funciones definidas sobre X que toman valores en $\mathbb{K}$. Claramente, con las operaciones de suma $(f + g)(x) = f(x) + g(x)$ y producto por escalares $(\lambda \cdot f)(x) = \lambda \cdot f(x)$, el conjunto $\mathcal{F}(X, \mathbb{K})$ es un espacio vectorial sobre el cuerpo $\mathbb{K}$.

Definición 1.1. *Sea X un conjunto arbitrario. Diremos que $\mathcal{H}$ es un espacio de Hilbert con núcleo reproductor (en adelante, EHNR) de X en $\mathbb{K}$, si:*

(i) *$\mathcal{H}$ es un subespacio vectorial de $\mathcal{F}(X, \mathbb{K})$.*
(ii) *$\mathcal{H}$ está dotado de un producto interior $\langle \cdot, \cdot \rangle$ que lo convierte en espacio de Hilbert.*

(iii) Para cada $y \in X$, el funcional evaluación $E_y : \mathcal{H} \to \mathbb{K}$, definido por $E_y(f) = f(y)$, está acotado.

Si $\mathcal{H}$ es un EHNR sobre X entonces, aplicando el teorema de representación de Fréchet-Riesz al funcional evaluación en $y \in X$, obtenemos una única función $k_y \in \mathcal{H}$ tal que para toda $f \in \mathcal{H}$, se tiene

$$E_y(f) = f(y) = \langle f, k_y \rangle.$$

Definición 1.2. *La función $k_y : X \to \mathbb{K}$ se llama núcleo reproductor del punto y.*

Definición 1.3. *La función de dos variables $K : X \times X \to \mathbb{K}$ definida por $K(x, y) = k_y(x)$ se llama núcleo reproductor de $\mathcal{H}$.*

Aplicando nuevamente el teorema de representación de Fréchet-Riesz, resulta:

$$K(x, y) = k_y(x) = E_x(k_y) = \langle k_y, k_x \rangle, \quad x, y \in X. \tag{1.1}$$

De la cadena de igualdades (1.1) podemos deducir sin dificultad las siguientes propiedades:

1. $\|E_y\|^2 = \|k_y\|^2 = \langle k_y, k_y \rangle = K(y, y)$.
2. Si $\mathbb{K} = \mathbb{C}$, $K(x, y) = \langle k_y, k_x \rangle = \overline{\langle k_x, k_y \rangle} = \overline{K(y, x)}$.
3. Si $\mathbb{K} = \mathbb{R}$, $K(x, y) = K(y, x)$.
4. $|K(x, y)| = |\langle k_y, k_x \rangle| \leq \|k_y\| \, \|k_x\| = K(x, x)^{\frac{1}{2}} K(y, y)^{\frac{1}{2}}$.

La primera de ellas es consecuencia del teorema de representación de Fréchet-Riesz; la cuarta, de la desigualdad de Cauchy-Schwarz.

Finalizaremos este primer capítulo dando algunos ejemplos y contraejemplos.

1.2. Ejemplos de EHNR

A continuación veremos algunos ejemplos con distinto nivel de complejidad, procedentes de distintas ramas del análisis, además de un ejemplo de espacio de Hilbert que no es EHNR.

1.2.1. Ejemplos básicos

Ejemplo 1.4 (Espacio vectorial complejo multidimensional). Sea $\mathbb{C}^n$ el espacio vectorial complejo n-dimensional, provisto del producto interior usual: si $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$ son dos elementos de $\mathbb{C}^n$, entonces

$$\langle v, w \rangle = \sum_{i=1}^{n} v_i \overline{w_i}.$$

Es bien conocido que el par $(\mathbb{C}^n, \langle\cdot,\cdot\rangle)$ constituye un espacio de Hilbert. Una n-upla $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$ puede ser contemplada como una función $v : X \to \mathbb{C}$, donde $X = \{1, \ldots, n\}$ y $v(j) = v_j$, $j = 1, \ldots, n$; con esta identificación, $\mathbb{C}^n$ se convierte en el espacio vectorial $\mathcal{F}(X, \mathbb{C})$. Sea $\{e_j\}_{j=1}^n$ la base ortonormal canónica de $\mathbb{C}^n$:

$$e_j(i) = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases}, \quad j = 1, \ldots, n.$$

Para cada $v \in \mathbb{C}^n$ se cumple que $v(j) = v_j = \langle v, e_j\rangle$, $j = 1, \ldots, n$. Por lo tanto, la base canónica de $\mathbb{C}^n$ es, precisamente, el conjunto de funciones núcleo para los funcionales evaluación, cuando consideramos $\mathbb{C}^n$ como espacio de funciones.

En virtud de (1.1), el núcleo reproductor de $\mathbb{C}^n$ viene dado por

$$K(i, j) = \langle e_j, e_i\rangle = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases}, \quad i, j = 1, \ldots, n,$$

que puede ser asimilado con la matriz identidad. Se concluye que $\mathbb{C}^n$ es un EHNR. □

Ejemplo 1.5 (Espacios ℓ^2). Dado un conjunto discreto X cualquiera, consideramos el espacio vectorial

$$\ell^2(X) = \left\{ f : X \to \mathbb{K} \mid \sum_{x \in X} |f(x)|^2 < +\infty \right\}$$

y definimos el producto interior

$$\langle f, g\rangle = \sum_{x \in X} f(x)\overline{g(x)}, \quad f, g \in \ell^2(X).$$

El par $(\ell^2(X), \langle\cdot,\cdot\rangle)$ es un espacio de Hilbert de funciones sobre X. Fijado un punto $y \in X$, denotamos por $e_y \in \ell^2(X)$ la función

$$e_y(x) = \begin{cases} 1, & \text{si } x = y \\ 0, & \text{si } x \neq y \end{cases}, \quad x \in X.$$

Es fácil ver que $\{e_y\}_{y \in X}$ constituye una base ortonormal de $\ell^2(X)$, y además $E_y(f) = f(y) = \langle f, e_y\rangle$. Por tanto, también en este caso, las funciones $\{e_y\}_{y \in X}$ son los núcleos reproductores para los puntos y. De forma análoga al Ejemplo 1.4, encontramos que el núcleo reproductor de $\ell^2(X)$ viene dado por:

$$K(x, y) = \langle e_y, e_x\rangle = \begin{cases} 1, & \text{si } x = y \\ 0, & \text{si } x \neq y \end{cases}, \quad x, y \in X.$$

Luego, $\ell^2(X)$ es un EHNR. □

A continuación, presentaremos un «no-ejemplo». Más concretamente, veremos que el espacio de Hilbert $L^2[-1,1]$ no es un EHNR.

Ejemplo 1.6 (Espacio $L^2[-1,1]$). Sea $L^2[-1,1]$ el espacio vectorial de las funciones medibles Lebesgue en el intervalo real $[-1,1]$ con cuadrado absolutamente integrable:

$$L^2[-1,1] = \left\{ f : [-1,1] \to \mathbb{K} \;\middle|\; f \text{ medible Lebesgue}, \int_{-1}^{1} |f(x)|^2 dx < +\infty \right\}.$$

Claramente, $L^2[-1,1]$ es un subespacio vectorial del espacio vectorial $\mathcal{F}([-1,1],\mathbb{K})$ formado por las funciones definidas sobre el intervalo cerrado $[-1,1]$, con valores en $\mathbb{K}$. Además, $L^2[-1,1]$ es un espacio de Hilbert con el producto interior

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)}dx, \quad f, g \in L^2[-1,1],$$

que induce la norma

$$\|f\|_2 = \left(\int_{-1}^{1} |f(x)|^2 dx \right)^{\frac{1}{2}}, \quad f \in L^2[-1,1].$$

Recordemos que, en sentido estricto, $L^2[-1,1]$ no es un espacio de funciones, sino un espacio de clases de equivalencia de funciones, en el que identificamos aquellas que son iguales c.t.p.. Este espacio coincide con la compleción del espacio $(C[-1,1], \|\cdot\|_2)$. Así, un primer escollo para que $L^2[-1,1]$ sea un EHNR es que las funciones que lo constituyen no están definidas unívocamente en todo $x \in [-1,1]$, por lo que no es posible definir funcionales evaluación para todos los puntos del intervalo. Estos funcionales sí están definidos en el subespacio $C[-1,1]$, pero no están acotados; concretamente, veamos que el funcional evaluación E_0 no lo está. A tal fin, definimos la sucesión $\{f_n\}_{n=1}^{\infty} \subset C[-1,1]$ por

$$f_n^2(x) = \begin{cases} 0, & \text{si } -1 \leq x < -n^{-1} \\ n^2 x + n, & \text{si } -n^{-1} \leq x < 0 \\ -n^2 x + n, & \text{si } 0 \leq x < n^{-1} \\ 0, & \text{si } n^{-1} \leq x \leq 1 \end{cases}, \quad n \in \mathbb{N}.$$

Se tiene que $E_0(f_n) = f_n(0) = \sqrt{n}$ y $\|f_n\|_2 = 1$ para todo $n \in \mathbb{N}$. Si E_0 fuese continuo, existiría $C > 0$ tal que, en particular,

$$\sqrt{n} = |E_0(f_n)| \leq C\|f_n\|_2 = C, \quad n \in \mathbb{N},$$

una contradicción. Como el funcional evaluación E_0 no está acotado sobre $C[-1,1]$, que es denso en $L^2[-1,1]$, no hay forma de extenderlo con continuidad a todo $L^2[-1,1]$, por lo que este espacio no será, entonces, un EHNR.

Utilizando la misma idea y siguiendo la misma secuencia de pasos se puede probar que el espacio de Hilbert $L^2[a,b]$ $(a, b \in \mathbb{R},\ a < b)$ no es EHNR. Resulta interesante observar que $L^2[-1,1]$ y $\ell^2(\mathbb{Z})$ son isomorfos como espacios de Hilbert —ya que ambos tienen bases ortonormales numerables—, si bien $\ell^2(\mathbb{Z})$ es EHNR (Ejemplo 1.5), mientras que $L^2[-1,1]$ no lo es. □

1.2.2. Ejemplos de matemática aplicada

Comenzamos con un ejemplo de EHNR que aparece en la teoría de las ecuaciones diferenciales.

Ejemplo 1.7 (Espacio de Sobolev en $[0,1]$). Recordemos que una función $f : [0,1] \to \mathbb{R}$ es absolutamente continua si dado $\varepsilon > 0$, existe $\delta > 0$ de modo que para cualquier colección finita $(x_1, y_1), \ldots, (x_n, y_n)$ de subintervalos no solapados de $[0,1]$ con $\sum_{i=1}^n |y_i - x_i| < \delta$, se tiene que $\sum_{i=1}^n |f(y_i) - f(x_i)| < \varepsilon$.

Un importante teorema de teoría de la medida [16, Theorem 7.18] asegura que f es absolutamente continua si, y solo si, existe $f'(x)$ en casi todo x, f' es integrable y, excepto por una constante aditiva, f es igual a la integral de f'. Dicho de otra manera, las funciones absolutamente continuas son aquellas a las que es posible aplicar el primer teorema fundamental del cálculo.

Sea $\mathcal{H}$ el conjunto formado por las funciones absolutamente continuas $f : [0,1] \to \mathbb{R}$ que satisfacen las condiciones de frontera $f(0) = f(1) = 0$ y cuya derivada f' es de cuadrado integrable:

$$\mathcal{H} = \Big\{ f : [0,1] \to \mathbb{R} \;\Big|\; f \text{ absolutamente continua}, \\ f(0) = f(1) = 0,\ \int_0^1 |f'(x)|^2 dx < +\infty \Big\}.$$

Provisto de las operaciones de suma y producto por escalares usuales, $\mathcal{H}$ es un subespacio vectorial de $\mathcal{F}([0,1], \mathbb{R})$. A continuación, dotaremos a $\mathcal{H}$ de estructura de espacio de Hilbert. Definimos la forma sequilineal no negativa

$$\langle f, g \rangle = \int_0^1 f'(t) g'(t) dt, \quad f, g \in \mathcal{H}.$$

Nótese que la integral del segundo miembro es, en realidad, el producto interior definido en $L^2[0,1]$. Esta forma sesquilineal induce una seminorma

$$\|f\| = \|f'\|_2, \quad f \in \mathcal{H}.$$

Para que $\langle \cdot, \cdot \rangle$ sea un producto interior y $\|\cdot\|$ una norma, sólo necesitamos probar que $\langle f, f \rangle = 0$ implica $f = 0$. A tal fin, sean $x \in [0,1]$, arbitrario, y $f \in \mathcal{H}$. Como f es absolutamente continua,

$$f(x) = \int_0^x f'(t)dt = \int_0^1 f'(t)\chi_{[0,x]}(t)dt.$$

La desigualdad de Cauchy-Schwarz permite escribir:

$$|f(x)| \leq \left(\int_0^1 f'(t)^2 dt\right)^{\frac{1}{2}} \left(\int_0^1 \chi_{[0,x]}(t)dt\right)^{\frac{1}{2}} = \|f'\|_2\sqrt{x} = \|f\|\sqrt{x}. \tag{1.2}$$

Si $\langle f, f\rangle = 0$, entonces $\|f'\|_2 = 0$ y, por (1.2), $|f(x)| = 0$; la arbitrariedad de x muestra que $f = 0$. Recíprocamente, si $f = 0$, es claro que $\langle f, f\rangle = 0$. Se concluye que $\langle \cdot, \cdot\rangle$ es un producto interior. Además, de nuevo por (1.2),

$$|E_x(f)| = |f(x)| \leq \|f\|\sqrt{x},$$

es decir, el funcional evaluación en x está acotado sobre $\mathcal{H}$, con $\|E_x\| \leq \sqrt{x}$.

Para ver que $\mathcal{H}$ es un espacio de Hilbert, resta probar que es completo con la norma inducida por este producto interior. Sea $\{f_n\}_{n=1}^{\infty}$ una sucesión de Cauchy en $\mathcal{H}$. Entonces $\{f_n'\}_{n=1}^{\infty}$ es de Cauchy en $L^2[0,1]$ y, por lo tanto, existe $g \in L^2[0,1]$ tal que $\lim\limits_{n\to\infty} \|f_n' - g\|_2 = 0$. La desigualdad (1.2) obliga a que $\{f_n\}_{n=1}^{\infty}$ sea puntualmente de Cauchy, lo que define una función $f(x) = \lim\limits_{n\to\infty} f_n(x)$, $x \in [0,1]$. Ahora,

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \int_0^x f_n'(t)dt = \int_0^x g(t)dt.$$

Así pues, $f' = g$ c.t.p., y f es absolutamente continua. Nótese que, aunque g sea una clase de equivalencia de funciones, $\int_0^x g(t)dt$ no depende del representante elegido en la clase; consecuentemente, $f' \in L^2[0,1]$. Por último,

$$f(0) = \lim_{n\to\infty} f_n(0) = 0 = \lim_{n\to\infty} f(1) = f(1),$$

y con esto $f \in \mathcal{H}$.

Hemos visto que $\mathcal{H}$ es un EHNR sobre $[0,1]$; ahora queremos encontrar las funciones núcleo. Sabemos que $f \in \mathcal{H}$ se puede escribir como

$$f(x) = \int_0^1 f'(t)\chi_{[0,x]}(t)dt.$$

Por tanto, si pudiéramos resolver el siguiente problema de valores de frontera:

$$g'(t) = \chi_{[0,x]}(t), \quad g(0) = g(1) = 0,$$

tendríamos que $g \in \mathcal{H}$ con $f(x) = \langle f, g\rangle$, es decir, $g = k_x$. La solución de este problema es

$$g(t) = \int_0^t \chi_{[0,x]}(s)ds = \begin{cases} t, & \text{si } 0 \leq t < x \\ x, & \text{si } x \leq t < 1 \end{cases}, \quad g(1) = 0,$$

que no es continua; sin embargo, la función $k_x(t)$ existe y es (absolutamente) continua, porque $\mathcal{H}$ es EHNR. Para hallarla, plantearemos formalmente un problema de valores de frontera distinto y *a posteriori* demostraremos que la solución de este problema pertenece a $\mathcal{H}$ y es el núcleo reproductor buscado.

Aplicando integración por partes, obtenemos:

$$\begin{aligned} f(x) = \langle f, k_x \rangle &= \int_0^1 f'(t)k_x'(t)dt \\ &= f(t)k_x'(t)\big|_0^1 - \int_0^1 f(t)k_x''(t)dt \\ &= -\int_0^1 f(t)k_x''(t)dt. \end{aligned}$$

Si δ_x denota la función delta de Dirac formal concentrada en x, entonces $f(x) = \int_0^1 f(t)\delta_x(t)dt$. Esto sugiere plantear el problema de valores de frontera

$$-k_x''(t) = \delta_x(t), \quad k_x(0) = k_x(1) = 0.$$

La solución de este problema es la llamada función de Green para la ecuación diferencial. Resolviéndolo formalmente, integrando dos veces e imponiendo las condiciones de frontera, encontramos que

$$K(t,x) = k_x(t) = \begin{cases} (1-x)t, & \text{si } 0 \leq t < x \\ (1-t)x, & \text{si } x \leq t \leq 1 \end{cases}.$$

Una vez obtenida esta solución formal, es fácil ver que

$$k_x'(t) = \begin{cases} 1-x, & \text{si } 0 \leq t < x \\ -x, & \text{si } x < t \leq 1 \end{cases}.$$

Así pues, k_x es derivable excepto en x —en casi todo punto—, es igual a la integral de su derivada —luego, es absolutamente continua—, $\int_0^1 |k_x(t)|^2 dt < +\infty$, y $k_x(0) = k_x(1) = 0$; por tanto, $k_x \in \mathcal{H}$. Finalmente, si $f \in \mathcal{H}$,

$$\begin{aligned} \langle f, k_x \rangle = \int_0^1 f'(t)k_x'(t)dt &= \int_0^x f'(t)(1-x)dt + \int_x^1 f'(t)(-x)dt \\ &= [f(x) - f(0)](1-x) - x[f(1) - f(x)] = f(x). \end{aligned}$$

Se concluye que $k_x(t) \in \mathcal{H}$ es el núcleo reproductor para el punto x.

Nótese que

$$\|E_x\|^2 = \|k_x\|^2 = K(x,x) = x(1-x),$$

lo cual mejora considerablemente la estimación inicial de $\|E_x\|$. La posibilidad de obtener el valor exacto de $\|E_x\|$ a partir de la existencia de un EHNR es un primer indicativo de la utilidad de esta teoría. □

A continuación, veremos un ejemplo aplicable al ámbito del procesamiento de señales.

Ejemplo 1.8 (Espacio de Paley-Wiener). Sea $f : \mathbb{R} \to \mathbb{R}$ una función, tal que $f(t)$ representa la amplitud de una onda sonora en el instante t. La transformada de Fourier de f,

$$\widehat{f}(x) = \int_{-\infty}^{\infty} f(t)e^{-2\pi ixt}dt,$$

es un intento de descomponer f en «frecuencias», donde el número $\widehat{f}(x)$ representa la magnitud de la contribución de $e^{2\pi ixt}$. La integral anterior no está definida para toda función f, sino que se requieren hipótesis adicionales, adecuadas a la situación de interés. Bajo el conjunto de hipótesis oportunas, se obtienen los llamados espacios de Paley-Wiener.

Los espacios de Paley-Wiener desempeñan un papel importante en el análisis de Fourier y la teoría del muestreo, especialmente cuando queremos estudiar las transformadas de Fourier de funciones con soporte compacto tanto en el dominio temporal como en el de la frecuencia.

Consideremos las funciones que no se anulan en un intervalo temporal finito $[-A, A]$, para algún $A > 0$. Supongamos también que $\int_{-A}^{A} |f(t)|^2 dt < +\infty$. Entonces existe la transformada de Fourier de f, y definimos el espacio de Paley-Wiener, PW_A, como el conjunto de las transformadas de Fourier de funciones de cuadrado integrable, soportadas en $[-A, A]$; es decir,

$$PW_A = \left\{ \widehat{f} \mid f \in L^2[-A, A] \right\}.$$

Como ha quedado dicho en anteriores ejemplos, los elementos de $L^2[-A, A]$ son clases de equivalencia de funciones. Pese a ello, el número $\widehat{f}(x)$ está bien definido y es independiente del representante elegido en la clase de equivalencia. Por lo tanto, aunque $L^2[-A, A]$ no es EHNR, el conjunto PW_A de sus transformadas de Fourier es una clase perfectamente definida de funciones sobre $\mathbb{R}$. Nos disponemos a probar que, provisto de una norma apropiada, PW_A es un EHNR sobre $\mathbb{R}$.

Dada $F \in PW_A$, existe una función $f \in L^2[-A, A]$, tal que

$$F(x) = \int_{-A}^{A} f(t)e^{-2\pi ixt}dt. \tag{1.3}$$

La función f es única salvo igualdad en casi todo punto; es decir, la aplicación lineal $\widehat{\cdot}: L^2[-A,A] \to PW_A$, dada por $\widehat{f} = F$ para $f \in L^2[-A,A]$, es inyectiva. En efecto, es sabido que las funciones $\{e^{2\pi int/A}\}_{n\in\mathbb{Z}}$ constituyen una base ortonormal de $L^2[-A,A]$, y que los coeficientes de Fourier respecto de esta base son

$$\int_{-A}^{A} f(t)e^{-2\pi i\frac{n}{A}t}dt = \widehat{f}\left(\frac{n}{A}\right) = F\left(\frac{n}{A}\right), \quad n \in \mathbb{Z}.$$

Luego, $F = 0$ implica $f = 0$ c.t.p.. Como la aplicación es claramente sobreyectiva, definiendo sobre PW_A la norma

$$\|F\| = \|\widehat{f}\| = \|f\|_2$$

resulta que PW_A es espacio de Hilbert, y $\widehat{\cdot}$, un isomorfismo de espacios de Hilbert. Además, el funcional evaluación es acotado respecto de esta norma: si $x \in \mathbb{R}$ y $F \in PW_A$, entonces

$$|F(x)| = \left|\int_{-A}^{A} f(t)e^{-2\pi ixt}dt\right| \leq \|f\|_2 \, \|e^{2\pi ixt}\|_2 = \sqrt{2A}\, \|F\|.$$

Por lo tanto, PW_A es EHNR sobre $\mathbb{R}$.

Para calcular el núcleo reproductor de PW_A usamos el isomorfismo $\widehat{\cdot}$ entre PW_A y $L^2[-A,A]$. Sean $F = \widehat{f}$ y $G = \widehat{g}$. Entonces,

$$\langle F, G\rangle_{PW_A} = \langle f, g\rangle_{L^2} = \int_{-A}^{A} f(t)\overline{g(t)}dt.$$

Como

$$\langle F, k_y\rangle_{PW_A} = F(y) = \langle f, e^{2\pi iyt}\rangle_{L^2} = \langle F, \widehat{e^{2\pi iyt}}\rangle_{PW_A},$$

obtenemos

$$k_y(x) = \widehat{e^{2\pi iyt}}(x) = \int_{-A}^{A} e^{2\pi i(y-x)t}dt.$$

Resolviendo la integral encontramos que

$$K(x,y) = \begin{cases} \dfrac{\operatorname{sen}(2\pi A(x-y))}{\pi(x-y)}, & \text{si } x \neq y \\ 2A, & \text{si } x = y \end{cases}, \quad x, y \in [-A, A].$$

□

1.2.3. Ejemplos de teoría de funciones

Ejemplo 1.9 (Espacio de Hardy sobre el disco unidad). Este espacio juega un papel clave en la teoría de funciones, la teoría de operadores y la teoría de procesos estocásticos.

Sea $H^2(\mathbb{D})$ el conjunto de las series de potencias complejas formales sobre el disco unidad $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, tales que la sucesión de sus coeficientes es de cuadrado absolutamente sumable:

$$f \sim \sum_{n=0}^{\infty} a_n z^n, \quad \sum_{n=0}^{\infty} |a_n|^2 < +\infty.$$

Veamos primero que cada una de estas series de potencias es absolutamente convergente en el disco unidad, y, en particular, define una función de $\mathbb{D}$ en $\mathbb{C}$. A tal fin, fijemos $z \in \mathbb{D}$. Entonces,

$$\begin{aligned} |f(z)| &= \left| \sum_{n=0}^{\infty} a_n z^n \right| \leq \sum_{n=0}^{\infty} |a_n|\, |z|^n \leq \left(\sum_{n=0}^{\infty} |a_n|^2 \right)^{\frac{1}{2}} \left(\sum_{n=0}^{\infty} |z|^{2n} \right)^{\frac{1}{2}} \\ &\leq C \, \frac{1}{\sqrt{1-|z|^2}} < +\infty, \end{aligned} \tag{1.4}$$

con $C \geq 0$. Ahora queremos ver que si dos series de potencias definen la misma función en el disco unidad entonces son la misma serie, es decir sus coeficientes coinciden. Supongamos que $f, g \in H^2(\mathbb{D})$, $f \sim \sum_{n=0}^{\infty} a_n z^n$, $g \sim \sum_{n=0}^{\infty} b_n z^n$, satisfacen $f(z) = g(z)$ para todo $z \in \mathbb{D}$. Las funciones definidas por series de potencias son infinitamente derivables en su círculo de convergencia, y sus derivadas se pueden calcular por derivación término a término. Encontramos así que $f^{(n)}(0) = g^{(n)}(0)$, o bien $n!a_n = n!b_n$, de donde $a_n = b_n$ para cada $n \in \mathbb{Z}^+$. Por tanto, $H^2(\mathbb{D}) \subset \mathcal{F}(\mathbb{D}, \mathbb{C})$. En particular, con la definición usual de la suma y el producto por escalares, $H^2(\mathbb{D})$ es un espacio vectorial.

Dadas $f, g \in H^2(\mathbb{D})$ como anteriormente, definimos el producto interior

$$\langle f, g \rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}$$

y, a partir del producto interior, la norma

$$\|f\| = \langle f, f \rangle^{\frac{1}{2}} = \left(\sum_{n=0}^{\infty} |a_n|^2 \right)^{\frac{1}{2}} < +\infty.$$

Consideramos ahora la aplicación $L : H^2(\mathbb{D}) \to \ell^2(\mathbb{Z}^+)$, definida por $L(f) = \{a_0, a_1, \ldots\}$. Toda serie de potencias está unívocamente determinada por la sucesión de sus coeficientes $a = \{a_0, a_1, \ldots\}$, y además

$$\|f\| = \left(\sum_{n=0}^{\infty} |a_n|^2 \right)^{\frac{1}{2}} = \langle a, a \rangle_{\ell^2}^{\frac{1}{2}} = \|a\|_{\ell^2}. \tag{1.5}$$

Luego, L es un isomorfismo de espacios vectoriales que preserva el producto interior y la norma. Podemos entonces identificar $H^2(\mathbb{D})$ con el espacio de Hilbert $\ell^2(\mathbb{Z}^+)$.

A pesar de que $\mathbb{Z}^+$ es un conjunto discreto y, consecuentemente (Ejemplo 1.5), $\ell^2(\mathbb{Z}^+)$ es un EHNR, ya vimos en el Ejemplo 1.6 que dos espacios de Hilbert podían ser isomorfos y aun así solo uno de ellos tener núcleo reproductor, por lo que, de momento, tan solo podemos afirmar que $H^2(\mathbb{D})$ cumple (i) y (ii) de la definición de EHNR. Veamos seguidamente que $H^2(\mathbb{D})$ también cumple (iii).

Si $z \in \mathbb{D}$, (1.4) y (1.5) muestran que

$$|E_z(f)| = |f(z)| \leq \|f\| \, \frac{1}{\sqrt{1-|z|^2}}, \qquad f \in H^2(\mathbb{D}).$$

Así pues, E_z está acotado, con

$$\|E_z\| \leq \frac{1}{\sqrt{1-|z|^2}}. \tag{1.6}$$

De la arbitrariedad de $z \in \mathbb{D}$ se concluye que $H^2(\mathbb{D})$ es un EHNR sobre $\mathbb{D}$.

Para calcular su núcleo reproductor, sean $w \in \mathbb{D}$, $k_w \sim \sum_{n=0}^{\infty} b_n z^n \in H^2(\mathbb{D})$ el núcleo reproductor del punto w, y $f \sim \sum_{n=0}^{\infty} a_n z^n \in H^2(\mathbb{D})$. Se verifica que

$$\sum_{n=0}^{\infty} a_n w^n = f(w) = \langle f, k_w \rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}.$$

Particularizando $f \sim z^n \in H^2(\mathbb{D})$, resulta $b_n = \overline{w^n}$, $n \in \mathbb{Z}^+$. Luego, $k_w \sim \sum_{n=0}^{\infty} \overline{w^n} z^n$, y

$$K(z,w) = k_w(z) = \sum_{n=0}^{\infty} \overline{w^n} z^n = \frac{1}{1-\overline{w}z}, \qquad z, w \in \mathbb{D}.$$

El núcleo reproductor del espacio de Hardy se llama núcleo de Szegö del disco unidad. Nótese que

$$\|E_z\| = K(z,z)^{\frac{1}{2}} = \frac{1}{\sqrt{1-|z|^2}},$$

de modo que la estimación (1.6) era la mejor posible. □

Ejemplo 1.10 (Espacios de Bergman en dominios complejos). Estudiaremos a continuación una clase de espacios introducidos por S. Bergman, quien desarrolló la teoría de EHNRs y fue pionero en la deducción de las propiedades de estos espacios a partir de sus núcleos.

Dado un dominio (conjunto abierto y conexo) $G \subset \mathbb{C}$, consideramos el espacio vectorial

$$B^2(G) = \Big\{ f : G \to \mathbb{C} \mid f \text{ analítica en } G, \iint_G |f(x+iy)|^2 dxdy < +\infty \Big\},$$

donde $dxdy$ denota la medida de área. Sobre $B^2(G)$ definimos la forma sesquilineal

$$\langle f, g\rangle = \iint_G f(x+iy)\overline{g(x+iy)}dxdy, \quad f, g \in B^2(G).$$

Si $f = 0$, entonces $\langle f, f\rangle = 0$. Si $f \in B^2(G)$ no es idénticamente nula, entonces, como f es analítica, también es continua, y consecuentemente existe un subconjunto abierto de G donde $|f| > 0$, de modo que $\langle f, f\rangle > 0$. Así pues, $\langle\cdot,\cdot\rangle$ es un producto interior en $B^2(G)$. Por otra parte, dos funciones continuas iguales c.t.p. son iguales en todo punto, así que $B^2(G)$ puede ser contemplado como un subespacio vectorial de $L^2(G)$ formado por clases de equivalencia unipuntuales, es decir, por funciones genuinas. En consecuencia, denotaremos por $\|\cdot\|_2$ la norma de $B^2(G)$.

Para que $B^2(G)$ sea EHNR, el funcional evaluación E_w en cualquier punto $w \in G$ debe estar acotado. Sea $w \in G$, y sea $R > 0$ tal que la bola cerrada $B(w; R)$ de centro w y radio R está contenida en G. Por la fórmula integral de Cauchy, para cualquier $0 \le r \le R$ se verifica

$$f(w) = \frac{1}{2\pi}\int_0^{2\pi} f(w + re^{i\theta})d\theta.$$

Teniendo en cuenta que

$$\frac{2}{R^2}\int_0^R rdr = \frac{2}{R^2}\frac{R^2}{2} = 1,$$

podemos escribir:

$$\begin{aligned} f(w) &= \frac{1}{\pi R^2}\int_0^R r[2\pi f(w)]dr = \frac{1}{\pi R^2}\int_0^R rdr\int_0^{2\pi} f(w+re^{i\theta})d\theta \\ &= \frac{1}{\pi R^2}\iint_{B(w;R)} f(x+iy)dxdy. \end{aligned}$$

Tomando valor absoluto y aplicando la desigualdad de Cauchy-Schwarz:

$$\begin{aligned} |f(w)| &= \frac{1}{\pi R^2}\left|\iint_{B(w;R)} f(x+iy)dxdy\right| \le \frac{1}{\pi R^2}\|f\|_2\left(\iint_{B(w;R)} dxdy\right)^{\frac{1}{2}} \\ &= \frac{1}{\pi R^2}\sqrt{\pi R^2}\|f\|_2 = \frac{1}{R\sqrt{\pi}}\|f\|_2. \end{aligned} \tag{1.7}$$

Por lo tanto, el funcional evaluación E_w está acotado.

Ahora queremos probar que $B^2(G)$ es un espacio de Hilbert. Sea $\{f_n\}_{n=1}^{\infty} \subset B^2(G)$ una sucesión de Cauchy. Dado $w \in G$, sea $R > 0$ como anteriormente, y sea $0 < \delta < d(B(w;R), G^c)$, donde $d(\cdot,\cdot)$ es la distancia entre ambos conjuntos. Entonces, para cualquier $z \in B(w;R)$, la bola cerrada de centro z y radio δ está contenida en G, y, por (1.7),

$$|f_n(z) - f_m(z)| \leq \frac{1}{\delta\sqrt{\pi}}\|f_n - f_m\|_2.$$

Así pues, la sucesión $\{f_n\}_{n=1}^{\infty}$ converge uniformemente en cualquier bola cerrada contenida en G. Denotando por $f(z) = \text{lím}_{n\to\infty} f_n(z)$ el límite puntual de esta sucesión, encontramos que $\{f_n\}_{n=1}^{\infty}$ converge uniformemente a f en cada bola cerrada contenida en G, y el teorema de Montel [9, Theorem 2.9] obliga a que f sea analítica. Como $B^2(G) \subset L^2(G)$ y $L^2(G)$ es completo, existe $h \in L^2(G)$ tal que $\|h - f_n\|_2 \to 0$ cuando $n \to \infty$. Además, podemos elegir una subsucesión $\{f_{n_k}\}_{k=1}^{\infty}$ tal que $h(z) = \text{lím}_{k\to\infty} f_{n_k}(z)$ c.t.p.. Pero esto implica que $h(z) = f(z)$ c.t.p., así que $\|f - f_n\|_2 \to 0$ cuando $n \to \infty$. Luego, $f \in B^2(G)$, y $B^2(G)$ es completo.

Nótese que (1.7) implica que $B^2(\mathbb{C}) = \{0\}$, pues si $G = \mathbb{C}$ podemos tomar R arbitrariamente grande y obtener que $|f(w)| = 0$ para cualquier $f \in B^2(\mathbb{C})$. Dicho de otro modo, la única función entera de cuadrado absolutamente integrable es la función idénticamente nula.

El núcleo reproductor de un espacio de Bergman $B^2(G)$ se denomina núcleo de Bergman de G. A continuación, vamos a calcular el núcleo de Bergman del disco unidad ($G = \mathbb{D}$). Con objeto de simplificar los cálculos, normalizamos el producto interior dividiendo la integral por el área de $\mathbb{D}$:

$$\langle f, g\rangle = \frac{1}{\pi}\iint_{\mathbb{D}} f(x+iy)\overline{g(x+iy)}dxdy.$$

Sea $w \in \mathbb{D}$, y sea

$$k_w(z) = \sum_{j=0}^{\infty} \frac{k_w^{(j)}(0)}{j!} z^j \in B^2(\mathbb{D}).$$

Si $f(z) = z^k$, $k = 0, 1, 2, \ldots$, entonces

$$w^k = \langle z^k, k_w\rangle = \sum_{j=0}^{\infty} \overline{\frac{k_w^{(j)}(0)}{j!}} \langle z^k, z^j\rangle.$$

Ahora bien,

$$\begin{aligned}\langle z^k, z^j\rangle &= \frac{1}{\pi}\iint_{\mathbb{D}} (x+iy)^k\overline{(x+iy)^j}dxdy = \frac{1}{\pi}\int_0^{2\pi} d\theta \int_0^1 r^k e^{ik\theta} r^j e^{-ij\theta} r dr \\ &= \frac{1}{\pi}\int_0^{2\pi} e^{i\theta(k-j)}d\theta \int_0^1 r^{k+j+1}dr = \frac{1}{\pi}\frac{1}{k+j+2}\int_0^{2\pi} e^{i\theta(k-j)}d\theta\end{aligned}$$

$$= \frac{2\pi}{\pi(k+j+2)}\delta_{k,j} = \begin{cases} \dfrac{1}{k+1}, & \text{si } j = k \\ 0, & \text{si } j \neq k \end{cases}.$$

Por tanto,

$$w^k = \frac{\overline{k_w^{(k)}(0)}}{k!}\frac{1}{k+1}$$

implica

$$\overline{k_w^{(k)}(0)} = w^k k!(k+1),$$

y se concluye que

$$k_w(z) = \sum_{j=0}^{\infty} \frac{\overline{w^j} j!(j+1)}{j!} z^j = \sum_{j=0}^{\infty}(j+1)(\overline{w}z)^j = \frac{1}{(1+\overline{w}z)^2}.$$

Una vez calculado el núcleo reproductor para un punto arbitrario $w \in \mathbb{D}$, podemos deducir el núcleo reproductor de $B^2(\mathbb{D})$:

$$K(z,w) = k_w(z) = \frac{1}{(1+\overline{w}z)^2}, \quad z, w \in \mathbb{D}.$$

□

Ejemplo 1.11 (Complejificación de un EHNR de funciones reales). Todo EHNR $\mathcal{H}$ de funciones reales sobre un conjunto X con núcleo reproductor $K(x,y)$ admite una extensión

$$\mathcal{W} = \{f_1 + if_2 \mid f_1, f_2 \in \mathcal{H}\}$$

como EHNR de funciones complejas sobre el mismo conjunto X y con el mismo núcleo reproductor $K(x,y)$.

En efecto: ya que $\mathcal{H}$ es un espacio vectorial, $\mathcal{W}$ también lo es. Definimos el producto interior en $\mathcal{W}$ por:

$$\langle f_1 + if_2, g_1 + ig_2\rangle_{\mathcal{W}} = \langle f_1, g_1\rangle_{\mathcal{H}} + \langle f_2, g_2\rangle_{\mathcal{H}} - i\langle f_1, g_2\rangle_{\mathcal{H}} + i\langle f_2, g_1\rangle_{\mathcal{H}},$$

que se corresponde con la norma

$$\|f_1 + if_2\|_{\mathcal{W}}^2 = \|f_1\|_{\mathcal{H}}^2 + \|f_2\|_{\mathcal{H}}^2.$$

Es fácil ver que, con esta norma, $\mathcal{W}$ es completo y, consecuentemente, un espacio de Hilbert. Solo resta comprobar que $\mathcal{W}$ y $\mathcal{H}$ tienen el mismo núcleo reproductor:

$$f_1(x) + if_2(x) = \langle f_1, k_x\rangle_{\mathcal{H}} + i\langle f_2, k_x\rangle_{\mathcal{H}} = \langle f_1 + if_2, k_x\rangle_{\mathcal{W}}, \quad x \in X.$$

Llamamos a $\mathcal{W}$ la complejificación de $\mathcal{H}$. Puesto que cualquier EHNR de funciones reales se puede complejificar preservando el núcleo reproductor, a partir de ahora se considerarán únicamente EHNR de funciones complejas. □

2

Teoría general

Sea X un conjunto, y sea $\mathcal{H}$ un EHNR sobre X, con núcleo K. Comenzamos este capítulo con algunos resultados que prueban que K determina completamente el espacio $\mathcal{H}$. Introduciremos el concepto de marco de Parseval y demostraremos que, dado un marco de Parseval para un EHNR, es posible construir el núcleo como una serie de potencias. Recíprocamente, los términos de cualquier serie que proporcione el núcleo de esta manera deben conformar un marco de Parseval para el EHNR.

Seguidamente probaremos el teorema de Moore-Aronszajn, que caracteriza aquellas funciones que son núcleos de algún EHNR. Tales funciones suelen ser denominadas en la literatura como definidas positivas o semidefinidas positivas. Así, cualquier función de este tipo proporciona, en virtud del teorema de Moore-Aronszajn, un EHNR, pero obtener una descripción concreta del EHNR inducido no es una tarea sencilla. Este es el denominado problema de reconstrucción, que ilustraremos con algunos ejemplos básicos.

2.1. Estructura de espacio de Hilbert

Proposición 2.1. *Sea $\mathcal{H}$ un EHNR sobre un conjunto X, con núcleo K. Entonces, el subespacio vectorial*

$$\mathcal{M} = \operatorname{span}\{k_y(\cdot) = K(\cdot, y) \mid y \in X\}$$

es denso en $\mathcal{H}$.

Demostración. Una función $f \in \mathcal{H}$ es ortogonal a $\mathcal{M}$ si, y solo si, $\langle f, k_y \rangle = f(y) = 0$ para todo $y \in X$, es decir, si, y solo si, $f = 0$. Por tanto, $\mathcal{M}^{\perp} = \{0\}$, probando que $\mathcal{M}$ es denso en $\mathcal{H}$. $\square$

En un EHNR, la convergencia en norma de una sucesión implica la convergencia puntual.

Lema 2.2. *Sea $\mathcal{H}$ un EHNR sobre el conjunto X, y sea $\{f_n\}_{n=1}^{\infty} \subset \mathcal{H}$. Si $\lim\limits_{n\to\infty} \|f_n - f\| = 0$, entonces $f(x) = \lim\limits_{n\to\infty} f_n(x)$ para todo $x \in X$.*

Demostración. La desigualdad de Cauchy-Schwarz implica que

$$|f_n(x) - f(x)| = |\langle f_n - f, k_x\rangle| \leq \|f_n - f\| \, \|k_x\| \longrightarrow 0$$

cuando $n \to \infty$. □

El siguiente resultado establece que un EHNR queda unívocamente determinado por su núcleo reproductor.

Proposición 2.3. *Sean $\mathcal{H}_i$, $i = 1, 2$, dos EHNRs sobre un conjunto X, con núcleos reproductores K_i, $i = 1, 2$. Denotamos por $\|\cdot\|_i$ la norma de $\mathcal{H}_i$, $i = 1, 2$. Si $K_1(x, y) = K_2(x, y)$ para cualesquiera $x, y \in X$, entonces $\mathcal{H}_1 = \mathcal{H}_2$ y $\|f\|_1 = \|f\|_2$ para toda $f \in \mathcal{H}$.*

Demostración. Sean $K(x, y) = K_1(x, y) = K_2(x, y)$ y

$$\mathcal{W}_i = \operatorname{span}\{k_x \in \mathcal{H}_i \mid x \in X\}, \quad i = 1, 2.$$

Por la Proposición 2.1, $\mathcal{W}_i$ es denso en $\mathcal{H}_i$, $i = 1, 2$. Una función $f \in \mathcal{W}_i$, $i = 1, 2$, puede ser expresada como una combinación lineal finita

$$f(x) = \sum_{j=1}^{n} \alpha_j k_{x_j}(x) = \sum_{j=1}^{n} \alpha_j K(x, x_j),$$

de modo que los valores de f son independientes de que la consideremos como elemento de $\mathcal{W}_1$ o de $\mathcal{W}_2$. Además, para tales f se tiene que

$$\begin{aligned}
\|f\|_1^2 &= \Big\langle \sum_{i=1}^{n} \alpha_i k_{x_i}(x), \sum_{j=1}^{n} \alpha_j k_{x_j}(x) \Big\rangle_1 = \sum_{i,j} \alpha_i \overline{\alpha_j} \langle k_{x_i}, k_{x_j}\rangle_1 \\
&= \sum_{i,j=1}^{n} \alpha_i \overline{\alpha_j} K(x_j, x_i) = \Big\langle \sum_{i=1}^{n} \alpha_i k_{x_i}(x), \sum_{j=1}^{n} \alpha_j k_{x_j}(x) \Big\rangle_2 = \|f\|_2^2.
\end{aligned}$$

Consecuentemente, $\|f\|_1 = \|f\|_2$ para toda $f \in \mathcal{W}_1 = \mathcal{W}_2$.

Si $f \in \mathcal{H}_1$, existe una sucesión $\{f_n\}_{n=1}^{\infty} \subset \mathcal{W}_1$ tal que $\|f - f_n\|_1 \to 0$ cuando $n \to \infty$. Entonces $\{f_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy en $\mathcal{W}_1$ y, por lo tanto, también en $\mathcal{W}_2$, así que existe una función $g \in \mathcal{H}_2$ tal que $\|g - f_n\|_2 \to 0$ cuando $n \to \infty$. En virtud del Lema 2.2,

$$f(x) = \lim_{n\to\infty} f_n(x) = g(x).$$

Hemos probado así que $\mathcal{H}_1 \subset \mathcal{H}_2$. Por simetría, $\mathcal{H}_2 \subset \mathcal{H}_1$. Luego, $\mathcal{H}_1 = \mathcal{H}_2$. Finalmente, como $\|f\|_1 = \|f\|_2$ para toda f en el subconjunto denso $\mathcal{W}_1 = \mathcal{W}_2$ y la norma es continua, se concluye que $\|f\|_1 = \|f\|_2$ cualquiera que sea $f \in \mathcal{H}_1 = \mathcal{H}_2$. □

Una segunda consecuencia del Lema 2.2 es otra forma de calcular el núcleo de un EHNR que se revela muy útil en la práctica.

Dado un conjunto de vectores $\{h_s \mid s \in S\}$ en un espacio normado $\mathcal{H}$, indexado en un conjunto arbitrario S, decimos que $h = \sum_{s\in S} h_s$ si para todo $\varepsilon > 0$, existe un subconjunto finito $F_0 \subset S$ tal que para cualquier otro conjunto finito F con $F_0 \subset F \subset S$, se cumple que $\|h - \sum_{s\in F} h_s\| < \varepsilon$. Dos ejemplos de este tipo de convergencia son las identidades de Parseval: cuando $\{e_s \mid s \in S\}$ es una base ortonormal de un espacio de Hilbert $\mathcal{H}$, para cualquier $h \in \mathcal{H}$ se tiene que

$$\|h\|^2 = \sum_{s\in S} |\langle h, e_s\rangle|^2 \qquad \text{y} \qquad h = \sum_{s\in S} \langle h, e_s\rangle e_s.$$

Nótese que para definir esta clase de sumas no es necesario que S sea un conjunto ordenado. Un buen ejemplo para entender mejor este modo de convergencia viene dado por

$$a_n = \frac{(-1)^n}{n}, \quad n \in \mathbb{N};$$

la serie $\sum_{n=1}^{\infty} a_n$ converge, pero $\sum_{n\in\mathbb{N}} a_n$, no. En el caso complejo, $\sum_{n\in\mathbb{N}} z_n$ converge si, y solo si, $\sum_{n=1}^{\infty} |z_n|$ converge, es decir, si, y solo si, $\sum_{n=1}^{\infty} z_n$ converge absolutamente.

A continuación, usando esta nueva definición de convergencia, daremos la primera de varias caracterizaciones del núcleo que, en muchas ocasiones, facilitarán su cálculo.

Teorema 2.4. *Sea $\mathcal{H}$ un EHNR sobre un conjunto X, con núcleo $K(x,y)$. Sea $\{e_s \mid s \in S\}$ una base ortonormal de $\mathcal{H}$. Entonces, $K(x,y) = \sum_{s\in S} e_s(x)\overline{e_s(y)}$, donde la serie converge puntualmente.*

Demostración. Dado $y \in X$, se tiene que

$$\langle k_y, e_s\rangle = \overline{\langle e_s, k_y\rangle} = \overline{e_s(y)}.$$

Aplicando la segunda identidad de Parseval resulta que $k_y = \sum_{s\in S} \overline{e_s(y)} e_s$, donde, puesto que $\{e_s \mid s \in S\}$ es una base ortonormal, la serie de Fourier converge en la norma de $\mathcal{H}$. En tal caso, el Lema 2.2 asegura que la serie también converge puntualmente. Se concluye que

$$K(x,y) = k_y(x) = \sum_{s\in S} e_s(x)\overline{e_s(y)}, \quad x \in X.$$

$\square$

Mediante el Teorema 2.4 es posible recalcular algunos de los núcleos del Capítulo 1.

Ejemplo 2.5 (Espacio de Hardy sobre el disco unidad). Recordemos que el espacio de Hardy (Ejemplo 1.9) consiste en las series de potencias formales sobre el disco unidad $\mathbb{D}$. Las funciones $e_n = z^n$, $n \in \mathbb{Z}^+$, constituyen una base ortonormal de este espacio, de modo que, por el Teorema 2.4, su núcleo reproductor es

$$K(z,w) = \sum_{n\in\mathbb{Z}^+} \overline{e_n(w)} e_n(z) = \sum_{n=0}^{\infty} (\overline{w}z)^n = \frac{1}{1-\overline{w}z}, \quad z, w \in \mathbb{D}.$$

□

Ejemplo 2.6 (Espacio de Sobolev en $[0,1]$). Las funciones

$$c_n(t) = \frac{1}{\sqrt{2}\pi n}[\cos(2\pi nt) - 1], \quad s_n(t) = \frac{1}{\sqrt{2}\pi n} \operatorname{sen}(2\pi nt), \quad n \in \mathbb{N},$$

constituyen una base ortonormal del espacio del Ejemplo 1.7:

$$\mathcal{H} = \Big\{ f : [0,1] \to \mathbb{R} \;\Big|\; f \text{ absolutamente continua}, \\ f(0) = f(1) = 0, \int_0^1 |f'(x)|^2 dx < +\infty \Big\}.$$

En efecto, la base ortonormal estándar de $L^2[0,1]$ está formada por la función constante 1 junto con las funciones $\sqrt{2}\cos(2\pi n\,\cdot)$ y $\sqrt{2}\operatorname{sen}(2\pi n\,\cdot)$, $n \in \mathbb{N}$; por tanto, el conjunto $\{c_n, s_n\}_{n\in\mathbb{N}}$ es un sistema ortonormal en $\mathcal{H}$. Para ver que es completo, supongamos que $\langle f, c_n\rangle = \langle f, s_n\rangle = 0$ cualquiera que sea $n \in \mathbb{N}$. Entonces f' es ortogonal a todos los elementos no constantes de la base ortonormal de $L^2[0,1]$. Sigue que f' es constante y, por lo tanto, que f es un polinomio de primer grado. Las condiciones de frontera $f(0) = f(1) = 0$ obligan a que este polinomio sea 0.

Atendiendo ahora al Teorema 2.4 y al cálculo del núcleo reproductor efectuado con anterioridad, resulta que

$$\sum_{n\in\mathbb{N}} \frac{\cos[2\pi n(x-y)] - \cos(2\pi nx) - \cos(2\pi ny) + 1}{2\pi^2 n^2} = \begin{cases} (1-y)x, & 0 \le x < y \\ y(1-x), & y \le x \le 1 \end{cases}.$$

□

Teorema 2.7. *Sea $\mathcal{H}$ un EHNR sobre un conjunto X, con núcleo reproductor K. Si $\mathcal{H}_0 \subset \mathcal{H}$ es un subespacio cerrado y $P_0 : \mathcal{H} \to \mathcal{H}_0$ es la proyección ortogonal de $\mathcal{H}$ sobre $\mathcal{H}_0$, entonces $\mathcal{H}_0$ es un EHNR sobre el conjunto X, con núcleo reproductor $K_0(x,y) = \langle P_0(k_y), k_x\rangle$.*

Demostración. Como $\mathcal{H}_0$ es un subespacio cerrado del espacio de Hilbert $\mathcal{H}$, se tiene que $\mathcal{H}_0$ es un espacio de Hilbert. Fijemos $x \in X$. Puesto que el funcional evaluación E_x está acotado sobre $\mathcal{H}$, también lo está sobre $\mathcal{H}_0$. Además, si $f \in \mathcal{H}_0$,

$$f(x) = \langle f, k_x \rangle = \langle f, P_0(k_x) \rangle.$$

Luego, $P_0(k_x)$ es el núcleo reproductor del punto x para $\mathcal{H}_0$ y, en consecuencia,

$$K_0(x, y) = \langle P_0(k_y), P_0(k_x) \rangle = \langle P_0(k_y), k_x \rangle.$$

□

Puede ocurrir que el conjunto de funciones cuya suma proporciona el núcleo como en el Teorema 2.4 no sea una base ortonormal del EHNR. Nos disponemos a caracterizar los conjuntos que tienen esta propiedad. La siguiente definición está inspirada en las identidades de Parseval.

Definición 2.8. *Sea $\mathcal{H}$ un espacio de Hilbert con un producto interior $\langle \cdot, \cdot \rangle$. Un marco de Parseval para $\mathcal{H}$ es un conjunto de vectores $A = \{f_s \mid s \in S\} \subset \mathcal{H}$ tal que para todo $h \in \mathcal{H}$, se verifica:*

$$\|h\|^2 = \sum_{s \in S} |\langle h, f_s \rangle|^2.$$

Un marco de Parseval es una base ortonormal si, y sólo si, es un conjunto ortonormal. En particular, toda base ortonormal es un marco de Parseval. También, si tenemos dos bases ortonormales $\{f_s \mid s \in S\}$ y $\{g_t \mid t \in T\}$ en un espacio de Hilbert $\mathcal{H}$, entonces tanto $\{f_s \mid s \in S\} \cup \{0\}$ como

$$\left\{ \frac{f_s}{\sqrt{2}} \;\middle|\; s \in S \right\} \cup \left\{ \frac{g_t}{\sqrt{2}} \;\middle|\; t \in T \right\}$$

constituyen marcos de Parseval para $\mathcal{H}$. En particular, los marcos de Parseval no tienen por qué ser conjuntos linealmente independientes.

La siguiente proposición muestra una de las formas más comunes de generar marcos de Parseval.

Proposición 2.9. *Sean $\mathcal{H}$ un espacio de Hilbert, $\mathcal{M} \subset \mathcal{H}$ un subespacio cerrado de $\mathcal{H}$, y P la proyección ortogonal de $\mathcal{H}$ sobre $\mathcal{M}$. Si $\{e_s \mid s \in S\}$ es una base ortonormal de $\mathcal{H}$, entonces $\{P(e_s) \mid s \in S\}$ es un marco de Parseval para $\mathcal{M}$.*

Demostración. Sea $h \in \mathcal{M}$. Como $h = P(h)$, se tiene:

$$\langle h, e_s \rangle = \langle P(h), e_s \rangle = \langle h, P(e_s) \rangle.$$

Luego,

$$\|h\|^2 = \sum_{s \in S} |\langle h, e_s \rangle|^2 = \sum_{s \in S} |\langle h, P(e_s) \rangle|^2,$$

de donde se sigue el resultado. □

A continuación probaremos una caracterización de los marcos de Parseval. En particular, veremos que los marcos de Parseval forman un sistema generador de $\mathcal{H}$, aunque no todo sistema generador es un marco de Parseval. Además, mostraremos que la definición de marco de Parseval es independiente de la identidad de Parseval elegida.

Sea
$$\ell^2(S) = \Big\{ f : S \to \mathbb{C} \mid \sum_{s \in S} |f(s)|^2 < +\infty \Big\}$$
el espacio de Hilbert de las funciones de cuadrado absolutamente sumable sobre S, y sea $\{e_s \mid s \in S\}$ la base ortonormal canónica de $\ell^2(S)$:
$$e_s : S \to \mathbb{C}, \quad e_s(t) = \begin{cases} 1, & \text{si } t = s \\ 0, & \text{si } t \neq s \end{cases}, \quad s \in S.$$

Proposición 2.10. *Sea $\mathcal{H}$ un espacio de Hilbert, y sea $\{f_s \mid s \in S\} \subset \mathcal{H}$. Los siguientes asertos son equivalentes.*

(i) El conjunto $\{f_s \mid s \in S\}$ es un marco de Parseval.

(ii) La función $V : \mathcal{H} \to \ell^2(S)$ dada por $(Vh)(s) = \langle h, f_s \rangle$ es una isometría bien definida.

(iii) Para todo $h \in \mathcal{H}$ se tiene que $h = \sum_{s \in S} \langle h, f_s \rangle f_s$. Además,
$$\langle h_1, h_2 \rangle = \sum_{s \in S} \langle h_1, f_s \rangle \overline{\langle h_2, f_s \rangle}$$
cualesquiera sean $h_1, h_2 \in \mathcal{H}$.

Demostración. $(i) \Rightarrow (ii)$ Es fácil comprobar que $V : \mathcal{H} \to \ell^2(S)$ dada por $(Vh)(s) = \langle h, f_s \rangle$ está bien definida y es lineal. Nótese que, en términos de la base ortonormal canónica $\{e_s\}_{s \in S}$ de $\ell^2(S)$, $Vh = \sum_{s \in S} \langle h, f_s \rangle e_s$. Puesto que $\{f_s \mid s \in S\}$ es un marco de Parseval, podemos escribir
$$\begin{aligned} \|Vh\|^2 = \langle Vh, Vh \rangle &= \Big\langle \sum_{s \in S} \langle h, f_s \rangle e_s, \sum_{t \in S} \langle h, f_t \rangle e_t \Big\rangle \\ &= \sum_{s \in S} \sum_{t \in S} \langle h, f_s \rangle \overline{\langle h, f_t \rangle} \langle e_s, e_t \rangle = \sum_{s \in S} |\langle h, f_s \rangle|^2 = \|h\|^2, \end{aligned}$$
probando que V es una isometría.

$(ii) \Rightarrow (iii)$ Supongamos que $V : \mathcal{H} \to \ell^2(S)$ así definida es una isometría; hemos de ver que $h = \sum_{s \in S} \langle h, f_s \rangle f_s$ para todo $h \in \mathcal{H}$.

Como V es una isometría, es continua, luego existe un único operador lineal continuo $V^* : \ell^2(S) \to \mathcal{H}$, el adjunto de V, tal que $\langle Vh, g \rangle = \langle h, V^* g \rangle$ para cualesquiera $h \in \mathcal{H}$ y $g \in \ell^2(S)$. Si $h \in \mathcal{H}$, entonces

$$\langle h, V^* e_t \rangle = \langle Vh, e_t \rangle = \Big\langle \sum_{s\in S} \langle h, f_s \rangle e_s, e_t \Big\rangle = \sum_{s\in S} \langle h, f_s \rangle \langle e_s, e_t \rangle = \langle h, f_t \rangle.$$

Así pues,

$$V^* e_t = f_t, \quad t \in S. \tag{2.1}$$

Ya que V es una isometría, $V^*V = I_{\mathcal{H}}$, y finalmente

$$h = V^*Vh = V^* \left(\sum_{s\in S} \langle h, f_s \rangle e_s \right) = \sum_{s\in S} \langle h, f_s \rangle V^* e_s = \sum_{s\in S} \langle h, f_s \rangle f_s$$

para todo $h \in \mathcal{H}$.
$(iii) \Rightarrow (i)$ Como, por hipótesis, $h = \sum_{s\in S} \langle h, f_s \rangle f_s$, se tiene:

$$\begin{aligned} \|h\|^2 = \langle h, h \rangle = \Big\langle \sum_{s\in S} \langle h, f_s \rangle f_s, h \Big\rangle &= \sum_{s\in S} \langle h, f_s \rangle \langle f_s, h \rangle \\ = \sum_{s\in S} \langle h, f_s \rangle \overline{\langle h, f_s \rangle} &= \sum_{s\in S} |\langle h, f_s \rangle|^2. \end{aligned}$$

Por lo tanto, el conjunto $\{f_s \mid s \in S\}$ es un marco de Parseval. Para completar la prueba, veamos que efectivamente se cumple la identidad de Parseval para el producto interior: si $h_1, h_2 \in \mathcal{H}$,

$$\langle h_1, h_2 \rangle = \Big\langle \sum_{s\in S} \langle h_1, f_s \rangle f_s, h_2 \Big\rangle = \sum_{s\in S} \langle h_1, f_s \rangle \overline{\langle h_2, f_s \rangle},$$

como se pretendía. □

Se estableció en la Proposición 2.9 que la proyección ortogonal de una base ortonormal de un espacio de Hilbert $\mathcal{H}$ es un marco de Parseval en el subespacio correspondiente. A continuación veremos que esta es la forma más general en la que se puede obtener un marco de Parseval.

Proposición 2.11 (Larson). *Sea $\{f_s \mid s \in S\}$ un marco de Parseval en un espacio de Hilbert $\mathcal{H}$. Entonces existen un espacio de Hilbert $\mathcal{K}$, que contiene a $\mathcal{H}$ como subespacio, y una base ortonormal $\{e_s \mid s \in S\}$ de $\mathcal{K}$, tales que $f_s = P_{\mathcal{H}}(e_s)$, $s \in S$, donde $P_{\mathcal{H}}$ denota la proyección ortogonal de $\mathcal{K}$ sobre $\mathcal{H}$.*

Demostración. Sean $\mathcal{K} = \ell^2(S)$ y $V : \mathcal{H} \to \mathcal{K}$ la isometría de la Proposición 2.10. Identificando $\mathcal{H}$ con $V(\mathcal{H})$ podemos contemplar $\mathcal{H}$ como un subespacio de $\mathcal{K}$. Sea ahora $P = VV^* : \mathcal{K} \to \mathcal{K}$. Como V y V^* son lineales, P también lo es; además, se tiene que $P = P^*$ y, puesto que $V^*V = I_{\mathcal{H}}$,

$$P^2 = (VV^*)(VV^*) = V(V^*V)V^* = VV^* = P.$$

Luego, P es la proyección ortogonal sobre algún subespacio de $\mathcal{K}$. Ya que, por (2.1),

$$P(e_s) = V(V^*e_s) = Vf_s \in V(\mathcal{H}),$$

se infiere que P es la proyección sobre $V(\mathcal{H})$. Identificando $h \equiv Vh$ se concluye que P es la proyección sobre $\mathcal{H}$, con $P(e_s) = Vf_s \equiv f_s$. $\square$

El siguiente resultado proporciona una forma de calcular núcleos reproductores más general que la del Teorema 2.4.

Teorema 2.12 (Papadakis). *Sea $\mathcal{H}$ un EHNR sobre un conjunto X, con núcleo reproductor $K(x,y)$. Entonces, $\{f_s \mid s \in S\} \subset \mathcal{H}$ es un marco de Parseval para $\mathcal{H}$ si, y solo si,*

$$K(x,y) = \sum_{s\in S} f_s(x)\overline{f_s(y)}, \tag{2.2}$$

donde la serie converge puntualmente.

Demostración. Supongamos que $\{f_s \mid s \in S\}$ es un marco de Parseval. En tal caso (Proposición 2.10),

$$\begin{aligned} K(x,y) = k_y(x) = \langle k_y, k_x\rangle &= \sum_{s\in S}\langle k_y, f_s\rangle\overline{\langle k_x, f_s\rangle} \\ &= \sum_{s\in S}\langle f_s, k_x\rangle\overline{\langle f_s, k_y\rangle} = \sum_{s\in S} f_s(x)\overline{f_s(y)}. \end{aligned}$$

Recíprocamente, supongamos que

$$K(x,y) = \langle k_y, k_x\rangle = \sum_{s\in S} f_s(x)\overline{f_s(y)};$$

hemos de ver que $\{f_s \mid s \in S\}$ es un marco de Parseval. A tal fin, sea $h \in \mathcal{M} = \operatorname{span}\{k_x \mid x \in X\}$, de modo que $h = \sum_{i=1}^n \alpha_i k_{x_i}$, con α_i escalares, es una combinación lineal finita de funciones núcleo. Entonces,

$$\begin{aligned} \|h\|^2 = \langle h, h\rangle = \Big\langle \sum_{i=1}^n \alpha_i k_{x_i}, \sum_{j=1}^n \alpha_j k_{x_j}\Big\rangle &= \sum_{i,j=1}^n \alpha_i\overline{\alpha_j}\langle k_{x_i}, k_{x_j}\rangle \\ = \sum_{i,j=1}^n \alpha_i\overline{\alpha_j}K(x_j, x_i) &= \sum_{i,j=1}^n \alpha_i\overline{\alpha_j}\sum_{s\in S} f_s(x_j)\overline{f_s(x_i)} \\ = \sum_{i,j=1}^n \alpha_i\overline{\alpha_j}\sum_{s\in S}\langle f_s, k_{x_j}\rangle\overline{\langle f_s, k_{x_i}\rangle} &= \sum_{s\in S}\Big\langle f_s, \sum_{j=1}^n \alpha_j k_{x_j}\Big\rangle\Big\langle \sum_{i=1}^n \alpha_i k_{x_i}, f_s\Big\rangle \\ = \sum_{s\in S}\langle f_s, h\rangle\langle h, f_s\rangle = \sum_{s\in S}\langle h, f_s\rangle\overline{\langle h, f_s\rangle} &= \sum_{s\in S}|\langle h, f_s\rangle|^2. \end{aligned}$$

Como $\mathcal{M}$ es denso en $\mathcal{H}$ (Proposición 2.1), para todo $h \in \mathcal{H}$ existe una sucesión $\{h_n\}_{n=1}^{\infty} \subset \mathcal{M}$ tal que $\lim\limits_{n\to\infty} h_n = h$. Aplicando la continuidad de la norma y del producto interior, podemos escribir

$$\|h\|^2 = \lim_{n\to\infty} \|h_n\|^2 = \lim_{n\to\infty} \sum_{s\in S} |\langle h_n, f_s\rangle|^2 = \sum_{s\in S} |\langle h, f_s\rangle|^2.$$

Por lo tanto, $\{f_s \mid s \in S\}$ es un marco de Parseval. □

2.2. Caracterización de los núcleos reproductores

Nos ocupamos ahora del problema de determinar condiciones necesarias y suficientes para que una función $K : X \times X \to \mathbb{C}$ sea el núcleo reproductor de algún EHNR sobre el conjunto X. Solventado este, nos plantearemos, siquiera someramente, el problema de reconstrucción, consistente en obtener una descripción concreta del EHNR inducido por una de estas funciones.

Para abordar el primer problema necesitamos recordar los conceptos de matriz positiva y función definida positiva.

Definición 2.13. *Sea $A = (a_{i,j})_{i,j=1}^{n}$ una matriz compleja cuadrada de orden n. Se dice que A es positiva, y se escribe $A \geq 0$, si para cada $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ se tiene que $\sum_{i,j=1}^{n} \overline{\alpha_i}\alpha_j a_{i,j} \geq 0$.*

Si $\langle\cdot,\cdot\rangle$ denota el producto interior usual de $\mathbb{C}^n$ entonces, en términos de este producto interior, $A \geq 0$ si, y sólo si, $\langle Ax, x\rangle \geq 0$ para cada $x \in \mathbb{C}^n$. En efecto, la suma que comparece en la definición es $\langle Ax, x\rangle$ para el vector x cuya i-ésima componente es α_i.

Asumiremos cierta familiaridad con algunas propiedades básicas de las matrices positivas. En particular, usaremos, sin demostración, que si $A \geq 0$ y $B \geq 0$ son ambas matrices cuadradas de orden n, entonces $A + B \geq 0$ y $rA \geq 0$ para cualquier $r \in \mathbb{R}^+$. También usaremos el hecho de que las matrices positivas pueden ser caracterizadas en términos de sus autovalores: se tiene que $A \geq 0$ si, y sólo si, $A = A^*$ y cada autovalor de A es no negativo. Por esta razón, algunos autores prefieren denominar a estas matrices semidefinidas positivas o no negativas. Nosotros nos adheriremos a la notación y terminología más comunes en teoría de operadores. En caso de que $A = A^*$ y todo autovalor de A sea estrictamente positivo, diremos que A es estrictamente positiva y escribiremos $A > 0$. Puesto que A es una matriz cuadrada de orden n, vemos que $A > 0$ equivale a que $A \geq 0$ y A es inversible.

Definición 2.14. *Sean X un conjunto y $K : X \times X \to \mathbb{C}$ una función de dos variables. Diremos que K es una función núcleo, y escribiremos $K \geq 0$, si para cada $n \in \mathbb{N}$ y cada elección de n puntos distintos $\{x_1, \ldots, x_n\} \subset X$, la matriz $(K(x_i, x_j))_{i,j=1}^{n}$ es positiva.*

De nuevo, esta terminología no es estándar. Algunos autores prefieren referirse a las funciones núcleo como funciones definidas positivas, mientras que otros las denominan funciones semidefinidas positivas, reservando el término «función definida positiva» para las funciones $K(x, y)$ tales que la matriz $(K(x_i, x_j))_{i,j=1}^n$ es estrictamente positiva. En algunas disciplinas, el término «función núcleo» significa una función simétrica de dos variables, esto es, tal que $K(x, y) = K(y, x)$. En breve probaremos que una función es una función núcleo si, y sólo si, es el núcleo reproductor de algún EHNR, lo que justifica la terminología adoptada.

Proposición 2.15. *Sea $\mathcal{H}$ un EHNR sobre un conjunto X, con núcleo reproductor K. Entonces K es una función núcleo.*

Demostración. Fijemos puntos $\{x_1, \ldots, x_n\} \subset X$ tales que $x_i \neq x_j$ siempre que $i \neq j$, y escalares $\{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{C}$. Entonces,

$$\sum_{i,j=1}^{n} \overline{\alpha_i}\alpha_j K(x_i, x_j) = \left\langle \sum_{j=1}^{n} \alpha_j k_{x_j}, \sum_{i=1}^{n} \alpha_i k_{x_i} \right\rangle = \left\| \sum_{i=1}^{n} \alpha_i k_{x_i} \right\|^2 \geq 0.$$

□

En general, dado un EHNR, la matriz $(K(x_i, x_j))_{i,j=1}^n$ no es estrictamente positiva. El cálculo en la prueba de la Proposición 2.15 muestra que $\{x_1, \ldots, x_n\} \subset X$ y $\{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{C}$ son tales que $\sum_{i,j=1}^n \overline{\alpha_i}\alpha_j K(x_i, x_j) = 0$ si, y solo si, $\|\sum_{i=1}^n \alpha_i k_{x_i}\| = 0$. Luego, para todo $f \in \mathcal{H}$ se verifica que $\sum_{i=1}^n \overline{\alpha_i} f(x_i) = \langle f, \sum_{i=1}^n \alpha_i k_{x_i} \rangle = 0$, lo que significa que existe una relación de dependencia lineal entre las imágenes de cualquier función $f \in \mathcal{H}$ en los puntos $\{x_1, \ldots, x_n\}$.

Un ejemplo de esta situación es el espacio de Sobolev en $[0, 1]$ (Ejemplo 1.7), donde la condición de frontera $f(0) = f(1)$ se traduce en que $k_0(t) = k_1(t)$.

Por otro lado, muchos espacios de funciones analíticas, como los de Hardy (Ejemplo 1.9) y Bergman (Ejemplo 1.10), contienen a todos los polinomios. Pero no existe una ecuación de la forma $\sum_{i=1}^n \beta_i p(x_i) = 0$, con escalares β_i no todos nulos, que sea satisfecha por todos los polinomios p. Luego, los núcleos reproductores de estos espacios siempre definen matrices estrictamente positivas e inversibles.

Así, por ejemplo, el núcleo de Szegö para el espacio de Hardy tiene la propiedad de que para cualesquiera $\alpha_1, \ldots, \alpha_n \in \mathbb{D}$ distintos, la matriz

$$\left(\frac{1}{1 - \overline{\alpha_i}\alpha_j} \right)_{i,j=1}^n$$

es inversible. Este es otro ejemplo de la potencia de la teoría de EHNRs; demostrar la misma propiedad por métodos estándar de álgebra lineal resulta mucho más difícil.

La Proposición 2.15 es bastante elemental. Pese a ello, admite un recíproco muy profundo que permite caracterizar las funciones que son núcleos reproductores.

Teorema 2.16 (Moore-Aronszajn). *Sean X un conjunto y $K : X \times X \to \mathbb{C}$ una función. Si K es una función núcleo, entonces existe un EHNR $\mathcal{H}$ de funciones complejas definidas sobre X cuyo núcleo reproductor es K.*

Demostración. Para cada $y \in X$, sea $k_y(x) = K(x, y)$, $x \in X$, y sea $\mathcal{W} = \operatorname{span}\{k_y \mid y \in X\}$. Definimos la forma $B : \mathcal{W} \times \mathcal{W} \to \mathbb{C}$ por

$$\begin{aligned} B(f,g) = B\left(\sum_{i=1}^n \alpha_i k_{x_i}, \sum_{j=1}^m \beta_j k_{x_j}\right) &= \sum_{i=1}^n \sum_{j=1}^m \alpha_i \overline{\beta_j} K(x_j, x_i) \\ &= \sum_{i=1}^n \sum_{j=1}^m \alpha_i \overline{\beta_j} k_{x_i}(x_j), \quad f, g \in \mathcal{W}, \end{aligned}$$

donde $f = \sum_{i=1}^n \alpha_i k_{x_i} \in \mathcal{W}$, $g = \sum_{j=1}^m \beta_j k_{x_j} \in \mathcal{W}$, con $\alpha_i, \beta_j \in \mathbb{C}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$.

Veamos que B está bien definida. A tal fin, supongamos que $(f, g) = (f', g') \in \mathcal{W} \times \mathcal{W}$, siendo $f = \sum_i \alpha_i k_{x_i} = \sum_r \lambda_r k_{x_r} = f'$ y $g = \sum_j \beta_j k_{x_j} = \sum_s \gamma_s k_{x_s} = g'$. Entonces:

$$\begin{aligned} B(f,g) &= \sum_j \overline{\beta_j} \sum_i \alpha_i k_{x_i}(x_j) = \sum_j \overline{\beta_j} f(x_j) = \sum_j \overline{\beta_j} f'(x_j) \\ &= \sum_j \overline{\beta_j} \sum_r \lambda_r k_{x_r}(x_j) = \sum_r \lambda_r \sum_j \overline{\beta_j k_{x_j}(x_r)} \\ &= \sum_r \lambda_r \overline{g(x_r)} = \sum_r \lambda_r \overline{g'(x_r)} = \sum_r \lambda_r \sum_s \overline{\gamma_s k_{x_s}(x_r)} \\ &= \sum_r \lambda_r \sum_s \overline{\gamma_s} k_{x_r}(x_s) = \sum_{r,s} \lambda_r \overline{\gamma_s} k_{x_r}(x_s) = B(f', g'). \end{aligned}$$

Luego, B está bien definida. Se comprueba rutinariamente que la forma B es lineal en la primera componente y lineal conjugada en la segunda. Observamos, además, que si $f \in \mathcal{W}$ es como anteriormente, entonces

$$B(f, k_x) = \sum_{i=1}^n \alpha_i k_{x_i}(x) = f(x), \quad x \in X.$$

Para que la forma sesquilineal B sea un producto interior, debe verificar que

$$B(f,f) \geq 0 \quad \text{y} \quad B(f,f) = 0 \text{ si, y solo si, } f = 0, \quad f \in \mathcal{W}.$$

Al ser K una función núcleo, por definición tenemos que

$$B(f,f)=\sum_{i,j=1}^{n}\alpha_i\overline{\alpha_j}k_{x_i}(x_j)\geq 0.$$

Si $f=0$, es obvio que $B(f,f)=0$. Recíprocamente, si $B(f,f)=0$ entonces, por la desigualdad de Cauchy-Schwarz para formas sequilineales no negativas,

$$|f(x)|^2=|B(f,k_x)|^2\leq B(f,f)B(k_x,k_x)=0,\quad x\in X,$$

obligando a que $f=0$. Hemos probado así que B es un producto interior sobre el espacio vectorial $\mathcal{W}$.

Es bien sabido que todo espacio con producto interior admite una compleción; por tanto, existe un espacio de Hilbert $(\mathcal{H},\langle\cdot,\cdot\rangle)$ tal que $\mathcal{W}$ puede ser contemplado como un subespacio denso de $\mathcal{H}$. Dada $h\in\mathcal{H}$, definimos la función

$$\widehat{h}(x)=\langle h,k_x\rangle,\quad x\in X.$$

Entonces,

$$\widehat{\mathcal{H}}=\left\{\widehat{h}\mid h\in\mathcal{H}\right\}\subset\mathcal{F}(X,\mathbb{C})$$

es un espacio vectorial con las operaciones algebraicas usuales, y la aplicación $L:\mathcal{H}\to\mathcal{F}(X,\mathbb{C})$ definida por $L(h)=\widehat{h}$ es lineal. Además, para cualquier $f\in\mathcal{W}$ se cumple que

$$\widehat{f}(x)=\langle f,k_x\rangle=B(f,k_x)=f(x),\quad x\in X.$$

Queremos ver que L es inyectiva, esto es, que $Lh=0$ implica $h=0$. Sea $h\in\mathcal{H}$ tal que $\widehat{h}(x)=0$ para todo $x\in X$. Entonces, $h\perp k_x$ para todo $x\in X$ y, por lo tanto, $h\perp\mathcal{W}$. Como $\mathcal{W}$ es denso en $\mathcal{H}$, necesariamente $h=0$. Concluimos así que $L:\mathcal{H}\to\widehat{\mathcal{H}}$ es un isomorfismo. Trasladamos el producto interior de $\mathcal{H}$ a $\widehat{\mathcal{H}}$ definiendo

$$\left\langle\widehat{h_1},\widehat{h_2}\right\rangle=\langle h_1,h_2\rangle,$$

con lo que convertimos $\widehat{\mathcal{H}}$ en un espacio de Hilbert de funciones complejas sobre X.

Dado $x\in X$, el funcional evaluación

$$E_x\big(\widehat{h}\big)=\widehat{h}(x)=\langle h,k_x\rangle=\left\langle\widehat{h},\widehat{k_x}\right\rangle$$

está acotado, con norma

$$\|E_x\|=\left\|\widehat{k_x}\right\|=\|k_x\|=K(x,x)<+\infty,$$

y además $\widehat{k_x}=k_x$ es el núcleo reproductor del punto x. Así pues, $\widehat{\mathcal{H}}$ es un EHNR sobre el conjunto X, siendo su núcleo reproductor

$$\widehat{K}(x,y)=\widehat{k_y}(x)=\langle\widehat{k_y},\widehat{k_x}\rangle=\langle k_y,k_x\rangle=K(x,y),\quad x,y\in X.$$

$\square$

El teorema de Moore-Aronszajn, junto con la Proposición 2.15, prueba que existe una correspondencia biunívoca entre los EHNRs sobre un conjunto y las funciones núcleo sobre ese conjunto.

Definición 2.17. *Dada una función núcleo $K : X \times X \to \mathbb{C}$, denotamos por $\mathcal{H}(K)$ el único EHNR sobre X con núcleo reproductor K.*

Como ya indicamos, uno de los problemas más arduos de la teoría de EHNRs es el problema de reconstrucción, consistente en dar una descripción concreta del espacio $\mathcal{H}(K)$ a partir de una función núcleo K. En ocasiones, este problema conduce a resultados contraintuitivos. Por ejemplo, si partimos del núcleo de Szegö en el disco unidad $\mathbb{D}$, donde $K(z,w) = 1/(1-\overline{w}z)$ (Ejemplo 1.9), entonces el espacio $\mathcal{W}$ que se construye en la demostración del teorema de Moore-Aronszajn consiste en combinaciones lineales de las funciones $k_w(z)$, que son funciones racionales con un único polo simple fuera de $\mathbb{D}$; así pues, $\mathcal{W}$ no contiene a los polinomios. Sin embargo, los polinomios son densos en $\mathcal{H}(K) = H^2(\mathbb{D})$.

Cerramos el capítulo con dos aproximaciones sencillas al problema de reconstrucción.

Proposición 2.18. *Sea X un conjunto, sea $f \neq 0$ una función sobre X, y definamos $K(x,y) = f(x)\overline{f(y)}$. Entonces K es una función núcleo, $\mathcal{H}(K)$ es el subespacio generado por f, y $\|f\| = 1$.*

Demostración. Se comprueba fácilmente que K es un núcleo:

$$\sum_{i,j=1}^{n} \alpha_i \overline{\alpha_j} K(x_i, x_j) = \left| \sum_{i=1}^{n} \alpha_i f(x_i) \right|^2 \geq 0.$$

Para determinar $\mathcal{H}(K)$, notemos que $k_x = \overline{f(x)} f$ para todo $x \in X$ y, por lo tanto, $\mathcal{W}$ no es más que el subespacio unidimensional generado por f. Puesto que todo espacio de dimensión finita es completo, $\mathcal{H}$ es precisamente el subespacio generado por f.

Finalmente, calculamos la norma de f. Fijemos cualquier punto x tal que $f(x) \neq 0$. Entonces,

$$|f(x)|^2 \, \|f\|^2 = \|\overline{f(x)} f\|^2 = \|k_x\|^2 = \langle k_x, k_x \rangle = K(x,x) = |f(x)|^2,$$

y se concluye que $\|f\| = 1$. $\square$

Teorema 2.19. *Sea X un espacio topológico y sea $K : X \times X \to \mathbb{C}$ una función núcleo, donde $X \times X$ está provisto de la topología producto. Si K es continua, entonces todas las funciones de $\mathcal{H}(K)$ son continuas.*

Demostración. Sea $f \in \mathcal{H}(K)$, y fijemos $y_0 \in X$. Dado $\varepsilon > 0$, debemos encontrar un entorno U de y_0 tal que $y \in U$ implique $|f(y) - f(y_0)| < \varepsilon$.

Puesto que K es continua, existe un entorno $V \subset X \times X$ de (y_0, y_0) tal que

$$|K(x,y) - K(y_0,y_0)| < \frac{\varepsilon^2}{3(\|f\|^2+1)}, \quad (x,y) \in V.$$

Por otra parte, como $X \times X$ está provisto de la topología producto, podemos elegir un entorno $U \subset X$ de y_0 tal que $U \times U \subset V$. Si $y \in U$, entonces

$$\begin{aligned}
&\|k_y - k_{y_0}\|^2 \\
&= \langle k_y - k_{y_0}, k_y - k_{y_0} \rangle \\
&= K(y,y) - K(y,y_0) - K(y_0,y) + K(y_0,y_0) \\
&= [K(y,y) - K(y_0,y_0)] - [K(y,y_0) - K(y_0,y_0)] - [K(y_0,y) - K(y_0,y_0)] \\
&< \frac{\varepsilon^2}{\|f\|^2+1}.
\end{aligned}$$

Consecuentemente,

$$|f(y) - f(y_0)| = |\langle f, k_y - k_{y_0} \rangle| \le \|f\| \, \|k_y - k_{y_0}\| < \varepsilon,$$

como se pretendía. □

3

Interpolación y aproximación

La resolución de problemas de interpolación y aproximación es una de las aplicaciones principales de la teoría de EHNR. Es relativamente sencillo dar fórmulas concretas para la interpolación y la aproximación en estos espacios.

Definición 3.1. *Sean X e Y dos conjuntos, sea $\{x_1, \ldots, x_n\} \subset X$ una colección de puntos distintos, y sea $\{\lambda_1, \ldots, \lambda_n\} \subset Y$. Diremos que una función $g : X \to Y$ interpola esos puntos si $g(x_i) = \lambda_i$, para todo $i = 1, \ldots, n$.*

Sea $\mathcal{H}$ un EHNR sobre X, con núcleo reproductor K. Supongamos que $\{x_1, \ldots, x_n\} \subset X$ es un conjunto de puntos distintos y que $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ es una colección de escalares no necesariamente distintos. En este capítulo comenzaremos dando condiciones necesarias y suficientes para que exista una función $g \in \mathcal{H}$ que interpole esos valores; seguidamente, probaremos la existencia de un único interpolante de norma mínima y daremos una expresión concreta de este interpolante minimal. Por último, utilizaremos la misma técnica para proponer una solución del problema de reconstrucción.

3.1. Interpolación en un EHNR

Antes de continuar, adoptaremos la siguiente notación. Dado un conjunto finito $F = \{x_1, \ldots, x_n\} \subset X$ de puntos distintos, denotaremos por $\mathcal{H}_F \subset \mathcal{H}$ el subespacio engendrado por las funciones núcleo $\{k_{x_1}, \ldots, k_{x_n}\}$:

$$\mathcal{H}_F = \operatorname{span}\{k_{x_1}, \ldots, k_{x_n}\}.$$

Nótese que $\dim(\mathcal{H}_F) \leq n$, y que se da la desigualdad estricta si, y sólo si, existe alguna relación de dependencia lineal entre estas funciones, con coeficientes no simultáneamente nulos. Más precisamente, supongamos que $\sum_{j=1}^{n} \alpha_j k_{x_j} = 0$. Entonces, para cada $f \in \mathcal{H}$,

$$0 = \left\langle f, \sum_{j=1}^{n} \alpha_j k_{x_j} \right\rangle = \sum_{j=1}^{n} \overline{\alpha_j} f(x_j).$$

Vemos así que $\dim(\mathcal{H}_F) < n$ si, y solo si, los valores de cualquier $f \in \mathcal{H}$ en los puntos de F guardan una dependencia lineal; equivalentemente, si la aplicación lineal $T_F : \mathcal{H} \to \mathbb{C}^n$ definida por

$$T_F(f) = (f(x_1), \ldots, f(x_n))$$

no es sobre, puesto que cualquier vector $(\alpha_1, \ldots, \alpha_n)$ que exprese una tal relación de dependencia será ortogonal al rango de T_F.

Por tanto, cuando $\dim(\mathcal{H}_F) < n$, existen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ que no pueden ser interpoladas por ninguna $f \in \mathcal{H}$.

La existencia de relaciones de dependencia lineal entre las funciones núcleo no solo es posible sino que, en ocasiones, es deseable. Recuérdese el caso del espacio de Sobolev (Ejemplo 1.7), donde la condición de frontera $f(0) = f(1) = 0$ de la ecuación diferencial requiere que $k_1 = k_0 = 0$. Un cambio de la condición de frontera a $f(0) = f(1)$ requeriría que $k_1 = k_0$.

Denotaremos por P_F la proyección ortogonal de $\mathcal{H}$ en $\mathcal{H}_F$.

Nótese que $g \in \mathcal{H}_F^{\perp}$ si, y solo si, $g(x_i) = \langle g, k_{x_i} \rangle = 0$ para todo $i = 1, \ldots, n$. Luego, cualquiera que sea $h \in \mathcal{H}$, se tiene

$$P_F(h)(x_i) = h(x_i), \quad i = 1, \ldots, n.$$

Proposición 3.2. *Sea $\{x_1, \ldots, x_n\} \subset X$ un conjunto de puntos distintos, y sea $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$. Si existe $g \in \mathcal{H}$ que interpola estos valores, entonces $P_F(g)$ es la única función de norma mínima que los interpola.*

Demostración. Si g_1 y g_2 son dos interpolantes, entonces $(g_1 - g_2) \in \mathcal{H}_F^{\perp}$. Por tanto, todas las soluciones del problema de interpolación son de la forma $g + h$, con $h \in \mathcal{H}_F^{\perp}$; en particular, $P_F(g)$ es un interpolante. Ahora, para cualquier $h \in \mathcal{H}_F^{\perp}$ se cumple

$$\|P_F(g)\| = \|P_F(g + h)\| \leq \|g + h\|,$$

así que $P_F(g)$ es el único interpolante de norma mínima. □

Seguidamente daremos condiciones necesarias y suficientes para la existencia de solución del problema interpolatorio.

Previamente, debemos hacer algunos comentarios sobre matrices y vectores. Si $A = (a_{i,j})_{i,j=1}^n$ es una matriz cuadrada de orden n y queremos escribir una ecuación vectorial-matricial $v = Aw$, entonces hemos de considerar v y w como vectores columna. Por razones tipográficas es más cómodo escribir vectores fila, así que típicamente representaremos un vector columna como $v = (v_1, \ldots, v_n)^t$, donde t denota transposición. La matriz A define una transformación lineal de $\mathbb{C}^n$ en $\mathbb{C}^n$. Para evitar confusiones, nos referiremos al espacio vectorial $\mathcal{N}(A) = \{w \in \mathbb{C}^n : Aw = 0\}$ como el espacio nulo de A, en vez de como su núcleo. El rango de A es el espacio vectorial $\mathcal{R}(A) = \{Aw : w \in \mathbb{C}^n\}$.

Proposición 3.3. *Sea X un conjunto, sea $\mathcal{H}$ un EHNR sobre X con núcleo K, y sea $\{x_1, \ldots, x_n\} \subset X$ un conjunto finito de puntos distintos. Si $w = (\alpha_1, \ldots, \alpha_n)^t$ es un vector del espacio nulo de la matriz $(K(x_i, x_j)) = (K(x_i, x_j))_{i,j=1}^n$, entonces la función $f = \sum_{j=1}^n \alpha_j k_{x_j}$ es idénticamente nula. En consecuencia, si $w_1 = (\alpha_1, \ldots, \alpha_n)^t$ y $w_2 = (\beta_1, \ldots, \beta_n)^t$ son dos vectores que satisfacen $(K(x_i, x_j)) w_1 = (K(x_i, x_j)) w_2$, entonces*

$$\sum_{j=1}^n \alpha_j k_{x_j}(y) = \sum_{j=1}^n \beta_j k_{x_j}(y)$$

para cada $y \in X$.

Demostración. Se tiene que $f = 0$ si, y solo si, $\|f\| = 0$. Sin más que calcular

$$\|f\|^2 = \sum_{i,j=1}^n \overline{\alpha_i}\alpha_j \left\langle k_{x_j}, k_{x_i}\right\rangle_{\mathcal{H}} = \sum_{i,j=1}^n \overline{\alpha_i}\alpha_j K(x_i, x_j) = \langle (K(x_i, x_j)) w, w\rangle_{\mathbb{C}^n} = 0,$$

se sigue el resultado.

Nótese que $w_1 - w_2$ está en el espacio nulo de la matriz $(K(x_i, x_j))_{i,j=1}^n$, de modo que la función $\sum_{j=1}^n (\alpha_j - \beta_j) k_{x_j}$ es idénticamente nula. □

Teorema 3.4 (Interpolación en un EHNR). *Sea $\mathcal{H}$ un EHNR sobre X con núcleo reproductor K, sea $F = \{x_1, \ldots, x_n\} \subset X$ un conjunto de n puntos distintos, y sea $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$. Existe $g \in \mathcal{H}$ que interpola estos valores si, y solo si, $v = (\lambda_1, \ldots, \lambda_n)^t$ está en el rango de la matriz $(K(x_i, x_j)) = (K(x_i, x_j))_{i,j=1}^n$. En tal caso, si elegimos cualquier vector $w = (\alpha_1, \ldots, \alpha_n)^t$ cuya imagen sea v, entonces $h = \sum_{j=1}^n \alpha_j k_{x_j}$ es el único interpolante de norma mínima en $\mathcal{H}$. Además, $\|h\|^2 = \langle v, w\rangle$.*

Demostración. Supongamos, en primer lugar, que existe $g \in \mathcal{H}$ tal que $g(x_i) = \lambda_i$, para todo $i = 1, \ldots, n$. Entonces, la solución de norma mínima es $P_F(g) = \sum_{j=1}^n \beta_j k_{x_j}$ para ciertos escalares $\beta_1, \ldots, \beta_n$. Como $\lambda_i = g(x_i) = P_F(g)(x_i) = \sum_{j=1}^n \beta_j k_{x_j}(x_i)$, $i = 1, \ldots, n$, resulta que $w_1 = (\beta_1, \ldots, \beta_n)^t$ es solución de $v = (K(x_i, x_j)) w$.

Recíprocamente, si $w = (\alpha_1, \ldots, \alpha_n)^t$ es cualquier solución de la ecuación vectorial-matricial $v = (K(x_i, x_j)) w$ y ponemos $h = \sum_{j=1}^n \alpha_j k_{x_j}$, entonces h será un interpolante.

Como $w - w_1$ está en el núcleo de la matriz $(K(x_i, x_j))$, en virtud de la Proposición 3.3, $P_F(g)$ y h son la misma función; consecuentemente, h es el interpolante de norma mínima. Por último,

$$\|h\|^2 = \sum_{i,j=1}^n \overline{\alpha_i}\alpha_j K(x_i, x_j) = \langle (K(x_i, x_j)) w, w\rangle = \langle v, w\rangle.$$

□

Corolario 3.5. *Sea $\mathcal{H}$ un EHNR sobre X con núcleo reproductor K, y sea $F = \{x_1, \ldots, x_n\} \subset X$ un conjunto de n puntos distintos. Si la matriz $(K(x_i, x_j)) = (K(x_i, x_j))_{i,j=1}^n$ es inversible entonces, dados $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$, existe una función que interpola estos valores, y el único interpolante de norma mínima es $g = \sum_{j=1}^n \alpha_j k_{x_j}$, donde $w = (\alpha_1, \ldots, \alpha_n)^t$ viene dado por $w = (K(x_i, x_j))^{-1} v$, con $v = (\lambda_1, \ldots, \lambda_n)^t$.*

3.2. Núcleos estrictamente positivos

Definición 3.6. *Dado un conjunto X, una función núcleo $K : X \times X \to \mathbb{C}$ se dice estrictamente positiva si para todo n y todo conjunto de puntos distintos $\{x_1, \ldots, x_n\} \subset X$, la matriz*

$$(K(x_i, x_j))_{i,j=1}^n$$

es estrictamente definida positiva, es decir,

$$\sum_{i,j=1}^n \overline{\alpha_i} \alpha_j K(x_i, x_j) > 0$$

siempre que $\{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{C}$ no son todos nulos.

Sabemos por álgebra lineal que una matriz es estrictamente definida positiva si, y sólo si, es (semidefinida) positiva e inversible. Muchos núcleos que se presentan en la práctica son estrictamente positivos, lo que justifica que recopilemos aquí algunas de las principales propiedades de los correspondientes EHNRs. Como veremos, la positividad estricta está íntimamente ligada a la interpolación y a la separación de puntos.

Teorema 3.7. *Sea X un conjunto, y sea $K : X \times X \to \mathbb{C}$ un núcleo. Equivalen los siguientes enunciados:*

(i) K es estrictamente positivo.

(ii) Para cualquier n y cualquier conjunto de puntos distintos $\{x_1, \ldots, x_n\}$, las funciones núcleo $k_{x_1}, \ldots, k_{x_n}$ son linealmente independientes.

(iii) Para cualquier n, cualquier conjunto de puntos distintos $\{x_1, \ldots, x_n\}$, y cualquier conjunto de escalares $\{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{C}$ no todos nulos, existe $f \in \mathcal{H}(K)$ tal que

$$\alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \neq 0.$$

(iv) Para cualquier n y cualquier conjunto de puntos distintos $\{x_1, \ldots, x_n\}$, existen funciones $g_1, \ldots, g_n \in \mathcal{H}(K)$, satisfaciendo:

$$g_i(x_j) = \begin{cases} 1, & si\ i = j \\ 0, & si\ i \neq j \end{cases}, \quad i, j = 1, \ldots, n.$$

Demostración. Puesto que

$$\sum_{i,j=1}^{n} \overline{\alpha_i}\alpha_j K(x_i, x_j) = \left\| \sum_{j=1}^{n} \alpha_j k_{x_j} \right\|^2,$$

(i) y (ii) son equivalentes.

Por otra parte, $\overline{\alpha_1}k_{x_1} + \cdots + \overline{\alpha_n}k_{x_n} = 0$ si, y solo si, $\langle f, \sum_{i=1}^{n} \overline{\alpha_i}k_{x_i}\rangle = 0$ para toda $f \in \mathcal{H}$, lo que ocurre si, y solo si, $\alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) = 0$ para toda $f \in \mathcal{H}$. En consecuencia, (ii) y (iii) son equivalentes.

Para ver que (iv) implica (iii), basta suponer que $\alpha_i \neq 0$ y elegir $f = g_i$.

Por último, si se cumple (i) entonces, para cada i, la aplicación del Corolario 3.5 al conjunto de números $\lambda_j = 0$, $j \neq i$, y $\lambda_i = 1$, proporciona la función g_i en (iv). $\square$

Definición 3.8. *Un conjunto de funciones que verifica el apartado (iv) del Teorema 3.7 se denomina partición de la unidad para $\{x_1, \ldots, x_n\}$.*

Si $g_1, \ldots, g_n$ es una partición de la unidad para $x_1, \ldots, x_n$, se obtiene $f \in \mathcal{H}(K)$ con $f(x_i) = \lambda_i$ sin más que tomar $f = \lambda_1 g_1 + \cdots + \lambda_n g_n$.

Definición 3.9. *Un EHNR $\mathcal{H}$ sobre X que satisface cualquiera de las condiciones equivalentes del Teorema 3.7 se dice totalmente interpolante.*

Nos disponemos a describir la forma de calcular una partición de la unidad, cuando existe. Supongamos que $F = \{x_1, \ldots, x_n\}$ y que $P = (K(x_i, x_j))_{i,j=1}^{n}$ es inversible, como en el Corolario 3.5, y escribamos $P^{-1} = (b_{i,j})_{i,j=1}^{n} = B$. Sean $e_j, j = 1, \ldots, n$, los vectores de la base canónica de $\mathbb{C}^n$. Las columnas de B son los únicos vectores $w_j, j = 1, \ldots, n$, solución de la ecuación $e_j = Pw_j$. Poniendo

$$g_j = \sum_{i=1}^{n} b_{i,j} k_{x_i}$$

encontramos que $g_j(x_i) = \delta_{i,j}$, donde $\delta_{i,j}$ denota la delta de Kronecker. Así pues, estas funciones constituyen una partición de la unidad para F. Además, si

$$g = \sum_{j=1}^{n} \lambda_j g_j,$$

entonces g es la única función en $\mathcal{H}_F$ cumpliendo que $g(x_i) = \lambda_i, i = 1, \ldots n$, de modo que g es también el interpolante de norma mínima. Esta partición de la unidad, que permite obtener fácilmente el interpolante de norma mínima para el conjunto en cuestión, recibe el nombre de partición de la unidad canónica de F.

3.3. Mejor aproximación por mínimos cuadrados

Como vimos en el Teorema 3.4, si $\mathcal{H}$ es un EHNR sobre X, si $F = \{x_1, \ldots, x_n\}$ es un conjunto finito de puntos distintos y si $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$, pero la matriz $(K(x_i, x_j))_{i,j=1}^n$ no es inversible, entonces podría no existir ninguna $f \in \mathcal{H}$ tal que $f(x_i) = \lambda_i$, para $i = 1, \ldots, n$. En estos casos interesa encontrar las funciones que minimizan el error cuadrático

$$J(f) = \sum_{i=1}^{n} |f(x_i) - \lambda_i|^2$$

y luego hallar, entre ellas, la de norma mínima. El Teorema 3.10 garantiza la existencia y unicidad de esta función, que denominaremos mejor aproximante por mínimos cuadrados, y proporciona una fórmula para calcularla.

Teorema 3.10. *Sea $\mathcal{H}$ un EHNR sobre X con núcleo K, sea $\{x_1, \ldots, x_n\} \subset X$ un conjunto finito de puntos distintos, sea $v = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{C}^n$, y sea $Q = (K(x_i, x_j))_{i,j=1}^n$. Entonces existe un vector $w = (\alpha_1, \ldots, \alpha_n)^t$ tal que $(v - Qw) \in \mathcal{N}(Q)$. La función*

$$g = \alpha_1 k_{x_1} + \cdots + \alpha_n k_{x_n}$$

minimiza el error cuadrático y es la única función de norma mínima entre todas las funciones de $\mathcal{H}$ con esta propiedad.

Demostración. En virtud del Teorema 3.4, dada cualquier función $f \in \mathcal{H}$ existe un vector $w \in \mathbb{C}^n$ tal que $Qw = (f(x_1), \ldots, f(x_n))^t$. Por lo tanto, se tiene que

$$J(f) = \sum_{i=1}^{n} |f(x_i) - \lambda_i|^2 = \|Qw - v\|^2.$$

Esta norma es minimizada por cualquier vector $w = (\alpha_1, \ldots, \alpha_n)^t$ tal que $Qw = P_{\mathcal{R}(Q)}(v) = v_1$, donde $\mathcal{R}(Q)$ denota el rango de la matriz Q y $P_{\mathcal{R}(Q)}$ denota la proyección ortogonal sobre $\mathcal{R}(Q)$.

Si $w' = (\alpha'_1, \ldots, \alpha'_n)^t$ es otro vector que resuelve $v_1 = Qw'$, tendremos que

$$\sum_{i=1}^{n} \alpha_i k_{x_i} = \sum_{i=1}^{n} \alpha'_i k_{x_i},$$

ya que $w - w' \in \mathcal{N}(Q)$. Al proyectar una función f sobre el subespacio generado por las funciones núcleo $k_{x_1}, \ldots, k_{x_n}$, disminuye la norma y no cambia el valor de f en los puntos $x_1, \ldots, x_n$, por lo que g es el único minimizante de J de norma mínima. $\square$

3.4. Los elementos de $\mathcal{H}(K)$

Aplicaremos ahora la interpolación a la tarea de dar una solución general del problema de reconstrucción para núcleos. Concretamente, el Teorema 3.13 caracteriza las funciones $f : X \to \mathbb{C}$ que pertenecen al EHNR $\mathcal{H}(K)$ determinado por un núcleo K. Este teorema se expresa mejor en términos de conjuntos dirigidos, redes y convergencia de redes, conceptos que procedemos a introducir.

Formalmente, un conjunto dirigido es cualquier conjunto parcialmente ordenado con la propiedad de que dados dos elementos cualesquiera del conjunto, siempre existe, al menos, otro elemento del conjunto que es mayor o igual que ambos. Para lo que sigue solo necesitaremos manejar un único conjunto dirigido. Dado un conjunto X, denotamos por $\mathcal{F}_X$ la colección de todos los subconjuntos finitos de X. Entonces $\mathcal{F}_X$ es un conjunto dirigido respecto al orden parcial determinado por la inclusión. En efecto, la relación $F_1 \leq F_2$ si $F_1 \subset F_2$ define un orden parcial en $\mathcal{F}_X$. Además, dados dos conjuntos finitos cualesquiera F_1, F_2, siempre existe un tercer conjunto finito $G = F_1 \cup F_2$ que es mayor que ambos.

El concepto de red generaliza el de sucesión, de tal manera que el conjunto de índices pasa a ser un conjunto dirigido arbitrario, en vez de $\mathbb{N}$. Una red en un espacio de Hilbert $\mathcal{H}$ es una colección de vectores $\{g_F\}_{F\in\mathcal{F}} \subset \mathcal{H}$, donde $(\mathcal{F}, \leq)$ es un conjunto dirigido. La convergencia de redes se define por analogía con la convergencia de sucesiones. Se dice que la red $\{g_F\}_{F\in\mathcal{F}}$ converge a $g \in \mathcal{H}$ si dado $\varepsilon > 0$, existe $F_0 \in \mathcal{F}$ tal que $F_0 \leq F$ implica $\|g - g_F\| < \varepsilon$. Similarmente, una red $\{g_F\}_{F\in\mathcal{F}}$ es de Cauchy en $\mathcal{H}$ si dado $\varepsilon > 0$, existe $F_0 \in \mathcal{F}$ tal que $F_0 \leq F$ y $F_0 \leq G$ implica $\|g_F - g_G\| < \varepsilon$. Las redes convergentes son de Cauchy; un espacio métrico es completo si, y solo si, toda red de Cauchy es convergente.

Proposición 3.11. *Sea $\mathcal{H}$ un EHNR sobre el conjunto X, sea $g \in \mathcal{H}$ y, para cada conjunto finito $F \subset X$, sea $g_F = P_F(g)$, donde P_F denota la proyección ortogonal de $\mathcal{H}$ sobre $\mathcal{H}_F$. Entonces, la red $\{g_F\}_{F\in\mathcal{F}_X}$ converge en norma a g.*

Demostración. Sea $K(x, y)$ el núcleo reproductor de $\mathcal{H}$, y sea $k_y(\cdot) = K(\cdot, y)$. Dado $\varepsilon > 0$, por la Proposición 2.1 existe una colección finita de puntos $F_0 = \{x_1, \ldots, x_n\}$, y de escalares $\{\alpha_1, \ldots, \alpha_n\}$, tales que $\|g - \sum_{i=1}^n \alpha_i k_{x_i}\| < \varepsilon$. Puesto que g_{F_0} es el punto de $\mathcal{H}_{F_0}$ más próximo a g, se tiene que $\|g - g_{F_0}\| < \varepsilon$.

Sea ahora F cualquier conjunto finito, con $F_0 \subset F$. Entonces $\mathcal{H}_{F_0} \subset \mathcal{H}_F$, y como g_F es el punto de $\mathcal{H}_F$ más próximo a g, mientras que $g_{F_0} \in \mathcal{H}_F$, necesariamente $\|g - g_F\| \leq \|g - g_{F_0}\| < \varepsilon$ para todo $F_0 \subset F$, lo que completa la demostración. □

Antes de abordar el Teorema 3.13 necesitaremos un resultado sobre matrices. Recordemos que si A y B son matrices autoadjuntas, escribimos $A \leq B$ ó $B \geq A$ para indicar que $B - A \geq 0$.

Proposición 3.12. *Sea $P \geq 0$ una matriz cuadrada de orden n, y sea $x = (x_1, \ldots, x_n)^t$ un vector en $\mathbb{C}^n$. Si $xx^* = (x_i \overline{x_j})_{i,j=1}^n \leq cP$ para algún escalar $c > 0$, entonces $x \in \mathcal{R}(P)$. Además, si y es cualquier vector tal que $x = Py$, entonces $0 \leq \langle x, y \rangle \leq c$.*

Demostración. Para cualquier matriz A, se tiene que $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp$. Como $P = P^*$, necesariamente $\mathcal{N}(P) = \mathcal{R}(P)^\perp$. Así, podemos escribir $x = v + w$, con $v \in \mathcal{R}(P)$ y $w \in \mathcal{N}(P)$.

Ahora, $\langle x, w \rangle = \langle w, w \rangle$ y, por tanto,

$$\begin{aligned} \|w\|^4 = \langle w, w \rangle \langle w, w \rangle = \langle w, x \rangle \langle x, w \rangle &= \sum_{i,j=1}^n \overline{x_j} w_j x_i \overline{w_i} \\ &= \langle (x_i \overline{x_j})\, w, w \rangle \leq \langle cPw, w \rangle = 0, \end{aligned}$$

porque $Pw = 0$. Esta estimación obliga a que $w = 0$, de donde $x = v \in \mathcal{R}(P)$.

Si escribimos $x = Py$, entonces $\langle x, y \rangle = \langle Py, y \rangle \geq 0$. Además, como anteriormente, se tiene que

$$\langle x, y \rangle^2 = \langle y, x \rangle \langle x, y \rangle = \langle (x_i \overline{x_j})\, y, y \rangle \leq \langle cPy, y \rangle = c \langle x, y \rangle.$$

Sin más que cancelar el factor $\langle x, y \rangle$ en esta desigualdad, se obtiene el resultado esperado. □

Quedamos ya en disposición de probar el teorema que caracteriza las funciones pertenecientes a un EHNR en términos del núcleo reproductor.

Teorema 3.13. *Sea $\mathcal{H}$ un EHNR sobre X con núcleo reproductor K, y sea $f : X \to \mathbb{C}$ una función. Equivalen:*

(i) $f \in \mathcal{H}$.
(ii) Para cierta constante $c \geq 0$ y todo conjunto finito $F = \{x_1, \ldots x_n\} \subset X$, existe una función $h \in \mathcal{H}$ con $\|h\| \leq c$ y $f(x_i) = h(x_i)$, $i = 1, \ldots, n$.
(iii) Existe una constante $c \geq 0$ tal que la función $c^2 K(x, y) - f(x)\overline{f(y)}$ es una función núcleo.

Además, si $f \in \mathcal{H}$ entonces $\|f\|$ es la menor constante c que satisface las desigualdades en (ii) y (iii).

Demostración. (i) $\Rightarrow$ (iii) Sea $F = \{x_1, \ldots, x_n\} \subset X$, sean $\alpha_1, \ldots, \alpha_n$ escalares, y sea $g = \sum_{j=1}^n \alpha_j k_{x_j}$. Entonces,

$$\begin{aligned} \sum_{i,j=1}^n \overline{\alpha_i} \alpha_j f(x_i)\, \overline{f(x_j)} = \left| \sum_{i=1}^n \overline{\alpha_i} f(x_i) \right|^2 &= |\langle f, g \rangle|^2 \leq \|f\|^2 \|g\|^2 \\ &= \|f\|^2 \sum_{i,j=1}^n \overline{\alpha_i} \alpha_j K(x_i, x_j)\,. \end{aligned}$$

Puesto que los escalares fueron elegidos arbitrariamente, resulta que

$$\left(f(x_i)\overline{f(x_j)}\right)_{i,j=1}^{n} \leq \|f\|^2 \left(K(x_i,x_j)\right)_{i,j=1}^{n},$$

de donde sigue (iii), con $c = \|f\|$.
$(iii) \Rightarrow (ii)$ Sea $F = \{x_1, \ldots, x_n\} \subset X$ un conjunto finito. Usamos la Proposición 3.12 para deducir que el vector v de componentes $\lambda_i = f(x_i)$, $i = 1, \ldots, n$, está en el rango de $(K(x_i,x_j))_{i,j=1}^{n}$. Aplicamos entonces el teorema de interpolación (Teorema 3.4) para encontrar $h = \sum_{i=1}^{n} \alpha_i k_{x_i} \in \mathcal{H}_F$ con $h(x_i) = f(x_i)$, $i = 1, \ldots, n$. Sea w el vector de componentes α_i, $i = 1, \ldots, n$. Aplicando de nuevo la Proposición 3.12, se concluye que $\|h\|^2 = \langle v, w\rangle \leq c^2$.
$(ii) \Rightarrow (i)$ Por hipótesis, dado cualquier conjunto finito F existe $h_F \in \mathcal{H}$ tal que $\|h_F\| \leq c$ y $h_F(x) = f(x)$ si $x \in F$. Pongamos $g_F = P_F(h_F)$. Entonces $g_F(x) = h_F(x) = f(x)$ para todo $x \in F$, y

$$\|g_F\| \leq \|h_F\| \leq c.$$

Afirmamos que la red $\{g_F\}_{F\in\mathcal{F}_X}$ es de Cauchy y converge a f.

Para ver que la red es de Cauchy, sea $M = \sup \|g_F\| \leq c$, y fijemos $\varepsilon > 0$. Elegimos un conjunto F_0 tal que $M - \varepsilon^2 < \|g_{F_0}\|$. Dado $F \in \mathcal{F}_X$, con $F_0 \subset F$, se tiene que $P_{F_0}(g_F) = g_{F_0}$ y, por lo tanto,

$$\langle g_F - g_{F_0}, g_{F_0}\rangle = 0.$$

Así,

$$\|g_F\|^2 = \|g_{F_0}\|^2 + \|g_F - g_{F_0}\|^2,$$

de modo que

$$M - \varepsilon^2 \leq \|g_{F_0}\| \leq \|g_F\| \leq M.$$

Consecuentemente, $0 \leq \|g_F\| - \|g_{F_0}\| \leq \varepsilon^2$, y podemos escribir

$$\begin{aligned}\|g_F - g_{F_0}\|^2 &= \|g_F\|^2 - \|g_{F_0}\|^2 \\ &= (\|g_F\| + \|g_{F_0}\|)(\|g_F\| - \|g_{F_0}\|) \leq 2M\varepsilon^2.\end{aligned}$$

Sigue que $\|g_F - g_{F_0}\| < \sqrt{2M}\varepsilon$. De este modo, $\|g_{F_1} - g_{F_2}\| < 2\sqrt{2M}\varepsilon$ para cualquier $F_1, F_2 \in \mathcal{F}_X$ con $F_0 \subset F_1, F_0 \subset F_2$, probando que la red es de Cauchy.

Existe entonces una función $g \in \mathcal{H}$ que es el límite de esta red, con $\|g\| \leq M \leq c$. Pero cualquier red que converge en norma también converge puntualmente (cf. Lema 2.2), por lo que $g(x) = f(x)$ para todo $x \in X$. Queda así demostrado que (ii) implica (i).

Finalmente, como $f \in \mathcal{H}$, la constante $c = \|f\|$ cumple las condiciones (ii) y (iii); luego, el menor c que las satisface es inferior a $\|f\|$. Recíprocamente, al probar que (iii) implica (ii) se vio que cualquier c que satisface (iii), también satisface (ii); pero en la demostración de que (ii) implica (i) encontramos que $\|f\| \leq c$. Se concluye que cualquier c que verifique las desigualdades en (ii) o (iii) debe ser mayor que $\|f\|$. $\square$

El resultado siguiente ilustra alguna de las consecuencias inesperadas del Teorema 3.13.

Corolario 3.14. *Sea $f : \mathbb{D} \to \mathbb{C}$ una función. Entonces, f es analítica en $\mathbb{D}$ y admite un desarrollo en serie de potencias de cuadrado absolutamente sumable si, y sólo si, existe $c > 0$ tal que*

$$K(z,w) = \frac{c^2}{1 - z\overline{w}} - f(z)\overline{f(w)}$$

es una función núcleo sobre $\mathbb{D}$.

Lo llamativo de este resultado es que la analiticidad de f se siga de la condición sobre la función núcleo, es decir, del hecho de que ciertas matrices sean positivas.

4

Núcleos condicionalmente definidos positivos y espacios de Pontryagin

El problema de reconstruir una función desconocida a partir de sus valores en datos dispersos es un problema clásico en teoría de la aproximación, estadística y aprendizaje automático. Un método muy potente para resolver este problema consiste en asumir que los aproximantes son combinaciones lineales de trasladadas de una cierta función definida positiva —esto es, una función $f : \mathbb{R} \to \mathbb{C}$ tal que la matriz $(f(x_i - x_j))_{i,j=1}^n$ es positiva para cualesquiera $x_1, \ldots, x_n \in \mathbb{R}$— o, más en general, de una función núcleo (Definición 2.14). La ventaja de este enfoque es que, como sabemos, cada función núcleo tiene asociado un EHNR de modo que la interpolación equivale a la interpolación de norma mínima en ese espacio (Teorema 3.4).

En general, el enfoque clásico carece de algunas propiedades de aproximación deseables, como la de reproducir polinomios de un grado determinado. Sin embargo, el método se puede generalizar, reemplazando la definición positiva por la noción, más débil, de definición positiva condicional.

En este capítulo nos proponemos ilustrar la conexión entre los núcleos condicionalmente definidos positivos y el análogo en este contexto de los EHNRs, los llamados espacios de Pontryagin con núcleo reproductor (en adelante, EPNRs). El concepto de EPNR es bien conocido en teoría de operadores y análisis armónico abstracto, pero quizá no tanto en teoría de la aproximación o estadística. Nos enfocaremos en los aspectos del método interpolatorio que son propios del núcleo reproductor, sin formular ninguna hipótesis adicional sobre el espacio del que provienen los datos.

El capítulo se estructura como sigue. Comenzaremos, en la siguiente sección, con las definiciones necesarias, para luego considerar los EPNRs. A continuación mostraremos la conexión entre los EPNRs y ciertos EHNRs. Tras focalizar nuestra atención en el problema de aproximación con datos dispersos, para el que establecemos una estimación local del error dentro del marco general, concluiremos con un ejemplo.

4.1. Preliminares

Definición 4.1. *Sea X un conjunto no vacío. Diremos que una aplicación $\Phi : X \times X \to \mathbb{C}$ es un núcleo hermitiano o hermítico si $\Phi(y,x) = \overline{\Phi(x,y)}$ para cualesquiera $x, y \in X$, donde, como habitualmente, $\overline{\Phi(x,y)}$ denota conjugación compleja.*

Definición 4.2. *Supongamos que $\kappa \in \mathbb{Z}^+$. Se dice que un núcleo hermítico $\Phi : X \times X \to \mathbb{C}$ es condicionalmente definido positivo de orden κ, si existe un espacio κ-dimensional $\mathcal{U}$ de funciones complejas definidas sobre X, tales que*

$$\sum_{i,j=1}^{n} c_i \overline{c_j} \Phi\left(x_i, x_j\right) \geq 0 \tag{4.1}$$

para cualesquiera $n \in \mathbb{N}, x_1, \ldots, x_n \in X$, y $c_1, \ldots, c_n \in \mathbb{C}$ satisfaciendo

$$\sum_{i=1}^{n} c_i u\left(x_i\right) = 0, \quad u \in \mathcal{U}. \tag{4.2}$$

El conjunto de todos los núcleos condicionalmente definidos positivos de orden κ será denotado por CDP_κ.

Nótese que todo núcleo condicionalmente definido positivo de orden κ es condicionalmente definido positivo de cualquier orden superior. Los núcleos condicionalmente definidos positivos de orden 0 son los llamados núcleos definidos positivos o funciones núcleo (Definición 2.14). Simbólicamente, $DP = CDP_0$.

Definición 4.3. *Para $x_1, \ldots, x_n \in X$, llamaremos a*

$$G = G\left(x_1, \ldots, x_n\right) = \left(\Phi\left(x_i, x_j\right)\right)_{i,j=1}^{n}$$

la matriz gramiana de Φ.

Los núcleos DP dan lugar a matrices positivas, que tienen autovalores no negativos. La matriz gramiana G de un núcleo CDP puede tener autovalores negativos. Por tanto, nos interesará conocer el número de autovalores negativos de G. Escribamos (4.1) y (4.2) en notación vectorial-matricial. Si $u \in \mathcal{U}$, denotamos por $\mathbf{u}$ el vector de componentes $u_i = \overline{u\left(x_i\right)}$, $i = 1, \ldots, n$. Entonces

$$\mathbf{c}^H \overline{G} \mathbf{c} \geq 0 \tag{4.3}$$

para cualquier $\mathbf{c} \in \mathbb{C}^n$ verificando

$$\mathbf{u}^H \mathbf{c} = 0, \quad u \in \mathcal{U}, \tag{4.4}$$

donde $\mathbf{c}^H$ denota la transpuesta conjugada compleja de $\mathbf{c} \in \mathbb{C}^n$, y la conjugación $\overline{G}$ opera sobre cada entrada. La relación de ortogonalidad (4.4) conduce a una descomposición del espacio vectorial $\mathbb{C}^n$:

$$\mathbb{C}^n = \left\{\mathbf{c} \mid \mathbf{u}^H\mathbf{c} = 0,\ u \in \mathcal{U}\right\} \oplus \{\mathbf{u} \mid u \in \mathcal{U}\} = P \oplus R.$$

Ya que $\overline{G}$ es semidefinida positiva sobre P, la matriz G puede tener, a lo sumo,

$$\dim R \leq \dim \mathcal{U} = \kappa$$

autovalores negativos.

Como sabemos, las funciones núcleo —también llamadas núcleos definidos positivos— están en correspondiencia biyectiva con los núcleos reproductores de EHNR (Proposición 2.15 y Teorema 2.16). Existe una conexión similar para la noción más general de núcleos condicionalmente definidos positivos. Antes de formular el resultado análogo, necesitamos otra definición.

Definición 4.4. *Supóngase que $\kappa \in \mathbb{Z}^+$. Se dice que un núcleo hermítico $\Phi : X \times X \to \mathbb{C}$ tiene κ cuadrados negativos, si para cada elección de $n \in \mathbb{N}$, $x_1, \ldots, x_n \in X$, la matriz gramiana de Φ tiene, a lo sumo, κ autovalores negativos, y existen $y_1, \ldots, y_n \in X$ tales que $G\,(y_1, \ldots, y_n)$ tiene exactamente κ autovalores negativos, contados según sus multiplicidades.*

El conjunto de los núcleos hermíticos que tienen a lo sumo κ cuadrados negativos será denotado por N_κ.

Las observaciones anteriores conducen inmediatamente a la siguiente conexión entre los núcleos condicionalmente definidos positivos y los núcleos con un número finito de cuadrados negativos.

Proposición 4.5. *Todo núcleo condicionalmente definido positivo de orden κ tiene, a lo sumo, κ cuadrados negativos.*

4.2. Espacios de Pontryagin con núcleo reproductor

Antes de proseguir con los espacios con núcleo reproductor, resumiremos algunos hechos básicos relativos a los espacios de Pontryagin. Para un estudio más exhaustivo remitimos a [7, 18].

Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{C}$ con un producto interior indefinido $(\cdot;\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$, es decir, tal que

(*i*) $(\alpha u + \beta v; w) = \alpha(u; w) + \beta(v; w)$, y
(*ii*) $(w; v) = \overline{(v; w)}$

para cualesquiera $u, v, w \in \mathcal{V}$, $\alpha, \beta \in \mathbb{C}$.

Un vector $v \in \mathcal{V}$ se dice negativo si $(v; v) < 0$. Un subespacio es negativo si todo elemento no nulo es negativo. Un subespacio negativo se dice maximal si no está propiamente contenido en ningún otro subespacio negativo. El espacio $(\mathcal{V}, (\cdot;\cdot))$ se dice no degenerado si no existe ningún vector no trivial $v \in \mathcal{V}$ tal que $(v; w) = 0$ para todo $w \in \mathcal{V}$.

Supóngase que $(\mathcal{V}, (\cdot;\cdot))$ es no degenerado y contiene un subespacio negativo maximal de dimensión κ, $\kappa \in \mathbb{Z}^+$. Se demuestra (cf. [7, IX.1]) que todos los subespacios negativos maximales $\mathcal{N}$ de $\mathcal{V}$ tienen la misma dimensión κ, y que existe una descomposición

$$\mathcal{V} = \mathcal{N} \oplus \mathcal{N}^\perp, \quad v = v^- + v^+,$$

donde $\mathcal{N}^\perp = \{v \in \mathcal{V} \mid (v; w) = 0,\ w \in \mathcal{N}\}$ es un espacio (pre-)Hilbert respecto de $(\cdot;\cdot)$. Cualquier descomposición de este tipo define un producto interior

$$[u; v]_{\mathcal{N}} = \big(u^+; v^+\big) - \big(u^-; v^-\big)$$

(y una norma $\|\cdot\|_{\mathcal{N}}$) sobre $\mathcal{V}$.

En general, los subespacios negativos maximales no están unívocamente determinados, pero, por [7, V.1], todas las normas inducidas $\|\cdot\|_{\mathcal{N}}$ son equivalentes.

Definición 4.6. *Se dice que $(\mathcal{V}, (\cdot;\cdot))$ es un espacio de Pontryagin de índice κ si es completo con respecto a una de las normas $\|\cdot\|_{\mathcal{N}}$ (y, por lo tanto, a todas ellas).*

Observación 4.7. *Si en la definición de espacio de Pontryagin se omite la restricción de finitud sobre la dimensión del subespacio negativo maximal, se obtienen los llamados espacios de Krein, ampliamente estudiados en teoría de operadores [10].*

Los espacios de Pontryagin de índice 0 son espacios de Hilbert, y recíprocamente; es decir, el concepto de espacio de Pontryagin es una generalización del concepto de espacio de Hilbert. Interesa, por tanto, extender a espacios de Pontryagin la rica teoría existente sobre EHNRs.

Definición 4.8. *Sea Π un espacio de Pontryagin de funciones sobre X. Un núcleo hermítico $\Phi : X \times X \to \mathbb{C}$ es un núcleo reproductor para Π, si*

(i) $\Phi_x = \Phi(\cdot, x) \in \Pi$ *para todo* $x \in X$, *y*
(ii) $g(x) = (g; \Phi_x)$ *para cualesquiera* $g \in \Pi$, $x \in X$.

La segunda condición es la llamada propiedad reproductora del núcleo Φ.

El papel que juegan las funciones núcleo en el contexto de los EHNRs es desempeñado por los núcleos con un número finito de cuadrados negativos en el contexto de los EPNRs.

Teorema 4.9. *Supóngase que* $\kappa \in \mathbb{Z}^+$.

(i) Sea $(\Pi, (\cdot;\cdot))$ *un espacio de Pontryagin de índice* κ*, con núcleo reproductor* Φ*. Entonces* Φ *tiene* κ *cuadrados negativos.*

(ii) Para todo núcleo hermítico $\Phi : X \times X \to \mathbb{C}$ *con* κ *cuadrados negativos, existe un (único) espacio de Pontryagin* Π *de índice* κ*, cuyo núcleo reproductor es* Φ*, tal que el conjunto* $\{\Phi_x \mid x \in X\}$ *es denso en* Π*.*

Demostración. Véase [20]. □

De la Proposición 4.5 se deduce inmediatamente:

Teorema 4.10. *Supóngase que* $\kappa \in \mathbb{Z}^+$*. Para todo núcleo condicionalmente definido positivo* Φ *de orden* κ*, existe un espacio de Pontryagin* Π *de índice* λ*,* $0 \le \lambda \le \kappa$*, tal que* Φ *es el núcleo reproductor de* Π*.*

El recíproco también vale, es decir, los núcleos condicionalmente definidos positivos pueden ser identificados con núcleos reproductores de EPNRs.

Teorema 4.11. *Supongamos que* $\kappa \in \mathbb{Z}^+$*. Todo núcleo hermítico con* κ *cuadrados negativos es condicionalmente definido positivo de orden* κ*.*

Demostración. Supongamos que $\Phi : X \times X \to \mathbb{C}$ tiene κ cuadrados negativos y denotemos por Π el espacio de Pontryagin asociado, proporcionado por el Teorema 4.9. Elijamos un subespacio negativo maximal $\mathcal{N}$ de Π. Dados $x_1, \ldots, x_n \in X$, $c_1, \ldots, c_n \in \mathbb{C}$ satisfaciendo (4.2) para cada $u \in \mathcal{N}$, la propiedad reproductora de Φ conduce a

$$0 = \sum_{i=1}^{n} c_i u(x_i) = \sum_{i=1}^{n} c_i (u; \Phi_{x_i}) = \Big(u; \sum_{i=1}^{n} \overline{c_i} \Phi_{x_i}\Big), \quad u \in \mathcal{N}.$$

Luego, $h = \sum_{i=1}^{n} \overline{c_i} \Phi_{x_i} \in \mathcal{N}^{\perp}$. Como $\mathcal{N}^{\perp}$ es un espacio de Hilbert,

$$0 \le (h;h) = \Big(\sum_{j=1}^{n} \overline{c_j} \Phi_{x_j}; \sum_{i=1}^{n} \overline{c_i} \Phi_{x_i}\Big) = \sum_{i,j=1}^{n} c_i \overline{c_j} \Phi(x_i, x_j).$$

Por definición, $\dim \mathcal{N} = \kappa$. Luego, Φ es condicionalmente definida positiva de orden κ. □

Combinando la Proposición 4.5 con el Teorema 4.11, podemos identificar los núcleos condicionalmente definidos positivos con los núcleos que tienen un número finito de cuadrados negativos.

Teorema 4.12. *Supongamos que $\kappa \in \mathbb{Z}^+$. El conjunto de los núcleos condicionalmente definidos positivos de orden κ, y el conjunto de los núcleos con, a lo sumo, κ cuadrados negativos, coinciden; simbólicamente,*

$$CDP_\kappa = N_\kappa.$$

Corolario 4.13. *Supongamos que $\kappa \in \mathbb{N}$ y que Φ es condicionalmente definido positivo de orden κ, pero no de orden $\kappa - 1$. Entonces, Φ tiene κ cuadrados negativos.*

4.3. Conexión con los EHNRs

Como se mencionó anteriormente, todo subespacio negativo maximal $\mathcal{N}$ de un espacio de Pontryagin Π da lugar a un producto interior $[\cdot;\cdot]_{\mathcal{N}}$ sobre Π, tal que $(\Pi, [\cdot;\cdot]_{\mathcal{N}})$ es un espacio de Hilbert.

Consideremos un núcleo condicionalmente definido positivo $\Phi : X \times X \to \mathbb{C}$ de orden κ, y sea Π el espacio de Pontryagin asociado de índice λ, $\lambda \leq \kappa$ (Teorema 4.10). Fijemos un subespacio negativo maximal $\mathcal{N}$ de Π. Nótese que $(\Pi, [\cdot;\cdot]_{\mathcal{N}})$ es un EHNR, porque los funcionales evaluación en puntos de X son continuos sobre este espacio. En efecto, si E_x denota el funcional evaluación en $x \in X$ entonces, usando la propiedad reproductora de Φ, resulta

$$\begin{aligned} E_x(f) &= E_x(f^+) + E_x(f^-) = (f^+;\Phi_x) + (f^-;\Phi_x) \\ &= (f^+;\Phi_x^+) + (f^-;\Phi_x^-) = [f^+;\Phi_x]_{\mathcal{N}} - [f^-;\Phi_x]_{\mathcal{N}}. \end{aligned}$$

La desigualdad de Cauchy-Schwarz y el teorema de Pitágoras implican

$$\begin{aligned} |E_x(f)| &\leq \left|[f^+;\Phi_x]_{\mathcal{N}}\right| + \left|[f^-;\Phi_x]_{\mathcal{N}}\right| \\ &\leq \left(\|f^+\|_{\mathcal{N}} + \|f^-\|_{\mathcal{N}}\right) \|\Phi_x\|_{\mathcal{N}} \leq 2\|f\|_{\mathcal{N}}\|\Phi_x\|_{\mathcal{N}}, \end{aligned}$$

probando la continuidad de E_x.

Queremos dar una representación del núcleo reproductor Ψ de $(\Pi, [\cdot;\cdot]_{\mathcal{N}})$ en términos de Φ. Puesto que Π es suma directa de dos espacios de Hilbert, Ψ se puede escribir (Teorema 2.7) como suma de los núcleos reproductores de $(\mathcal{N}, [\cdot;\cdot]_{\mathcal{N}})$ y $\left(\mathcal{N}^\perp, [\cdot;\cdot]_{\mathcal{N}}\right)$. Denotemos el núcleo reproductor de $\mathcal{N}^\perp$ por h. Para $g \in \mathcal{N}^\perp$, obtenemos

$$g(x) = (g;\Phi_x) = \left(g;\Phi_x^+\right) = \left[g;\Phi_x^+\right]_{\mathcal{N}}, \quad x \in X.$$

Por la unicidad de los representantes de los funcionales lineales sobre un espacio de Hilbert, necesariamente

$$h(x,y) = \left[\Phi_y^+;\Phi_x^+\right]_{\mathcal{N}}, \quad x, y \in X.$$

Escribiendo $\Phi_x^+ = \Phi_x - \Phi_x^-$ encontramos que

$$
\begin{aligned}
h(x,y) &= \left[\Phi_y^+; \Phi_x^+\right]_{\mathcal{N}} = \left(\Phi_y^+; \Phi_x^+\right) = \left(\Phi_y - \Phi_y^-; \Phi_x - \Phi_x^-\right) \\
&= (\Phi_y; \Phi_x) - (\Phi_y - \Phi_y^-; \Phi_x^-) - \left(\Phi_y^-; \Phi_x\right) \\
&= (\Phi_y; \Phi_x) - \Phi_y^-(x), \quad x, y \in X. \qquad (4.5)
\end{aligned}
$$

Para construir el núcleo reproductor de $\mathcal{N}$, elegimos una base ortonormal $\{p_1, \ldots, p_\lambda\}$ de $\mathcal{N}$ respecto de $[\cdot;\cdot]_{\mathcal{N}} = -(\cdot;\cdot)$, es decir, $(p_i; p_j) = -\delta_{i,j}$. En virtud del Teorema 2.4, el núcleo reproductor de $\mathcal{N}$ viene dado por

$$(x,y) \mapsto \sum_{j=1}^{\lambda} p_j(x)\overline{p_j(y)}.$$

Además, podemos expresar la proyección ortogonal v^- de cualquier $v \in \Pi$ sobre $\mathcal{N}$ en la forma

$$v^- = \sum_{j=1}^{\lambda} [v; p_j]_{\mathcal{N}}\, p_j = -\sum_{j=1}^{\lambda} (v; p_j)\, p_j.$$

Insertando esta representación en (4.5), resulta

$$h(x,y) = \Phi(x,y) + \sum_{j=1}^{\lambda} p_j(x)\overline{p_j(y)}, \quad x, y \in X.$$

Como $\Psi(x,y) = h(x,y) + \sum_{j=1}^{\lambda} p_j(x)\overline{p_j(y)}$, $x, y \in X$, queda probado el teorema siguiente.

Teorema 4.14. *El núcleo reproductor Ψ de $(\Pi, [\cdot;\cdot]_{\mathcal{N}})$ está dado por*

$$\Psi(x,y) = \Phi(x,y) + 2\sum_{j=1}^{\lambda} p_j(x)\overline{p_j(y)}, \quad x, y \in X.$$

Puesto que los núcleos reproductores de espacios de Hilbert son definidos positivos, obtenemos una descomposición de los núcleos condicionalmente definidos positivos.

Corolario 4.15. *Todo núcleo condicionalmente definido positivo de orden κ puede ser escrito como diferencia de núcleos definidos positivos.*

4.4. Aproximación con datos dispersos

Mantenemos la notación de la sección anterior y suponemos que Φ es estrictamente condicionalmente definida positiva de orden κ, es decir, para puntos distintos $x_1, \ldots, x_m \in X$, se da la igualdad en (4.1) solamente cuando

$c_1 = \ldots = c_m = 0$. Así, la matriz gramiana G correspondiente a esta elección es definida positiva.

Dentro del marco interpolatorio, pensamos en los puntos x_j como en posiciones ξ_j donde se nos proporcionan los valores de una función f que queremos recuperar. Denotamos por $\Xi = \{\xi_j : 1 \leq j \leq M\}$ el conjunto de estas posiciones, y por $f_1, \ldots, f_M \in \mathbb{C}$ los valores de $f \in \Pi$ en los puntos $\xi_1, \ldots, \xi_M \in X$.

Bajo estas condiciones, podemos aplicar a Π los resultados de las secciones 3.1, 3.2 y 3.3. El interpolante de norma mínima a f en Π es

$$s_{f,\Xi}(x) = \sum_{j=1}^{M} \alpha_j \Psi_{\xi_j}(x), \quad x \in X,$$

de modo que

$$s_{f,\Xi}(\xi_i) = f_i, \quad 1 \leq i \leq M.$$

Como Φ es estrictamente condicionalmente definida positiva, los coeficientes α_j están unívocamente determinados por el sistema lineal

$$\sum_{j=1}^{M} \alpha_j \Psi(\xi_i, \xi_j) = f_i, \quad 1 \leq i \leq M.$$

Ahora, el Teorema 4.14 nos faculta para escribir

$$s_{f,\Xi}(x) = \sum_{j=1}^{M} \alpha_j \Psi_{\xi_j}(x) = \sum_{j=1}^{M} \alpha_j \Phi_{\xi_j}(x) + 2 \sum_{i=1}^{\lambda} \sum_{j=1}^{M} \alpha_j \overline{p_i(\xi_j)} p_i(x).$$

Poniendo

$$\beta_i = \sum_{j=1}^{M} \alpha_j \overline{p_i(\xi_j)},$$

obtenemos

$$s_{f,\Xi}(x) = \sum_{j=1}^{M} \alpha_j \Phi_{\xi_j}(x) + \sum_{i=1}^{\lambda} \beta_i p_i(x), \quad x \in X.$$

La estimación del error puntual estándar sigue de la desigualdad de Cauchy-Schwarz:

$$\begin{aligned} |s_{f,\Xi}(x) - f(x)|^2 &= \left| [s_{f,\Xi} - f; \Psi_x]_{\mathcal{N}} \right|^2 \leq \|s_{f,\Xi} - f\|_{\mathcal{N}}^2 \, \|\Psi_x\|_{\mathcal{N}}^2 \\ &\leq \Psi(x,x) \|f\|_{\mathcal{N}}^2 = \left(\Phi(x,x) + 2 \sum_{j=1}^{\lambda} |p_j(x)|^2 \right) \|f\|_{\mathcal{N}}^2. \end{aligned}$$

Proposición 4.16. *Sean* $f_1, \dots, f_M \in \mathbb{C}$ *los valores de la función* $f \in \Pi$ *en los puntos* $\xi_1, \dots, \xi_M \in X$*, y definamos* $s_{f,\Xi}$ *como anteriormente. La estimación del error local*

$$|s_{f,\Xi}(x) - f(x)| \leq \left(\Phi(x,x) + 2 \sum_{j=1}^{\lambda} |p_j(x)|^2 \right)^{1/2} \|f\|_{\mathcal{N}} \tag{4.6}$$

vale para todo $x \in X$*.*

Definición 4.17. *El término entre paréntesis en la estimación* (4.6) *se denomina habitualmente función potencial.*

4.5. Ejemplo

Sea $M \in \mathbb{N}$ y sea $A \in \mathbb{C}^{M \times M}$ una matriz hermítica arbitraria, con $\kappa \leq M$ autovalores negativos. La aplicación

$$\Phi : \{1, \dots, M\} \times \{1, \dots, M\} \to \mathbb{C}, \quad (r,s) \mapsto a_{rs}$$

define un núcleo hermítico con κ cuadrados negativos. El EPNR asociado Π de índice κ puede ser identificado con span $\{a_1, \dots, a_M\}$, donde a_i denota la columna i-ésima de A. Se obtiene un producto interior indefinido sobre Π mediante la relación

$$(u; v) = v^H \overline{A} u.$$

Puesto que A es hermitiana, existe una base ortonormal $\{v_1, \dots, v_M\}$ formada por autovectores de A. El subespacio negativo maximal $\mathcal{N}$ viene dado como suma directa de autoespacios correspondientes a autovalores negativos de A,

$$\mathcal{N} = \bigoplus_{\substack{\lambda \in \sigma(A) \\ \lambda < 0}} \mathrm{Eig}(\lambda, A).$$

Si (tras una renumeración) $v_1, \dots, v_\kappa$ generan $\mathcal{N}$, entonces

$$\Psi(r,s) = \Phi(r,s) + 2 \sum_{j=1}^{\kappa} \overline{v_j(r)} v_j(s) = a_{rs} + 2 \sum_{j=1}^{\kappa} \overline{v_j(r)} v_j(s)$$

para $r, s \in \{1, \dots, M\}$.

Bibliografía

[1] D. Alpay (ed.): *Operator theory*. Springer, 2015.

[2] D. Alpay, A. Dijksma, J. Rovnyak, H.S.V. de Snoo: *Reproducing kernel Pontryagin spaces*. En S. Axler, J.E. McCarthy, D. Sarason (eds.): *Holomorphic spaces*, Math. Sci. Res. Inst. Publ. **33**, Cambridge University Press, 1998, pp. 425–444.

[3] N. Aronszajn: Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** (1950), 337–404.

[4] H. Begehr, S. Saitoh: One century since Bergman, Szegö and Bochner on reproducing kernels. *Int. Jr. Repro. Kernels* **1** (2022), 1–7.

[5] S. Bergman: *The kernel function and conformal mapping*. American Mathematical Society, 1950.

[6] G. Berschneider, W. zu Castell: *Conditionally positive definite kernels and Pontryagin spaces*. En M. Neamtu, L.L. Schumaker (eds.): *Approximation Theory XII*, Proceedings of the 12th International Conference, San Antonio, TX, USA, March 4-8, 2007, Nashboro Press, 2008, pp. 27–37.

[7] J. Bognár: *Indefinite inner product spaces*. Springer, 1974.

[8] E.W. Cheney, W.A. Light: *A course in approximation theory*. Brooks/Cole, 1999.

[9] J.B. Conway: *Functions of one complex variable*. Springer, 1973.

[10] M.A. Dritschel, J. Rovnyak: *Operators on indefinite inner product spaces*. En A. Böttcher, A. Dijksma, H. Langer, M.A. Dritschel, J. Rovnyak, M.A. Kaashoek (eds.): *Lectures on operator theory and its applications*, Fields Institute for Research in Mathematical Sciences, Toronto, Canada 1994, Fields Institute Monographs, vol. **3**, American Mathematical Society, 1996, pp. 143–232.

[11] J. Mercer: Functions of positive and negative type, and their connection with the theory of integral equations. *Lond. Phil. Trans. (A)* **209** (1909), 415–446.

[12] E.H. Moore: *General analysis, I*. The American Philosophical Society, 1935.

[13] V.I. Paulsen: *An introduction to the theory of reproducing kernel Hilbert spaces*, 2009. Disponible en https://www.math.uh.edu/~vern/rkhs.pdf.

[14] V.I. Paulsen, M. Raghupathi: *An introduction to the theory of reproducing kernel Hilbert spaces*. Cambridge University Press, 2016.

[15] S. Pereverzyev: *An introduction to artificial intelligence based on reproducing kernel Hilbert spaces*. Birkhäuser, 2022.

[16] W. Rudin: *Real and complex analysis*. McGraw-Hill, 1987.

[17] S. Saitoh, Y. Sawano: *Theory of reproducing kernels and applications*. Springer, 2016.

[18] Z. Sasvári: *Positive definite and definitizable functions*. Akademie, 1994.

[19] L. Schwartz: Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). *J. Analyse Math.* **13** (1964), 115–256.

[20] P. Sorjonen: Pontrjaginräume mit einem reproduzierenden Kern. *Ann. Acad. Sci. Fenn, Ser. A I Math.* **594** (1975), 30pp.

[21] S. Zaremba: *Zarys pierwszych zasad teoryi liczb calkowitych*. Akademia Umiejetności, 1907.

Reproducing kernel Hilbert spaces

Pedro Manuel de San Gil González
Facultad de Ciencias • Sección de Matemáticas
Universidad de La Laguna
alu0101122892@ull.edu.es

Abstract

A REPRODUCING KERNEL *Hilbert space (RKHS) is a Hilbert space of functions on which all point functionals are continuous. Via the Fréchet-Riesz representation theorem, every RKHS is associated with a kernel that allows to evaluate any function in the space by performing its inner product against another function in the space, determined by the kernel. The Moore-Aronszajn theorem establishes a bijective correspondence between the reproducing kernels of Hilbert spaces and the so-called kernel functions, or positive (semi)definite kernels, and it can be said that the theory of RKHSs amounts to studying the consequences derived from such a bijection.*

RKHSs find applications in a wide variety of fields. Here, its relevance in approximation theory is illustrated by proving that for every kernel function there exists an RKHS so that interpolation equals minimal norm interpolation within this space. Furthermore, a variant of such an interpolation scheme is addressed that, unlike the original, allows to reproduce polynomials of a given degree. In this new setting, the role of the kernel functions and of the RKHSs are respectively played by the so-called conditionally positive definite kernels and reproducing kernel Pontryagin spaces.

1. Definition, first properties and examples

LET $\mathbb{K}$ DENOTE $\mathbb{R}$ or $\mathbb{C}$, and let X be a set. The class $\mathcal{F}(X,\mathbb{K})$ of all functions from X to $\mathbb{K}$ is a vector space over the field $\mathbb{K}$ with the operations of addition, $(f+g)(x) = f(x)+g(x)$, and scalar multiplication, $(\lambda \cdot f)(x) = \lambda \cdot f(x)$.

Definition 1.1 *A subset $\mathcal{H} \subset \mathcal{F}(X,\mathbb{K})$ is called a reproducing kernel Hilbert space (RKHS, for short) on X provided that:*

1. *$\mathcal{H}$ is a vector subspace of $\mathcal{F}(X,\mathbb{K})$;*
2. *$\mathcal{H}$ is endowed with an inner product $\langle \cdot,\cdot \rangle$, which makes it into a Hilbert space;*
3. *for every $x \in X$, the (linear) evaluation functional $E_x : \mathcal{H} \to \mathbb{K}$, defined by $E_x(f) = f(x)$ whenever $f \in \mathcal{H}$, is bounded.*

If $\mathcal{H}$ is an RKHS on X, then the Fréchet-Riesz representation theorem shows that the evaluation functional can be retrieved as the inner product against a unique vector in $\mathcal{H}$. Therefore, given $x \in X$ there exists a unique function $k_x \in \mathcal{H}$ such that $f(x) = E_x(f) = \langle f, k_x \rangle$, for all $f \in \mathcal{H}$.

Definition 1.2 *The function k_x is called the reproducing kernel for the point x. The function $K : X \times X \to \mathbb{K}$ defined by $K(x,y) = k_y(x)$, with $x, y \in \mathcal{H}$, is the reproducing kernel for $\mathcal{H}$.*

2. General theory

THE FOLLOWING CORRESPONDENCE between kernel functions and RKHSs is a cornerstone of the theory.

Definition 2.1 *A two-variable function $K : X \times X \to \mathbb{C}$ is said to be a kernel function if for every $n \in \mathbb{N}$ and every choice of n distinct points $\{x_1,\dots,x_n\} \subset X$, the $n \times n$ matrix $\left(K(x_i,x_j)\right)_{i,j=1}^n$ is positive, that is, for each $\alpha_1,\dots,\alpha_n \in \mathbb{C}$, one has $\sum_{i,j=1}^n \overline{\alpha_i}\alpha_j K(x_i,x_j) \geq 0$.*

Proposition 2.2 *Let $\mathcal{H}$ be an RKHS on a set X, with reproducing kernel K. Then K is a kernel function.*

Theorem 2.3 (Moore-Aronszajn) *If $K : X \times X \to \mathbb{C}$ is a kernel function, then there exists an RKHS, $\mathcal{H}$, of functions on X such that K is the reproducing kernel for $\mathcal{H}$.*

3. Interpolation and approximation

ONE OF THE primary applications of RKHSs is to problems of interpolation and approximation.

Definition 3.1 *Let X and Y be sets, let $\{x_1,\dots,x_n\} \subset X$ be a collection of n distinct points, and let $\{\lambda_1,\dots,\lambda_n\} \subset Y$. A function $g : X \to Y$ is said to interpolate these points provided that $g(x_i) = \lambda_i$, for all $i = 1,\dots,n$.*

Theorem 3.2 *Let $\mathcal{H}$ be an RKHS on X with reproducing kernel K, let $F = \{x_1,\dots,x_n\} \subset X$ consist of distinct points, and let $\{\lambda_1,\dots,\lambda_n\} \subset \mathbb{C}$. There exists $g \in \mathcal{H}$ that interpolates these values if, and only if, $v = (\lambda_1,\dots,\lambda_n)^t$ is in the range of the matrix $\left(K(x_i,x_j)\right)_{i,j=1}^n$. In this case, if $w = (\alpha_1,\dots,\alpha_n)^t$ is any vector whose image is v, then $h = \sum_{i=1}^n \alpha_i k_{x_i}$ is the unique function of minimal norm in $\mathcal{H}$ that interpolates these points. Moreover, $\|h\|^2 = \langle v, w\rangle$.*

4. Conditionally positive definite kernels and Pontryagin spaces

CONDITIONALLY POSITIVE DEFINITE kernels provide a powerful tool for scattered data approximation. Theorem 4.4 below yields a theoretical framework which allows to study approximation with conditionally positive definite kernels in associated Pontryagin spaces. Recall that a map $\Phi : X \times X \to \mathbb{C}$ is a Hermitian kernel if $\Phi(y,x) = \overline{\Phi(x,y)}$, for all $x, y \in X$.

Definition 4.1 *Suppose $\kappa \in \mathbb{Z}^+$. A Hermitian kernel $\Phi : X \times X \to \mathbb{C}$ is called conditionally positive definite of order κ, if there exists a κ-dimensional space $\mathcal{U}$ of complex-valued functions on X, such that $\sum_{i,j=1}^n c_i\overline{c_j}\Phi(x_i,x_j) \geq 0$ for all choices of $n \in \mathbb{N}$, $x_1,\dots,x_n \in X$, and $c_1,\dots,c_n \in \mathbb{C}$, satisfying $\sum_{i=1}^n c_i u(x_i) = 0$ whenever $u \in \mathcal{U}$.*

Definition 4.2 *The antispace of an indefinite inner product space $(\mathfrak{H}, (\cdot;\cdot))$ is $(\mathfrak{H}, -(\cdot;\cdot))$. A Pontryagin space is an indefinite inner product space Π which can be written as the orthogonal direct sum $\Pi = \Pi_+ \oplus \Pi_-$ of a Hilbert space Π_+ and the antispace Π_- of a finite-dimensional Hilbert space. The (negative) index κ of Π is the maximum dimension of a subspace $\Pi-$ which is the antispace of a Hilbert space in the inner product of Π, and every such κ-dimensional subspace occurs in the above decomposition.*

Definition 4.3 *Let $(\Pi, (\cdot;\cdot))$ be a Pontryagin space of functions on X. A Hermitian kernel $\Phi : X \times X \to \mathbb{C}$ is called a reproducing kernel for Π, if*

1. *$\Phi_x = \Phi(\cdot, x) \in \Pi$ for all $x \in X$, and*
2. *$g(x) = (g; \Phi_x)$ for all $g \in \Pi$, $x \in X$.*

Theorem 4.4 *Suppose $\kappa \in \mathbb{Z}^+$. For every conditionally positive definite kernel Φ of order κ there exists a Pontryagin space Π of index λ, $0 \leq \lambda \leq \kappa$, such that Φ is the reproducing kernel for Π. The converse statement also holds true, i.e. conditionally positive definite kernels can be identified with reproducing kernels in reproducing kernel Pontryagin spaces.*

References

[1] G. Berschneider, W. zu Castell: *Conditionally positive definite kernels and Pontryagin spaces.* En M. Neamtu, L.L. Schumaker (eds.): *Approximation Theory XII*, Proceedings of the 12th International Conference, San Antonio, TX, USA, March 4-8, 2007, Nashboro Press, 2008, pp. 27–37.

[2] V.I. Paulsen, M. Raghupathi: *An introduction to the theory of reproducing kernel Hilbert spaces.* Cambridge University Press, 2016.