

Métodos de funciones de Base Radial para la solución de EDP

Pedro González-Casanova¹

Alexei Gazca¹



¹Instituto de Matemáticas
Universidad Nacional Autónoma de México

Estas notas se realizan como parte del proyecto *Métodos de Funciones de Base Radial para la Solución de Ecuaciones Diferenciales Parciales*. Constituyen un primer esfuerzo por presentar algunos de los resultados más significativos de la teoría, así como de los métodos numéricos para la solución de ecuaciones diferenciales parciales mediante funciones de base radial. En su estado actual, estas notas pretenden recolectar algunos de los elementos teóricos y algorítmicos más relevantes.

Mayo 2016
Notas de curso.

Índice

1. Preliminares	1
1.1. Espacios de Hilbert	1
1.2. Espacios de Sobolev	4
1.2.1. Encajes de Sobolev	6
1.3. Transformada de Fourier y Distribuciones	7
1.4. Kernel Reprodutor	10
2. Interpolación Multivariada con RBF	11
2.1. Enfoque Abstracto	11
2.1.1. Splines Cúbicos	13
2.1.2. Spline de Placa Delgada	17
2.2. Funciones Positivas Definidas	21
2.3. Espacios Nativos	26
2.4. Funciones Completamente Monótonas	31
2.5. Estimación del Error	33
3. Solución de Ecuaciones Diferenciales Parciales con RBF	40
3.1. Colocación Asimétrica de Kansa	40
3.2. Colocación Simétrica	41
3.2.1. Interpolación Generalizada de Hermite-Birkhoff	42
3.3. Cuadratura Diferencial	46
3.3.1. Normalización de la Región de Soporte para RBFs con Parámetro de Forma	49
3.4. Interpolación Local de Hermite	49
4. Métodos Iterativos	52
4.1. El Método de Jacobi	52
4.2. El Método de Gauss-Seidel	54
4.3. El Método SOR	55
4.4. Métodos de Descenso	56
4.5. Métodos de Subespacio de Krylov para Matrices no Simétricas	59
5. Referencias	62

1. Preliminares

Este capítulo no es de ninguna manera un estudio completo de los temas expuestos; más bien presenta lo *mínimo* necesario para entender el resto de las notas. Si estás interesado en conocer más puedes consultar los siguientes libros:

1. Espacios de Hilbert.

- E. Zeidler, *Applied Functional Analysis. Applications to Mathematical Physics*. Springer Verlag, 1995.
- C. Aliprantis, O. Burkinshaw, *Principles of Real Analysis*. Academic Press, 1998.

2. Espacios de Sobolev.

- R.A. Adams, *Sobolev Spaces*. Academic Press, 1975.
- L.C. Evans, *Partial Differential Equations*. Amer. Math. Soc., 1998.
- M. Renardy, R.C. Rogers, *An Introduction to Partial Differential Equations*. Springer Verlag, 2004.

3. Transformada de Fourier y distribuciones.

- L. Schwartz, *Mathematics for the Physical Sciences*. Addison-Wesley, 1966.
- S. Salsa, *Partial Differential Equations in Action. From Modelling to Theory*. Universitext, Springer Verlag, 2008.

1.1. Espacios de Hilbert

En esta sección se repasarán algunos conceptos que serán útiles más adelante.

Sea V un espacio vectorial sobre \mathbb{R} . Una norma es una función real

$$\| \cdot \| : V \rightarrow \mathbb{R}$$

tal que para $\lambda \in \mathbb{R}$, $u, v \in V$, se cumplen las siguientes propiedades:

1. $\|u\| \geq 0$; $\|u\| = 0$ si y sólo si $u = 0$.
2. $\|\lambda u\| = |\lambda| \|u\|$.
3. $\|u + v\| \leq \|u\| + \|v\|$.

Un *espacio normado* es un espacio vectorial V equipado con una norma $\|\cdot\|$. Utilizando la norma se puede definir la distancia entre dos vectores como

$$d(u, v) = \|u - v\|,$$

lo cual hace de V un *espacio métrico* y permite definir una *topología en V* y una noción de convergencia de una manera simple.

Decimos que una sucesión $\{u_n\} \subset V$ converge a u en V , y lo denotamos $u_n \rightarrow u$ en V , si

$$d(u_n, u) = \|u_n - u\| \rightarrow 0, \quad \text{cuando } n \rightarrow \infty.$$

Un concepto importante es el de *sucesión de Cauchy*. Una sucesión $\{u_n\} \subset V$ es de Cauchy si

$$d(u_n, u_m) = \|u_n - u_m\| \rightarrow 0, \quad \text{cuando } n, m \rightarrow \infty,$$

en particular una sucesión convergente siempre es de Cauchy.

Definición 1.1.1. *Un espacio de Banach es un espacio normado y completo; esto es, un espacio normado en el que toda sucesión de Cauchy converge a un elemento del espacio.*

Ejemplo 1.1.2. *Sean Ω un subconjunto abierto de \mathbb{R}^d y $p \in [1, \infty)$. Sea $V = L^p(\Omega)$ el conjunto de funciones f tales que $|f|^p$ es Lebesgue integrable en Ω . Identificando dos funciones f y g cuando son iguales c.d.¹ en Ω , $L^p(\Omega)$ se convierte en un espacio de Banach equipado con la norma*

$$\|f\|_{L^p(\Omega)} = \left(\int_{\Omega} |f|^p \right)^{1/p}.$$

Ejemplo 1.1.3. *Sea $V = L^\infty(\Omega)$ el conjunto de las funciones esencialmente acotadas en Ω . Recordemos que una función $f : \Omega \rightarrow \mathbb{R}$ es esencialmente acotada si existe $M > 0$ de forma que*

$$|f(x)| \leq M \quad \text{c.d. en } \Omega.$$

Podemos definir para $f \in L^\infty(\Omega)$

$$\|f\|_{L^\infty(\Omega)} = \inf\{M \in \mathbb{R} : |f(x)| \leq M \text{ p.c.t. en } \Omega\}.$$

Identificando dos funciones de $L^\infty(\Omega)$ cuando coinciden c.d. en Ω , lo anterior define una norma con la cual $L^\infty(\Omega)$ es un espacio de Banach.

Sea V un espacio vectorial sobre \mathbb{R} . Un producto interior en V es una función

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R},$$

tal que para $\lambda, \mu \in \mathbb{R}$, $u, v, w \in V$ se cumple

¹Se dice que una propiedad se cumple *casi dondequiera* (c.d.) en un conjunto Ω , o bien que se cumple *para casi todo* (p.c.t.) $x \in \Omega$, si se cumple para todos los puntos de Ω salvo un conjunto de medida cero

1. $\langle \lambda u + \mu v, w \rangle = \lambda \langle u, w \rangle + \mu \langle v, w \rangle$.
2. $\langle u, v \rangle = \langle v, u \rangle$
3. $\langle u, u \rangle \geq 0$; $\langle u, u \rangle = 0$ si y sólo si $u = 0$.

Un producto interior induce una norma, la cual está dada por

$$\|v\| = \sqrt{\langle v, v \rangle}, \quad (1.1.1)$$

y se cumplen las siguientes relaciones.

Proposición 1.1.4. Sean $u, v \in V$. Entonces se tiene:

1. Desigualdad de Cauchy-Schwarz

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

2. Identidad del paralelogramo

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$$

Definición 1.1.5. Un espacio de Hilbert H es un espacio vectorial con producto interior $\langle \cdot, \cdot \rangle$ que es completo respecto a la norma inducida (1.1.1).

Ejemplo 1.1.6. El espacio $L^2(\Omega)$ con el producto interior

$$\langle u, v \rangle_{L^2(\Omega)} = \int_{\Omega} uv,$$

es un espacio de Hilbert.

Proposición 1.1.7. Un subespacio vectorial V de un espacio de Hilbert H es un espacio de Hilbert si y sólo si V es cerrado en H .

Sean H un espacio de Hilbert con producto interior $\langle \cdot, \cdot \rangle$ y V un subespacio vectorial de H . El espacio ortogonal a V en H se define como

$$V^\perp = \{v \in H : \langle u, v \rangle = 0 \quad \forall u \in V\}.$$

El espacio ortogonal se puede caracterizar de la siguiente forma.

Proposición 1.1.8. Sean V un subespacio vectorial de H y $w \in H$. Entonces

$$w \in V^\perp \iff \|w\| = \inf_{v \in V} \|w - v\|.$$

Si V es cerrado, el ínfimo se alcanza y es único.

Teorema 1.1.9 (Proyección Ortogonal). *Sea V un subespacio vectorial cerrado de H . Entonces para cada $u \in H$, existe un único $v \in V$ tal que*

$$\|u - v\| = \inf_{w \in V} \|u - w\|.$$

Sea H un espacio de Hilbert. Un *funcional* en H es una función $L : H \rightarrow \mathbb{R}$ lineal. Al conjunto de todos los funcionales acotados (i.e. continuos) en H se le denomina *espacio dual* y se denota por H^* . El siguiente resultado da una descripción completa del espacio dual de un espacio de Hilbert.

Teorema 1.1.10 (de Representación de Riesz). *Sea H un espacio de Hilbert. Para cualquier $T \in H^*$ existe un único $u_T \in H$ tal que*

1. $Tu = \langle u_T, u \rangle$ para todo $u \in H$.
2. $\|T\| = \|u_T\|$.

Decimos que u_T es el elemento de Riesz de T asociado al producto interior $\langle \cdot, \cdot \rangle$. Más aún, H^* equipado con el producto interior

$$\langle T_1, T_2 \rangle_{H^*} = \langle u_{T_1}, u_{T_2} \rangle,$$

es claramente un espacio de Hilbert. Así, el teorema de representación de Riesz permite la *identificación de un espacio de Hilbert con su dual*.

1.2. Espacios de Sobolev

Un vector de la forma $\alpha = (\alpha_1, \dots, \alpha_d)$, donde cada componente α_i es un entero no negativo, se denomina un multi-índice de orden $|\alpha| = \alpha_1 + \dots + \alpha_d$. Si $|\alpha| \leq k$, $x \in \mathbb{R}^d$, y $u \in C^k(\Omega)$, se pueden definir la derivada α -ésima de u y la potencia α -ésima de x como

$$D^\alpha u(x) := \frac{\partial^{|\alpha|} u(x)}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}} = \partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d} u(x) \quad \text{y} \quad x^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d}. \quad (1.2.1)$$

Si k es un entero no negativo,

$$D^k u(x) := \{D^\alpha u(x) : |\alpha| = k\},$$

es el conjunto de todas las derivadas parciales de u de orden k . En el caso particular $k = 1$, se considera a los elementos de Du arreglados en forma de vector (gradiente):

$$Du = (\partial_{x_1} u, \dots, \partial_{x_d} u).$$

Sea $C_c^\infty(\Omega)$ el espacio de las funciones infinitamente diferenciables $\phi : \Omega \rightarrow \mathbb{R}$ con soporte compacto en Ω . A los elementos de este espacio normalmente se les llama funciones de prueba. Supongamos que $u \in C^1(\Omega)$ y $\phi \in C_c^\infty(\Omega)$; por la fórmula de integración por partes

$$\int_{\Omega} u \partial_{x_i} \phi \, dx = - \int_{\Omega} \phi \partial_{x_i} u \, dx, \quad i \in \{1, \dots, d\}, \quad (1.2.2)$$

de forma más general, si se tiene un multi-índice α , aplicando la fórmula anterior $|\alpha|$ veces se tiene

$$\int_{\Omega} u D^{\alpha} \phi dx = (-1)^{|\alpha|} \int_{\Omega} \phi D^{\alpha} u dx, \quad i \in \{1, \dots, d\}. \quad (1.2.3)$$

Nótese que el lado derecho de (1.2.3) tiene sentido incluso si u no pertenece a $C^k(\Omega)$; solamente es necesario que $u \in L^1_{loc}(\Omega)$ (i.e. es localmente integrable). Esto motiva la siguiente definición

Definición 1.2.1. Sean $u, v \in L^1_{loc}(\Omega)$ y α un multi-índice. Decimos que v es la α -ésima derivada débil de u , denotada $D^{\alpha}u = v$, si se cumple

$$\int_{\Omega} u D^{\alpha} \phi dx = (-1)^{|\alpha|} \int_{\Omega} v \phi dx.$$

para todas las funciones de prueba $\phi \in C_c^{\infty}(\Omega)$.

La α -ésima derivada débil de una función u , si existe, es única (módulo una redefinición en un conjunto de medida cero) y claramente coincide con la α -ésima derivada clásica de u , si es que esta última existe.

Ejemplo 1.2.2. Sea $f \in C(\mathbb{R})$ la función dada por

$$f(x) = x_+ = \begin{cases} x, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

Entonces f es débilmente diferenciable y su derivada débil está dada por la función de Heavyside

$$H(x) = \begin{cases} 1, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

La elección del valor de $f'(x)$ en $x = 0$ es irrelevante, ya que la derivada débil se define módulo conjuntos de medida cero. Para probarlo nota que para cualquier $\phi \in C_c^{\infty}(\mathbb{R})$, utilizando integración por partes se tiene

$$\int_{\mathbb{R}} f \phi' = \int_0^{\infty} x \phi'(x) dx = - \int_0^{\infty} \phi(x) dx = \int_{\mathbb{R}} H \phi.$$

Ejemplo 1.2.3. La función de Heavyside definida en el ejemplo anterior no es débilmente diferenciable. Para demostrarlo supongamos lo contrario, esto es, que existe una función $v \in L^1_{loc}(\mathbb{R})$ tal que

$$\int_{\mathbb{R}} H \phi' = - \int_{\mathbb{R}} v \phi, \quad \forall \phi \in C_c^{\infty}(\mathbb{R}).$$

Entonces se cumple, en particular, que

$$\begin{aligned} \int_0^{\infty} v \phi &= - \int_0^{\infty} \phi' = 0, & \forall \phi \in C_c^{\infty}(0, \infty). \\ \int_{-\infty}^0 v \phi &= - \int_{-\infty}^0 0 = 0 & \forall \phi \in C_c^{\infty}(-\infty, 0). \end{aligned}$$

Esto implica que $v = 0$ c.d. en \mathbb{R} , por lo que

$$\phi(0) = - \int_0^\infty \phi' = - \int_{\mathbb{R}} H\phi' = 0 \quad \forall \phi \in C_c^\infty(\mathbb{R}),$$

lo cual es una contradicción.

Definición 1.2.4. Sean $k \in \mathbb{N}$, $p \in [1, \infty]$. El espacio de Sobolev $W^{k,p}(\Omega)$ se define como

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) : D^\alpha u \text{ existe y } D^\alpha u \in L^p(\Omega) \text{ para } |\alpha| \leq k\}$$

En el caso particular $p = 2$ usualmente se escribe $W^{k,2}(\Omega) = H^k(\Omega)$. Nota que $H^0(\Omega) = L^2(\Omega)$. Al igual que en los espacios $L^p(\Omega)$, en los espacios de Sobolev se identifica a dos funciones cuando coinciden casi dondequiera.

Para $u \in W^{k,p}(\Omega)$ se pueden definir las normas

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p}, \quad p \in [1, \infty),$$

y

$$\|u\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)},$$

con las que $W^{k,p}(\Omega)$ se convierte en un espacio de Banach. Más aún, $H^k(\Omega)$ es un espacio de Hilbert con producto interior

$$\langle u, v \rangle_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \langle D^\alpha u, D^\alpha v \rangle_{L^2(\Omega)}.$$

1.2.1. Encajes de Sobolev

Sean V y W dos espacios de Banach con $V \subseteq W$. Decimos que V está encajado continuamente en W (se escribe $V \hookrightarrow W$) si existe una constante $c > 0$ tal que

$$\|v\|_W \leq c\|v\|_V \quad \forall v \in V,$$

en otras palabras, la identidad es un operador acotado.

Teorema 1.2.5 (Encajes de Sobolev). Sea $\Omega \in \mathbb{R}^d$ un dominio Lipschitz. Entonces se cumple lo siguiente:

- (a) Si $k < \frac{d}{p}$, entonces $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ para toda $q \leq p^*$, donde $\frac{1}{p^*} = \frac{1}{p} - \frac{k}{d}$.
- (b) Si $k = \frac{d}{p}$, entonces $W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ para toda $q < \infty$.

(c) Si $k > \frac{d}{p}$, entonces $W^{k,p}(\Omega) \hookrightarrow C^{k - [\frac{d}{p}] - 1, \beta}(\Omega)$, donde

$$\beta = \begin{cases} [\frac{d}{p}] + 1 - \frac{d}{p}, & \text{si } \frac{d}{p} \notin \mathbb{Z} \\ \text{cualquier número positivo} < 1, & \text{si } \frac{d}{p} \in \mathbb{Z}. \end{cases}$$

Como ayuda para recordar este resultado, nota que entre más grande el producto kp , más suaves son las funciones de $W^{k,p}(\Omega)$. Hay un valor crítico d (la dimensión del espacio), tal que si $kp > d$, se tiene que las funciones de $W^{k,p}(\Omega)$ son continuas². Para ver qué tan diferenciables son las funciones se puede aplicar este resultado sucesivamente, i.e. cuando $k - l > \frac{d}{p}$ tenemos que $W^{k,p}(\Omega) \hookrightarrow C^l(\Omega)$.

Los espacios de Sobolev están basados en los espacios L^p , por lo que las funciones están definidas de manera única solamente c.d. en Ω . Como la frontera $\partial\Omega$ tiene medida cero en \mathbb{R}^d , parece que no tiene mucho sentido definir el valor en la frontera de una función en un espacio de Sobolev. Para resolver esto, es necesario generalizar la noción de restricción a la frontera.

Teorema 1.2.6 (Traza). Sean Ω un dominio Lipschitz en \mathbb{R}^d , $p \in [1, \infty)$. Entonces existe un operador lineal y continuo $\gamma : W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega)$, tal que

(a) $\gamma v = v|_{\partial\Omega}$ si $v \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$.

(b) Hay una constante $c > 0$ tal que $\|\gamma v\|_{L^p(\partial\Omega)} \leq c\|v\|_{W^{1,p}(\Omega)}$.

(c) El mapeo $\gamma : W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega)$ es compacto.

A γ se le conoce como el *operador de traza*, y γv se puede considerar como una generalización de la restricción a la frontera de v . El operador de traza no es en general ni inyectivo ni sobre. El rango $\gamma(W^{1,p}(\Omega))$ es un subespacio más pequeño que $L^p(\partial\Omega)$, concretamente $W^{1-\frac{1}{p},p}(\Omega)$. En particular, es común tratar problemas de valores a la frontera en los que hay que trabajar con trazas de funciones de $H^1(\Omega)$, por lo que es útil recordar que³

$$H^{1/2}(\partial\Omega) = \gamma(H^1(\Omega)).$$

A este espacio se le puede equipar con la norma $\|g\|_{H^{1/2}(\partial\Omega)} = \inf_{\substack{v \in H^1(\Omega) \\ \gamma v = g}} \|v\|_{H^1(\Omega)}$.

1.3. Transformada de Fourier y Distribuciones

Para todas las funciones $f \in L^1(\mathbb{R}^d)$ se define la transformada de Fourier \mathcal{F} a través de

$$\mathcal{F}f(\xi) = \hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} dx.$$

Por la desigualdad de Hölder, se tiene $\mathcal{F} : L^1(\mathbb{R}^d) \rightarrow L^\infty(\mathbb{R}^d)$. De hecho, es posible demostrar que \hat{f} es uniformemente continua.

²Más precisamente, iguales a una función continua casi en todas partes.

³Los espacios de Sobolev fraccionarios sobre \mathbb{R}^d se definirán más adelante.

Nota que la transformada de Fourier de una función integrable no es necesariamente integrable. Por esta razón es más conveniente trabajar en el llamado *espacio de Schwartz* o de *funciones de descenso rápido*

$$\mathcal{S}(\mathbb{R}^d) := \{u \in C^\infty(\mathbb{R}^d) : x^\beta D^\alpha u \in L^\infty(\mathbb{R}^d) \quad \forall \alpha, \beta \in \mathbb{N}^d\}.$$

De la definición se sigue entonces que

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d),$$

y además que para multi-índices α, β cualquiera

$$\mathcal{F}(x^\beta f) = (i)^{|\beta|} D_\xi^\beta \hat{f} \quad \text{y} \quad \mathcal{F}(D_x^\alpha f) = (i)^{|\alpha|} \xi^\alpha \hat{f},$$

es decir, la transformada de Fourier transforma polinomios en derivadas y viceversa.

Para $f \in L^1(\mathbb{R}^d)$ definimos el operador \mathcal{F}^* como

$$\mathcal{F}^* f(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(\xi) e^{ix \cdot \xi} d\xi.$$

Un resultado importante es el siguiente.

Teorema 1.3.1 (Plancherel). *La transformada $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ es un isomorfismo con inversa \mathcal{F}^* que preserva el producto interior de $L^2(\mathbb{R}^d)$.*

A los elementos del espacio dual $\mathcal{S}(\mathbb{R}^d)^*$ se les llama distribuciones temperadas. Para $p \in [1, \infty]$ hay una inyección natural de $L^p(\mathbb{R}^d)$ en $\mathcal{S}(\mathbb{R}^d)^*$ dada por

$$\langle f, u \rangle = \int_{\mathbb{R}^d} f(x) u(x) dx \quad \forall u \in \mathcal{S}(\mathbb{R}^d).$$

De manera similar, a cualquier medida finita en \mathbb{R}^d se le puede asociar un elemento de $\mathcal{S}(\mathbb{R}^d)^*$. El ejemplo básico es la delta de Dirac, definida por

$$\langle \delta, u \rangle = u(0) \quad , \text{ o de manera más general } \quad \langle \delta_x, u \rangle = u(x) \quad \forall u \in \mathcal{S}(\mathbb{R}^d).$$

En el espacio $\mathcal{S}(\mathbb{R}^d)^*$ se puede definir la α -ésima derivada distribucional

$$D^\alpha : \mathcal{S}(\mathbb{R}^d)^* \rightarrow \mathcal{S}(\mathbb{R}^d)^*,$$

a partir de la relación

$$\langle D^\alpha T, u \rangle = (-1)^\alpha \langle T, D^\alpha u \rangle,$$

para $T \in \mathcal{S}(\mathbb{R}^d)^*$, $u \in \mathcal{S}(\mathbb{R}^d)$. De manera similar,

$$\langle fT, u \rangle = \langle T, fu \rangle$$

define el producto fT con $T \in \mathcal{S}(\mathbb{R}^d)^*$, siempre y cuando f y sus derivadas crezcan a lo más como un polinomio.

Ejemplo 1.3.2. Sea H la función de Heavyside

$$H(x) = \begin{cases} 1, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

Entonces si $u \in \mathcal{S}(\mathbb{R})$, se tiene que

$$\langle \delta, u \rangle = u(0) = - \int_0^\infty \frac{du}{dx} dx = - \langle H, \frac{du}{dx} \rangle = \langle \frac{dH}{dx}, u \rangle,$$

lo cual implica que

$$\frac{dH}{dx} = \delta \quad \text{en} \quad \mathcal{S}(\mathbb{R})^*.$$

La transformada de Fourier $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}(\mathbb{R}^d)$ se puede extender al espacio de distribuciones $\mathcal{S}(\mathbb{R}^d)^*$ mediante

$$\langle \mathcal{F}T, u \rangle = \langle T, \mathcal{F}u \rangle, \quad \forall u \in \mathcal{S}(\mathbb{R}^d),$$

con una fórmula análoga para la transformada inversa \mathcal{F}^* .

Existe una generalización de la transformada de Fourier que será útil más adelante. Para definirla consideremos a los subespacios \mathcal{S}_m de \mathcal{S} definidos como

$$\mathcal{S}_m := \{ \phi \in \mathcal{S} : \phi(x) = \mathcal{O}(\|x\|^m) \text{ para } \|x\| \rightarrow 0, m \in \mathbb{N} \},$$

recordando que $\phi(x) = \mathcal{O}(\|x\|^m)$ significa que $|\phi(x)|/\|x\|^m$ está acotada cerca de cero. Entonces si f es una función continua que crece a lo más como un polinomio, se dice que $\hat{f} : \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}$ es su *transformada de Fourier generalizada* si existe un número $m \in \mathbb{N}$ tal que

$$\int_{\mathbb{R}^d} f(x) \hat{\phi}(x) dx = \int_{\mathbb{R}^d} \hat{f}(x) \phi(x) dx,$$

para cualquier $\phi \in \mathcal{S}_{2m}$. Al número m más pequeño que cumpla lo anterior se le denomina el *orden* de \hat{f} .

Si se tiene una función $f \in L^1(\mathbb{R}^d) \cup L^2(\mathbb{R}^d)$, la transformada de Fourier generalizada coincide con la transformada de Fourier clásica (el orden es cero). De igual forma, la transformada de Fourier generalizada y la transformada de Fourier en el sentido de distribuciones coinciden en el conjunto \mathcal{S}_{2m} .

Los espacios de Sobolev basados en $L^2(\mathbb{R}^d)$ poseen una caracterización muy útil en términos de la transformada de Fourier. Utilizando el teorema de Plancherel no es difícil demostrar que $u \in H^k(\mathbb{R}^d)$ si y sólo si $(1 + |\xi|^2)\hat{u} \in L^2(\mathbb{R}^d)$ y las normas

$$u \mapsto \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^2(\mathbb{R}^d)}^2 \right)^{1/2} \quad y \quad u \mapsto \left(\int_{\mathbb{R}^d} |\hat{u}(\xi)|^2 (1 + |\xi|^2)^k d\xi \right)^{1/2},$$

son equivalentes. Esto sugiere una generalización de $H^k(\mathbb{R}^d)$ en la que k se pueda reemplazar por un número real s :

$$H^s(\mathbb{R}^d) := \{f \in \mathcal{S}(\mathbb{R}^d)^* : \hat{f} \text{ es una función y } \|f\|_s^2 = \int_{\mathbb{R}^d} |\hat{u}(\xi)|^2 (1 + |\xi|^2)^s d\xi < \infty\}.$$

1.4. Kernel Reprodutor

Supongamos que H es un espacio de Hilbert compuesto de funciones $f : \Omega \rightarrow \mathbb{R}$, donde Ω es un conjunto no vacío. Se dice que una función $K : \Omega \times \Omega \rightarrow \mathbb{R}$ es un kernel reproductor de H si

1. $K(\cdot, y) \in H$ para cualquier $y \in \Omega$.
2. $f(y) = \langle f, K(\cdot, y) \rangle_H$ para cualquier $f \in H$ y $y \in \Omega$.

De la segunda propiedad se sigue fácilmente que si un espacio de Hilbert tiene un kernel reproductor, éste es único. Una caracterización de los espacios de Hilbert que tienen un kernel reproductor es la siguiente

Proposición 1.4.1. *Supongamos que H es un espacio de Hilbert compuesto de funciones $f : \Omega \rightarrow \mathbb{R}$. Entonces son equivalentes:*

1. Los funcionales de evaluación $\delta_y : f \mapsto f(y)$, $f \in H$, son continuos para toda $y \in \Omega$; en otras palabras, $\delta_y \in H^* \quad \forall y \in \Omega$.
2. El espacio H tiene kernel reproductor.

El siguiente teorema enuncia algunas propiedades importantes de un espacio de Hilbert con kernel reproductor.

Teorema 1.4.2. *Supongamos que H es un espacio de Hilbert compuesto de funciones $f : \Omega \rightarrow \mathbb{R}$ con kernel reproductor K . Entonces*

1. $K(x, y) = \langle K(\cdot, x), K(\cdot, y) \rangle_H = \langle \delta_x, \delta_y \rangle_{H^*}$, para $x, y \in \Omega$
2. $K(x, y) = K(y, x)$, para $x, y \in \Omega$.
3. Si $f \in H$ y f_n es una sucesión que converge a f en la norma de H , entonces f_n converge puntualmente a f .
4. H es continuamente encajable en $C^0(\Omega)$.

2. Interpolación Multivariada con RBF

2.1. Enfoque Abstracto

En muchas disciplinas científicas uno se enfrenta al siguiente problema: encontrar para cierto conjunto de datos (mediciones y las ubicaciones en las que se obtuvieron las mediciones) una regla que permita deducir información sobre el proceso que se estudia en lugares donde no se realizaron mediciones. Esto es, se busca una función s que ajuste “adecuadamente” a los datos; hay muchas maneras de decidir el significado de “adecuadamente”, en nuestro caso el único criterio es que la función s coincida con los datos de las mediciones. A este enfoque se le conoce como interpolación. Una formulación precisa del problema es la siguiente

Problema. Dado un conjunto de datos $(\mathbf{x}_j, u_j) \in \Omega \times \mathbb{R}$, $j \in \{1, \dots, N\}$, donde $\Omega \subseteq \mathbb{R}^d$ es una región, encontrar una función continua $s : \Omega \rightarrow \mathbb{R}$ tal que

$$s(\mathbf{x}_j) = u_j, \quad j \in \{1, \dots, N\}$$

Es un resultado conocido que en el caso unidimensional se puede escoger s como un polinomio de grado a lo más $N - 1$. Lo notable en este caso es que el espacio de funciones interpolantes sólo depende del número de puntos en $X = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$, mas no de los valores explícitos de (\mathbf{x}_j, u_j) . Así, suena razonable buscar espacios con esa propiedad en dimensiones mayores; el concepto que generaliza lo anterior es el de espacio de Haar:

Definición 2.1.1. Supongamos que $\Omega \subseteq \mathbb{R}^d$ contiene al menos N puntos y sea $V \subseteq C(\Omega)$ un espacio vectorial N -dimensional. Entonces se dice que V es un espacio de Haar de dimensión N sobre Ω si para cualesquiera puntos distintos $\mathbf{x}_1, \dots, \mathbf{x}_N \in \Omega$ y cualesquiera $u_1, \dots, u_N \in \mathbb{R}$ existe una única función $s \in V$ tal que $s(\mathbf{x}_j) = u_j$, $j \in \{1, \dots, N\}$.

Desafortunadamente, Mairhuber y Curtis demostraron el siguiente resultado:

Teorema 2.1.2. Supongamos que $d \geq 2$ y que $\Omega \subseteq \mathbb{R}^d$ contiene un punto interior. Entonces no existe ningún espacio de Haar sobre Ω de dimensión $N \geq 2$.

Por esta razón, si se trabaja en dimensión mayor a uno, es imposible fijar un espacio N -dimensional de funciones que sea apropiado para todos los conjuntos de N nodos. Esto es, si se busca un esquema apropiado de interpolación en \mathbb{R}^d es necesario imponer condiciones sobre los nodos de forma que sea posible interpolación estable con polinomios, o bien que el espacio de funciones dependa explícitamente de los nodos. Utilizando la teoría de espacios de Hilbert, el problema de interpolación se puede tratar de la siguiente manera. Sea $X : a < x_1 < x_2 < \dots < x_N < b$ un conjunto de nodos en \mathbb{R}^d y V un espacio de Hilbert de funciones sobre \mathbb{R}^d . Buscamos entonces una solución al siguiente problema:

Problema. Encontrar una función $\sigma \in V$ que cumpla las siguientes condiciones:

1. $\sigma \in X_I$ donde $X_I = \{u \in V \mid u(x_i) = u_i, \quad i = 1, \dots, N\}$, con $u_i \in \mathbb{R}^d$ datos.
2. $\sigma = \inf_{u \in X_I} \|u\|_V$

Supóngase ahora que el espacio V tiene un kernel reproductor K . Entonces no es difícil ver que:

$$X_I = \{u \in V \mid \ll u \mid K(\cdot, x_i) \gg_V = u_i, \quad i = 1, \dots, N\},$$

es una variedad lineal o coset del subespacio dado por :

$$X_I^0 = \{u \in V \mid \ll u \mid K(\cdot, x_i) \gg_V = 0, \quad i = 1, \dots, N\},$$

esto es, si $v, w \in X_I$, entonces $v - w \in X_I^0$.

Debido a que la función σ interpola, pertenece al conjunto X_I ; además como σ tiene norma mínima (i.e. distancia mínima al cero), el teorema 1.1.9 (Proyección Ortogonal) nos dice que es la única función que pertenece a $X_I \cap X_I^{0\perp}$, donde $X_I^{0\perp}$ es el espacio ortogonal a X_I^0 .

Recordemos de álgebra lineal que el espacio generado por el conjunto $\{K(\cdot, x_i)\}_{i=1}^N$ se define como

$$\text{span}\{K(\cdot, x_i)\}_{i=1}^N = \left\{ \sum_{i=1}^N \lambda_i K(\cdot, x_i) : \lambda_i \in \mathbb{R} \quad \forall i \in \{1, \dots, N\} \right\}.$$

Si tomamos $v \in X_I^0$, entonces $\ll v, \sum_{i=1}^N \lambda_i K(\cdot, x_i) \gg = 0$ para cualesquiera $\lambda_i \in \mathbb{R}$, es decir $v \in \text{span}\{K(\cdot, x_i)\}_{i=1}^N \perp$. De igual manera si $v \in \text{span}\{K(\cdot, x_i)\}_{i=1}^N \perp$, se tiene en particular que $v(x_j) = \ll v, K(\cdot, x_j) \gg = 0$ para toda $j \in \{1, \dots, N\}$, lo cual implica que $v \in X_I^0$. Por esta razón se cumple

$$X_I^{0\perp} = \text{span}\{K(\cdot, x_i)\}_{i=1}^N,$$

en otras palabras, el interpolante σ es de la forma

$$\sigma(x) = \sum_{j=1}^N \lambda_j K(x, x_j), \tag{2.1.1}$$

por lo que al conocer la forma explícita del kernel reproductor y los valores λ_i puede determinarse el spline interpolante σ . Sustituyendo en la ecuación anterior $x = x_i$ y utilizando que $\sigma(x_i) = u_i$, resulta el sistema lineal

$$\begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_N) \\ K(x_2, x_1) & \cdots & \cdots & K(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_N, x_1) & K(x_N, x_2) & \cdots & K(x_N, x_N) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} \tag{2.1.2}$$

del cual se obtienen los valores λ_j . A la matriz de este sistema se le conoce como la matriz de Gram, y es invertible por el teorema 1.1.9. Es decir, para un conjunto fijo de nodos $X : a < x_1 < x_2 < \dots < x_N < b$, el problema tiene solución única.

2.1.1. Splines Cúbicos

A continuación aplicaremos los resultados anteriores al caso de los llamados splines cúbicos.

Definición 2.1.3. *El conjunto de splines cúbicos correspondiente a la descomposición $X : a < x_1 < x_2 < \dots < x_N < b$ está dado por*

$$S_3(X) := \{s \in C^2[a, b] : s|_{[x_i, x_{i+1}]} \in \mathbb{P}_3(\mathbb{R}), 0 \leq i \leq N\},$$

donde $x_0 = a$, $x_{N+1} = b$. En otras palabras, son las funciones con dos derivadas continuas que coinciden con polinomios cúbicos en los intervalos definidos por X .

El espacio $S_3(X)$ tiene dimensión $N + 4$, por lo que las condiciones de interpolación $s(\mathbf{x}_j) = f_j$ no pueden garantizar un único interpolante. Una estrategia posible para solucionar esto es trabajar con los llamados splines naturales, los cuales consisten en los splines cúbicos que son rectas en los intervalos exteriores $[a, x_1]$, $[x_N, b]$:

$$\mathcal{N}S_3(X) := \{s \in S_3(X) : s|_{[a, x_1]}, s|_{[x_N, b]} \in \mathbb{P}_1(\mathbb{R})\}.$$

Este nuevo espacio tiene N dimensiones y es posible demostrar que el problema inicial de interpolación tiene una única solución en $\mathcal{N}S_3(X)$.

En este caso se tomaremos a V como $H^2(0, 1)$, el espacio estándar de Sobolev equipado con el producto interior

$$\ll u, v \gg_{H^2(0,1)} = u(0)v(0) + u(1)v(1) + \langle u'', v'' \rangle_{L^2(0,1)}. \quad (2.1.3)$$

Nota que este es un producto interior equivalente al definido en la sección 1.2. Veremos que los interpolantes asociados con este espacio son justamente los splines cúbicos naturales. En este contexto el problema de interpolación se puede expresar como

Problema. *Encontrar una función $\sigma \in H^2(0, 1)$ que cumpla las siguientes condiciones:*

1. $\sigma \in X_I$ donde $X_I = \{u \in H^2(0, 1) | u(x_i) = u_i, \quad i = 1, \dots, N\}$, con $u_i \in \mathbb{R}$ datos.
2. $\sigma = \inf_{u \in X_I} \|u\|_{H^2(0,1)}$

Una pregunta natural sería: ¿Por qué uno buscaría una función σ que minimice $\int_0^1 |u''(x)|^2 dx$? Hay dos respuestas:

- La curvatura de la función σ es $\sigma''/(1 + \sigma'^2)^{3/2}$, lo que la hace óptima en el sentido de que interpola con la menor curvatura total si σ' es pequeña.
- Esta función aproxima la solución de un problema de la física, en el que una barra uniforme, delgada, elástica y lineal es deformada para interpolar los puntos especificados en ausencia de fuerzas externas. La forma de dicha barra está determinada por un principio de energía mínima, en este caso energía potencial elástica, y la forma funcional de la energía se aproxima a primer orden por la integral $\int_0^1 |u''(x)|^2 dx$.

Supondremos ahora que el espacio $H^2(0, 1)$ tiene un kernel reproductor K y definimos a los conjuntos

$$X_I = \{u \in H^2(0, 1) \mid \ll u \mid K(\cdot, x_i) \gg_{H^2(0,1)} = u_i, \quad i = 1, \dots, N\},$$

$$X_I^0 = \{u \in H^2(0, 1) \mid \ll u \mid K(\cdot, x_i) \gg_{H^2(0,1)} = 0, \quad i = 1, \dots, N\}.$$

Como se vió en la sección anterior, si $v, w \in X_I$, entonces $v - w \in X_I^0$. Además el interpolante es de la forma

$$\sigma(x) = \sum_{j=1}^N \lambda_j K(x, x_j), \quad (2.1.4)$$

donde los coeficientes λ_j se encuentran resolviendo el sistema (2.1.2). Para encontrar la forma explícita del Kernel reproductor, nota que utilizando dos veces integración por partes, el producto interior (2.1.3) se puede escribir como

$$\begin{aligned} \ll u, v \gg_{H^2(0,1)} &= u(0)v(0) + u(1)v(1) + [u''(x)v'(x) - u'''(x)v(x)] \Big|_0^1 + \int_0^1 u^{(4)}(x)v(x) dx \\ &= u(0)v(0) + u'''(0)v(0) + u(1)v(1) - u'''(1)v(1) - u''(0)v'(0) + u''(1)v'(1) + \int_0^1 u^{(4)}(x)v(x) dx. \end{aligned} \quad (2.1.5)$$

La forma del producto interior sugiere que el kernel reproductor debe ser de la forma $K = f + p$, donde f es la solución fundamental del operador $\frac{d^4}{dx^4}$, es decir, solución de la ecuación

$$\frac{d^4 f(x, y)}{dx^4} = \delta(x - y),$$

y p es una función perteneciente al espacio nulo del operador $\frac{d^4}{dx^4}$ (i.e. un polinomio cúbico) tal que los términos de frontera en (2.1.5) sean nulos. Así, al sustituir u por la función K en la ecuación (2.1.5) claramente se satisface la propiedad reproductora.

El objetivo es ahora encontrar la forma explícita de las funciones f y p . Por el ejemplo 1.3.2, tenemos que

$$\frac{d^3 f(x, y)}{dx^3} = H(x - y),$$

con H la función de Heavyside; de esta forma, integrando 3 veces se obtiene la solución fundamental

$$f(x, y) = \frac{1}{6}(x - y)_+^3,$$

donde $(x - y)_+ := \max\{x - y, 0\}$. La función p se propone como

$$p(x, y) = a_0 + a_1(x - y) + a_2(x - y)^2 + a_3(x - y)^3,$$

donde los coeficientes se escogen de forma que se anulen los términos de frontera en (2.1.5). Si escribimos $K_y(x) = K(x, y)$, estas condiciones se traducen en

$$\begin{aligned}
K_y(0) + K_y'''(0) &= 0 \\
K_y(1) - K_y'''(1) &= 0. \\
K_y''(0) &= 0 \\
K_y''(1) &= 0
\end{aligned}$$

Al sustituir la propuesta del Kernel, lo anterior se convierte en un sistema de 4×4 para los coeficientes del polinomio:

$$\begin{aligned}
a_0 - a_1y + a_2y^2 + a_3(6 - y^3) &= 0 \\
a_0 + a_1(1 - y) + a_2(1 - y)^2 + a_3((1 - y)^3 - 6) &= 1 - \frac{1}{6}(1 - y)^3 \\
a_2 - 3a_3y &= 0 \\
2a_2 + 6a_3(1 - y) &= y - 1
\end{aligned}$$

el cual tiene por solución

$$\begin{aligned}
a_0 &= \frac{1}{3}(y^4 - 2y^3 + 7y^2 - 6y + 3) \\
a_1 &= \frac{1}{3}(2y^3 - 3y^2 + 7y - 3) \\
a_2 &= -\frac{1}{2}y(1 - y) \\
a_3 &= -\frac{1}{6}(1 - y)
\end{aligned}$$

Así, juntando términos encontramos finalmente la expresión del kernel reproductor

$$K(x, y) = \frac{1}{6}((x - y)_+^3 + (y - 1)x^3 + (y^3 - 3y^2 + 2y)x) + xy + (1 - y)(1 - x).$$

Dada esta expresión, es evidente entonces que el interpolante dado por (2.1.4) es un polinomio cúbico por piezas tal que $\sigma \in C^2(0, 1)$, es decir, σ es un spline cúbico.

A continuación encontraremos una expresión alternativa para los splines cúbicos naturales, que será útil para motivar la teoría en dimensiones mayores. Por los resultados anteriores, un spline cúbico cualquiera en el intervalo $[a, b]$ se puede escribir de la forma

$$\sigma(x) = \sum_{j=1}^N \alpha_j (x - x_j)_+^3 + \sum_{j=0}^3 \beta_j x^j, \quad x \in [a, b]. \quad (2.1.6)$$

En el intervalo $[a, x_1]$ esta expresión se reduce a $\sigma(x) = \sum_{j=0}^3 \beta_j x^j$, y debido a que un spline cúbico natural es una recta en los intervalos exteriores, necesariamente $\beta_2 = \beta_3 = 0$. Así, (2.1.6) se reduce a

$$\sigma(x) = \sum_{j=1}^N \alpha_j (x - x_j)_+^3 + \beta_0 + \beta_1 x, \quad x \in [a, b]. \quad (2.1.7)$$

Para encontrar la expresión de σ en el intervalo $[x_N, b]$, simplemente hay que remover los símbolos $+$ de las funciones truncadas $(\cdot - x_j)_+^3$ en (2.1.7). Utilizando el teorema del binomio para expandir la potencia cúbica se tiene

$$\sum_{k=0}^3 \binom{3}{k} (-1)^{3-k} \left(\sum_{j=1}^N \alpha_j x_j^{3-k} \right) x^k + \beta_0 + \beta_1 x, \quad x \in [x_N, b], \quad (2.1.8)$$

y como esto representa una recta, los coeficientes tienen que cumplir

$$\sum_{j=1}^N \alpha_j = \sum_{j=1}^N \alpha_j x_j = 0. \quad (2.1.9)$$

Esta es una caracterización más de los splines cúbicos naturales, pero podemos decir aún más; utilizando la identidad $x_+^3 = (|x|^3 + x^3)/2$ y las restricciones anteriores tenemos que

$$\begin{aligned} \sigma(x) &= \sum_{j=1}^N \frac{\alpha_j}{2} |x - x_j|^3 + \sum_{j=1}^N \frac{\alpha_j}{2} (x - x_j)^3 + \beta_0 + \beta_1 x \\ &= \sum_{j=1}^N \frac{\alpha_j}{2} |x - x_j|^3 + \sum_{k=0}^3 \frac{1}{2} \binom{3}{k} (-1)^{3-k} \sum_{j=1}^N \alpha_j x_j^{3-k} x^k + \beta_0 + \beta_1 x \\ &= \sum_{j=1}^N \tilde{\alpha}_j |x - x_j|^3 + \tilde{\beta}_0 + \tilde{\beta}_1 x, \end{aligned}$$

donde los nuevos coeficientes están dados por

$$\begin{aligned} \tilde{\alpha}_j &= \frac{\alpha_j}{2}, \quad j \in \{1, \dots, N\} \\ \tilde{\beta}_0 &= \beta_0 - \frac{1}{2} \sum_{j=1}^N \alpha_j x_j^3 \\ \tilde{\beta}_1 &= \beta_1 + \frac{3}{2} \sum_{j=1}^N \alpha_j x_j^2 \end{aligned}$$

Teorema 2.1.4. *Todo spline cúbico natural σ se puede representar como*

$$\sigma(x) = \sum_{j=1}^N \alpha_j \phi(|x - x_j|) + p(x), \quad x \in \mathbb{R}, \quad (2.1.10)$$

donde $\phi(r) = r^3$, $r \geq 0$, y $p \in \mathbb{P}_1(\mathbb{R})$; los coeficientes α_j tienen que satisfacer (2.1.9). Inversamente, para cualquier conjunto $X = \{x_1, \dots, x_N\} \subset \mathbb{R}$ de puntos distintos y cualquier $u \in \mathbb{R}^N$, existe una función σ de la forma (2.1.10), que cumple (2.1.9) e interpola, i.e. $\sigma(x_j) = u_j$, $1 \leq j \leq N$.

2.1.2. Splines de Placa Delgada

En adelante supondremos que $m > d/2$ ⁴. Es posible generalizar el tratamiento de los splines cúbicos de la sección anterior a mayores dimensiones. En este caso, el problema de minimización no se trabaja en espacios Sobolev, sino en los llamados espacios de Beppo-Levi:

$$V^m(\mathbb{R}^d) = \{u \in C(\mathbb{R}^d) : D^\alpha u \in L^2(\mathbb{R}^d), \quad |\alpha| = m\}.$$

A diferencia de los espacios de Sobolev, en los espacios de Beppo-Levi sólo se requiere que las derivadas de orden igual a m estén en $L^2(\mathbb{R}^d)$. Para este espacio definimos el semi-producto interior

$$\langle u, v \rangle_{V^m(\mathbb{R}^d)} = \sum_{|\alpha|=m} \frac{m!}{\alpha!} \int_{\mathbb{R}^d} D^\alpha u D^\alpha v, \quad (2.1.11)$$

donde el factorial de un multi-índice α se define como $\alpha! = \alpha_1! \cdots \alpha_d!$. Nota que el espacio nulo de la semi-norma inducida por (2.1.11) es simplemente el espacio $\mathbb{P}_{m-1}(\mathbb{R}^d)$ de polinomios en d variables de grado menor a m .

Los coeficientes en (2.1.11) se escogieron de esa forma tomando como motivación el teorema del multinomio, el cual nos dice que podemos expresar $\|x\|^{2m}$ como $\sum_{|\alpha|=m} m! x^{2\alpha}/\alpha!$. De igual manera se puede expresar al laplaciano iterado Δ^m como $\sum_{|\alpha|=m} m! D^{2\alpha}/\alpha!$. Estos dos hechos serán importantes más adelante.

Por simplicidad, analizaremos el caso $m = d = 2$. En este caso, el problema de interpolación se puede enunciar como:

Problema. *Encontrar una función σ que cumpla las siguientes condiciones:*

1. $\sigma \in X_I$ donde $X_I = \{u \in V^2(\mathbb{R}^2) \mid u(x_i) = u_i, \quad i \in \mathcal{I}\}$, con $u_i \in \mathbb{R}$ datos.
2. $\sigma = \inf_{u \in X_I} |u|_{V^2(\mathbb{R}^2)}$

Al igual que en la sección anterior, la seminorma $|\cdot|_{V^2(\mathbb{R}^2)}$ es proporcional a la energía elástica, en este caso de una lámina delgada de metal. Por eso se puede interpretar a la gráfica del interpolante como la forma que tomaría una placa infinita de metal si se obliga a pasar por los puntos (x_i, u_i) . Por esta razón se les llama a las soluciones del problema splines de placa delgada.

Supondremos también que el conjunto de nodos $\{x_i\}_{i \in \mathcal{I}}$ contiene un conjunto de puntos $\{x_j\}_{j \in \mathcal{J}}$ que es $\mathbb{P}_1(\mathbb{R}^2)$ -solvente, esto quiere decir que si $p \in \mathbb{P}_1(\mathbb{R}^2)$ y $p(x_j) = 0$ para toda $j \in \mathcal{J}$, entonces $p = 0$. Además supondremos que $\{x_j\}_{j \in \mathcal{J}}$ ya tiene el menor número de elementos posible; en nuestro caso son 3 elementos ya que $\dim(\mathbb{P}_1(\mathbb{R}^2)) = 3$. Así, podemos modificar (2.1.11) para convertirlo en un producto interior genuino:

⁴Esta condición nos permite suponer que las funciones con las que se trabajarán son continuas

$$\ll u, v \gg_{V^2(\mathbb{R}^2)} = \langle u, v \rangle_{\mathbb{P}_1} + \langle u, v \rangle_{V^2(\mathbb{R}^2)} = \sum_{j \in \mathcal{J}} u(x_j)v(x_j) + \int_{\mathbb{R}^2} \left(\frac{\partial^2 u}{\partial x^2} \frac{\partial^2 v}{\partial x^2} + 2 \frac{\partial^2 u}{\partial x \partial y} \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} \frac{\partial^2 v}{\partial y^2} \right). \quad (2.1.12)$$

Aplicando dos veces la fórmula de integración por partes, y utilizando que las funciones decaen a cero en infinito, este producto interior se puede expresar como

$$\ll u, v \gg_{V^2(\mathbb{R}^2)} = \langle u, v \rangle_{\mathbb{P}_1} + \int_{\mathbb{R}^2} (\Delta^2 u)v, \quad (2.1.13)$$

donde Δ^2 es el bilaplaciano (segunda iteración del laplaciano), i.e. $\Delta^2 = \Delta\Delta = \partial_{xxxx} + 2\partial_{xyxy} + \partial_{yyyy}$. Es posible demostrar que el espacio nulo del operador Δ^2 es justamente $\mathbb{P}_1(\mathbb{R}^2)$, por lo que si V_0 es el subespacio de $V^2(\mathbb{R}^2)$ dado por

$$V_0 = \{v \in V^2(\mathbb{R}^2) : v(x_j) = 0 \quad j \in \mathcal{J}\},$$

entonces usando (2.1.13), se tiene que $V_0^\perp = \mathbb{P}_1(\mathbb{R}^2)$, o equivalentemente, que el espacio $V^2(\mathbb{R}^2)$ se puede descomponer como

$$V^2(\mathbb{R}^2) = V_0 \oplus \mathbb{P}_1(\mathbb{R}^2). \quad (2.1.14)$$

Ahora, por la unisolvencia de $\{x_j\}_{j \in \mathcal{J}}$ podemos escoger $p_j \in \mathbb{P}_1(\mathbb{R}^2)$, $j \in \mathcal{J}$, de manera que $p_j(x_k) = \delta_{jk}$ (i.e. $p_j(x_k)$ es 1 cuando $k = j$ y cero en otro caso) y $\text{span}\{p_j\}_{j \in \mathcal{J}} = \mathbb{P}_1(\mathbb{R}^2)$. A este conjunto de polinomios se le llama una base de Lagrange de $\mathbb{P}_1(\mathbb{R}^2)$ asociada a $\{x_j\}_{j \in \mathcal{J}}$. Así, la proyección $\Pi : V^2(\mathbb{R}^2) \rightarrow \mathbb{P}_1(\mathbb{R}^2)$ se puede expresar simplemente como

$$(\Pi v)(x) = \sum_{j \in \mathcal{J}} p_j(x)v(x_j), \quad v \in V^2(\mathbb{R}^2). \quad (2.1.15)$$

Por otro lado, se tiene

$$v(x) = (I - \Pi)v(x) + (\Pi v)(x) = \langle K_0(\cdot, x), v \rangle_{V_0} + \langle K_1(\cdot, x), v \rangle_{\mathbb{P}_1}, \quad (2.1.16)$$

donde K_0, K_1 son los kernels reproductores de V_0 y \mathbb{P}_1 respectivamente. Por la forma del producto interior $\langle \cdot, \cdot \rangle_{\mathbb{P}_1}$ es fácil ver que el kernel reproductor K_1 está dado por

$$K_1(y, x) = \sum_{j \in \mathcal{J}} p_j(y)p_j(x), \quad x, y \in \mathbb{R}^2, \quad (2.1.17)$$

en efecto, $K_1(\cdot, x) \in \mathbb{P}_1(\mathbb{R}^2)$ para toda $x \in \mathbb{R}^2$, y si $q \in \mathbb{P}_1(\mathbb{R}^2)$, se cumple

$$\langle q, K(\cdot, x) \rangle_{\mathbb{P}_1} = \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J}} q(x_i)p_j(x_i)p_j(x) = \sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{J}} \delta_{ij}q(x_i)p_j(x) = \sum_{j \in \mathcal{J}} q(x_j)p_j(x) = q(x), \quad \forall x \in \mathbb{R}^2. \quad (2.1.18)$$

Utilizando las ecuaciones (2.1.13) y (2.1.16) se sigue que el kernel reproductor K_0 debe satisfacer la ecuación diferencial

$$\Delta_y^2 h(y, x) = \delta(y - x) - \sum_{j \in \mathcal{J}} p_j(x)\delta(y - x_j), \quad (2.1.19)$$

la cual por ser lineal tiene como solución a

$$h(y, x) = \Phi(y - x) - \sum_{j \in \mathcal{J}} p_j(x) \Phi(y - x_j), \quad (2.1.20)$$

donde $\Phi(x) = \frac{1}{8\pi} \|x\|^2 \log(\|x\|)$ es la solución fundamental del bilaplaciano. Recuerda que la solución fundamental es la función que satisface $\Delta^2 \Phi(x) = \delta(x)$. Es posible demostrar que esta solución cumple $h(\cdot, x) \in V^2(\mathbb{R}^2)$ para $x \in \mathbb{R}^2$, por lo que para encontrar el kernel k_0 sóloamente hay que calcular la proyección de h sobre espacio V_0 :

$$K_0(\cdot, x) = (I - \Pi)h(\cdot, x) = h(\cdot, x) - \sum_{j \in \mathcal{J}} h(x_j, x) p_j(\cdot).$$

En consecuencia, el kernel reproductor del espacio $V^2(\mathbb{R}^2)$ se escribe como

$$\begin{aligned} K(y, x) &= K_0(y, x) + K_1(y, x) \\ &= \Phi(y - x) - \sum_{j \in \mathcal{J}} p_j(x) \Phi(y - x) - \sum_{j \in \mathcal{J}} p_j(y) \Phi(x - x_j) + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} p_i(x) p_j(y) \Phi(x - x_j) + \sum_{j \in \mathcal{J}} p_j(x) p_j(y). \end{aligned} \quad (2.1.21)$$

Al igual que en la sección anterior, se tiene que el spline interpolante σ es la única función que pertenece a $X_I \cap X_I^{0\perp}$. Además se cumple también $X_I^{0\perp} = \text{span}\{K(\cdot, x_i)\}_{i \in \mathcal{I}}$, pero nota que en este caso

$$\text{span}\{K(\cdot, x_i)\}_{i \in \mathcal{I}} = \text{span}\{K(\cdot, x_i)\}_{i \in \mathcal{I} \setminus \mathcal{J}} \oplus \text{span}\{K(\cdot, x_j)\}_{j \in \mathcal{J}} = \text{span}\{K_0(\cdot, x_i)\}_{i \in \mathcal{I} \setminus \mathcal{J}} \oplus \mathbb{P}_1(\mathbb{R}^2).$$

Por esta razón, se sigue que σ es de la forma

$$\sigma(x) = \sum_{i \in \mathcal{I} \setminus \mathcal{J}} \gamma_i K_0(x, x_i) + \sum_{j \in \mathcal{J}} u_j p_j(x), \quad (2.1.22)$$

donde los coeficientes γ_i , se encuentran resolviendo el sistema

$$\sum_{j \in \mathcal{I} \setminus \mathcal{J}} K_0(x_i, x_j) \gamma_j = u_i - \sum_{j \in \mathcal{J}} u_j p_j(x_i), \quad \forall i \in \mathcal{I} \setminus \mathcal{J}. \quad (2.1.23)$$

Aunque anterior resuelve por completo el problema de interpolación, no se trabaja en la práctica con ese sistema debido a que la expresión del kernel reproductor (2.1.21) es complicada. De las ecuaciones (2.1.22) y (2.1.21) se sigue que el spline interpolante σ tiene una única representación de la forma

$$\sigma(x) = \sum_{i \in \mathcal{I}} \lambda_i \Phi(x - x_i) + q(x), \quad \forall x \in \mathbb{R}^2, \quad (2.1.24)$$

donde $q \in \mathbb{P}_1(\mathbb{R}^2)$. Nota que en esta nueva representación hay $N + 3$ coeficientes (N en la suma y 3 del polinomio) y sóloamente N condiciones de interpolación, por lo que hay que encontrar 3 condiciones extra para los coeficientes. Para ello, observa que por la propiedad reproductora, se tiene

$$\langle v, w \rangle_{V^2(\mathbb{R}^2)} = \sum_{i \in \mathcal{I}} \lambda_i v(x_i), \quad \forall v \in V^2(\mathbb{R}^2), w \in V_0^\perp,$$

lo cual implica que

$$\sum_{i \in \mathcal{I}} \lambda_i p(x_i) = 0, \quad \forall p \in \mathbb{P}_1(\mathbb{R}^2). \quad (2.1.25)$$

La discusión anterior se puede resumir en el siguiente teorema

Teorema 2.1.5. *Todo spline de placa delgada σ se puede representar como*

$$\sigma(x) = \sum_{i=1}^N \lambda_i \phi(\|x - x_i\|) + q(x), \quad x \in \mathbb{R}^2, \quad (2.1.26)$$

donde $\phi(r) = r^2 \log(r)$, $r \geq 0$, y $q \in \mathbb{P}_1(\mathbb{R}^2)$; los coeficientes λ_i tienen que satisfacer (2.1.25). Inversamente, para cualquier conjunto $X = \{x_1, \dots, x_N\} \subset \mathbb{R}^2$ de puntos que contenga un conjunto $\mathbb{P}_1(\mathbb{R}^2)$ -solvente y cualquier $u \in \mathbb{R}^N$, existe una función σ de la forma (2.1.26), que cumple (2.1.25) e interpola, i.e. $\sigma(x_i) = u_i$, $1 \leq i \leq N$.

En dimensiones mayores la discusión anterior es válida, sólomente haciendo pequeñas modificaciones. En la ecuación (2.1.13) es necesario sustituir el bilaplaciano Δ^2 por la m -ésima iteración del laplaciano Δ^m ; el espacio nulo de este operador ya no es $\mathbb{P}_1(\mathbb{R}^2)$ sino $\mathbb{P}_{m-1}(\mathbb{R}^d)$. En consecuencia hay que sustituir la solución fundamental Φ en (2.1.21) por la función

$$\Phi(x) = \begin{cases} \alpha_{m,d} \|x\|^{2m-d} \log(\|x\|), & \text{si } 2m \geq d \text{ y } d \text{ es par,} \\ \beta_{m,d} \|x\|^{2m-d}, & \text{en otro caso,} \end{cases} \quad (2.1.27)$$

para $x \in \mathbb{R}^d$, y con coeficientes dados por

$$\alpha_{m,d} = \frac{(-1)^{d/2+1}}{2^{2m-1} \pi^{d/2} (m-1)! (m-d/2)!}, \quad (2.1.28)$$

$$\beta_{m,d} = \frac{(-1)^m \Gamma(d/2 - m)}{2^{2m} \pi^{d/2} (m-1)!}. \quad (2.1.29)$$

Es muy interesante notar que al utilizar estas funciones con $m = d = 1$, se reduce todo al caso de la sección anterior, es decir a los splines cúbicos naturales.

Es importante resaltar que en esta sección, a diferencia de la anterior, se trabajó en todo el espacio; esto es porque podemos encontrar fácilmente una solución a la ecuación (2.1.19) en todo \mathbb{R}^2 . Si quisiéramos en contraste trabajar en un conjunto abierto y acotado $\Omega \subseteq \mathbb{R}^2$ cualquiera, para encontrar el kernel reproductor de $V^2(\Omega)$ tendríamos que resolver el problema

$$\Delta^2 U(x) = 0, \quad x \in \Omega,$$

con las condiciones

$$\begin{aligned}\frac{\partial^2}{\partial x_1^2}\Psi(x) &= \frac{\partial^3}{\partial x_1^3}\Psi(x) = \frac{\partial^3}{\partial x_2\partial x_1^2}\Psi(x) = 0, \\ \frac{\partial^4}{\partial x_2^2\partial x_1^2}\Psi(x) &= \frac{\partial^2}{\partial x_2^2}\Psi(x) = \frac{\partial^3}{\partial x_2^3}\Psi(x) = 0,\end{aligned}\quad x \in \partial\Omega,$$

donde $\Psi = \Phi + U$, con Φ la solución fundamental (2.1.27). Es virtualmente imposible encontrar una solución analítica a dicho problema para un conjunto arbitrario Ω y por esta razón fue necesario trabajar en todo \mathbb{R}^2 .

2.2. Funciones Positivas Definidas

Definición 2.2.1. Una función $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ es llamada radial, si existe una función $\phi : [0, \infty) \rightarrow \mathbb{R}$ tal que:

$$\Phi(x) = \phi(\|x\|),$$

donde $\|\cdot\|$ es la norma Euclidiana en \mathbb{R}^d .

Tomando como motivación las secciones anteriores, podríamos proponer para un problema de interpolación con centros $\{x_1, \dots, x_N\} \subseteq \mathbb{R}^d$, un interpolante de la forma

$$\sigma(x) = \sum_{j=1}^N \lambda_j \Phi(x - x_j), \quad (2.2.1)$$

donde $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ es, por ejemplo, una función radial. Las condiciones de interpolación $\sigma(x_j) = u_j$, $j \in \{1, \dots, N\}$, conducen a un sistema lineal, el cual tiene solución si y sólo si la matriz de interpolación con componentes

$$A_{jk} = \Phi(x_j - x_k), \quad j, k \in \{1, \dots, N\}, \quad (2.2.2)$$

es invertible. Un concepto importante para analizar este problema es el de función positiva definida. Antes de dar la definición, recordemos del álgebra lineal el concepto de matriz positiva definida.

Definición 2.2.2. Diremos que una matriz real simétrica A es semidefinida positiva si su forma cuadrática asociada es no-negativa, i.e.,

$$\sum_{j=1}^N \sum_{k=1}^N c_j c_k a_{jk} \geq 0 \quad (2.2.3)$$

para $c = (c_1, \dots, c_N)^T \in \mathbb{R}^N$. En particular, si el único vector c que satisface la igualdad en (2.2.3) es el vector cero, entonces diremos que A es definida positiva.

Definición 2.2.3. Una función continua $\Phi : \mathbb{R}^d \rightarrow \mathbb{C}$ es positiva definida si

$$\sum_{j=1}^N \sum_{k=1}^N c_j \bar{c}_k \Phi(x_j - x_k) \geq 0, \quad (2.2.4)$$

para cualquier conjunto $\{x_1, \dots, x_N\} \subset \mathbb{R}^d$ de N puntos distintos por pares y $c = (c_1, \dots, c_N)^T \in \mathbb{C}^N$. La función será estrictamente positiva definida si (2.2.4) es cero sólo cuando $c = 0$.

Una propiedad importante de las matrices positivas definidas, es que sus eigenvalores son positivos, y por lo tanto son matrices invertibles. Así, si en (2.2.1) la función Φ es estrictamente positiva definida, el problema de interpolación tendrá solución.

Ejemplo 2.2.4. La función dada por $\Phi(x) = e^{ix \cdot y}$ para $y \in \mathbb{R}^d$ fijo, es positiva definida en \mathbb{R}^d . En efecto, si $c = (c_1, \dots, c_N)^T \in \mathbb{C}^N$ es un vector cualquiera, se tiene

$$\begin{aligned} \sum_{j=1}^N \sum_{k=1}^N c_j \overline{c_k} \Phi(x_j - x_k) &= \sum_{j=1}^N \sum_{k=1}^N c_j \overline{c_k} e^{i(x_j - x_k) \cdot y} \\ &= \sum_{j=1}^N c_j e^{ix_j \cdot y} \sum_{k=1}^N \overline{c_k} e^{-ix_k \cdot y} \\ &= \left| \sum_{j=1}^N c_j e^{ix_j \cdot y} \right|^2 \geq 0. \end{aligned}$$

Aunque nosotros estamos interesados en problemas con datos y coeficientes reales, la definición de función positiva definida incluye el caso de funciones que toman valores en los complejos. Esto simplifica algunas demostraciones y permite más naturalmente el uso de técnicas como la transformada de Fourier. Algunas propiedades importantes de las funciones positivas definidas se enuncian en el siguiente teorema.

Teorema 2.2.5. Supongamos que Φ es una función positiva definida. Entonces se cumple:

1. $\Phi(0) \geq 0$.
2. $\Phi(-x) = \overline{\Phi(x)}$ para cualquier $x \in \mathbb{R}^d$.
3. Φ está acotada. En particular, $|\Phi(x)| \leq \Phi(0)$ para toda $x \in \mathbb{R}^d$.
4. $\Phi(0) = 0$ si y sólo si Φ es la función cero.
5. Si Φ_1, \dots, Φ_n son funciones positivas definidas y $c_j \geq 0, 1 \leq j \leq n$, entonces $\Phi = \sum_{j=1}^n c_j \Phi_j$ es también positiva definida. Si alguna Φ_j es estrictamente positiva definida y su correspondiente c_j es positiva, entonces Φ es también estrictamente positiva definida.
6. El producto de funciones positivas definidas es a su vez una función positiva definida.

Ejemplo 2.2.6. La función coseno es positiva definida en \mathbb{R} . Esto se sigue del ejemplo 2.2.4, la propiedad 5 del teorema anterior y la identidad

$$\cos(x) = \frac{1}{2}(e^{ix} + e^{-ix}).$$

La segunda propiedad en el teorema 2.2.5 implica que una función positiva definida que toma valores reales es necesariamente par. La siguiente proposición nos dice que también nos podemos restringir a vectores reales en la forma cuadrática.

Proposición 2.2.7. Una función continua $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ es positiva definida si y sólo si es par y se cumple

$$\sum_{j=1}^N \sum_{k=1}^N c_j c_k \Phi(x_j - x_k) \geq 0, \quad (2.2.5)$$

para cualquier conjunto $\{x_1, \dots, x_N\} \subset \mathbb{R}^d$ de N puntos distintos por pares y $c = (c_1, \dots, c_N)^T \in \mathbb{R}^N$.

Supongamos ahora que $\Phi \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ es una función con transformada de Fourier integrable. Entonces por la fórmula de inversión se tiene

$$\Phi(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{\Phi}(\xi) e^{ix \cdot \xi} d\xi.$$

Así, una forma cuadrática que involucre a Φ se puede expresar como

$$\begin{aligned} \sum_{j=1}^N \sum_{k=1}^N c_j \overline{c_k} \Phi(x_j - x_k) &= (2\pi)^{-d/2} \sum_{j=1}^N \sum_{k=1}^N c_j \overline{c_k} \int_{\mathbb{R}^d} \hat{\Phi}(\xi) e^{i(x_j - x_k) \cdot \xi} d\xi \\ &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{\Phi}(\xi) \left| \sum_{j=1}^N c_j e^{ix_j \cdot \xi} \right|^2 d\xi. \end{aligned}$$

Por esto, si $\hat{\Phi}$ es no negativa, entonces la función Φ es positiva definida. El siguiente teorema muestra que en realidad estas condiciones son equivalentes; fue demostrado por Bochner en 1933 y es uno de los resultados más célebres sobre funciones positivas definidas.

Teorema 2.2.8 (Bochner). Una función continua $\Phi : \mathbb{R}^d \rightarrow \mathbb{C}$ es positiva definida si y sólo si es la transformada de Fourier de una función $\omega \in L^1(\mathbb{R}^d)$ no negativa y acotada ⁵, i.e.

$$\Phi(x) = \hat{\omega}(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \omega(\xi) e^{-ix \cdot \xi} d\xi.$$

⁵En realidad se trata de una medida de Borel finita y no negativa, pero éste es un detalle técnico que no será importante aquí.

Observación 2.2.9. Por teorema de Bochner se puede interpretar a la función Φ del ejemplo 2.2.4 como la **función positiva definida fundamental**, ya que cualquier otra función positiva definida se obtiene como una combinación lineal (infinita) de Φ .

En un problema de interpolación nos interesa obviamente con funciones estrictamente positivas definidas. De la simetría entre la transformada de Fourier y la transformada inversa se puede encontrar un criterio para verificar si una función es estrictamente positiva definida, el cual se enuncia en el siguiente teorema.

Teorema 2.2.10. Supongamos que $\Phi \in L^1(\mathbb{R}^d)$ es una función continua. Entonces Φ es estrictamente positiva definida si y sólo si está acotada y su transformada de Fourier es no negativa y no es idénticamente cero.

Ejemplo 2.2.11. La gaussiana $\Phi(x) = e^{-\alpha\|x\|^2}$, $\alpha > 0$, es estrictamente positiva definida en \mathbb{R}^d . Esto se debe a que la transformada de Fourier de una gaussiana es esencialmente otra gaussiana. De forma más explícita se tiene

$$\hat{\Phi}(\xi) = (2\alpha)^{-d/2} e^{-\frac{\|\xi\|^2}{4\alpha}} > 0.$$

El teorema 2.2.10, aunque útil, tiene dos grandes limitaciones:

1. No contempla funciones que son muy importantes en la teoría de funciones de base radial (e.g. spline de placa delgada), simplemente por no ser estrictamente positivas definidas. De hecho, las gaussianas del ejemplo anterior son las únicas funciones positivas que son radiales y estrictamente positivas definidas en \mathbb{R}^d para toda d .
2. El teorema supone que la función es integrable en \mathbb{R}^d , lo cual es muy restrictivo y deja fuera a funciones de base radial muy importantes, tales como los multicuádricos $\Phi(x) = \sqrt{c^2 + \|x\|^2}$.

La solución a estos problemas es trabajar con la transformada generalizada de Fourier en el espacio de Schwartz, además de generalizar el concepto de función positiva definida. Esto nos permitirá de manera más general a (2.2.1) formar interpolantes de la forma

$$\sigma(x) = \sum_{j=1}^N \lambda_j \Phi(x - x_j) + \sum_{k=1}^Q \alpha_k p_k(x), \quad (2.2.6)$$

donde $Q = \dim(\mathbb{P}_{m-1}(\mathbb{R}^d))$ y p_1, \dots, p_Q es una base de $\mathbb{P}_{m-1}(\mathbb{R}^d)$. Las condiciones de interpolación

$$\sigma(x_j) = u_j, \quad j \in \{1, \dots, N\},$$

se completan con las condiciones de momentos

$$\sum_{j=1}^N \lambda_j p_k(x_j) = 0, \quad k \in \{1, \dots, Q\}.$$

Resolver este problema es equivalente a resolver el sistema lineal

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \alpha \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix}, \quad (2.2.7)$$

donde A y P son matrices de $N \times N$ y $N \times Q$ respectivamente, cuyas componentes son

$$A_{ij} = \Phi(x_i - x_j), \quad P_{ij} = p_j(x_i).$$

Definición 2.2.12. *Se dice que una función par y continua $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ es condicionalmente positiva definida de orden m , si la forma cuadrática*

$$\sum_{j=1}^N \sum_{k=1}^N c_j c_k \Phi(x_j - x_k),$$

es positiva para cualquier conjunto de centros $\{x_1, \dots, x_N\} \subset \mathbb{R}^d$ distintos por pares, y cualquier $c \in \mathbb{R}^N \setminus \{0\}$ que satisfaga

$$\sum_{j=1}^N c_j p(x_j) = 0,$$

para todos los polinomios de grado menor a m .

Observa que una función condicionalmente positiva definida de orden m , es también condicionalmente positiva definida de orden $l \geq m$. También se cumple que si una función es condicionalmente positiva definida de orden m en \mathbb{R}^d , entonces es condicionalmente positiva definida de orden m en \mathbb{R}^k , para $k \leq d$.

El que una función Φ sea condicionalmente positiva definida de orden m , se puede interpretar como que la matriz A de componentes $A_{jk} = \Phi(x_j - x_k)$ es positiva definida en el espacio de vectores $c \in \mathbb{R}^N$ tales que

$$\sum_{j=1}^N c_j p_l(x_j) = 0, \quad 1 \leq l \leq Q = \dim \mathbb{P}_{m-1}(\mathbb{R}^d).$$

En este sentido, A es positiva definida en el espacio de vectores c que son “perpendiculares” a los polinomios. Así, si en (2.2.6) la función Φ es condicionalmente estrictamente positiva definida de orden m y el conjunto de centros contiene un subconjunto $\mathbb{P}_{m-1}(\mathbb{R}^d)$ -solvente, el problema de interpolación tendrá solución (la condición de unisolvencia es para asegurar la unicidad).

El resultado análogo al teorema de Bochner para funciones condicionalmente positivas definidas se expresa en términos de la transformada de Fourier generalizada.

Teorema 2.2.13 (Madych & Nelson). *Supongamos que Φ es una función continua que crece a lo más como un polinomio y que posee una transformada de Fourier generalizada $\hat{\Phi}$ de orden m , que es continua en $\mathbb{R}^d \setminus \{0\}$. Entonces Φ es estrictamente condicionalmente positiva definida de orden m si y sólo si $\hat{\Phi}$ no se anula y es no negativa.*

2.3. Espacios Nativos

En las secciones anteriores se explicó el procedimiento para encontrar el kernel reproductor (y por consiguiente el interpolante) dado un espacio de Hilbert. En esta sección se analizará el otro caso, es decir dado un kernel radial construir su espacio de Hilbert asociado: su *espacio nativo*.

Primero enunciaremos un resultado que muestra la conexión entre espacios de Hilbert con kernel reproductor y los kernels positivos definidos. El concepto de kernel positivo definido es una pequeña generalización de la noción de función positiva definida. Esencialmente hay que sustituir $\Phi(x_j - x_k)$ por $\Phi(x_j, x_k)$ en la definición 2.2.3.

Teorema 2.3.1. *Si \mathcal{H} un espacio de Hilbert con kernel reproductor $\Phi : \Omega \times \Omega \rightarrow \mathbb{R}$, entonces Φ es positivo definido. Más aún, Φ es estrictamente positivo definido si y sólo si los funcionales de evaluación δ_x son linealmente independientes en \mathcal{H}^* .*

Demostración. Como el kernel Φ toma valores reales y es simétrico, es posible considerar sólo coeficientes reales en la forma cuadrática. Para $x_1, \dots, x_N \in \Omega$ distintos y $\alpha \in \mathbb{R}^N$ tenemos

$$\sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \Phi(x_j, x_k) = \left(\sum_{j=1}^N \alpha_j \delta_{x_j}, \sum_{k=1}^N \alpha_k \delta_{x_k} \right)_{\mathcal{H}^*} = \left\| \sum_{j=1}^N \alpha_j \delta_{x_j} \right\|_{\mathcal{H}^*}^2 \geq 0.$$

La última expresión sólo puede ser cero si los funcionales de evaluación son linealmente dependientes. \square

Por definición de kernel reproductor, sabemos que \mathcal{H} contiene a todas las funciones de la forma $f = \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j)$ siempre y cuando $x_j \in \Omega$. Además, por el teorema 1.4.2 sabemos que

$$\|f\|_{\mathcal{H}}^2 = \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k (\Phi(\cdot, x_j), \Phi(\cdot, x_k))_{\mathcal{H}} = \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \Phi(x_j, x_k).$$

Tomando lo anterior como motivación, definimos el espacio

$$H_{\Phi}(\Omega) := \text{span}\{\Phi(\cdot, y) : y \in \Omega\},$$

y lo dotamos de la forma bilineal

$$\left(\sum_{j=1}^N \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^M \beta_k \Phi(\cdot, y_k) \right)_{\Phi} := \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \Phi(x_j, y_k).$$

Teorema 2.3.2. *Si Φ es un kernel simétrico y positivo definido entonces $(\cdot, \cdot)_{\Phi}$ define un producto interno en $H_{\Phi}(\Omega)$. Más aún, $H_{\Phi}(\Omega)$ es un espacio pre-Hilbert con kernel reproductor Φ .*

Debido a que el teorema anterior establece que $H_{\Phi}(\Omega)$ es un espacio pre-Hilbert, i.e. no es necesariamente completo, el primer candidato para un espacio de Hilbert con kernel reproductor Φ es la completación $\mathcal{H}_{\Phi}(\Omega)$ de $H_{\Phi}(\Omega)$ respecto a la norma $\|\cdot\|_{\Phi}$. El problema es que los elementos

de la completación quedan simplemente como elementos abstractos de un espacio de Hilbert. Para representar a estos elementos como funciones, consideremos primero el mapeo lineal

$$R : \mathcal{H}_\Phi(\Omega) \rightarrow C(\Omega), \quad R(f)(x) := (f, \Phi(\cdot, x))_\Phi.$$

Para ver que el mapeo está bien definido, nota que

$$|Rf(x) - Rf(y)| = (f, \Phi(\cdot, x) - \Phi(\cdot, y))_\Phi \leq \|f\|_\Phi \|\Phi(\cdot, x) - \Phi(\cdot, y)\|_\Phi,$$

y

$$\|\Phi(\cdot, x) - \Phi(\cdot, y)\|_\Phi^2 = \Phi(x, x) + \Phi(y, y) - 2\Phi(x, y),$$

por lo que $Rf(x) \rightarrow Rf(y)$ cuando $x \rightarrow y$, i.e. $R(f) \in C(\Omega)$. Además, tenemos que este mapeo es inyectivo y $Rf(x) = f(x)$ para toda $x \in \Omega$ y $f \in H_\Phi(\Omega)$. Así, podemos finalmente dar la definición de un espacio nativo.

Definición 2.3.3. *El espacio nativo correspondiente al kernel simétrico y positivo definido $\Phi : \Omega \times \Omega \rightarrow \mathbb{R}$ se define como*

$$\mathcal{N}_\Phi(\Omega) := R(\mathcal{H}_\Phi(\Omega)).$$

Este espacio está equipado con el producto interior

$$(f, g)_{\mathcal{N}_\Phi(\Omega)} := (R^{-1}f, R^{-1}g)_\Phi.$$

Ciertamente, el espacio nativo así definido es un espacio de Hilbert de funciones continuas en Ω con kernel reproductor Φ . Debido a que $\Phi(\cdot, x)$ es un elemento de $H_\Phi(\Omega)$ para $x \in \Omega$, no cambia bajo R , por lo que

$$f(x) = (R^{-1}f, \Phi(\cdot, x))_\Phi = (f, \Phi(\cdot, x))_{\mathcal{N}_\Phi(\Omega)}. \quad (2.3.1)$$

Teorema 2.3.4. *Si $\Phi : \Omega \times \Omega \rightarrow \mathbb{R}$ es un kernel simétrico y positivo definido entonces su espacio nativo asociado $\mathcal{N}_\Phi(\Omega)$ es un espacio de Hilbert con kernel reproductor Φ . Este espacio es además único, i.e. si \mathcal{G} es un espacio de Hilbert de funciones $f : \Omega \rightarrow \mathbb{R}$ con kernel reproductor Φ , entonces \mathcal{G} es el espacio nativo $\mathcal{N}_\Phi(\Omega)$ y los productos interiores coinciden.*

En el caso especial en el que $\Omega = \mathbb{R}^d$ y Φ es invariante bajo traslaciones, el resultado anterior nos permite dar la siguiente caracterización del espacio nativo utilizando la transformada de Fourier.

Teorema 2.3.5. *Supongamos que $\Phi \in C(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ es una función real y positiva definida. Definimos al espacio*

$$\mathcal{G} := \left\{ f \in L^2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \frac{\hat{f}}{\sqrt{\hat{\Phi}}} \in L^2(\mathbb{R}^d) \right\},$$

y lo equipamos con la forma bilineal

$$(f, g)_\mathcal{G} := (2\pi)^{-d/2} \left(\frac{\hat{f}}{\sqrt{\hat{\Phi}}}, \frac{\hat{g}}{\sqrt{\hat{\Phi}}} \right)_{L^2(\mathbb{R}^d)} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega) \overline{\hat{g}(\omega)}}{\hat{\Phi}(\omega)} d\omega.$$

Entonces \mathcal{G} es un espacio de Hilbert con producto interior $(\cdot, \cdot)_\mathcal{G}$ y kernel reproductor $\Phi(\cdot - \cdot)$. Por lo tanto, \mathcal{G} es el espacio nativo de Φ en \mathbb{R}^d , i.e. $\mathcal{G} = \mathcal{N}_\Phi(\mathbb{R}^d)$. En particular, cualquier $f \in \mathcal{N}_\Phi(\mathbb{R}^d)$ se puede recuperar de su transformada de Fourier $\hat{f} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$.

Ejemplo 2.3.6. Recordemos que el espacio de Sobolev de orden s para $s > d/2$ ⁶ se puede definir como

$$H^s(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \hat{u}(\cdot)(1 + \|\cdot\|^2)^{s/2} \in L^2(\mathbb{R}^d)\}.$$

Supongamos ahora que $\Phi \in L^1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ satisface

$$c_1(1 + \|\omega\|^2)^{-s} \leq \hat{\Phi}(\omega) \leq c_2(1 + \|\omega\|^2)^{-s}, \quad \omega \in \mathbb{R}^d,$$

con $s > d/2$ y $c_1 \leq c_2$ positivos, es decir Φ tiene una transformada de Fourier que decae algebraicamente. Entonces por el teorema 2.3.5 el espacio nativo $\mathcal{N}_\Phi(\mathbb{R}^d)$ correspondiente a Φ corresponde con el espacio de Sobolev $H^s(\mathbb{R}^d)$ y las normas son equivalentes. Por esta razón los espacios nativos son en cierta forma una generalización de los espacios de Sobolev.

A continuación procederemos a construir el espacio nativo para un kernel condicionalmente positivo definido. De manera análoga al caso anterior consideramos primero el espacio

$$H_\Phi(\Omega) := \left\{ \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) : N \in \mathbb{N}, \alpha \in \mathbb{R}^N, x_1, \dots, x_N \in \Omega, \text{ con } \sum_{j=1}^N \alpha_j p(x_j) \text{ para todo } p \in \mathcal{P}^7 \right\},$$

el cual se convierte en un espacio pre-Hilbert al introducir el producto interior

$$\left(\sum_{j=1}^N \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^M \beta_k \Phi(\cdot, y_k) \right)_\Phi := \sum_{j=1}^N \sum_{k=1}^M \alpha_j \beta_k \Phi(x_j, y_k).$$

Para interpretar a los elementos de la completación $\mathcal{H}_\Phi(\Omega)$ del espacio $H_\Phi(\Omega)$ como funciones, escojamos a un conjunto \mathcal{P} -unisolvante $\{\xi_j\}_{j \in \mathcal{J}} \subseteq \Omega$ con $|\mathcal{J}| = \dim \mathcal{P}$, y sea $\{p_j\}_{j \in \mathcal{J}}$ su base de Lagrange asociada. Definimos primero la proyección $\Pi_{\mathcal{P}} : C(\Omega) \rightarrow \mathcal{P}$

$$\Pi_{\mathcal{P}}(f) = \sum_{j \in \mathcal{J}} f(\xi_j) p_j,$$

y a la función $G : \Omega \times \Omega \rightarrow \mathbb{R}$

$$G(y, x) = ((I - \Pi_{\mathcal{P}})\Phi(y, \cdot))(x) = \Phi(y, x) - \sum_{j \in \mathcal{J}} p_j(x) \Phi(y, \xi_j). \quad (2.3.2)$$

Con esto podemos definir el mapeo lineal

$$R : \mathcal{H}_\Phi(\Omega) \rightarrow C(\Omega), \quad R(f)(x) := (f, G(\cdot, x))_\Phi,$$

el cual está bien definido ya que

⁶Esta condición es sólo para asegurar que las funciones son continuas.

⁷El espacio \mathcal{P} en el caso de un kernel correspondiente a una función condicionalmente positiva definida de orden m es simplemente $\mathbb{P}_{m-1}(\mathbb{R}^d)$.

$$|Rf(x) - Rf(y)| \leq \|f\|_{\Phi} \|G(\cdot, x) - G(\cdot, y)\|_{\Phi},$$

y

$$\begin{aligned} \|G(\cdot, x) - G(\cdot, y)\|_{\Phi}^2 &= \Phi(x, x) + \Phi(y, y) - 2\Phi(x, y) - 2 \sum_{j \in \mathcal{J}} (p_j(x) - p_j(y)) (\Phi(x, \xi_j) - \Phi(y, \xi_j)) \\ &\quad + \sum_{j, k \in \mathcal{J}} (p_j(x) - p_j(y)) (p_k(x) - p_k(y)) \Phi(\xi_j, \xi_k), \end{aligned}$$

lo cual implica que $Rf(x) \rightarrow Rf(y)$ cuando $x \rightarrow y$. Al igual que en el caso anterior, el mapeo R es inyectivo, sin embargo en este caso se cumple para $f \in H_{\Phi}(\Omega)$

$$Rf(x) = (f, G(\cdot, x))_{\Phi} = f(x) - \sum_{j \in \mathcal{J}} p_j(x) f(\xi_j) = f(x) - \Pi_{\mathcal{P}} f(x),$$

por lo que hay que modificar ligeramente la definición de espacio nativo.

Definición 2.3.7. *El espacio nativo correspondiente a un kernel simétrico $\Phi : \Omega \times \Omega \rightarrow \mathbb{R}$ que es condicionalmente positivo definido respecto a \mathcal{P} se define como*

$$\mathcal{N}_{\Phi}(\Omega) := R(\mathcal{H}_{\Phi}(\Omega)) \oplus \mathcal{P}.$$

Este espacio está equipado con el semi-producto interior

$$(f, g)_{\mathcal{N}_{\Phi}(\Omega)} = (R^{-1}(f - \Pi_{\mathcal{P}} f), R^{-1}(g - \Pi_{\mathcal{P}} g))_{\Phi},$$

y en particular se cumple

$$f(x) = \Pi_{\mathcal{P}} f(x) + (f, G(\cdot, x))_{\mathcal{N}_{\Phi}(\Omega)}, \quad f \in \mathcal{N}_{\Phi}(\Omega),$$

donde G es la función definida en (2.3.2).

La forma bilinear $(\cdot, \cdot)_{\mathcal{N}_{\Phi}(\Omega)}$ no es un producto interior; su espacio nulo es precisamente \mathcal{P} . Así, es posible definir un producto en el espacio nativo $\mathcal{N}_{\Phi}(\Omega)$ mediante

$$(f, g) = (f, g)_{\mathcal{N}_{\Phi}(\Omega)} + \sum_{j \in \mathcal{J}} f(\xi_j) g(\xi_j). \quad (2.3.3)$$

El siguiente teorema proporciona más información acerca de la conexión que hay con los kernels reproductores

Teorema 2.3.8. *Sea $X_{\Phi}(\Omega)$ el subespacio del espacio nativo $\mathcal{N}_{\Phi}(\Omega)$ dado por*

$$X_{\Phi}(\Omega) = \{f \in \mathcal{N}_{\Phi}(\Omega) : f(\xi_j) = 0, \quad j \in \mathcal{J}\} = R(\mathcal{H}_{\Phi}(\Omega)).$$

Entonces en este espacio la forma bilinear $(\cdot, \cdot)_{\mathcal{N}_{\Phi}(\Omega)}$ se convierte en un producto interior y el espacio cuenta con el kernel reproductor

$$K(x, y) = \Phi(x, y) - \sum_{j \in \mathcal{J}} p_j(x) \Phi(\xi_j, y) - \sum_{j \in \mathcal{J}} p_j(y) \Phi(x, \xi_j) + \sum_{j, k \in \mathcal{J}} p_j(x) p_k(y) \Phi(\xi_j, \xi_k).$$

Más aún, si \mathcal{G} es un espacio con kernel reproductor K cuyo producto interior tenga espacio nulo \mathcal{P} entonces \mathcal{G} es el espacio nativo correspondiente a Φ en Ω , i.e. el espacio nativo es único.

El siguiente teorema es el resultado análogo al teorema 2.3.5 para funciones condicionalmente positivas definidas.

Teorema 2.3.9. *Supongamos que Φ es una función par, condicionalmente positiva definida de orden m , cuya transformada de Fourier generalizada de orden m es continua en $\mathbb{R}^d \setminus \{0\}$. Sea \mathcal{G} el espacio de las funciones $f \in C(\mathbb{R}^d)$ que crecen a lo más como un polinomio y tienen transformada generalizada de Fourier \hat{f} de orden $m/2$ que satisface $\hat{f}/\sqrt{\hat{\Phi}} \in L^2(\mathbb{R}^d)$. Definimos en este espacio la forma bilineal*

$$(f, g)_{\mathcal{G}} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega) \overline{\hat{g}(\omega)}}{\hat{\Phi}(\omega)} d\omega.$$

Entonces \mathcal{G} es el espacio nativo correspondiente a Φ , i.e. $\mathcal{G} = \mathcal{N}_{\Phi}(\mathbb{R}^d)$, y los productos interiores coinciden. Además, toda $f \in \mathcal{N}_{\Phi}(\Omega)$ se puede representar como

$$f(x) = \Pi_{\mathcal{P}} f(x) + (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega) \left(e^{ix \cdot \omega} - \sum_{j \in \mathcal{J}} p_j(x) e^{i\xi_j \cdot \omega} \right) d\omega.$$

Ejemplo 2.3.10. *En este ejemplo encontraremos el espacio nativo correspondiente al spline de placa delgada $\Phi(x) = \|x\|^2 \log \|x\|$, $x \in \mathbb{R}^2$. Sabemos de la sección 2.1.2 que Φ satisface*

$$\Delta^2 \Phi(x) = \sum_{|\alpha|=2} \frac{2!}{\alpha!} D^{2\alpha} \Phi(x) = \delta(x).$$

Tomando transformada de Fourier a ambos lados de la ecuación se tiene

$$\sum_{|\alpha|=2} \frac{2}{\alpha!} (i\xi)^{2\alpha} \hat{\Phi}(\xi) = \frac{1}{2\pi},$$

despejando obtenemos

$$\hat{\Phi}(\xi) = \frac{1}{2\pi \|\xi\|^4}.$$

Del teorema anterior, el semi-producto interior correspondiente al espacio nativo $\mathcal{N}_{\Phi}(\mathbb{R}^2)$ se ve como

$$\begin{aligned}
(f, g)_{\mathcal{N}_\Phi(\mathbb{R}^2)} &= \frac{1}{2\pi} \int_{\mathbb{R}^2} \hat{f}(\xi) \overline{\hat{g}(\xi)} 2\pi \|\xi\|^4 d\xi = \sum_{|\alpha|=2} \frac{2!}{\alpha!} \int_{\mathbb{R}^2} \xi^{2\alpha} \hat{f}(\xi) \overline{\hat{g}(\xi)} d\xi \\
&= \sum_{|\alpha|=2} \frac{2!}{\alpha!} \int_{\mathbb{R}^2} D^{\hat{\alpha}} f(\xi) \overline{D^{\hat{\alpha}} g(\xi)} d\xi = \sum_{|\alpha|=2} \frac{2!}{\alpha!} \int_{\mathbb{R}^2} D^\alpha f(x) \overline{D^\alpha g(x)} dx \\
&= (f, g)_{V^2(\mathbb{R}^2)},
\end{aligned}$$

donde en la penúltima igualdad se utilizó la identidad de Parseval, y $(\cdot, \cdot)_{V^2(\mathbb{R}^2)}$ es el semi-producto interior del espacio de Beppo-Levi $V^2(\mathbb{R}^2)$. En conclusión, el espacio nativo correspondiente al spline de placa delgada es el espacio $V^2(\mathbb{R}^2)$, lo cual ya conocíamos de la sección 2.1.2.

2.4. Funciones Completamente Monótonas

No siempre es sencillo calcular transformadas de Fourier, por lo que es necesario encontrar otros criterios para saber si una función es positiva definida. A continuación definiremos las llamadas funciones completamente monótonas, las cuales serán útiles para este fin.

Definición 2.4.1. Una función $\phi : [0, \infty) \rightarrow \mathbb{R}$ que pertenece a $C[0, \infty) \cap C^\infty(0, \infty)$ y satisface

$$(-1)^l \phi^{(l)}(r) \geq 0, \quad r > 0, \forall l \in \mathbb{N},$$

es llamada una función completamente monótona en $[0, \infty)$.

Ejemplo 2.4.2. La función $\phi(r) = e^{-\varepsilon r}$, con $\varepsilon \geq 0$, es completamente monótona en $[0, \infty)$ ya que

$$(-1)^l \phi^{(l)}(r) = \varepsilon^l e^{-\varepsilon r} \geq 0, \quad l \in \mathbb{N}.$$

Ejemplo 2.4.3. La función $\phi(r) = \frac{1}{(1+r)^\beta}$, con $\beta > 0$, es completamente monótona en $[0, \infty)$ ya que

$$(-1)^l \phi^{(l)}(r) = (-1)^{2l} \beta(\beta+1) \cdots (\beta+l-1) (1+r)^{-\beta-l} \geq 0, \quad l \in \mathbb{N}.$$

De manera análoga al teorema de Bochner, el siguiente teorema establece para las funciones completamente monótonas una caracterización en términos de una representación integral.

Teorema 2.4.4 (Hausdorff-Bernstein-Widder). Una función $\phi : [0, \infty) \rightarrow \mathbb{R}$ es completamente monótona en $[0, \infty)$ si y sólo si es la transformada de Laplace de una función $\omega \in L^1[0, \infty)$ no negativa y acotada⁸, i.e.

$$\phi(r) = \mathcal{L}\omega(r) = \int_0^\infty e^{-rt} \omega(t) dt.$$

⁸Al igual que en el teorema de Bochner, en realidad se trata de una medida de Borel finita y no negativa en $[0, \infty)$.

En el mismo estilo que en la observación 2.2.9, podemos entonces considerar a la función ϕ del ejemplo 2.4.2 como la **función completamente monótona fundamental**.

La conexión entre funciones completamente monótonas y funciones positivas definidas fue señalada por Schoenberg en 1938.

Teorema 2.4.5 (Schoenberg). *Una función ϕ es completamente monótona en $[0, \infty)$ si y sólo si $\Phi = \phi(\|\cdot\|^2)$ es positiva definida y radial en \mathbb{R}^d para toda d . Más aún, si ϕ no es constante, se tiene que Φ es estrictamente positiva definida.*

Ejemplo 2.4.6. *Del teorema anterior y los ejemplos 2.4.2 y 2.4.3, se sigue que las gaussianas $\Phi(x) = e^{-\varepsilon\|x\|^2}$, $\varepsilon > 0$, y los multicuádricos inversos $\Psi(x) = 1/(1 + \|x\|^2)^\beta$, $\beta \geq 0$, son funciones radiales y estrictamente positivas definidas en \mathbb{R}^d para toda d .*

Claramente este criterio de monotonidad es mucho más simple que calcular transformadas de Fourier. Entonces podría uno preguntarse por qué no se tomó este criterio como punto de partida y si realmente era necesario estudiar el teorema de Bochner. La respuesta es que el teorema de Bochner sí es necesario, entre otras cosas para estudiar los espacios nativos y hacer estimaciones del error.

El criterio análogo al teorema de Schoenberg para el caso de funciones condicionalmente positivas definidas se enuncia a continuación.

Teorema 2.4.7 (Michelli). *Sea $\phi \in C[0, \infty) \cap C^\infty(0, \infty)$. Entonces la función $\Phi = \phi(\|\cdot\|^2)$ es radial y condicionalmente positiva definida de orden m en \mathbb{R}^d para toda d si y sólo si $(-1)^m \phi^{(m)}$ es completamente monótona en $(0, \infty)$.*

En caso de que ϕ no sea un polinomio de grado a lo más m , entonces Φ es estrictamente condicionalmente positiva definida de orden m .

Ejemplo 2.4.8. *Sea ϕ la función dada por*

$$\phi(r) = (-1)^{\lceil \beta \rceil} (1+r)^\beta, \quad 0 < \beta \notin \mathbb{N}.$$

Entonces se tiene

$$\phi^{(l)}(r) = (-1)^{\lceil \beta \rceil} \beta(\beta-1) \cdots (\beta-l+1)(1+r)^{\beta-l},$$

por lo que

$$(-1)^{\lceil \beta \rceil} \phi^{(\lceil \beta \rceil)}(r) = \beta(\beta-1) \cdots (\beta - \lceil \beta \rceil + 1)(1+r)^{\beta - \lceil \beta \rceil}$$

es completamente monótona. Más aún, $m = \lceil \beta \rceil$ es el número más pequeño posible tal que $(-1)^m \phi^{(m)}$ es completamente monótona. Como $\beta \notin \mathbb{N}$, ϕ no es un polinomio, y por lo tanto el multicuádrico

$$\Phi(x) = (-1)^{\lceil \beta \rceil} (1 + \|x\|^2)^\beta,$$

es estrictamente condicionalmente positivo definido de orden $\lceil \beta \rceil$ y radial en \mathbb{R}^d para toda d .

Ejemplo 2.4.9. Sea ϕ la función dada por

$$\phi(r) = (-1)^{\lceil \beta \rceil / 2} r^{\beta/2}, \quad 0 < \beta \notin 2\mathbb{N}.$$

Entonces se tiene

$$\phi^{(l)}(r) = (-1)^{\lceil \beta \rceil / 2} \frac{\beta}{2} \left(\frac{\beta}{2} - 1\right) \cdots \left(\frac{\beta}{2} - l + 1\right) r^{\beta/2 - l},$$

por lo que $(-1)^{\lceil \beta \rceil / 2} \phi^{(\lceil \beta \rceil / 2)}$ es completamente monótona y $m = \lceil \beta \rceil / 2$ es el número más pequeño para el cual esto se cumple. Como β no es par, ϕ no es un polinomio, y por lo tanto las potencias

$$\Phi(x) = (-1)^{\lceil \beta \rceil / 2} \|x\|^\beta, \quad 0 < \beta \notin 2\mathbb{N}$$

son estrictamente condicionalmente positivas definidas de orden $\lceil \beta \rceil / 2$ y radiales en \mathbb{R}^d para toda d .

Ejemplo 2.4.10. Los splines de placa delgada

$$\Phi(x) = (-1)^{\beta+1} \|x\|^{2\beta} \log(\|x\|), \quad \beta \in \mathbb{N},$$

son estrictamente condicionalmente positivos definidos de orden $\beta + 1$ y radiales en \mathbb{R}^d para toda d . Para ver esto observa que

$$2\Phi(x) = (-1)^{\beta+1} \|x\|^{2\beta} \log(\|x\|^2).$$

Esto sugiere que definamos la función ϕ

$$\phi(r) = (-1)^{\beta+1} r^\beta \log(r), \quad \beta \in \mathbb{N},$$

la cual obviamente no es un polinomio. Derivando se tiene

$$\phi^{(l)}(r) = (-1)^{\beta+1} \beta(\beta-1) \cdots (\beta-l+1) r^{\beta-l} \log(r) + p_l(r), \quad l \in \{1, \dots, \beta\},$$

donde p_l es un polinomio de grado $\beta - l$. En particular se tiene

$$\phi^{(\beta)}(r) = (-1)^\beta \beta! \log(r) + C,$$

y finalmente

$$\phi^{(\beta+1)}(r) = (-1)^{\beta+1} \frac{\beta!}{r},$$

la cual es completamente monótona. Esto es una prueba alternativa del teorema 2.1.5.

2.5. Estimación del Error

El objetivo de esta sección es realizar estimaciones del error para interpolación con funciones (condicionalmente) positivas definidas. Introduciremos primero el concepto de *distancia de llenado* asociada a un conjunto $\mathcal{X} \subset \Omega$ de nodos aleatorios:

$$h = h_{\mathcal{X}, \Omega} = \sup_{x \in \Omega} \min_{x_j \in \mathcal{X}} \|x - x_j\|_2.$$

Esta distancia es justamente el radio de la bola más grande que puede ser colocada entre los datos. Lo que buscamos ahora es conocer, para una sucesión $\{\mathcal{P}^{(h)}\}_h$ de operadores de interpolación, si el error de interpolación

$$\|f - \mathcal{P}_f^{(h)}\|,$$

tiende a cero cuando $h \rightarrow 0$, y qué tan rápido. La idea ahora es expresar al interpolante en su *forma de Lagrange*, utilizando las llamadas *funciones cardinales*.

Sean $A = (\Phi(x_i, x_j)) \in \mathcal{M}_{n \times n}(\mathbb{R})$, con Φ un kernel condicionalmente positivo definido respecto a \mathcal{P} , $P = (p_j(x_i)) \in \mathcal{M}_{N \times Q}(\mathbb{R})$, con p_1, \dots, p_Q una base de \mathcal{P} . Definamos a las funciones $R(x) = (\Phi(x, x_1), \dots, \Phi(x, x_N))^T \in \mathbb{R}^N$ y $S(x) = (p_1(x), \dots, p_Q(x))^T \in \mathbb{R}^Q$. Si \mathcal{X} es un conjunto \mathcal{P} -unisolvante, entonces el sistema lineal

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} \alpha^{(j)} \\ \beta^{(j)} \end{pmatrix} = \begin{pmatrix} e_j \\ 0 \end{pmatrix},$$

donde e_j es el j -ésimo vector canónico, tiene solución única. Las funciones asociadas

$$u_j^* = \sum_{i=1}^N \alpha_i^{(j)} \Phi(\cdot, x_i) + \sum_{k=1}^Q \beta_k^{(j)} p_k,$$

se denominan *funciones cardinales*, y claramente satisfacen $u_j^*(x_i) = \delta_{ij}$. Además estas funciones pertenecen al espacio

$$V_{\mathcal{X}} = \mathcal{P} + \left\{ \sum_{j=1}^N \alpha_j \Phi(\cdot, x_j) : \sum_{j=1}^N \alpha_j p(x_j) = 0, \quad p \in \mathcal{P} \right\}.$$

Teorema 2.5.1. *Sea Φ un kernel condicionalmente positivo definido respecto a \mathcal{P} en $\Omega \subset \mathbb{R}^d$. Supongamos que $\mathcal{X} = \{x_1, \dots, x_N\} \in \Omega$ es \mathcal{P} -unisolvante. Entonces existen funciones $u_j^* \in V_{\mathcal{X}}$ tales que $u_j^*(x_i) = \delta_{ij}$. Además también existen funciones v_1^*, \dots, v_Q^* tales que*

$$\begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} u^*(x) \\ v^*(x) \end{pmatrix} = \begin{pmatrix} R(x) \\ S(x) \end{pmatrix}.$$

Observa que las funciones v_j^* tienen la propiedad $v_k^*(x_l) = 0$. Como consecuencia, cualquier interpolante se puede escribir de la forma

$$\mathcal{P}_{f, \mathcal{X}}(x) = \sum_{j=1}^N f(x_j) u_j^*(x). \tag{2.5.1}$$

Observa que la función $\sigma_{f, \mathcal{X}}$ es tan suave como las funciones u_j^* , y éstas a su vez heredan la suavidad de \mathcal{P} y Φ en su primer argumento. Así, Si Φ es de clase C^k en su primer argumento y $\mathcal{P} \subseteq C^k(\Omega)$, el interpolante será también de clase C^k .

Otro ingrediente importante para obtener las estimaciones del error, es la llamada *función potencia*. Para este fin, considera un dominio $\Omega \subset \mathbb{R}^d$, un kernel positivo definido $\Phi \in C(\Omega \times \Omega)$, y un conjunto de nodos $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$. Para un vector $u \in \mathbb{R}^N$ definimos la forma cuadrática

$$\begin{aligned}
Q(u) &= \Phi(x, x) - 2 \sum_{j=1}^N u_j \Phi(x, x_j) + \sum_{j=1}^N \sum_{i=1}^N u_i u_j \Phi(x_i, x_j) \\
&= \langle \Phi(\cdot, x), \Phi(\cdot, x) \rangle_{\mathcal{N}_\Phi} - 2 \sum_{j=1}^N u_j \langle \Phi(\cdot, x_j), \Phi(\cdot, x_j) \rangle_{\mathcal{N}_\Phi} + \sum_{j=1}^N \sum_{i=1}^N u_i u_j \langle \Phi(\cdot, x_i), \Phi(\cdot, x_j) \rangle_{\mathcal{N}_\Phi} \\
&= \langle \Phi(\cdot, x) - \sum_{j=1}^N u_j \Phi(x, x_j), \Phi(\cdot, x) - \sum_{j=1}^N u_j \Phi(x, x_j) \rangle_{\mathcal{N}_\Phi} \\
&= \left\| \Phi(\cdot, x) - \sum_{j=1}^N u_j \Phi(x, x_j) \right\|_{\mathcal{N}_\Phi}^2.
\end{aligned}$$

En el caso de que Φ sea condicionalmente positivo definido el cálculo anterior sigue siendo válido, pero es necesario sustituir en la última igualdad Φ por el kernel modificado (2.3.2).

Definición 2.5.2. Sea $\Omega \subset \mathbb{R}^d$ abierto y $\Phi \in C(\Omega \times \Omega)$ un kernel condicionalmente positivo definido en Ω respecto a $\mathcal{P} \in C(\Omega)$. Si $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$ es un conjunto \mathcal{P} -solvente, se define a la función potencia a través de

$$[P_{\Phi, \mathcal{X}}(x)]^2 = Q(u^*(x)),$$

donde u^* es el vector de funciones cardinales del teorema 2.5.1.

Teorema 2.5.3. Sean $\Omega \in \mathbb{R}^d$ abierto, $\Phi \in C(\Omega \times \Omega)$ condicionalmente positivo en Ω respecto a $\mathcal{P} \in C(\Omega)$, y $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$ un conjunto de nodos \mathcal{P} -unisolvante. Denotemos por $\mathcal{P}_{f, \mathcal{X}}$ al interpolante de f en $\mathcal{N}_\Phi(\Omega)$. Entonces para toda $x \in \Omega$ se cumple

$$|f(x) - \mathcal{P}_{f, \mathcal{X}}(x)| \leq P_{\Phi, \mathcal{X}}(x) \|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

Si $\Phi \in C^{2k}(\Omega)$ y $\mathcal{P} \in C^k(\Omega)$, es posible aplicar este resultado con ligeras modificaciones a las derivadas para obtener

$$|D^\alpha f(x) - D^\alpha \mathcal{P}_{f, \mathcal{X}}(x)| \leq P_{\Phi, \mathcal{X}}^{(\alpha)}(x) \|f\|_{\mathcal{N}_\Phi(\Omega)}, \quad |\alpha| \leq k,$$

donde $[P_{\Phi, \mathcal{X}}^{(\alpha)}(x)]^2 = Q(D^\alpha u^*(x))$, y la forma cuadrática modificada es

$$Q(u) = D_1^\alpha D_2^\alpha \Phi(x, x) - 2 \sum_{j=1}^N u_j D_1^\alpha \Phi(x, x_j) + \sum_{j=1}^N \sum_{i=1}^N u_i u_j \Phi(x_i, x_j).$$

Con el resultado anterior se logró separar al error entre f y su interpolante en dos componentes, un término independiente de f (la función de potencia), y un término independiente de los nodos \mathcal{X} . La tarea ahora es estimar a la función de potencia en términos de la distancia de llenado.

En análisis numérico, la estrategia más común para encontrar estimaciones del error, es aprovechar la precisión polinomial de un método (al menos localmente), y luego realizar un desarrollo de Taylor. Para poder usar esta estrategia, será necesario que el dominio satisfaga la llamada *condición de cono interior*.

Definición 2.5.4. *Se dice que una región $\Omega \in \mathbb{R}^d$ satisface una condición de cono interior si hay un ángulo $\theta \in (0, \pi/2)$ y un radio $r > 0$, tales que para cualquier $x \in \Omega$ existe un vector unitario $\xi(x)$, de forma que el cono*

$$C = \{x + \lambda y : y \in \mathbb{R}^d, \|y\|_2 = 1, \quad y^T \xi(x) \geq \cos \theta, \quad \lambda \in [0, r]\},$$

queda contenido en Ω .

La condición de cono interior impone cierta regularidad sobre el dominio Ω . Si un dominio satisface la condición de cono interior, contiene bolas de radio controlable. Esto es lo que permite acotar el residuo en el desarrollo de Taylor. El siguiente teorema garantiza la existencia de un esquema con precisión polinomial (local).

Teorema 2.5.5. *Sea $\Omega \in \mathbb{R}^d$ un dominio acotado que satisface la condición de cono interior. Sea $l \in \mathbb{N}$ y $\alpha \in \mathbb{N}^d$ con $|\alpha| \leq l$. Entonces existen constantes $h_0, c_1^{(\alpha)}, c_2^\alpha > 0$ tales que para cualquier $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$ con $h_{\mathcal{X}, \Omega} \leq h_0$ y cualquier $x \in \Omega$ existen números $\tilde{u}_1^{(\alpha)}, \dots, \tilde{u}_N^{(\alpha)}$ con*

$$(a) \sum_{j=1}^N \tilde{u}_j^{(\alpha)}(x) p(x_j) = D^\alpha p(x), \text{ para cualquier } p \in \mathbb{P}_{m-1}(\mathbb{R}^d),$$

$$(b) \sum_{j=1}^N |\tilde{u}_j^{(\alpha)}(x)| \leq c_1^{(\alpha)} h_{\mathcal{X}, \Omega}^{-|\alpha|},$$

$$(c) \tilde{u}_j^{(\alpha)}(x) = 0, \text{ si } \|x - x_j\| \geq c_2^{(\alpha)} h_{\mathcal{X}, \Omega}.$$

Utilizando lo anterior, ya es posible obtener un estimado del error en términos de la distancia de llenado $h_{\mathcal{X}, \Omega}$.

Teorema 2.5.6. *Sea $\Omega \in \mathbb{R}^d$ un dominio abierto y acotado que satisface la condición de cono interior. Supóngase que $\Phi \in C^{2k}(\Omega \times \Omega)$ es condicionalmente positivo definido respecto a $\mathbb{P}_{m-1}(\mathbb{R}^d)$. Si $\mathcal{P}_{f, \mathcal{X}}$ denota al interpolante de $f \in \mathcal{N}_\Phi(\Omega)$ respecto al conjunto $\mathbb{P}_{m-1}(\mathbb{R}^d)$ -solvente de nodos $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$, entonces para $\alpha \in \mathbb{N}^d$, con $|\alpha| \leq k$, existen constantes $h_0, C > 0$ tales que*

$$|D^\alpha f(x) - D^\alpha \mathcal{P}_{f, \mathcal{X}}(x)| \leq C C_\Phi(x)^{1/2} h_{\mathcal{X}, \Omega}^{k-|\alpha|} |f|_{\mathcal{N}_\Phi(\Omega)}, \quad x \in \Omega,$$

si $h_{\mathcal{X}, \Omega} \leq h_0$. El número $C_\Phi(x)$ se define como

$$C_\Phi(x) := \max_{\substack{\beta, \nu \in \mathbb{N}^d \\ |\alpha| + |\beta| = 2k}} \max_{z, w \in \Omega \cap B(x, c_2^{(\alpha)} h_{\mathcal{X}, \Omega})} |D_1^\beta D_2^\nu \Phi(z, w)|.$$

La constante C es independiente de x, f y Φ .

La constante $C_\Phi(x)$ en el teorema anterior todavía depende del kernel Φ , i.e. ,la estimación del error en el teorema anterior sigue siendo genérica. Es necesario trabajar por separado cada kernel radial para obtener estimaciones más específicas. Por ejemplo, el siguiente teorema muestra que, bajo ciertas condiciones, el orden de convergencia llega a ser exponencial.

Teorema 2.5.7. Sea $\Omega \subset \mathbb{R}^d$ un cubo. Sea $\Phi = \phi(\|\cdot\|)$ una función radial condicionalmente positiva definida tal que $\psi = \phi(\sqrt{\cdot})$ satisface $|\psi^{(l)}(r)| \leq l!M^l$ para todos los enteros $l \geq l_0$ y todo $r > 0$, donde M es una constante positiva. Entonces existe una constante c tal que para cualquier $f \in \mathcal{N}_\Phi(\Omega)$

$$\|f - \mathcal{P}_f\|_{L^\infty(\Omega)} \leq e^{\frac{-c}{h_{\mathcal{X},\Omega}}} \|f\|_{\mathcal{N}_\Phi(\Omega)},$$

para cualquier conjunto de nodos \mathcal{X} con distancia de llenado $h_{\mathcal{X},\Omega}$ lo suficientemente pequeña. Más aún, si ψ satisface $|\psi^{(l)}(r)| \leq M^l$, entonces

$$\|f - \mathcal{P}_f\|_{L^\infty(\Omega)} \leq e^{\frac{-c|\log h_{\mathcal{X},\Omega}|}{h_{\mathcal{X},\Omega}}} \|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

para $h_{\mathcal{X},\Omega}$ suficientemente pequeña.

Ejemplo 2.5.8. Para las gaussianas $\Phi(x) = e^{-\varepsilon^2\|x\|^2}$, $\varepsilon > 0$, se tiene $\psi(r) = e^{-\varepsilon^2 r}$, por lo que $\psi^{(l)}(r) = (-1)^l \varepsilon^{2l} e^{-\varepsilon^2 r}$ para $l \geq l_0 = 0$. Tomando $M = \varepsilon^2$ podemos aplicar el teorema anterior y obtener el estimado

$$\|f - \mathcal{P}_f\|_{L^\infty(\Omega)} \leq e^{\frac{-c|\log h_{\mathcal{X},\Omega}|}{h_{\mathcal{X},\Omega}}} \|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

Ejemplo 2.5.9. Para los multicuádricos inversos $\Phi(x) = (1 + \|x\|^2)^\beta$, con $\beta < 0$ o $\beta \notin \mathbb{N}$, se tiene $\psi(r) = (1 + r)^\beta$. En este caso se puede probar que $|\psi^{(l)}(r)| \leq l!M^l$, siempre que $l \geq \lceil \beta \rceil$, con $M = 1 + \lceil \beta \rceil + 1$. Aplicando el teorema anterior se obtiene entonces el estimado

$$\|f - \mathcal{P}_f\|_{L^\infty(\Omega)} \leq e^{\frac{-c}{h_{\mathcal{X},\Omega}}} \|f\|_{\mathcal{N}_\Phi(\Omega)}.$$

En otros casos se puede probar lo siguiente.

Ejemplo 2.5.10. Para las potencias $\Phi(x) = (-1)^{\lceil \frac{\beta}{2} \rceil} \|x\|^\beta$, con $0 < \beta \notin 2\mathbb{N}$, se obtiene

$$|D^\alpha f(x) - D^\alpha \mathcal{P}_f(x)| \leq Ch_{\mathcal{X},\Omega}^{\frac{\beta}{2} - \alpha} |f|_{\mathcal{N}_\Phi(\Omega)},$$

siempre que $\alpha < \frac{\lceil \beta \rceil - 1}{2}$ y $f \in \mathcal{N}_\Phi(\Omega)$.

Ejemplo 2.5.11. Para los spline de placa delgada $\Phi(x) = (-1)^{k+1} \|x\|^{2k} \log \|x\|$, se tiene

$$|D^\alpha f(x) - D^\alpha \mathcal{P}_f(x)| \leq Ch_{\mathcal{X},\Omega}^{k-\alpha} |f|_{\mathcal{N}_\Phi(\Omega)},$$

siempre que $|\alpha| < k - 1$ y $f \in \mathcal{N}_\Phi(\Omega)$.

Todas las estimaciones anteriores son válidas siempre y cuando la función f pertenezca al espacio nativo de la función radial utilizada en la interpolación. Hay algunos resultados que proporcionan cotas para el error aún cuando f no pertenezca al espacio nativo. El caso en el que f pertenece a algún espacio de Sobolev es de especial interés:

Teorema 2.5.12. Sea Ω un conjunto compacto con frontera Lipschitz que satisface la condición de cono y $\mathcal{X} \subset \Omega$ un conjunto de nodos. Sean k un entero positivo, $0 < \sigma \leq 1$, $1 \leq p < \infty$, $1 \leq q \leq \infty$, y α un multi-índice que satisface $k > \alpha + s/p$ (con \geq en caso de que $p = 1$). Si un proceso de interpolación $\mathcal{P} : W^{k+\sigma,p}(\Omega) \rightarrow V$ correspondiente a \mathcal{X} , con V un subespacio de dimensión finita, satisface la propiedad $|\mathcal{P}_f|_{W^{k+\sigma,p}(\Omega)} \leq |f|_{W^{k+\sigma,p}(\Omega)}$, entonces se tiene que

$$|f - \mathcal{P}_f|_{W^{k+\sigma,p}(\Omega)} \leq ch^{k+\sigma-|\alpha|-s(1/p-1/q)_+} |f|_{W^{k+\sigma,p}(\Omega)},$$

con c una constante independiente de f y h , $y(x)_+ = \max\{x, 0\}$.

La propiedad adicional $|\mathcal{P}_f|_{W^{k+\sigma,p}(\Omega)} \leq |f|_{W^{k+\sigma,p}(\Omega)}$ se satisface en particular si el espacio nativo es un espacio de Sobolev. Así, el teorema anterior nos da una alternativa al enfoque de la función potencia del teorema 2.5.6, con la ventaja de que no es necesario lidiar con la constante $C_{\Phi}(x)$.

Además, se puede demostrar también que si la transformada de Fourier de Φ satisface

$$c_1(1 + \|\omega\|_2^2)^{-\tau} \leq \hat{\Phi}(\omega) \leq c_2(1 + \|\omega\|_2^2)^{-\tau}, \quad \|\omega\| \rightarrow \infty, \omega \in \mathbb{R}^d, \quad (2.5.2)$$

entonces la estimación anterior se satisface con $\tau = k + \sigma$ y $p = 2$, para h suficientemente pequeña. También, para funciones f fuera del espacio nativo de Φ que satisfacen (2.5.2) se tiene el siguiente teorema.

Teorema 2.5.13. Sean k y n enteros con $0 \leq n < k \leq \tau$ y $k \geq d/2$. Sea $f \in C^k(\bar{\Omega})$ y supóngase que $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$ satisface $\text{diam}(\mathcal{X}) \leq 1$ con distancia de llenado suficientemente pequeña. Entonces para cualquier $1 \leq q \leq \infty$ se tiene

$$|f - \mathcal{P}_f|_{W^{n,q}(\Omega)} \leq c\rho_{\mathcal{X}}^{\tau-k} h^{k-n-s(1/2-1/q)_+} |f|_{C^k(\bar{\Omega})},$$

donde $\rho_{\mathcal{X}} = \frac{h}{q_{\mathcal{X}}}$, con $q_{\mathcal{X}} = \frac{1}{2} \min_{i \neq j} \|x_i - x_j\|_2$ la distancia de separación.

Ejemplo 2.5.14. Para el spline de placa delgada $\Phi(x) = \|x\|^{2\beta} \log \|x\|$, se satisface (2.5.2) con $\tau = 2\beta$. Si tomamos $k = \tau$, $n = 0$ y $q = \infty$ en el teorema anterior, se obtiene el estimado

$$|f - \mathcal{P}_f|_{L^\infty} \leq ch^{2\beta-s/2} \|f\|_{C^{2\beta}(\bar{\Omega})}.$$

Para las potencias $\Phi(x) = \|x\|^\beta$, se satisface (2.5.2) con $\tau = \beta$. Tomando de igual forma $k = \tau$, $n = 0$ y $q = \infty$ en el teorema anterior, se obtiene

$$|f - \mathcal{P}_f|_{L^\infty} \leq ch^{\beta-s/2} \|f\|_{C^\beta(\bar{\Omega})}$$

Aunque las estimaciones anteriores son útiles, la teoría está lejos de estar terminada. Estos estimados no son del todo satisfactorios, ya que, por ejemplo, no toman en cuenta al parámetro de forma (para gaussianas, multicuádricos o multicuádricos inversos). Este parámetro tiene una influencia muy grande en el resultado de la aproximación, y hasta la fecha no hay una teoría matemática completa para determinar el parámetro de forma óptimo. En la mayoría de los casos la determinación del parámetro de forma es dependiente del problema, y sigue siendo un problema abierto. Se sabe que el estimado del error está compuesto por dos términos que compiten, una

parte que crece exponencialmente, y otra que decrece exponencialmente, cuando el parámetro de forma c aumenta:

$$\epsilon \sim O(e^{ac} \lambda^{c/h}) \quad 0 < \lambda < 1, a > 0,$$

o

$$\epsilon \sim O(e^{ac^2} \lambda^{c/h}) \quad 0 < \lambda < 1, a > 0,$$

Del último estimado, por ejemplo, se podría obtener una aproximación para el parámetro de forma óptimo:

$$c_{opt} \sim O(-\log \lambda / 2ah).$$

Fórmulas de este estilo sólo están disponibles en casos muy especiales, por lo que usualmente el parámetro de forma óptimo se determina numéricamente. Además, también hay que tener en cuenta que el número de condición de la matriz de interpolación aumenta con el parámetro de forma, de manera que éste no puede ser muy grande para no afectar la precisión de la aproximación.

3. Solución de Ecuaciones Diferenciales Parciales con RBF

En este capítulo se mostrará cómo se pueden aplicar las técnicas de interpolación discutidas en el capítulo anterior a la solución numérica de ecuaciones diferenciales parciales. Supondremos en las secciones siguientes que se tiene un dominio $\Omega \subseteq \mathbb{R}^d$ y el problema

$$\begin{cases} \mathcal{L}u = f, & \Omega, \\ \mathcal{B}u = g, & \partial\Omega, \end{cases} \quad (3.0.3)$$

donde las funciones f, g , el operador diferencial parcial \mathcal{L} y el operador de frontera \mathcal{B} están dados, y se busca la solución u .

3.1. Colocación Asimétrica de Kansa

Sean $X_I = \{x_i\}_{i=1}^{N_I} \subset \Omega$ y $X_B = \{x_i\}_{i=N_I+1}^N \subset \partial\Omega$ con $N = N_I + N_B$ conjuntos de nodos aleatorios en el interior y en la frontera respectivamente. Sea Φ una función de base radial condicionalmente positiva definida respecto $\mathbb{P}_{m-1}(\mathbb{R}^d)$. Considera el ansatz dado por

$$u_\Phi(\cdot) = \sum_{j=1}^N \lambda_j \Phi(\cdot - x_j) + \sum_{j=1}^{N_P} \alpha_j P_{m-1}^j(\cdot), \quad (3.1.1)$$

donde P_{m-1}^j es el j -ésimo término de un polinomio de grado $N_P = \dim(\mathbb{P}_{m-1}(\mathbb{R}^d))$, el cual está sujeto a las condiciones de momento

$$\sum_{j=1}^N \lambda_j P_{m-1}^k(x_j) = 0, \quad k \in \{1, \dots, N_P\}. \quad (3.1.2)$$

Colocando a u_Φ en el conjunto de nodos y utilizando 3.0.3, se obtiene el sistema

$$\begin{aligned} \sum_{j=1}^N \lambda_j \mathcal{L}\Phi(x_i - x_j) + \sum_{j=1}^{N_P} \alpha_j \mathcal{L}P_{m-1}^j(x_i) &= f_i, \quad i \in \{1, \dots, N_I\} \\ \sum_{j=1}^N \lambda_j \mathcal{B}\Phi(x_i - x_j) + \sum_{j=1}^{N_P} \alpha_j \mathcal{B}P_{m-1}^j(x_i) &= g_i, \quad i \in \{1 + N_I, \dots, N\}. \end{aligned}$$

Este sistema, junto con las condiciones de momento, se puede escribir en forma matricial como

$$\begin{bmatrix} W_{\mathcal{L}} & P_{\mathcal{L}} \\ W_{\mathcal{B}} & P_{\mathcal{B}} \\ P^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha \end{bmatrix} = \begin{bmatrix} f \\ g \\ 0 \end{bmatrix}, \quad (3.1.3)$$

donde

$$\begin{aligned}
(W_{\mathcal{L}})_{ij} &= \mathcal{L}\Phi(x_i - x_j), & i \in \{1, \dots, N_I\}, j \in \{1, \dots, N\}, \\
(W_{\mathcal{B}})_{i-N_I, j} &= \mathcal{B}\Phi(x_i - x_j), & i \in \{N_I + 1, \dots, N\}, j \in \{1, \dots, N\}, \\
(P_{\mathcal{L}})_{ij} &= \mathcal{L}P_{m-1}^j(x_i), & i \in \{1, \dots, N_I\}, j \in \{1, \dots, N_P\}, \\
(P_{\mathcal{B}})_{i-N_I, j} &= \mathcal{B}P_{m-1}^j(x_i), & i \in \{N_I + 1, \dots, N\}, j \in \{1, \dots, N_P\}, \\
(P)_{ij} &= P_{m-1}^j(x_i), & i \in \{1, \dots, N\}, j \in \{1, \dots, N_P\}.
\end{aligned}$$

Resolviendo el sistema de ecuaciones se encuentra la aproximación a la solución u_{Φ} . La principal ventaja de este método es su simplicidad. Programarlo requiere sólo unas cuantas líneas de código. Sin embargo, tiene la desventaja de que la matriz resultante 3.1.3 no es simétrica. De hecho, la invertibilidad de dicha matriz no se ha demostrado y a la fecha es un problema abierto.

3.2. Colocación Simétrica

En este método se propone un nuevo ansatz:

$$u_{\Phi}(\cdot) = \sum_{j=1}^{N_I} \lambda_j \bar{\mathcal{L}}\Phi(\cdot - x_j) + \sum_{j=N_I+1}^N \lambda_j \bar{\mathcal{B}}\Phi(\cdot - x_j) + \sum_{j=1}^{N_P} \alpha_j P_{m-1}^j(\cdot), \quad (3.2.1)$$

donde los operadores $\bar{\mathcal{L}}$ y $\bar{\mathcal{B}}$ actúan de la misma manera que \mathcal{L} y \mathcal{B} , pero sobre la segunda entrada de $\Phi(\cdot - \cdot)$. Colocando de la misma manera que en el método de Kansa, resulta el sistema lineal

$$\begin{bmatrix} W_{\mathcal{L}\bar{\mathcal{L}}} & W_{\mathcal{L}\bar{\mathcal{B}}} & P_{\mathcal{L}} \\ W_{\mathcal{L}\bar{\mathcal{B}}}^T & W_{\mathcal{B}\bar{\mathcal{B}}} & P_{\mathcal{B}} \\ P_{\mathcal{L}}^T & P_{\mathcal{B}}^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha \end{bmatrix} = \begin{bmatrix} f \\ g \\ 0 \end{bmatrix}, \quad (3.2.2)$$

donde

$$\begin{aligned}
(W_{\mathcal{L}\bar{\mathcal{L}}})_{ij} &= \mathcal{L}\bar{\mathcal{L}}\Phi(x_i - x_j), & i, j \in \{1, \dots, N_I\}, \\
(W_{\mathcal{L}\bar{\mathcal{B}}})_{i, j-N_I} &= \mathcal{L}\bar{\mathcal{B}}\Phi(x_i - x_j), & i \in \{1, \dots, N_I\}, j \in \{N_I + 1, \dots, N\}, \\
(W_{\mathcal{B}\bar{\mathcal{B}}})_{i-N_I, j-N_I} &= \mathcal{B}\bar{\mathcal{B}}\Phi(x_i - x_j), & i, j \in \{N_I + 1, \dots, N\},
\end{aligned}$$

y las matrices $P_{\mathcal{L}}$ y $P_{\mathcal{B}}$ son iguales que en el método asimétrico. La ventaja principal de esta formulación es que la matriz resultante es simétrica y no singular, como se demostrará a continuación. No obstante, se requiere que las RBF utilizadas en 3.2.2 tengan mayor diferenciabilidad que en el método asimétrico. Para la colocación no lineal es posible usar este método, pero aumenta la complejidad.

3.2.1. Interpolación Generalizada de Hermite-Birkhoff

El objetivo de esta sección es demostrar que el sistema (3.2.2) tiene solución. Para ello consideremos primero un problema más general, conocido en la literatura como interpolación de Hermite-Birkhoff. Sea H es un espacio de Hilbert y supongamos que $\Lambda = \{\lambda_1, \dots, \lambda_N\} \subset H^*$ es un conjunto de funcionales linealmente independientes. El problema de interpolación generalizado consiste en, dado un conjunto de valores $\{f_1, \dots, f_N\} \subset \mathbb{R}$, encontrar $\sigma \in H$ tal que

$$\lambda_j(\sigma) = f_j, \quad j \in \{1, \dots, N\}. \quad (3.2.3)$$

A σ se le conoce como un *interpolante generalizado*. Ahora, recuerda que el interpolante definido por un kernel radial condicionalmente positivo definido es óptimo en el siguiente sentido: el interpolante minimiza la semi-norma del espacio nativo entre todas las funciones en el espacio que satisfagan las condiciones de interpolación. Generalizando al contexto de interpolación de Hermite-Birkhoff, lo que tendríamos que buscar es un $\sigma^* \in H$ que satisfaga

$$\|\sigma^*\| = \min\{\|s\| : s \in H, \lambda_j(\sigma) = f_j, 1 \leq j \leq N\}. \quad (3.2.4)$$

El siguiente teorema nos ayudará a analizar este problema de interpolación generalizado

Teorema 3.2.1. *Sean H un espacio de Hilbert y $\lambda_1, \dots, \lambda_N \in H^*$ funcionales linealmente independientes con representantes de Riesz $v_1, \dots, v_N \in H$ respectivamente. Para $f_1, \dots, f_N \in \mathbb{R}$ dados, la solución única de (3.2.4) está dada por*

$$\sigma^* = \sum_{j=1}^N \alpha_j v_j, \quad (3.2.5)$$

donde los coeficientes α_j se determinan a partir de las condiciones de interpolación $\lambda_j(\sigma^*) = f_j$, $1 \leq j \leq N$.

Demostración. Antes de comenzar nota que σ^* está bien definido por las condiciones de interpolación. Como los funcionales λ_j son linealmente independientes, también lo son los representantes v_j , lo cual implica que la matriz de interpolación con entradas $(\lambda_i(v_j)) = ((v_i, v_j)_H)(\lambda_i, \lambda_j)_{H^*}$ es invertible. Para ver que el interpolante satisface la condición de optimalidad, supongamos que $v \in H$ satisface $\lambda_j(v) = 0$, $1 \leq j \leq N$, lo cual implica que

$$(\sigma^*, v)_H = \sum_{j=1}^N \alpha_j (v, v_j)_H = \sum_{j=1}^N \alpha_j \lambda_j(v) = 0.$$

Esto muestra que

$$\|\sigma^*\|^2 = (\sigma^*, \sigma^* - \sigma + \sigma)_H = (\sigma^*, \sigma)_H \leq \|\sigma^*\| \|\sigma\|,$$

para cualquier $\sigma \in H$ tal que $\lambda_j(\sigma) = f_j$, $1 \leq j \leq N$. □

Nota que si además se cumple que $f_j = \lambda_j(f)$, para alguna $f \in H$, entonces σ^* es la mejor aproximación en $V = \text{span}\{v_1, \dots, v_N\}$ de f , i.e.

$$\|f - \sigma^*\| = \min\{\|f - s\| : s \in V\}.$$

Por otro lado, se puede demostrar que si $\Omega \subset \mathbb{R}^d$ es abierto y $\Phi \in C^{2k}(\Omega \times \Omega)$ es un kernel condicionalmente positivo definido respecto a $\mathcal{P} \subset C^k$, entonces el espacio nativo $\mathcal{N}_\Phi(\Omega)$ está encajado en $C^k(\Omega)$ ⁹. En otras palabras, esto significa que los funcionales de la forma $\lambda_j = \delta_{x_j} \circ D^{\alpha^{(j)}}$, i.e. $\lambda_j(u) = D^{\alpha^{(j)}}u(x_j)$, con $|\alpha^{(j)}| \leq k$ pertenecen a $\mathcal{N}_\Phi(\Omega)^*$. Este caso es de especial interés, ya que justamente este tipo de operadores aparecen en los métodos de colocación. Los representantes de Riesz en este caso son de la forma $v_j = \overline{\lambda_j}\Phi(\cdot, \cdot)$, donde $\overline{\lambda_j}$ denota al funcional λ_j aplicado en la segunda entrada. Así, los interpolantes serán de la forma

$$\sigma(x) = \sum_{j=1}^N \alpha_j \overline{\lambda_j}\Phi(x, \cdot) = \sum_{j=1}^N \alpha_j \overline{D^{\alpha^{(j)}}}\Phi(x - x_j). \quad (3.2.6)$$

La matriz de interpolación asociada con (3.2.6) tiene entradas $D^{\alpha^{(i)}}\overline{D^{\alpha^{(j)}}}\Phi(x_i, x_j)$ y por lo tanto es simétrica. Se puede demostrar que también será invertible siempre y cuando los funcionales λ_j sean linealmente independientes. En el caso que nos interesa la independencia lineal se tiene con condiciones nada restrictiva, como lo demuestra el siguiente teorema.

Teorema 3.2.2. *Supóngase que $\Phi \in L^1(\mathbb{R}^d) \cap C^{2k}(\mathbb{R}^d)$ es positiva definida. Si los funcionales $\lambda_j := \delta_{x_j} \circ D^{\alpha^{(j)}}$, con $|\alpha^{(j)}|, 1 \leq j \leq N$, son distintos por pares, i.e. $\alpha^{(j)} \neq \alpha^{(k)}$ o $x_j \neq x_k$ si $j \neq k$, entonces son linealmente independientes sobre $\mathcal{N}_\Phi(\Omega)$.*

Demostración. Supongamos que se tienen números reales c_j tal que $\sum_{j=1}^N c_j \lambda_j = 0$ sobre $\mathcal{N}_\Phi(\Omega)$. Esto significa en particular que

$$0 = \left\| \sum_{j=1}^N c_j \lambda_j \right\|_{\mathcal{N}_\Phi(\Omega)^*} = \left\| \sum_{j=1}^N c_j \overline{\lambda_j}\Phi(\cdot - y) \right\|_{\mathcal{N}_\Phi(\Omega)}.$$

Del teorema 2.3.5 sabemos que esta norma se puede evaluar con la transformada de Fourier. Para esto, nota primero que

$$\overline{\lambda_j}\Phi(\cdot - y) = (-1)^{|\alpha^{(j)}|} D^{\alpha^{(j)}}\Phi(\cdot - x_j),$$

y que

$$(\overline{\lambda_j}\Phi(\cdot - y))^\wedge(\omega) = (-i\omega)^{\alpha^{(j)}} e^{-ix_j \cdot \omega} \hat{\Phi}(\omega) = \overline{\lambda_j}(e^{-iy \cdot \omega}) \hat{\Phi}(\omega).$$

Esto implica que

$$\left\| \sum_{j=1}^N c_j \overline{\lambda_j}\Phi(\cdot - y) \right\|_{\mathcal{N}_\Phi(\Omega)} = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \left\| \sum_{j=1}^N c_j \overline{\lambda_j}(e^{-iy \cdot \omega}) \right\|^2 \hat{\Phi}(\omega) d\omega.$$

Como Φ es positiva definida, hay un abierto $U \subset \mathbb{R}^d$ donde $\hat{\Phi}$ es positiva. Por lo tanto, se debe cumplir

$$\sum_{j=1}^N c_j (i\omega)^{\alpha^{(j)}} e^{ix_j \cdot \omega} = 0, \quad \text{si } -\omega \in U.$$

⁹Esto es una generalización del teorema de encajes de Sobolev. Para ver la demostración puedes consultar [2]

Por continuación analítica se puede ver que la ecuación anterior es válida para todo $\omega \in \mathbb{R}^d$. Así, para cualquier función de prueba en el espacio de Schwartz $f \in \mathcal{S}$ se tiene que

$$0 = \sum_{j=1}^N c_j (i\omega)^{\alpha^{(j)}} e^{ix_j \cdot \omega} \hat{f}(\omega) = \left(\sum_{j=1}^N c_j (D^{\alpha^{(j)}} f)(\cdot + x_j)^\wedge(\omega) \right), \quad \omega \in \mathbb{R}^d,$$

lo que implica que

$$\sum_{j=1}^N c_j D^{\alpha^{(j)}} f(x + x_j), \quad x \in \mathbb{R}^d, \quad (3.2.7)$$

y en particular para $x = 0$,

$$\sum_{j=1}^N c_j \lambda_j(f) = 0, \quad f \in \mathcal{S}. \quad (3.2.8)$$

Para concluir hay que escoger apropiadamente a una función f en particular. Sea f_0 una función de prueba con soporte contenido en la bola centrada en cero de radio $0 < \epsilon < \min_{j \neq k} \|x_j - x_k\|_2$ y $f_0(x) = 1$ si $\|x\|_2 < \epsilon/2$. Esto último implica en particular que $D^\alpha f_0(0) = 0$ para $\alpha \neq 0$. Para $1 \leq l \leq N$ definamos entonces a la función $f = f_l$ como

$$f_l(x) = \frac{(x - x_l)^{\alpha^{(l)}}}{\alpha^{(l)}!} f_0(x - x_l), \quad x \in \mathbb{R}^d.$$

Aplicando la regla de Leibniz obtenemos que $\lambda_j(f_k) = \delta_{jk}$, por lo que se concluye que $c_j = 0$, para $1 \leq j \leq N$, i.e. los funcionales λ_j son linealmente independientes. \square

En el teorema anterior se tomó $\Phi \in L^1(\mathbb{R}^d)$ sólomente por simplicidad y poder utilizar la transformada de Fourier clásica, pero el resultado es válido en general si Φ es positiva definida; sólomente sería necesario utilizar la transformada de Fourier generalizada. Nota además que si Φ es positiva definida entonces podemos definir un producto interior en $\mathcal{L} := \text{span}\{\lambda_1, \dots, \lambda_N\} \subset \mathcal{N}_\Phi(\Omega)^*$ utilizando la expresión

$$(\mu, \nu)_\mathcal{L} := \mu \bar{\nu} \Phi(x - y), \quad \mu, \nu \in \mathcal{L}. \quad (3.2.9)$$

De esta manera, por el teorema 3.2.2, si los funcionales $\lambda_1, \dots, \lambda_N$ son linealmente independientes, la matriz con entradas $D^{\alpha^{(i)}} \overline{D^{\alpha^{(j)}}} \Phi(x_i, x_j)$ es positiva definida y por lo tanto invertible.

El teorema 3.2.2 también se puede adaptar al caso condicionalmente positivo y, con un razonamiento análogo al expuesto anteriormente, se puede probar que el problema de interpolación de Hermite-Birkhoff tiene solución si se utiliza el kernel K del teorema 2.3.8 en lugar de Φ , ya que éste es positivo definido en el espacio nativo. No obstante, con los funcionales que hemos estado considerando es más natural formar un interpolante de la forma

$$\sigma(x) = \sum_{j=1}^N c_j \overline{D^{\alpha^{(j)}}} \Phi(x, x_j) + \sum_{l=1}^Q d_l p_l, \quad (3.2.10)$$

donde $\{p_1, \dots, p_Q\}$ es una base de \mathcal{P} . Las condiciones de interpolación (3.2.3) junto a las condiciones adicionales

$$\sum_{j=1}^N c_j D^{\alpha^{(j)}} p_l(x_j) = 0, \quad l \in \{1, \dots, Q\}, \quad (3.2.11)$$

conducen al sistema

$$\begin{pmatrix} A_{\Phi, \Lambda} & \Lambda(P) \\ \Lambda(P)^T & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \quad (3.2.12)$$

donde

$$(A_{\Phi, \Lambda})_{ij} = \lambda_i \overline{\lambda_j} \Phi(x_i, x_j), \quad (\Lambda(P))_{ij} = \lambda_i(p_j). \quad (3.2.13)$$

La invertibilidad del sistema anterior la proporciona el siguiente teorema.

Teorema 3.2.3. *Sea $\Phi \in C^{2k}(\Omega \times \Omega)$ un kernel condicionalmente positivo definido respecto a $\mathcal{P} \subset C^k(\Omega)$. Supongamos que los funcionales $\lambda_j = \delta_{x_j} \circ D^{\alpha^{(j)}}$ son linealmente independientes y que $\lambda_j(p) = 0$ para toda $1 \leq j \leq N$ y $p \in \mathcal{P}$ implica $p = 0$. Entonces hay un único interpolante de la forma (3.2.10) que satisface las condiciones (3.2.3) y (3.2.11).*

Demostración. Debido a que se trabaja en dimensión finita, para demostrar la invertibilidad es suficiente probar que la matriz en (3.2.12) es inyectiva. De esta manera, supongamos que

$$\begin{aligned} A_{\Phi, \Lambda} c + \Lambda(P) d &= 0 \\ \Lambda(P)^T c &= 0 \end{aligned}$$

Multiplicando la primera ecuación por c^T y utilizando la segunda ecuación se obtiene que

$$c^T A_{\Phi, \Lambda} c + c^T \Lambda(P) d = c^T A_{\Phi, \Lambda} c = 0.$$

Esto implica que

$$\begin{aligned} c^T A_{\Phi, \Lambda} c &= \sum_{i,j=1}^N c_i c_j D^{\alpha^{(i)}} \overline{D^{\alpha^{(j)}}} \Phi(x_i, x_j) \\ &= \sum_{i,j=1}^N c_i c_j D^{\alpha^{(i)}} \overline{D^{\alpha^{(j)}}} K(x_i, x_j), \end{aligned}$$

para todo c tal que $c^T \Lambda(P) d = 0$. Por lo discutido anteriormente la última forma cuadrática es positiva definida, lo que implica que $c = 0$. Las condiciones adicionales impuestas sobre los funcionales implican también que $d = 0$, lo que concluye la demostración. \square

Es importante notar que el teorema anterior es válido en un contexto mucho más general, e.g. los funcionales λ_j pueden ser operadores integrales. Los únicos requisitos para la validez del teorema son que los funcionales pertenezcan al espacio dual $\mathcal{N}_\Phi(\Omega)$ y que sean linealmente independientes. En el caso que nos interesa, estas condiciones se traducen en pedir que Φ sea suficientemente diferenciable y los nodos de colocación sean distintos, respectivamente. Si te interesa ver la versión más general, puedes consultar [4, 5, 6, 7]. La referencia [5] contiene además estimaciones del error.

A continuación aplicaremos las ideas discutidas anteriormente a la solución de la EDP (3.0.3). Si suponemos que \mathcal{L} es un operador diferencial lineal de orden k , es necesario escoger una función condicionalmente positiva definida en $\Phi \in C^{2k}(\Omega \times \Omega)$. Para realizar la colocación utilizaremos los nodos $X_I = \{x_1, \dots, x_{N_I}\} \subset \Omega$ y $X_B = \{x_{N_I+1}, \dots, x_N\} \subset \partial\Omega$, lo que sugiere definir a los funcionales como

$$\lambda_j = \begin{cases} \delta_{x_j} \circ \mathcal{L}, & j \in \{1, \dots, N_I\}, \\ \delta_{x_j} \circ \mathcal{B}, & j \in \{N_I + 1, \dots, N\}. \end{cases} \quad (3.2.14)$$

Incluso si no se ha mencionado, estamos suponiendo que el operador de frontera se comporta bien, i.e. está definido en $\mathcal{N}_\Phi(\Omega)$. Utilizando los teoremas 3.2.2 y 3.2.3 se tiene finalmente el siguiente resultado:

Corolario 3.2.4. *Sea $\Phi \in C^{2k}(\Omega \times \Omega)$ un kernel radial condicionalmente positivo definido respecto a \mathcal{P} y supongamos que $\lambda_j(p) = 0$ para toda $1 \leq j \leq N$ con $p \in \mathcal{P}$ implica $p = 0$ ¹⁰, donde λ_j está definido por (3.2.14). Entonces existe un único interpolante de la forma (3.2.1) que satisface*

$$\begin{cases} \mathcal{L}u_\Phi(x_j) = f(x_j), & j \in \{1, \dots, N_I\}, \\ \mathcal{B}u_\Phi(x_j) = g(x_j), & j \in \{N_I + 1, \dots, N\}. \end{cases}$$

3.3. Cuadratura Diferencial

El método de cuadratura diferencial es una técnica de discretización numérica para la aproximación de derivadas. Supongamos que se desea aproximar la derivada α -ésima de la función u en el punto x_i , y se tiene un conjunto $\mathcal{X}_i = \{x_1, \dots, x_N\}$ de puntos aleatorios cercanos¹¹ a x_i (llamados nodos de soporte o stencil). La idea es aproximar la derivada como

$$D^\alpha u|_{x=x_i} = \sum_{j=1}^N \omega_{ij}^\alpha u(x_j), \quad (3.3.1)$$

donde ω_{ij}^α son pesos a calcular. Nota que si trabajamos en una dimensión y se utilizan polinomios para interpolar u , (3.3.1) se reduce al método usual de diferencias finitas. En dimensiones mayores con nodos aleatorios ya no es posible utilizar polinomios, ya que, como se vió en la sección 2, el problema de interpolación polinomial en dimensiones mayores no está bien planteado. Es por esta razón que en dimensiones mayores se utilizan principalmente mallas cartesianas, aplicando el método de diferencias finitas en cada dirección por separado.

¹⁰Esta condición sólo se requiere para obtener la unicidad. En el caso de condiciones de frontera Dirichlet se reduce a pedir que los nodos formen un conjunto \mathcal{P} -solvente.

¹¹El método se puede aplicar de manera global a todos los nodos en el dominio, pero es mucho más útil si se aplica de manera local.

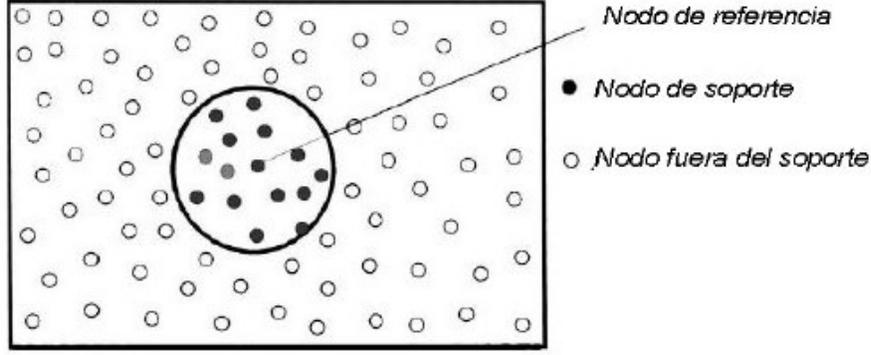


Figura 1: Nodos de soporte

El método de cuadratura diferencial consiste en utilizar interpolación basada en funciones de base radial para calcular los pesos ω_{ij}^α . Formalmente se puede escribir al interpolante utilizando funciones cardinales (la base de Lagrange):

$$\mathcal{P}_{u, \mathcal{X}_i}(x) = \sum_{j=1}^N u_j^*(x) u(x_j),$$

de donde se puede ver aplicando D^α a ambos lados de la ecuación que los pesos están simplemente dados por

$$\omega_{ij}^\alpha = D^\alpha u_j^*(x_i).$$

Aunque lo anterior sirve para aclarar las ideas, no es una manera práctica de calcular los pesos. Para encontrar una manera más práctica consideremos al interpolante escrito de la forma usual

$$\mathcal{P}_{u, \mathcal{X}_i}(x) = \sum_{j=1}^N \lambda_j \Phi(x - x_j). \quad (3.3.2)$$

Sabemos de las secciones anteriores que la matriz de Gram $\mathbf{\Phi} = (\Phi(x_i - x_j))$ es invertible, por lo que podemos escribir

$$\lambda_k = \sum_{j=1}^N \Psi_{kj} u(x_j),$$

con Ψ_{kj} las entradas de la matriz inversa $\mathbf{\Phi}^{-1}$. Aplicando el operador D^α al interpolante (3.3.2) y sustituyendo los valores de λ_k encontramos que

$$D^\alpha u(x_i) = \sum_{j=1}^N \left(\sum_{k=1}^N \Psi_{kj} D^\alpha \Phi(x_i - x_k) \right) u(x_j), \quad (3.3.3)$$

por lo que comparando con (3.3.1) se concluye que

$$\omega_{ij}^\alpha = \sum_{k=1}^N \Psi_{kj} D^\alpha \Phi(x_i - x_k),$$

lo cual es equivalente a la ecuación matricial

$$\begin{bmatrix} \Phi(x_1 - x_1) & \Phi(x_2 - x_1) & \cdots & \Phi(x_N - x_1) \\ \Phi(x_1 - x_2) & \Phi(x_2 - x_2) & \cdots & \Phi(x_N - x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi(x_1 - x_N) & \Phi(x_2 - x_N) & \cdots & \Phi(x_N - x_N) \end{bmatrix} \begin{bmatrix} \omega_{i1}^\alpha \\ \omega_{i2}^\alpha \\ \vdots \\ \omega_{iN}^\alpha \end{bmatrix} = \begin{bmatrix} D^\alpha \Phi(x - x_1)|_{x=x_i} \\ D^\alpha \Phi(x - x_2)|_{x=x_i} \\ \vdots \\ D^\alpha \Phi(x - x_N)|_{x=x_i} \end{bmatrix} \quad (3.3.4)$$

En el caso de que Φ no sea estrictamente positivo definido, la extensión es directa. Por ejemplo, si Φ es un kernel radial condicionalmente positivo definido respecto a $\mathbb{P}_1(\mathbb{R}^2)$, el sistema a resolver sería

$$\begin{bmatrix} & & & 1 & x_{1,1} & x_{1,2} \\ & \Phi^T & & \vdots & \vdots & \vdots \\ & & & 1 & x_{N,1} & x_{N,2} \\ 1 & \cdots & 1 & & & \\ x_{1,1} & \cdots & x_{N,1} & & 0 & \\ x_{1,2} & \cdots & x_{N,2} & & & \end{bmatrix} \begin{bmatrix} \omega_{i1}^\alpha \\ \vdots \\ \omega_{iN}^\alpha \\ \omega_{1,N+1}^\alpha \\ \omega_{1,N+2}^\alpha \\ \omega_{1,N+3}^\alpha \end{bmatrix} = \begin{bmatrix} D^\alpha \Phi(x - x_1)|_{x=x_i} \\ \vdots \\ D^\alpha \Phi(x - x_N)|_{x=x_i} \\ D^\alpha 1|_{x=x_i} \\ D^\alpha x|_{x=x_i} \\ D^\alpha y|_{x=x_i} \end{bmatrix}, \quad (3.3.5)$$

donde los pesos $\omega_{1,N+1}^\alpha$, $\omega_{1,N+2}^\alpha$ y $\omega_{1,N+3}^\alpha$ se agregan para que el stencil sea exacto en los polinomios $\mathbb{P}_1(\mathbb{R}^2)$, pero no se utilizan en la expresión final (3.3.1).

Supongamos ahora que se busca resolver la ecuación diferencial (3.0.3), donde por simplicidad se considerarán condiciones de frontera Dirichlet ($\mathcal{B} = Id$). Sean $\mathcal{X} = \{x_1, \dots, x_N\} \subset \Omega$ los nodos donde se calculará la solución y denotemos por \mathcal{I}_i al conjunto de índices asociados a los nodos en el soporte de x_i . Utilizando el procedimiento descrito anteriormente para calcular los pesos correspondientes al operador \mathcal{L} podemos escribir la discretización

$$\sum_{j \in \mathcal{I}_i} \omega_{ij}^\mathcal{L} u(x_j) = f(x_i), \quad i \in \{1, \dots, N\}. \quad (3.3.6)$$

Sustituyendo el valor de $u(x_j)$ por $g(x_j)$ cuando el nodo soporte x_j pertenezca a $\partial\Omega$, obtenemos un sistema de ecuaciones que puede ser escrito de la forma

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \quad (3.3.7)$$

donde \mathbf{u} es el vector con los valores de u en los nodos y \mathbf{f} es el término fuente que también incluye los términos de frontera. Nota que la matriz \mathbf{A} es *sparse* y puede ser invertida de manera eficiente.

3.3.1. Normalización de la Región de Soporte para RBFs con Parámetro de Forma

El número de nodos soporte en el método de cuadratura diferencial usualmente es fijado para una aplicación. Ya que los nodos son aleatoriamente generados, la escala de la región de soporte para cada nodo podría ser diferente y el parámetro de forma óptimo c podría también ser diferente. Usualmente es muy difícil asignar diferentes valores de c para diferentes nodos. Sin embargo, esta dificultad puede ser removida por una normalización de la escala en la región de soporte.

La idea esencial es transformar el soporte local en un círculo unitario en dos dimensiones o una esfera unitaria en 3 dimensiones. Para ilustrar lo anterior, supongamos que se trabaja con el multicuádrico en 2 dimensiones espaciales. El cambio de coordenadas tiene entonces la forma

$$\bar{x} = \frac{x}{D_i}, \quad \bar{y} = \frac{y}{D_i}, \quad (3.3.8)$$

donde (x, y) son las coordenadas en la región de soporte del espacio físico, (\bar{x}, \bar{y}) denota las coordenadas en el círculo unitario y D_i es el radio mínimo del círculo que contiene a todos los nodos en la región de soporte del nodo (x_i, y_i) . Los multicuádricos se convierten entonces en

$$\Phi_i(\bar{x}, \bar{y}) = \sqrt{\left(\bar{x} - \frac{x_i}{D_i}\right)^2 + \left(\bar{y} - \frac{y_i}{D_i}\right)^2 + \bar{c}^2}, \quad i \in \{1, \dots, N\}, \quad (3.3.9)$$

donde N es el número total de nodos en el soporte. Comparado con la forma tradicional del multicuádrico, podemos ver que el parámetro c es equivalente a $\bar{c}D_i$. La transformación de coordenadas (3.3.8) también cambia los pesos en la aproximación de cuadratura diferencial; para ver cómo cambian simplemente hay que usar la regla de la cadena. Por ejemplo, la primera derivada respecto a x se puede escribir como

$$\frac{du}{dx} = \frac{du}{d\bar{x}} \frac{d\bar{x}}{dx} = \frac{1}{D_i} \sum_{j=1}^N \omega_j^{(1x)} u = \sum_{j=1}^N \frac{\omega_j^{(1x)}}{D_i} u, \quad (3.3.10)$$

donde $\omega_j^{(1x)}$ son los pesos calculados en el círculo unitario y $\omega_j^{(1x)}/D_i$ son los pesos en el dominio físico. Claramente cuando D_i es cambiado, el equivalente c en el espacio físico es automáticamente cambiado. En las aplicaciones se escoge \bar{c} como una constante. El valor óptimo depende del número de nodos en el soporte.

3.4. Interpolación Local de Hermite

Supongamos que se busca resolver el problema (3.0.3) y se tienen nodos aleatorios distribuidos sobre el dominio $\bar{\Omega}$ separados en 3 categorías: nodos-solución $\{x_i\}_{i=1}^{N_S} \subset \Omega$, nodos-PDE $\{x_i\}_{i=N_S+1}^{N_S+N_L} \subset \Omega$ y nodos-frontera $\{x_i\}_{i=N_S+N_L+1}^N \subset \partial\Omega$. Al igual que en el método de cuadratura diferencial, se encuentran los $N^{(k)}$ nodos más cercanos a un nodo solución x_k dado para formar un subsistema (stencil). Denotemos por $\mathcal{I}_S^{(k)}$, $\mathcal{I}_L^{(k)}$ y $\mathcal{I}_B^{(k)}$ a los conjuntos de índices asociados a los nodos-solución, nodos-PDE y nodos-frontera en el subsistema k respectivamente. El método de interpolación local hermitiana comienza imponiendo en el subsistema al ansatz simétrico

$$\hat{u}^{(k)}(x) = \sum_{j \in \mathcal{I}_S^{(k)}} \alpha_j^{(k)} \Phi(x - x_j) + \sum_{j \in \mathcal{I}_L^{(k)}} \alpha_j^{(k)} \bar{\mathcal{L}}\Phi(x - x_j) + \sum_{j \in \mathcal{I}_B^{(k)}} \alpha_j^{(k)} \bar{\mathcal{B}}\Phi(x - x_j) + \sum_{j=1}^{N_P^{(k)}} \alpha_{j+N}^{(k)} P_{m-1}^j(x), \quad (3.4.1)$$

donde Φ es un kernel radial condicionalmente positivo definido de orden m , $N_P^{(k)}$ es el número de términos en un polinomio de grado $m - 1$ en \mathbb{R}^d y los operadores $\bar{\mathcal{L}}$, $\bar{\mathcal{B}}$ actúan sobre la segunda variable.

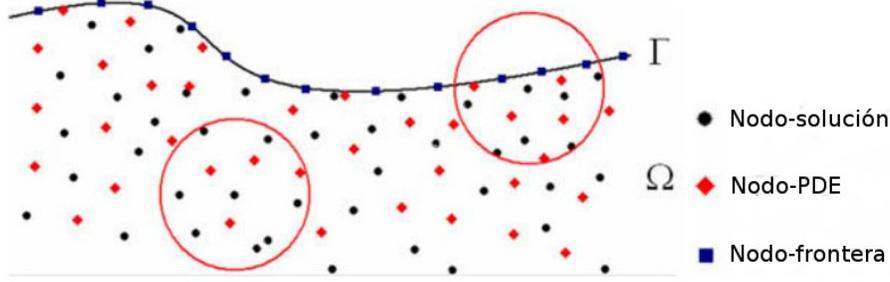


Figura 2: Nodos de soporte

En el subsistema se satisfacen las condiciones:

$$\begin{cases} \hat{u}^{(k)}(x_i) = h_i, & i \in \mathcal{I}_S^{(k)}, \\ \mathcal{L}\hat{u}^{(k)}(x_i) = f(x_i), & i \in \mathcal{I}_L^{(k)} \\ \mathcal{B}\hat{u}^{(k)}(x_i) = g(x_i), & i \in \mathcal{I}_B^{(k)}, \end{cases} \quad (3.4.2)$$

donde h_i son valores a determinar. Sustituyendo las condiciones anteriores en el Ansatz (3.4.1) se obtiene el sistema local $\mathbf{A}^{(k)}\alpha^{(k)} = \mathbf{b}^{(k)}$, donde

$$\mathbf{A}^{(k)} = \begin{bmatrix} \Phi_{ij} & \bar{\mathcal{L}}\Phi_{ij} & \bar{\mathcal{B}}\Phi_{ij} & P_{m-1} \\ \mathcal{L}\Phi_{ij} & \mathcal{L}\bar{\mathcal{L}}\Phi_{ij} & \mathcal{L}\bar{\mathcal{B}}\Phi_{ij} & \mathcal{L}P_{m-1} \\ \mathcal{B}\Phi_{ij} & \mathcal{B}\bar{\mathcal{L}}\Phi_{ij} & \mathcal{B}\bar{\mathcal{B}}\Phi_{ij} & \mathcal{B}P_{m-1} \\ P_{m-1}^T & \bar{\mathcal{L}}P_{m-1}^T & \bar{\mathcal{B}}P_{m-1}^T & 0 \end{bmatrix}, \quad \mathbf{b}^{(k)} = \begin{bmatrix} h \\ f \\ g \\ 0 \end{bmatrix}. \quad (3.4.3)$$

Invirtiendo la matriz $\mathbf{A}^{(k)}$, podemos escribir a cada componente de $\alpha^{(k)}$ de la forma

$$\alpha_i^{(k)} = \sum_{j \in \mathcal{I}_S^{(k)}} a_{ij}^{(k)} h_j + c_j^{(k)}, \quad j \in \mathcal{I}_S^{(k)} \cup \mathcal{I}_L^{(k)} \cup \mathcal{I}_B^{(k)} \cup \{N+1, \dots, N+N_P^{(k)}\}. \quad (3.4.4)$$

Sustituyendo la ecuación anterior en el ansatz (3.4.1) y utilizando que en cada nodo solución x_k se cumple $\mathcal{L}\hat{u}^{(k)}(x_k) = f(x_k)$, se obtiene el sistema global *sparse* $\mathbf{S}h = \hat{\mathbf{f}}$, donde

$$S_{kl} = \begin{cases} \sum_{j \in \mathcal{I}_S^{(k)}} a_{jl}^{(k)} \mathcal{L}\Phi(x_k - x_j) + \sum_{j \in \mathcal{I}_L^{(k)}} a_{jl}^{(k)} \mathcal{L}\bar{\mathcal{L}}\Phi(x_k - x_j) + \sum_{j \in \mathcal{I}_B^{(k)}} a_{jl}^{(k)} \mathcal{L}\bar{\mathcal{B}}\Phi(x_k - x_j) + \sum_{j=1}^{N_P^{(k)}} a_{j+N,l}^{(k)} P_{m-1}^j(x_k), & l \in \mathcal{I}_S^{(k)} \\ 0, & j \notin \mathcal{I}_S^{(k)}. \end{cases}$$

$$\hat{\mathbf{f}}_k = f(x_k) - \sum_{j \in \mathcal{I}_S^{(k)}} c_j^{(k)} \mathcal{L}\Phi(x_k - x_j) - \sum_{j \in \mathcal{I}_L^{(k)}} c_j^{(k)} \mathcal{L}\bar{\mathcal{L}}\Phi(x_k - x_j) - \sum_{j \in \mathcal{I}_B^{(k)}} c_j^{(k)} \mathcal{L}\bar{\mathcal{B}}\Phi(x_k - x_j) - \sum_{j=1}^{N_P^{(k)}} c_{j+N}^{(k)} P_{m-1}^j(x_k).$$

Resolviendo este sistema se encuentran los valores de u en los nodos-solución. Nota que para formar el renglón k de la matriz \mathbf{S} sólo se necesita la inversa de la matriz local $\mathbf{A}^{(k)}$, por lo que después de formar el renglón se puede liberar memoria. El sistema global, al ser *sparse*, puede ser invertido eficientemente por distintos métodos, tales como GMRES. Una de las ventajas del método de interpolación local de Hermite es que las condiciones de frontera ya están integradas desde el ansatz (3.4.1). A continuación se resume el algoritmo descrito anteriormente

Algorithm 1 Algoritmo LHI

- 1: **procedure** LHI(nodos-solución, nodos-PDE, nodos-frontera)
 - 2: **for** $k < N_s$ **do**
 - 3: Construir $\mathbf{A}^{(k)}$
 - 4: Calcular la inversa de $\mathbf{A}^{(k)}$
 - 5: Calcular $\alpha^{(k)} = \mathbf{A}^{(k)^{-1}} \mathbf{b}^{(k)}$
 - 6: Construir el k -ésimo renglón de \mathbf{S}
 - 7: $k = k + 1$
 - 8: **end for**
 - 9: Calcular $h = \mathbf{S}^{-1} \hat{\mathbf{f}}$
 - 10: **return** h ▷ solución en los nodos-solución
 - 11: **end procedure**
-

4. Métodos Iterativos

Como se vió en secciones anteriores, al discretizar una ecuación diferencial parcial lineal utilizando funciones de base radial, se obtiene un sistema de ecuaciones de la forma

$$A\mathbf{x} = \mathbf{b}. \quad (4.0.5)$$

Además, cuando se utiliza un método local (e.g. cuadratura diferencial), la matriz A resulta ser *sparse*, i.e. muchas de sus entradas son cero. En la práctica es conveniente tener esta propiedad en mente, tanto al almacenar las entradas de la matriz, como al escoger el método para resolver el sistema lineal. Para este último punto resultan ser muy útiles los llamados métodos iterativos.

Los métodos iterativos construyen la solución del sistema (4.0.5) como el límite de una sucesión $\{\mathbf{x}^k\}$, y usualmente involucran a la matriz A sólomente a través de multiplicaciones con vectores dados. De manera general, cualquier método iterativo está basado en una descomposición adecuada de la matriz A :

$$A = N - M, \quad (4.0.6)$$

donde N es invertible. La sucesión se genera entonces a partir de la relación

$$N\mathbf{x}^{k+1} = M\mathbf{x}^k + \mathbf{b}, \quad k \geq 0, \quad (4.0.7)$$

con \mathbf{x}^0 dado. Observa que si la sucesión converge, digamos a un vector \mathbf{x} , entonces se tiene que $N\mathbf{x} = M\mathbf{x} + \mathbf{b}$, i.e. \mathbf{x} es solución del sistema lineal.

La utilidad de un método iterativo depende básicamente de 2 cosas: qué tan rápido converge la sucesión $\{\mathbf{x}^k\}$ y qué tan fácil es invertir la matriz N . Para analizar el primer punto, si definimos

$$\mathbf{e}^k := \mathbf{x}^k - \mathbf{x}, \quad B := N^{-1}M, \quad (4.0.8)$$

se puede ver que el error \mathbf{e}^k satisface

$$\mathbf{e}^k = B\mathbf{e}^{k-1} = B^k\mathbf{e}^0. \quad (4.0.9)$$

De esta manera, la sucesión $\{\mathbf{x}^k\}$ converge a la sucesión del problema lineal \mathbf{x} para cualquier vector inicial \mathbf{x}^0 , si y sólo si el radio espectral¹² de la *matriz de iteración* B satisface

$$\rho(B) < 1. \quad (4.0.10)$$

4.1. El Método de Jacobi

El método de Jacobi es tal vez el método iterativo más simple. Se puede aplicar a matrices con elementos en la diagonal distintos de cero. Primero escribamos a la matriz a invertir como

¹²Recuerda que el radio espectral de una matriz se define como $\rho(A) = \max_{i \in \{1, \dots, n\}} |\lambda_i(A)|$, donde $\{\lambda_i\}_i$ son los eigenvalores de A

$A = L + D + U$, donde

$$L = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ A_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ A_{n-1,1} & \ddots & \ddots & \ddots & \vdots \\ A_{n1} & A_{n,2} & \cdots & A_{n,n-1} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} A_{11} & 0 & \cdots & 0 & 0 \\ 0 & A_{22} & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & A_{nn} \end{pmatrix}, \quad (4.1.1)$$

$$U = \begin{pmatrix} 0 & A_{12} & \cdots & A_{1,n-1} & A_{1,n} \\ 0 & 0 & \cdots & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & A_{n-1,n} \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

El método se basa en la factorización

$$N = D, \quad M = -(L + U), \quad (4.1.2)$$

y así, la matriz de iteración será

$$B_J = -D^{-1}(L + U). \quad (4.1.3)$$

Si definimos al residuo como $\mathbf{r}^k = b - A\mathbf{x}^k$, cada iteración del Método de Jacobi se puede escribir como

$$\mathbf{x}^{k+1} - \mathbf{x}^k = D^{-1}\mathbf{r}^k. \quad (4.1.4)$$

En componentes este algoritmo se escribe como

$$x_i^{k+1} = \frac{1}{A_{ii}} \left(b_i - \sum_{j \neq i}^n A_{ij} x_j^k \right) \quad i \in \{1, \dots, n\}. \quad (4.1.5)$$

Para analizar la convergencia del método de Jacobi, notemos que¹³

$$\rho(B_J) \leq \|D^{-1}(E + F)\|_\infty = \max_{i \in \{1, \dots, n\}} \sum_{j \neq i}^n \left| \frac{A_{ij}}{A_{ii}} \right| < 1, \quad (4.1.6)$$

es decir, si la matriz A es estrictamente diagonal dominante¹⁴ el método converge. Esta condición es bastante restrictiva; por ejemplo, la matriz resultante de una discretización de cinco puntos del laplaciano no satisface esta condición. Para relajar un poco esta condición necesitamos un par de definiciones.

¹³Se puede probar que para cualquier matriz B se cumple $\rho(B) \leq \|B\|$, donde $\|\cdot\|$ es una norma matricial natural, i.e. es de la forma $\|B\| = \max_{\|x\|=1} \|Bx\|$

¹⁴Una matriz A es estrictamente diagonal dominante si $|A_{ii}| > \sum_{j \neq i}^n |A_{ij}|$, para $i \in \{1, \dots, n\}$.

Definición 4.1.1. Una matriz A es irreducible si no hay una matriz de permutación Π (i.e. asociada a cambios de renglón o columna) tal que $\Pi^T A \Pi$ sea de la forma

$$\begin{pmatrix} M_1 & M_2 \\ 0 & M_3 \end{pmatrix}$$

Definición 4.1.2. Una matriz A es débilmente diagonal dominante si para cualquier $i \in \{1, \dots, n\}$ se tiene

$$|A_{ii}| \geq \sum_{j \neq i}^n |A_{ij}|$$

y existe al menos algún $k \in \{1, \dots, n\}$ tal que la desigualdad es estricta.

Utilizando las dos definiciones anteriores se puede demostrar el siguiente resultado ([9])

Teorema 4.1.3. Si A es una matriz irreducible y débilmente diagonal dominante, entonces el método de Jacobi converge.

4.2. El Método de Gauss-Seidel

El método de Gauss-Seidel se aplica, como en el caso del método de Jacobi, en matrices con elementos en la diagonal distintos de cero y resulta de escoger a las matrices N y M como

$$N := D + L, \quad M := -U, \quad (4.2.1)$$

Así, la matriz de iteración de Gauss-Seidel es

$$B_{GS} = -(D + L)^{-1}U. \quad (4.2.2)$$

En componentes, la iteración de Gauss-Seidel se puede escribir como

$$x_i^{k+1} = \frac{1}{A_{ii}} \left(b_i - \sum_{j=1}^{i-1} A_{ij} x_j^{k+1} - \sum_{j=i+1}^n A_{ij} x_j^k \right) \quad i \in \{1, \dots, n\}. \quad (4.2.3)$$

Para obtener un resultado acerca de la convergencia del método de Gauss-Seidel, recordemos el siguiente teorema.

Teorema 4.2.1. (Householder-John) Sea A una matriz simétrica y positiva definida. Si $N + N^T - A$ es positiva definida, entonces $\rho(N^{-1}M) < 1$.

De esta manera, si la matriz A es simétrica y positiva definida, entonces la iteración de Gauss-Seidel converge. En efecto,

$$N + N^T - A = (D + L) + (D + L)^T - (D + L + U) = D,$$

la cual es positiva definida. Además, también es posible demostrar que el método de Gauss-Seidel converge en los casos considerados en la sección anterior, i.e. cuando A es irreducible y diagonal

dominante, o estrictamente diagonal dominante.

Existen en la literatura algunos resultados que sirven para comparar los radios espectrales de las matrices de iteración de Jacobi y Gauss-Seidel, de los cuales se observa que Gauss-Seidel normalmente converge más rápido. Uno de esos resultados es el siguiente ([9])

Teorema 4.2.2. *Sea A una matriz tridiagonal. Si λ es un eigenvalor de B_J , entonces λ^2 será un eigenvalor de B_{GS} . De manera similar, si μ es un eigenvalor de B_{GS} , las raíces cuadradas de μ serán eigenvalores de B_J . De esta manera, el método de Jacobi converge si y sólo si el método de Gauss-Seidel converge, y además*

$$\rho(B_{GS}) = [\rho(B_J)]^2$$

4.3. El Método SOR

El método SOR (*successive over-relaxation* en inglés) es un procedimiento iterativo para acelerar el método de Gauss-Seidel con la ayuda de un parámetro positivo ω . En este caso se escoge

$$N := \frac{D}{\omega} + L, \quad M := \left(\frac{1}{\omega} - 1\right) D - U, \quad (4.3.1)$$

por lo que la matriz de iteración resulta

$$B_\omega = (D + \omega L)^{-1}[(1 - \omega)D - \omega U]. \quad (4.3.2)$$

Observa que si $\omega = 1$, lo anterior coincide con el método de Gauss-Seidel. En componentes, la iteración del método SOR toma la forma

$$x_i^{k+1} = (1 - \omega)x_i^k + \frac{\omega}{A_{ii}} \left(b_i - \sum_{j=1}^{i-1} A_{ij}x_j^{k+1} - \sum_{j=i+1}^n A_{ij}x_j^k \right). \quad (4.3.3)$$

El siguiente teorema proporciona una condición necesaria para la convergencia de este método.

Teorema 4.3.1. (Kahan) *La matriz de iteración del método SOR satisface*

$$\rho(B_\omega) \geq |\omega - 1|, \quad (4.3.4)$$

por lo que una condición necesaria para la convergencia es que

$$0 < \omega < 2. \quad (4.3.5)$$

Demostración. Si $\{\lambda_i\}_{i=1}^n$ es el conjunto de eigenvalores de la matriz de iteración B_ω , de (4.3.2) se tiene que

$$\prod_{i=1}^n |\lambda_i| = |\det B_\omega| = |\det D|^{-1} |\omega - 1|^n |\det D|.$$

Así, $\rho(B_\omega)^n \geq |\omega - 1|^n$, de donde se sigue el resultado. \square

Se puede probar que si A es irreducible y débilmente diagonal dominante, entonces el método SOR converge para $0 < \omega \leq 1$ (ver [9]). Por otro lado, la condición (4.3.5) es suficiente para las matrices simétricas y positivas definidas, como lo muestra el siguiente teorema.

Teorema 4.3.2. (Ostrowski-Reich) *Si A es una matriz simétrica y positiva definida, el método SOR converge si y sólo si se satisface la condición (4.3.5).*

Demostración. Notemos primero que

$$N + N^T - A = \left(\frac{D}{\omega} + L\right)^T + \left(\frac{1}{\omega} - 1\right) D - U = \left(\frac{2}{\omega} - 1\right) D,$$

por lo que utilizando la condición (4.3.5) y el teorema 4.2.1 se obtiene el resultado. \square

4.4. Métodos de Descenso

En esta sección supondremos que la matriz A es simétrica y positiva definida. Los métodos de descenso o gradiente se basan en la observación crucial de que resolver el sistema lineal (4.0.5) es equivalente a resolver el problema:

$$\text{Minimizar : } f(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, A\mathbf{x}) - (\mathbf{b}, \mathbf{x}). \quad (4.4.1)$$

Para ver esto, basta simplemente notar que un punto crítico (y por ser A positiva definida, por lo tanto un mínimo) tiene que satisfacer la condición

$$\nabla f(\mathbf{x}) = A\mathbf{x} - \mathbf{b} = 0.$$

Así, los métodos de descenso, en vez de buscar la solución de (4.0.5) directamente, buscan encontrar el mínimo de la función f . En general, un método de descenso sigue la estructura:

- Define una dirección de búsqueda $\mathbf{d}^{(k)}$.
- Minimiza $\alpha \mapsto \tilde{f}(\alpha) := f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$, y denota la solución por α_k .
- Define $\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} + \alpha_k\mathbf{d}^{(k)}$

Al segundo paso se le conoce normalmente como *búsqueda sobre líneas*, y el mínimo se puede calcular explícitamente. Para ello, nota que

$$\tilde{f}(\alpha) = \frac{1}{2}\alpha^2(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)}) + \alpha(\mathbf{d}^{(k)}, \mathbf{r}^{(k)}) + f(\mathbf{x}^{(k)}),$$

donde $\mathbf{r}^{(k)} := A\mathbf{x}^{(k)} - \mathbf{b} = \nabla f(\mathbf{x}^{(k)})$ es el residuo. Derivando se encuentra entonces que

$$\alpha_k = -\frac{\mathbf{d}^{(k)}\mathbf{r}^{(k)}}{(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)})}.$$

Debido a que A es positiva definida, la cantidad

$$\|\mathbf{d}\|_A := (\mathbf{d}, A\mathbf{d})^{1/2},$$

define una norma. Así, α_k es un número positivo bien definido.

Es conocido del cálculo que el vector $-\mathbf{r}^{(k)} = -\nabla f(\mathbf{x}^{(k)})$ apunta en la dirección hacia donde la función f decrece localmente más rápido. Esto motiva escoger la dirección de búsqueda como

$$\mathbf{d}^{(k)} := -\mathbf{r}^{(k)}. \quad (4.4.2)$$

Este es conocido como el *método del descenso más pronunciado*. Para este método se puede probar la siguiente estimación del error.

Teorema 4.4.1. *Para el método del descenso más pronunciado se satisface*

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_A \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|\mathbf{x}^{(0)} - \mathbf{x}\|_A, \quad (4.4.3)$$

donde κ es el número de condición espectral de la matriz A ¹⁵.

Este método presenta una velocidad de convergencia pobre, parecida a la del método de Jacobi, lo cual no es muy adecuado para utilizar en las aplicaciones. La razón de esto es que, aunque se puede probar que $(\mathbf{r}^{(k+1)}, \mathbf{d}^{(k)}) = (\mathbf{r}^{(k+1)}, -\mathbf{r}^{(k)}) = 0$, en general no se tiene necesariamente que $(\mathbf{r}^{(k+2)}, \mathbf{r}^{(k)}) = 0$. Por el contrario, estas direcciones pueden ser casi paralelas, como se puede ver en la Figura 3. Esto ocurre cuando la matriz A es mal condicionada, lo que implica que las curvas de nivel de f son elipses muy alargados en una dirección (las curvas de nivel siempre son elipses, porque A es positiva definida).

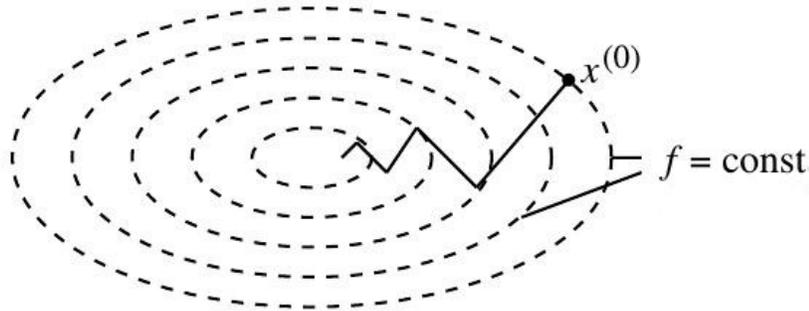


Figura 3: Comportamiento en zig-zag del método de descenso más pronunciado

El método de gradiente conjugado previene este tipo de comportamiento, y para presentarlo necesitamos primero introducir la noción de *vectores conjugados*.

Definición 4.4.2. *Se dice que dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ son conjugados, o A -ortogonales, si*

$$(\mathbf{x}, \mathbf{y})_A := (\mathbf{x}, A\mathbf{y}) = 0. \quad (4.4.4)$$

¹⁵Recuerda que el número de condición espectral es $\kappa = |\lambda_{max}|/|\lambda_{min}|$.

Si en el método se escogen direcciones de búsqueda que sean conjugadas, diremos que se trata de un *método de direcciones conjugadas*.

Supongamos que A es una matriz de $n \times n$, y sean $\mathbf{d}^{(0)}, \dots, \mathbf{d}^{(n-1)}$ direcciones conjugadas. Estos vectores son linealmente independientes y forman entonces una base. Así, podemos expresar a la solución como

$$\mathbf{x} = \sum_{k=0}^{n-1} \gamma_k \mathbf{d}^{(k)}.$$

Usando que las direcciones son conjugadas y que $A\mathbf{x} = \mathbf{b}$, se obtiene fácilmente que

$$\gamma_k = \frac{(\mathbf{d}^{(k)}, \mathbf{b})}{(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)})}, \quad (4.4.5)$$

y los coeficientes γ_k pueden entonces calcularse sin conocimiento alguno sobre \mathbf{x} . Si las direcciones de búsqueda están dadas, e.g. por algún proceso de ortogonalización de una base respecto a $(\cdot, \cdot)_A$, entonces la solución queda determinada.

Si aplicamos este mismo razonamiento al vector $\mathbf{x} - \mathbf{x}^{(0)}$, i.e. buscamos los coeficientes en la expresión

$$\mathbf{x} - \mathbf{x}^{(0)} = \sum_{k=0}^{n-1} \gamma_k \mathbf{d}^{(k)},$$

encontramos que

$$\gamma_k = -\frac{(\mathbf{d}^{(k)}, \mathbf{r}^{(0)})}{(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)})}. \quad (4.4.6)$$

Usando la expresión de la búsqueda sobre líneas, encontramos que

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + \sum_{i=0}^{k-1} \alpha_i \mathbf{d}^{(i)},$$

por lo que se tiene

$$\mathbf{r}^{(k)} = \mathbf{r}^{(0)} + \sum_{i=0}^{k-1} \alpha_i A\mathbf{d}^{(i)},$$

Para un método de direcciones conjugadas, esto implica que

$$(\mathbf{r}^{(k)}, \mathbf{d}^{(k)}) = (\mathbf{r}^{(0)}, \mathbf{d}^{(k)}),$$

y por lo tanto

$$\gamma_k = -\frac{\mathbf{d}^{(k)} \mathbf{r}^{(k)}}{(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)})} = \alpha_k,$$

lo que significa que $\mathbf{x} = \mathbf{x}^{(n)}$. Es decir, un método de direcciones conjugadas es *exacto* en a lo más n iteraciones. Por esta razón, durante mucho tiempo no se les consideró métodos iterativos. En la práctica n puede llegar a ser muy grande y el método se aplica de forma iterativa hasta que el residuo $\mathbf{r}^{(k)}$ sea lo suficientemente pequeño. El siguiente teorema presenta una propiedad de minimización muy importante de este tipo de métodos.

Teorema 4.4.3. *Los vectores $\mathbf{x}^{(k)}$ determinados por un método de direcciones conjugadas minimizan tanto la función f de (4.4.1), como el error $\|\mathbf{x}^{(k)} - \mathbf{x}\|_A$ en el subespacio afín $\mathbf{x}^{(0)} + \mathcal{K}_k(A, \mathbf{r}^{(0)})$, donde*

$$\mathcal{K}_k(A, \mathbf{r}^{(0)}) := \text{span}\{\mathbf{r}^{(0)}, A\mathbf{r}^{(0)}, \dots, A\mathbf{r}^{(k-1)}\} = \text{span}\{\mathbf{d}^{(0)}, \dots, \mathbf{d}^{(k-1)}\},$$

es el subespacio de Krylov de dimensión k de A respecto a $\mathbf{r}^{(0)}$.

En general, métodos que busquen minimizar el error o el residuo en $\mathcal{K}_k(A, \mathbf{r}^{(0)})$ respecto a alguna norma $\|\cdot\|$, son conocidos como *métodos de subespacios de Krylov*. En el caso aquí considerado se minimiza respecto a la norma $\|\cdot\|_A$.

El método de direcciones conjugadas más conocido es el llamado *método de gradiente conjugado*. En este método se determinan las direcciones de búsqueda durante la iteración utilizando el ansatz:

$$\mathbf{d}^{(k+1)} := -\mathbf{r}^{(k+1)} + \beta_k \mathbf{d}^{(k)}. \quad (4.4.7)$$

La condición necesaria $(\mathbf{d}^{(k+1)}, \mathbf{d}^{(k)})_A = 0$, conduce a

$$\beta_k = \frac{(\mathbf{r}^{(k+1)}, A\mathbf{d}^{(k)})}{(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)})}. \quad (4.4.8)$$

Es posible demostrar que los coeficientes α_k, β_k tienen otras expresiones equivalentes:

$$\alpha_k = \frac{(\mathbf{r}^{(k)}, \mathbf{r}^{(k)})}{(\mathbf{d}^{(k)}, A\mathbf{d}^{(k)})}, \quad \beta_k = \frac{(\mathbf{r}^{(k+1)}, \mathbf{r}^{(k+1)})}{(\mathbf{r}^{(k)}, \mathbf{r}^{(k)})}. \quad (4.4.9)$$

De esta manera, el número de operaciones requerido en una iteración se puede reducir a un producto de matriz con vector, dos productos escalares y tres operaciones de la forma $\mathbf{z} = \mathbf{x} + \alpha\mathbf{y}$, lo cual resulta computacionalmente mucho más adecuado para las aplicaciones. Además, la velocidad de convergencia resulta mejor que la del método de descenso más pronunciado, como lo indica el siguiente teorema.

Teorema 4.4.4. *Para el método de gradiente conjugado, se satisface*

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_A \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k \|\mathbf{x}^{(0)} - \mathbf{x}\|_A, \quad (4.4.10)$$

4.5. Métodos de Subespacio de Krylov para Matrices no Simétricas

En esta sección trataremos el caso en el que no es posible aplicar el método de gradiente conjugado; esto puede ocurrir cuando la matriz A no es simétrica, o si lo es, no es positiva definida. En cualquiera de estos casos, ya no es posible caracterizar la solución del sistema lineal (4.0.5)

como el mínimo de la función f en (4.4.1). No obstante, la solución \mathbf{x} sigue siendo el único vector que logra anular el cuadrado de la norma del residuo, y por lo tanto lo minimiza. En otras palabras, podemos analizar el problema:

$$\text{Encontrar } x \in \mathbb{R}^n \text{ tal que } g(\mathbf{x}) = \min_{\mathbf{y} \in \mathbb{R}^n} g(\mathbf{y}) = \|\mathbf{A}\mathbf{y} - \mathbf{b}\|_2^2. \quad (4.5.1)$$

En analogía al método del gradiente conjugado, la k -ésima iteración $\mathbf{x}^{(k)}$ se debe poder representar de la forma

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + \mathbf{z},$$

donde \mathbf{z} pertenece a un subespacio S_k , que esté generado por los residuos $\mathbf{r}^{(0)}, \dots, \mathbf{r}^{(k-1)}$, y cuya dimensión crezca con cada iteración. Si se escoge S_k como el subespacio de Krylov $\mathcal{K}_k(A, \mathbf{r}^{(0)})$, se obtienen los métodos denominados MINRES (Minimal Residual) y GMRES (Generalized Minimal Residual).

Ambos métodos consisten en encontrar una base ortonormal del subespacio de Krylov, y resolver el problema de minimización. El método MINRES se aplica cuando la matriz A es simétrica, lo cual simplifica varios cálculos; al igual que en el caso del método de gradiente conjugado, en el método MINRES es posible generar la base ortonormal utilizando una relación de recurrencia de 3 términos, lo cual es menos costoso computacionalmente. En cambio, el método GMRES se utiliza cuando la matriz no es simétrica, y es necesario utilizar todos los elementos calculados anteriormente.

Estos métodos (y muchos otros) están basados en el *método de Arnoldi* para construir iterativamente una base de $\mathcal{K}_k(A, \mathbf{r}^{(0)})$. El método combina la generación de una base de acuerdo a (4.4.7) y el proceso de ortogonalización de Gram-Schmidt. La idea del algoritmo se puede resumir de la siguiente forma:

- Sea $\mathbf{r}^{(0)} \in \mathbb{R}^n$, $\mathbf{r}^{(0)} \neq 0$ un vector dado. Define

$$\mathbf{v}^{(1)} := \mathbf{r}^{(0)} / \|\mathbf{r}^{(0)}\|_2.$$

- Para $j \in \{1, \dots, k\}$, calcula

$$\begin{aligned} h_{ij} &:= \langle \mathbf{v}^{(i)}, \mathbf{A}\mathbf{v}^{(j)} \rangle, \text{ para } i \in \{1, \dots, j\}, \\ \mathbf{w}^{(j)} &:= \mathbf{A}\mathbf{v}^{(j)} - \sum_{i=1}^j h_{ij} \mathbf{w}^{(i)}, \\ h_{j+1,j} &:= \|\mathbf{w}^{(j)}\|_2. \end{aligned}$$

- Si $h_{j+1,j} = 0$, termina. En otro caso, define

$$\mathbf{v}^{(j+1)} := \mathbf{w}^{(j)} / h_{j+1,j}$$

Si el método de Arnoldi se puede llevar a cabo hasta el índice k , definamos:

$$\begin{aligned} h_{ij} &:= 0, \quad \text{para } j \in \{1, \dots, k\}, i \in \{j+2, \dots, k+1\}, \\ H_k &:= (h_{ij})_{ij} \in \mathcal{M}_{k \times k}(\mathbb{R}), \\ \tilde{H}_k &:= (h_{ij})_{ij} \in \mathcal{M}_{(k+1) \times k}(\mathbb{R}), \\ V_{k+1} &:= (\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(k+1)}). \end{aligned}$$

El siguiente teorema enuncia las propiedades fundamentales del método GMRES.

Teorema 4.5.1. *Si el método de Arnoldi se puede llevar a cabo hasta el índice k , entonces:*

1. Los vectores $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(k+1)}$ forman una base ortonormal de $\mathcal{K}_{k+1}(A, \mathbf{r}^{(0)})$.

2. Se tiene

$$AV_k = V_k H_k + \mathbf{w}^{(k)} \mathbf{e}^{(k)T} = V_{k+1} \tilde{H}_k,$$

donde $\mathbf{e}^{(k)} = (0, \dots, 0, 1)^T \in \mathbb{R}^k$,

$$V_k^T AV_k = H_k.$$

3. El problema

$$\text{Encuentra } \mathbf{x}^{(k)} \text{ t.q. } g(\mathbf{x}^{(k)}) = \min_{y \in \mathbf{x}^{(0)} + \mathcal{K}_k(A, \mathbf{r}^{(0)})} g(\mathbf{y}) = \|\mathbf{A}\mathbf{y} - \mathbf{b}\|_2,$$

es equivalente a

$$\text{Encuentra } \xi^{(k)} \text{ t.q. } \tilde{g}(\xi^{(k)}) = \min_{\xi \in \mathbb{R}^k} \tilde{g}(\xi) = \|\tilde{H}_k \xi - \beta \mathbf{e}^{(1)}\|_2,$$

donde $\beta := -\|\mathbf{r}^{(0)}\|_2$, y además

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + V_k \xi^{(k)}.$$

Si el método de Arnoldi termina en el índice k , entonces

$$\mathbf{x}^{(k)} = \mathbf{x} = A^{-1} \mathbf{b}.$$

En la implementación del algoritmo se usa la propiedad (3) del teorema anterior, porque el minimizador $\xi^{(k)}$ es computacionalmente más barato de obtener, ya que requiere sólo la solución de un sistema de mínimos cuadrados de $(k+1) \times k$, donde k es típicamente mucho más pequeño que n .

5. Referencias

Referencias Básicas

- [1] Fasshauer ,G.E., *Meshfree Approximation Methods with MATLAB*, World Scientific, 2007.
- [2] Wendland, H., *Scattered Data Approximation*, Cambridge University Press, 2005.

Referencias

- [3] Chen, W., Zhuo-Jia, F., Chen, C.S., *Recent Advances in Radial Basis Function Collocation Methods* Springer, 2014.
- [4] Franke, C., Schaback, R., Solving Partial Differential Equations by Collocation Using Radial Basis Functions *Appl. Math. Comp.* **93**, pp. 73-82. 1998
- [5] Franke, C., Schaback, R., Convergence Orders of Meshless Collocation Methods Using Radial Basis Functions *Adv. in Comput. Math.* **8**, pp. 381-399. 1998
- [6] Iske, A., Reconstruction of Functions from Generalized Hermite-Birkhoff Data *Approximation Theory VII, Vol 1: Approximation and Interpolation*, C. Chui, and L. Schumaker (eds.), World Scientific Publishing(Singapore), pp. 257-264. 1995
- [7] Wu, Z., Hermite-Birkhoff Interpolation of Scattered Data by Radial Basis Functions *Approx. Theory Appl.* **8**, pp. 1-10. 1992
- [8] Quarteroni, A., Valli, A., *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, Springer 2008.
- [9] Young, D.M., *Iterative Solution of Large Linear Systems*, Academic Press 1971.
- [10] Knabner, P., Angermann, L., *Numerical Methods for Elliptic and Parabolic Partial Differential Equations*, Springer 2003.