4. Reducción a dimensión finita de problemas para EDPs elípticas lineales

Los problemas que interesa resolver, tales como (26), (30), etc. tienen la estructura siguiente:

$$\begin{cases} \text{Hallar} \quad u \in V \quad \text{tal que} \\ a(u,v) = \ell(v) \quad \forall v \in V, \end{cases} \tag{38}$$

donde V es un subespacio cerrado de $H^1(\Omega)$ de dimensión infinita, $a(\cdot\,,\cdot) : V \times V \mapsto \mathbf{R}$ es una forma bilineal continua y coerciva y $\ell : V \mapsto \mathbf{R}$ es una forma lineal continua.

Obviamente, el cálculo exacto de u está fuera de nuestro alcance. Por tanto, tiene sentido formular y resolver problemas aproximados.

En este contexto, la aproximación más natural se consigue considerando problemas totalmente análogos donde V queda sustitutido por un subespacio V_h de dimensión tal vez grande pero finita:

$$\begin{cases} \text{Hallar} \quad u_h \in V_h \quad \text{tal que} \\ a(u_h,v_h) = \ell(v_h) \quad \forall v_h \in V_h. \end{cases} \tag{39}$$

Con carácter general, el paso de (40) a (39) se suele denominar *procedimiento de Galerkin.*

La sección siguiente está dedicada a un método de construcción de V_h que posee un gran número de ventajas.

Observación 4.1 Cuando la forma $a(\cdot\,,\cdot)$ es simétrica, (40) se puede reformular como un problema de mínimos:

$$\begin{cases} \text{Hallar} \quad u \in V \quad \text{tal que} \\ J(u) \le J(v) \quad \forall v \in V, \end{cases} \tag{40}$$

donde tenemos

$$J(v) := \frac{1}{2} a(v,v) - \ell(v) \quad \forall v \in V.$$

Por tanto, otro modo completamente natural de aproximar (40) consiste en tratar de minimizar J en V_h. Este método suele denominarse *procedimiento de Ritz.* □

4.1. Construcción de un subespacio de $H^1(\Omega)$: elementos finitos de tipo P_k-Lagrange

La herramienta de partida será una *triangulación* de Ω.

Por definición, una triangulación $\mathcal{T}_h$ es una familia finita de N-símplices K (intervalos cerrados si $N = 1$, triángulos cerrados si $N = 2$, tetraedros cerrados si $N = 3$) con las propiedades siguientes:

(I) $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$.

(II) Si $K, K' \in \mathcal{T}_h$ y $K \neq K'$, entonces $K \cap K'$ sólo puede ser el vacío o un vértice, lado o cara común a K y K'.

Fijada una triangulación $\mathcal{T}_h$ de $\overline{\Omega}$, pondremos

$$W_h := \{\, v_h \in C^0(\overline{\Omega}) : v_h\big|_K \in \mathbb{P}_1(K) \quad \forall K \in \mathcal{T}_h \,\},$$

donde, para cada K, $\mathbb{P}_1(K)$ es el espacio vectorial de las funciones polinómicas de grado ≤ 1 en las variables $x_1, \dots, x_N$. Nótese que la dimensión de $\mathbb{P}_1(K)$ es $N + 1$.

El conjunto W_h es un espacio vectorial para las operaciones habituales. Obviamente, su dimensión es finita. Tenemos además el resultado siguiente:

Teorema 4.1 *W_h es un subespacio de $H^1(\Omega)$ cuya dimensión coincide con el número de vértices de $\mathcal{T}_h$. De hecho, si denotamos $\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^I$ los vértices de $\mathcal{T}_h$, fijados los números reales $\alpha^1, \alpha^2, \dots, \alpha^I$, existe una única $v_h \in W_h$ que verifica*

$$v_h(\mathbf{a}^i) = \alpha^i \quad \forall i = 1, \dots, I. \tag{41}$$

Prueba: Por brevedad, probaremos el teorema cuando $N = 3$. Para $N = 1$ o $N = 2$, la prueba es análoga e incluso más sencilla.

Veamos en primer lugar que $W_h \subset H^1(\Omega)$.

Sea $v_h \in W_h$ y consideremos la función w_h, definida c.p.d. en Ω como sigue:

$$w_h = \partial_1 \left(v_h\big|_K \right) \text{ en } \overset{\circ}{K} \quad \forall K \subset \mathcal{T}_h.$$

Veamos que w_h es la derivada generalizada de v_h respecto de x_1.

En efecto, si $\varphi \in \mathcal{D}(\Omega)$, tenemos:

$$\begin{aligned}
\int_\Omega w_h \, \varphi \, d\mathbf{x} &= \sum_{K \in \mathcal{T}_h} \int_K w_h \, \varphi \, d\mathbf{x} = \sum_{K \in \mathcal{T}_h} \int_K \partial_1 \left(v_h\big|_K \right) \varphi \, d\mathbf{x} \\
&= \sum_{K \in \mathcal{T}_h} \left(- \int_K v_h \, \partial_1 \varphi \, d\mathbf{x} + \int_{\partial K} v_h \, \varphi \, n_1^K \, d\Gamma \right) \\
&= - \int_\Omega v_h \, \partial_1 \varphi \, d\mathbf{x} + \sum_{K \in \mathcal{T}_h} \int_{\partial K} v_h \, \varphi \, n_1^K \, d\Gamma,
\end{aligned}$$

donde $n_1^K = n_1^K(\mathbf{x})$ es, para cada $K \in \mathcal{T}_h$, la primera componente del vector normal exterior a K en los puntos de ∂K.

En la igualdad precedente, el último término se puede escribir como una suma de integrales sobre las caras de los tetraedros $K \in \mathcal{T}_h$. Algunas caras están sobre $\partial\Omega$; las correspondientes integrales se anulan, dado que $\varphi = 0$ sobre $\partial\Omega$. Las otras caras son siempre comunes a dos de los tetraedros, por lo que aparecen dos veces en la suma. Como $v_h \in C^0(\overline{\Omega})$, vemos que los correspondientes integrandos se cancelan (si la cara en cuestión coincide con $K \cap K'$, tenemos $n_1^K = -n_1^{K'}$). Por tanto,

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} v_h \, \varphi \, n_1^K \, d\Gamma = 0$$

y tenemos que

$$\int_\Omega w_h \, \varphi \, d\mathbf{x} = - \int_\Omega v_h \, \partial_1 \varphi \, d\mathbf{x}.$$

Dado que φ es arbitraria en $\mathcal{D}(\Omega)$, obtenemos lo anunciado.

De igual modo se puede proceder con las otras $\partial_i v_h$. La conclusión es que $v_h \in H^1(\Omega)$ (y, de hecho, las derivadas generalizadas de v_h pertenecen a $L^\infty(\Omega)$).

Sea $\{ \mathbf{a}^1, \dots, \mathbf{a}^I \}$ el conjunto de los vértices de $\mathcal{T}_h$ y consideremos la aplicación $M : W_h \mapsto \mathbf{R}^I$ definida como sigue:

$$M v_h = (v_h(\mathbf{a}^1), \dots, v(\mathbf{a}^I)) \quad \forall v_h \in W_h.$$

Se trata de una aplicación lineal del espacio vectorial W_h en $\mathbf{R}^I$. Veamos que M es inyectiva y sobre, con lo cual quedará demostrado el teorema.

Si $M v_h = 0$, entonces, en cada tetraedro $K \in \mathcal{T}_h$, $v_h\big|_K$ es una función polinómica de grado ≤ 1 que se anula en los cuatro vértices. Se observa fácilmente que esto implica $v_h\big|_K \equiv 0$. Por tanto, $v_h = 0$ y este argumento muestra que M es inyectiva.

Por otra parte, supongamos dado $(\alpha^1, \dots, \alpha^I)$ en $\mathbf{R}^I$. Fijemos K en $\mathcal{T}_h$ y supongamos que los vértices de K son $\mathbf{a}^i$, $\mathbf{a}^j$, $\mathbf{a}^k$ y $\mathbf{a}^m$. Es inmediato que existe una única función p^K en $\mathbb{P}_1(K)$ verificando

$$p^K(\mathbf{a}^i) = \alpha^i,\ p^K(\mathbf{a}^j) = \alpha^j,\ p^K(\mathbf{a}^k) = \alpha^k \ \text{ y } \ p^K(\mathbf{a}^m) = \alpha^m.$$

Pondremos entonces $v_h = p^K$ en K. Repitiendo este proceso en todo otro tetredro de $\mathcal{T}_h$, seremos así capaces de definir v_h en todo el conjunto $\overline{\Omega}$. Observemos que esta definición es correcta: en efecto, sobre una cara común a dos tetraedros K y K', p^K y $p^{K'}$ coinciden, dado que se trata de dos funciones polinómicas en dos variables que son idénticas en tres puntos no alineados. Además, por construcción, $v_h \in W_h$ y $Mv_h = (\alpha^1, \dots, \alpha^I)$. Esto muestra que M es sobre. □

Por definición, W_h es un espacio de *elementos finitos.* Más precisamente, se dice que para cada $K \in \mathcal{T}_h$ la terna $\{K, \mathbb{P}_1(K), \Sigma_K\}$, donde Σ_K es el conjunto de los vértices de K, es un elemento finito *de tipo* P_1*-Lagrange.*

Se puede generalizar este procedimiento de construcción de subespacios de $H^1(\Omega)$ en varias direcciones. Por ejemplo, podemos definir

$$Z_h := \{\, v_h \in C^0(\overline{\Omega}) : v_h\big|_K \in \mathbb{P}_2(K) \ \ \forall K \in \mathcal{T}_h \,\},$$

donde ahora $\mathbb{P}_2(K)$ es el espacio vectorial de las funciones polinómicas de grado ≤ 2 en las variables $x_1, \dots, x_N$; obsérvese que la dimensión de $\mathbb{P}_2(K)$ es $\frac{1}{2}(N+1)(N+2)$.

Con una prueba similar a la del Teorema 4.1, tenemos:

Teorema 4.2 *Z_h es un subespacio de $H^1(\Omega)$ cuya dimensión coincide con la suma del número de vértices y de puntos medios de los lados de $\mathcal{T}_h$. De hecho, si denotamos $\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^I$ los vértices y $\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^H$ los puntos medios de los lados, fijados los números reales $\alpha^1, \alpha^2, \dots, \alpha^I$ y $\beta^1, \beta^2, \dots, \beta^H$, existe una única $z_h \in Z_h$ que verifica*

$$z_h(\mathbf{a}^i) = \alpha^i \quad \forall i = 1, \dots, I, \quad z_h(\mathbf{b}^j) = \beta^j \quad \forall j = 1, \dots, H. \tag{42}$$

Obtenemos así el nuevo espacio de elementos finitos Z_h. Específicamente, se dice que, para cada $K \in \mathcal{T}_h$, la terna $\{K, \mathbb{P}_2(K), \Lambda_K\}$, donde Λ_K es el conjunto de los vértices y puntos medios de los lados de K, es un elemento finito *de tipo* P_2*-Lagrange.*

4.2. La aproximación numérica

Volvamos a la situación descrita en (40). Supongamos elegida una de las dos estrategias P_1-Lagrange o P_2-Lagrange y sea $\{\mathcal{T}_h\}_{h>0}$ una familia de triangulaciones. Para cada $\mathcal{T}_h$, consideraremos el correspondiente problema (39), donde $V_h = W_h \cap V$ o $V_h = Z_h \cap V$. En esta sección mostraremos que, cuando la familia $\{\mathcal{T}_h\}_{h>0}$ cumple condiciones adecuadas, las soluciones asociadas u_h convergen hacia la solución u de (40).

En primer lugar, tenemos el resultado siguiente:

Teorema 4.3 *Existe una constante $C > 0$ tal que, para todo $h > 0$, se tiene:*

$$\|u - u_h\|_{H^1} \leq C \inf_{v_h \in V_h} \|u - v_h\|_{H^1} \tag{43}$$

Prueba: Para cada $v_h \in V_h$ tenemos que

$$\begin{aligned} \alpha\|u - u_h\|_{H^1}^2 &\leq a(u - u_h, u - u_h) \\ &= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\ &= a(u - u_h, u - v_h) \\ &\leq M\|u - u_h\|_{H^1}\|u - v_h\|_{H^1}. \end{aligned}$$

Por tanto,

$$\|u - u_h\|_{H^1} \leq \frac{M}{\alpha}\|u - v_h\|_{H^1} \quad \forall v_h \in V_h,$$

de donde resulta (43). □

Así pues, el análisis de la convergencia de u_h hacia u se reduce a un problema de aproximación: fijada la familia $\{\mathcal{T}_h\}_{h>0}$, ¿cómo se comporta la cantidad

$$\text{dist.}\,(u, V_h) := \inf_{v_h \in V_h} \|u - v_h\|_{H^1}$$

cuando $h \to 0$?

Utilizaremos la notación siguiente:

$$\delta(K) := \sup_{\mathbf{x},\mathbf{x}' \in K} |\mathbf{x} - \mathbf{x}'| \quad \text{(el “diámetro” de } K),$$

$$\rho(K) := \max\{\, \delta(B) : B \ \text{ es una bola, } \ B \subset K \,\} \quad \text{(el “grosor” de } K).$$

Teorema 4.4 *Supongamos que $\{\mathcal{T}_h\}_{h>0}$ es una familia regular de triangulaciones, es decir:*

(i) $\delta_h := \max_{K \in \mathcal{T}_h} \delta(K) \to 0$ cuando $h \to 0$.

(ii) Existe una constante $C > 0$ tal que

$$\max_{K \in \mathcal{T}_h} \frac{\delta(K)}{\rho(K)} \leq C \quad \forall h > 0.$$

Entonces $u_h \to u$ en V cuando $h \to 0$.

Puede encontrarse la demostración por ejemplo en [1, 2]. Se basa en la existencia y propiedades de los operadores de interpolación asociados a los V_h.

Observación 4.2 Si u es suficientemente regular, se pueden decir más cosas. Así, en el caso de la aproximación P_1-Lagrange, existe una constante $C > 0$ (que depende de u pero es independiente de h) tal que

$$\|u - u_h\|_{H^1} \leq Ch.$$

Para la aproximación P_2-Lagrange, se puede probar un resultado análogo con una estimación mejor:

$$\|u - u_h\|_{H^1} \leq Ch^2.$$

Obervamos por tanto (una vez más) un fenómeno típico del análisis numérico: se puede conseguir una mejor aproximación a costa de aplicar un método más costoso, pero el trabajo computacional extra es necesario. □

4.3. Aspectos prácticos de la resolución numérica

En la práctica, el problema (39) se puede re-escribir como un sistema lineal de ecuaciones que sabemos que posee solución única.

En efecto, supongamos por ejemplo que $V = H^1(\Omega)$, hemos fijado la triangulación $\mathcal{T}_h$ de vértices los $\mathbf{a}^1, \dots, \mathbf{a}^I$ y hemos elegido la aproximación P_1-Lagrange ($V_h = W_h$). Sea $\{ \varphi^1, \dots, \varphi^I \}$ una base de V_h. Entonces, la tarea consiste en hallar $\xi_1, \dots, \xi_I \in \mathbf{R}$ tales que

$$\sum_{i=1}^{I} a(\varphi^i, \varphi^j)\xi_i = \ell(\varphi_j), \quad j = 1, \dots, I \tag{44}$$

y después tomar

$$u_h = \sum_{i=1}^{I} \xi_i \varphi^i.$$

En consecuencia, interesa una base tal que la matriz $A_h := \{a(\varphi^i, \varphi^j)\}$ tenga, a ser posible, las propiedades siguientes:

- Las componentes de A_h son fáciles de calcular.
- A_h es una matriz *hueca,* es decir, casi todos los $a(\varphi^i, \varphi^j)$ son cero.
- A_h es una matriz *banda,* esto es, sólo las componentes de A_h próximas a la diagonal principal son distintas de cero.

Es fácil comprender que las dos primeras condiciones se cumplen cuando se elige la *base canónica* de V_h. Por definición, se trata de la base caracterizada por lo siguiente:

$$\varphi^i \in V_h, \quad \varphi^i(\mathbf{a}^j) = \delta_{ij} \quad \forall i, j = 1, \dots, I. \tag{45}$$

En efecto, está claro que las funciones φ^i que verifican (45) constituyen una base de V_h.

Cada φ^i se anula en todos los $K \in \mathcal{T}_h$ a los que no pertenece el vértice $\mathbf{a}^i$. Dicho de otro modo, el soporte de φ^i coincide con la unión de los intervalos, triángulos o tetraedros de $\mathcal{T}_h$ que contienen al vértice $\mathbf{a}^i$. Por tanto, el cálculo de los $a(\varphi^i, \varphi^j)$ es sencillo; cada una de estas componentes se puede escribir como la suma de un número reducido de integrales extendidas a intervalos, triángulos o tetraedros.

Por ejemplo, en el caso del problema (30), tenemos

$$a(\varphi^i, \varphi^j) = \sum_{K \in \mathcal{K}_{ij}} \int_K \left(\sum_{m,n=1}^{N} a_{mn}(\mathbf{x}) \partial_m \varphi^i \, \partial_n \varphi^j + c(\mathbf{x}) \varphi^i \varphi^j \right) d\mathbf{x},$$

donde, por definición, $\mathcal{K}_{ij}$ es la familia de los $K \in \mathcal{T}_h$ tales que $\mathbf{a}^i, \mathbf{a}^j \in K$. Esta igualdad muestra también que, casi siempre, $a(\varphi^i, \varphi^j) = 0$ (para que sea $\neq 0$, debe existir un K que contenga a la vez $\mathbf{a}^i$ y $\mathbf{a}^j$).

Además, observemos que es relativamente fácil conseguir que la matriz A_h posea estructura banda. En efecto, para que esto sea cierto basta con que se dé la propiedad siguiente:

$$\mathbf{a}^i \text{ es próximo a } \mathbf{a}^j \;\Leftrightarrow\; i \text{ es próximo a } j.$$

Y esto se obtiene con una enumeración adecuada de los vértices de $\mathcal{T}_h$.

Consideraciones análogas pueden hacerse para los otros problemas presentados en la Sección 3. También es sencillo extender y adaptar estos comentarios al método P_2-Lagrange; para más detalles, véase [3].

La principal conclusión es que los métodos de elementos finitos expuestos son computacionalmente eficientes y apropiados para la resolución de problemas elípticos lineales en dimensión 2 y 3.

Bibliografía

[1] Brenner S.C., Scott L.R., *The mathematical theory of finite element methods, 2nd edition.* Texts in Applied Mathematics, 15. Springer-Verlag, New York, 2002.

[2] Ciarlet, Ph.G., *The finite element method for elliptic problems.* Classics in Applied Mathematics, 40. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[3] Hecht, F., *Freefem++, 3rd edition, Version 3-19.1.* University of Paris VI, Laboratoire Jacques-Louis Lions, `http://www.freefem.org/ff++/index.htm`.