

Elementos finitos avanzados

Francisco–Javier Sayas

Versión 2008

Introducción

Aquí están, en su versión n -ésima y todavía sin completar, las notas de la segunda parte de mis cursos de doctorado de elementos finitos, que he impartido durante varios años en la Universidad de Zaragoza. El material corresponde a un curso con el título *Elementos Finitos Avanzados* del programa de doctorado Mecánica Computacional. El curso estaba diseñado para un público mixto de matemáticos, ingenieros y físicos y contenía una parte de implementación (que desarrollamos conjuntamente Andrés Riaguas y un servidor), que no está contemplada en estas notas.

El lector que se enfrente a estas notas debe tener ya unos ciertos conocimientos sobre el Método de los Elementos Finitos tal y como lo explicamos y entendemos los matemáticos, con formulaciones variacionales, con formas bilineales, bases nodales, etc. Las notas del curso previo a este se pueden encontrar también en mi web, en una versión algo más elaborada (en inglés), bajo el título *A gentle introduction to the Finite Element Method*. Al igual que en esas notas, aquí se trata de aprender sobre formulaciones, sobre convergencia y sobre algoritmos, pero no se da el análisis que lleva a los resultados matemáticos. Se trata de un intento de reconciliar la literatura más práctica con la forma de escribir más matemática, sin renunciar al rigor, pero sin incluir el muchas veces complejo análisis numérico del método de los elementos finitos.

Este curso cubre una serie de temas avanzados que no se suelen dar en las presentaciones más básicas sobre el método. Hago notar que el punto de vista va a ser siempre el de ir de lo particular a lo general (bueno, hay una excepción parcial en la lección sobre formulaciones mixtas) y que en muchas ocasiones no voy a abordar los problemas en su plena generalidad. Desgraciadamente, en la última revisión que acometí de estas notas no llegué a concluir de escribir todas las secciones (faltan solo dos) y no es probable que haga esta tarea en breve, de manera que he preferido ‘postear’ estas notas como están. Y así va a quedar el asunto por el momento.

Minneapolis, septiembre de 2008

Lección 1

Elementos lagrangianos de orden alto

Las presentaciones elementales del Método de Elementos Finitos con elementos triangulares o sobre cuadriláteros se centran en los elementos basados en polinomios de orden bajo, para los cuales las bases nodales (basadas en interpolación de Lagrange sobre redes uniformes de puntos en el elemento) son más que suficientes. No obstante, si uno intenta emplear elementos de grado más alto o quiere emplear una estrategia p ó $h-p$ (esto quiere decir que vamos a incrementar el grado polinómico en vez de reducir la triangulación o que vamos a hacer las dos cosas), conviene tener bases construidas de otra manera. En esta línea presentamos aquí una de las varias posibles organizaciones de las bases de elementos finitos polinómicos libres-de-nodos, que permiten ordenaciones jerárquicas con las cuales es fácil subir el grado o combinar elementos con distintos grados.

1. Funciones de Lobatto

Consideremos la sucesión de los **polinomios de Legendre**:

$$L_0(\xi), \quad L_1(\xi), \quad \dots, \quad L_k(\xi), \quad \dots$$

que se pueden definir mediante una recurrencia a tres términos:

$$\begin{aligned} L_0(\xi) &:= 1 \\ L_1(\xi) &:= \xi \\ L_k(\xi) &:= \frac{2k-1}{k} \xi L_{k-1}(\xi) - \frac{k-1}{k} L_{k-2}(\xi), \quad k \geq 2. \end{aligned}$$

El polinomio L_k tiene grado exactamente k . Por esta razón

$$\mathbb{P}_k := \{ \alpha_0 + \alpha_1 \xi + \dots + \alpha_k \xi^k \mid \alpha_i \in \mathbb{R}, \quad \forall i \}$$

tiene como base a

$$\{L_0, L_1, \dots, L_{k-1}, L_k\}.$$

Nota. Los polinomios de Legendre constituyen una sucesión de polinomios ortogonales en $(-1, 1)$, ya que cumplen las condiciones

$$\int_{-1}^1 L_i(\xi) L_j(\xi) d\xi = 0, \quad i \neq j.$$

Además cumplen una condición de normalización

$$L_k(1) = 1.$$

Los polinomios de Legendre con índice k impar son impares. Los de índice par, son pares. \square

Se introducen las **funciones de Lobatto** de la siguiente manera. Primero se definen

$$\ell_0(\xi) := \frac{1 - \xi}{2}, \quad \ell_1(\xi) := \frac{\xi + 1}{2},$$

que cumplen

$$\ell_0(-1) = 1, \quad \ell_0(1) = 0, \quad \ell_1(-1) = 0, \quad \ell_1(1) = 1.$$

Para $k \geq 2$ se definen

$$\ell_k(\xi) := \frac{1}{\sqrt{\int_{-1}^1 |L_{k-1}(t)|^2 dt}} \int_{-1}^{\xi} L_{k-1}(t) dt.$$

Es también fácil comprobar que

$$\{\ell_0, \ell_1, \dots, \ell_k\}$$

es una base de \mathbb{P}_k y que

$$\ell_k(\pm 1) = 0, \quad k \geq 2.$$

Por esta razón se pueden factorizar las funciones de la siguiente manera

$$\ell_k(\xi) = \ell_0(\xi) \ell_1(\xi) \phi_{k-2}(\xi) = \frac{1 - \xi^2}{4} \phi_{k-2}(\xi),$$

donde ϕ_{k-2} es un polinomio de grado $k - 2$.

Nota. Las funciones $\{\ell_2, \dots, \ell_k\}$ forman una base del conjunto

$$\{p \in \mathbb{P}_k \mid p(1) = p(-1) = 0\}.$$

Las funciones $\{\phi_0, \dots, \phi_k\}$ vuelven a constituir una base de \mathbb{P}_k y tienen la misma propiedad de paridad que los polinomios de Legendre, esto es, $\phi_k(-\xi) = (-1)^k \phi_k(\xi)$. \square

2. Bases libres de nodos

En dimensión uno, las funciones de Lobatto permiten dar bases de espacios de elementos finitos ignorando los nodos de interpolación. Con ellas se ve fácilmente cómo la transición de grado entre elementos consecutivos es totalmente trivial.

Si vamos a discretizar el intervalo $[0, L]$, lo podemos hacer partiéndolo en intervalos como es habitual. Consideramos una partición

$$0 = x_0 < x_1 < \dots < x_{n-1} < x_n = L.$$

En cada elemento $I_j := [x_{j-1}, x_j]$ con $j = 1, \dots, n$ consideramos en primera instancia el espacio de polinomios de grado menor o igual que p . Empleando la transformación de $[-1, 1]$ en I_j

$$x = F_j(\xi) := \frac{x_j + x_{j-1}}{2} + \xi \frac{x_j - x_{j-1}}{2} = x_{j-1} \frac{1 - \xi}{2} + x_j \frac{\xi + 1}{2} = x_{j-1} \ell_0(\xi) + x_j \ell_1(\xi)$$

se puede construir una base de \mathbb{P}_p

$$\ell_{k,j}(x) = \ell_k(\xi), \quad k = 0, \dots, p,$$

donde las dos primeras funciones son lineales y las restantes son funciones internas (se las suele llamar *burbuja*s, aunque sean funciones oscilantes). La notación anterior es una forma simplificada de escribir que

$$\ell_{k,j}(F_j(\xi)) = \ell_k(\xi)$$

o equivalentemente

$$\ell_{k,j}(x) = \ell_k(F_j^{-1}(x)),$$

o en notación más analítica

$$\ell_{k,j} \circ F_j = \ell_k, \quad \ell_{k,j} = \ell_k \circ F_j^{-1}.$$

Ahora, del espacio global de elementos finitos

$$X_h^p := \left\{ u \in \mathcal{C}[0, L] \mid u|_{I_j} \in \mathbb{P}_p, \forall j \right\}$$

se puede construir una base pegando los trozos anteriores. Obtenemos así:

- Una función *gorro* por cada nodo. Por tanto, hay $n + 1$ de estas funciones, cuyo soporte son dos intervalos consecutivos y que son lineales.
- $p - 1$ funciones internas por intervalo.

La diferencia entre esta base y la base lagrangiana normal de este espacio estriba en dos hechos:

- La parte que conecta dos elementos consecutivos es fija, e independiente del grado p . Siempre es lineal. Con una base nodal típica, la función asociada a un extremo de intervalo tiene ceros sobre nodos internos de cada intervalo y tiene grado p .
- Las $p - 1$ funciones internas a cada intervalo I_j son las $p - 1$ primeras funciones internas si subimos el grado.

Este último hecho nos permite emplear todos los cálculos de un ensamblado para un problema X_h^p cuando cambiamos al espacio X_h^{p+1} , es decir, cuando fijamos el mallado y subimos el grado. Los llamados **métodos p** consisten precisamente en fijar el mallado y subir el grado de los elementos finitos, en lugar de refinar la partición.

No hay ninguna dificultad adicional en considerar que a la partición del intervalo van asociados unos grados polinómicos:

$$\begin{aligned} I_1 &= [x_0, x_1], & p_1 &\geq 1, \\ &\dots \\ I_n &= [x_{n-1}, x_n], & p_n &\geq 1. \end{aligned}$$

El espacio de elementos finitos se define entonces como

$$X_h^{\mathbb{P}} := \left\{ u \in \mathcal{C}[0, L] \mid u|_{I_j} \in \mathbb{P}(I_j) \forall j \right\}$$

siendo $\mathbb{P}(I_j) = \mathbb{P}_{p_j}$. Esto permite cómodamente tener distintos grados en distintos elementos. No obstante, al igual que ocurre con muchas otras cuestiones, la extensión de este tipo de técnicas a dimensión dos o tres, dista de ser simple. Este va a ser el objetivo de las siguientes secciones.

3. Construcción de bases de \mathbb{Q}_p

En dimensión dos, el conjunto \mathbb{Q}_p se define como el conjunto de las combinaciones lineales de los monomios

$$\xi^i \eta^j, \quad 0 \leq i, j \leq p.$$

Tiene, por tanto, dimensión $(p+1)^2$. El objetivo de lo que sigue es dar una construcción de una base especial de \mathbb{Q}_p . Después, mediante un cambio de variables afín, trasvasaremos esta base a la de un espacio $\mathbb{Q}_p(K)$, donde K es un paralelogramo arbitrario.

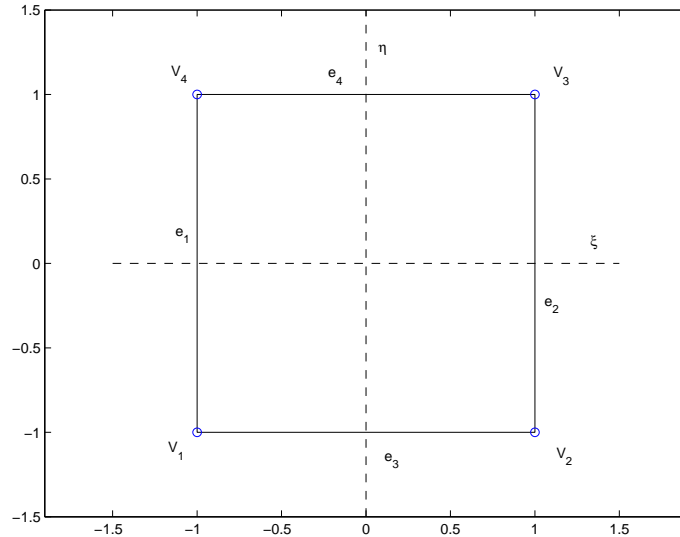


Figura 1.1: El cuadrado de referencia con la numeración de vértices y lados seguida para la construcción de las bases sin nodos.

Consideramos cuatro funciones especiales,

$$\begin{aligned}\lambda_1(\xi, \eta) &:= \ell_1(\xi) = (\xi + 1)/2, \\ \lambda_2(\xi, \eta) &:= \ell_0(\xi) = (1 - \xi)/2, \\ \lambda_3(\xi, \eta) &:= \ell_1(\eta) = (\eta + 1)/2, \\ \lambda_4(\xi, \eta) &:= \ell_0(\eta) = (1 - \eta)/2,\end{aligned}$$

de forma que $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ es base de \mathbb{Q}_1 . Notemos que

$$\xi = \lambda_1(\xi, \eta) - \lambda_2(\xi, \eta), \quad \eta = \lambda_3(\xi, \eta) - \lambda_4(\xi, \eta).$$

Por fin, damos una construcción de la base de \mathbb{Q}_p , organizada en tres clases de funciones:

- (1) Funciones de vértice: son cuatro funciones de \mathbb{Q}_1 , dadas por las expresiones

$$\begin{aligned}\ell_0(\xi) \ell_0(\eta) &= \lambda_2 \lambda_4 \\ \ell_1(\xi) \ell_0(\eta) &= \lambda_1 \lambda_4 \\ \ell_1(\xi) \ell_1(\eta) &= \lambda_1 \lambda_3 \\ \ell_0(\xi) \ell_1(\eta) &= \lambda_2 \lambda_3.\end{aligned}$$

Cada una de estas cuatro funciones se anula sobre dos lados consecutivos del cuadrado $[-1, 1] \times [-1, 1]$ y vale uno en el vértice opuesto.

- (2) Funciones de lado: a cada lado le asociamos $p - 1$ funciones. La lista completa es

$$\begin{aligned}\ell_0(\xi) \ell_k(\eta) & \quad 2 \leq k \leq p \\ \ell_1(\xi) \ell_k(\eta) & \quad 2 \leq k \leq p \\ \ell_k(\xi) \ell_0(\eta) & \quad 2 \leq k \leq p \\ \ell_k(\xi) \ell_1(\eta) & \quad 2 \leq k \leq p\end{aligned}$$

Cada función se anula sobre tres lados completos del cuadrado $[-1, 1] \times [-1, 1]$. Restringidas a los lados, simplemente tenemos las funciones de Lobatto.

- (3) Funciones internas: son las funciones

$$\ell_j(\xi) \ell_k(\eta), \quad 2 \leq j, k \leq p.$$

Todas estas funciones son internas a $[-1, 1] \times [-1, 1]$, ya que se anulan sobre los cuatro lados.

En total tenemos

$$4 + 4(p - 1) + (p - 1)(p - 1) = (p + 1)^2$$

polinomios. Entre todos forman una base de \mathbb{Q}_p . Cuando $p = 2$ hay 4 funciones de vértice, 4 de lado y 1 interior. Cuando $p = 3$ hay 4 funciones de vértice (estas son siempre las mismas), 8 de lado y 4 interiores.

El número de funciones internas aumenta proporcionalmente al grado, mientras que el número de funciones asociadas a lados aumenta siempre en número constante.

Jerarquía. Si disponemos de la base de \mathbb{Q}_p , para obtener la de \mathbb{Q}_{p+1} basta añadir:

- una función por lado (la correspondiente en la lista a $k = p + 1$)
- las internas correspondientes, tomando uno de los índices igual a $p + 1$ (esto es, añadimos, $2p + 1$ funciones).

Esto resulta ventajoso ya que, a diferencia de las bases nodales habituales, basadas en una distribución uniforme de nodos sobre el cuadrado, estas bases sí forman una jerarquía.

Paso a un paralelogramo cualquiera. Si K es un paralelogramo no degenerado y

$$F_K : [-1, 1] \times [-1, 1] \rightarrow K$$

es una transformación afín inversible, el espacio $\mathbb{Q}_p(K) := \{p : K \rightarrow \mathbb{R} \mid p \circ F_K \in \mathbb{Q}_p\}$ contiene a todos los polinomios de grado menor o igual que p más unos cuantos polinomios de grado menor o igual que p^2 , de forma que

$$\mathbb{P}_p \subset \mathbb{Q}_p(K) \subset \mathbb{P}_{p^2}$$

y

$$\dim \mathbb{Q}_p(K) = \dim \mathbb{Q}_p = (p + 1)^2.$$

Recordemos que este espacio es independiente de qué transformación concreta se haya efectuado para pasar del elemento de referencia a K . Por composición con F_K^{-1} , podemos construir una base de $\mathbb{Q}_p(K)$ con la misma estructura geométrica:

- Cuatro funciones de vértices, todas ellas bilineales, de forma que restringidas a los lados en los que se apoya el vértice son lineales.
- $p - 1$ funciones asociadas a cada lado, que se anulan sobre los tres lados restantes.
- $(p - 1)^2$ funciones internas.

Es importante notar que el valor de un polinomio sobre uno de los lados de K , si hemos expresado el polinomio en la base recién construida, depende sólo de los coeficientes de los vértices correspondientes y de las funciones de lado.

Bases de espacios de elementos finitos. Si \mathcal{T}_h es una triangulación de un polígono Ω formada por paralelogramos en las condiciones habituales, entonces se puede dar una base de

$$X_h^p := \left\{ u_h \in \mathcal{C}(\Omega) \mid u_h|_K \in \mathbb{Q}_p(K), \forall K \in \mathcal{T}_h \right\},$$

uniendo las funciones de base nodales anteriores. Esta base se organiza como la que hemos indicado elemento a elemento.

- Hay una función de base por cada vértice de la triangulación. Su soporte es el conjunto de elementos que tienen al vértice como tal.

- Hay $p - 1$ funciones de base por cada lado. El soporte de cada una se limita a dos elementos (uno si el lado está en la frontera).
- Hay $(p - 1)^2$ grados de libertad internos a cada elemento.

Este tipo de bases queda muy bien organizada para cálculos asociados a ensamblado para resolver problemas de segundo orden:

- Los grados internos se pueden tratar por condensación estática, lo cual reduce el tamaño del sistema que hay que resolver.
- Si se quieren quitar los tests asociados a un lado Dirichlet, basta eliminar las funciones asociadas a los dos extremos del lado y al lado en sí.
- La información que debe dar el mallador es muy simple: (a) numeración de vértices; (b) numeración de lados respecto de los vértices; (c) numeración de elementos. Si se emplea condensación estática, conforme se hace el ensamblado elemento a elemento, basta tener una numeración adecuada de los lados y los vértices para realizar todo el ensamblado de forma eficaz.

4. Construcción de bases en triángulos

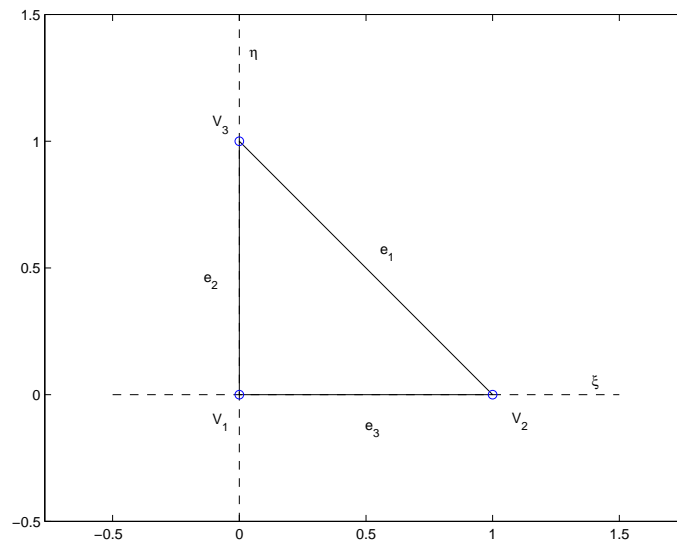


Figura 1.2: El triángulo de referencia con la numeración de vértices y lados seguida para la construcción de las bases sin nodos.

Las funciones

$$\begin{aligned} \lambda_1(\xi, \eta) &:= 1 - \xi - \eta, \\ \lambda_2(\xi, \eta) &:= \xi, \\ \lambda_3(\xi, \eta) &:= \eta, \end{aligned}$$

forman la base nodal típica del elemento \mathbb{P}_1 sobre el triángulo de referencia \widehat{K} . Cada una de ellas vale la unidad en un vértice y se anula sobre los otros dos. Además, son las coordenadas baricéntricas de los puntos del triángulo, es decir, si tenemos un punto de coordenadas (ξ, η) , entonces

$$\begin{aligned} \lambda_j(\xi, \eta) \geq 0 \\ j = 1, 2, 3 \end{aligned} \iff (\xi, \eta) \in \widehat{K}.$$

Notemos que

$$\lambda_1 + \lambda_2 + \lambda_3 = 1.$$

Una de las ventajas de emplear coordenadas baricéntricas para escribir expresiones polinómicas en un triángulo es que se preservan bajo transformaciones afines. Si K es un triángulo cualquiera, de vértices

$$(x_1, y_1), \quad (x_2, y_2), \quad (x_3, y_3),$$

entonces

$$F_K(\xi, \eta) = (1 - \xi - \eta) \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \xi \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + \eta \begin{bmatrix} x_3 \\ y_3 \end{bmatrix}$$

es la transformación afín que lleva \widehat{K} a K , respetando la numeración de los vértices. Las funciones

$$\lambda_j \circ F_K^{-1}$$

son las coordenadas baricéntricas en el triángulo K : suman la unidad, valen uno en un vértice y cero en los otros dos; indican si un punto está dentro del triángulo (todas son positivas) o no.

Consideremos además las funciones

$$\lambda_2(\xi, \eta) - \lambda_1(\xi, \eta) = 2\xi + \eta - 1,$$

$$\lambda_3(\xi, \eta) - \lambda_1(\xi, \eta) = 2\eta + \xi - 1,$$

$$\lambda_3(\xi, \eta) - \lambda_2(\xi, \eta) = \eta - \xi.$$

La siguiente tabla de valores es inmediata:

	(0, 0)	(1, 0)	(0, 1)
$\lambda_2 - \lambda_1$	-1	1	0
$\lambda_3 - \lambda_1$	-1	0	1
$\lambda_3 - \lambda_2$	0	-1	1

Vamos seguidamente a construir una base de \mathbb{P}_p , organizándola de forma similar a la de la sección precedente.

(1) Funciones de vértice: son las tres coordenadas baricéntricas

$$\lambda_1, \quad \lambda_2, \quad \lambda_3.$$

Son funciones lineales que valen uno sobre un vértice y cero en los otros dos.

(2) Funciones de lado: a cada lado le asociamos $p - 1$ funciones. La lista completa es

$$\begin{aligned} \lambda_2 \lambda_3 \left(\phi_{k-2} \circ (\lambda_3 - \lambda_2) \right), & \quad 2 \leq k \leq p, \\ \lambda_1 \lambda_3 \left(\phi_{k-2} \circ (\lambda_3 - \lambda_1) \right), & \quad 2 \leq k \leq p, \\ \lambda_1 \lambda_2 \left(\phi_{k-2} \circ (\lambda_2 - \lambda_1) \right), & \quad 2 \leq k \leq p. \end{aligned}$$

(Recordemos que $\phi_{k-2} = \ell_k / (\ell_0 \ell_1)$ han sido definidos en la Sección 1). Cada función se anula sobre dos lados completos del triángulo de referencia.

(3) Funciones internas: son las funciones

$$\lambda_1 \lambda_2 \lambda_3 \left(\phi_{j-1} \circ (\lambda_2 - \lambda_1) \right) \left(\phi_{k-1} \circ (\lambda_3 - \lambda_1) \right)$$

con índices

$$1 \leq k, j, \quad k + j \leq p - 1.$$

Todas ellas se anulan sobre los tres lados. Sólo aparecen a partir de $p = 3$.

Notemos que, por ejemplo

$$\lambda_2 \lambda_3 \left(\phi_{k-2} \circ (\lambda_3 - \lambda_2) \right) = \xi \eta \phi_{k-2}(\eta - \xi),$$

luego los polinomios ϕ_{k-2} se evalúan en la variable $\eta - \xi$, que toma valores de -1 a 1 . Las expresiones en función de las coordenadas baricéntricas son más útiles, puesto que se preservan al pasar a otros triángulos. Notemos también que

$$\ell_k(2\eta - 1) = (1 - \eta) \eta \phi_{k-2}(2\eta - 1).$$

Esto quiere decir que, cuando restringimos cualquiera de las funciones de lado al lado correspondiente, lo que obtenemos es la función de Lobatto resultante. Así, las funciones de vértice determinan el valor del polinomio en los vértices y el resto del valor sobre un lado es determinado por las funciones de lado correspondientes.

Observación. La elección de las diferencias $\lambda_2 - \lambda_1$ y $\lambda_3 - \lambda_1$ en lugar de otros dos pares (y con los signos en otros órdenes) es totalmente arbitraria y produce bases esencialmente iguales. \square

Las mismas ideas que hemos expuesto sobre cómo construir bases de espacios de elementos finitos sirven para esta situación. Aparece la base de elementos lineales asociados a los vértices, que es la que mayores soportes tiene. Las funciones de lado comunican dos triángulos adyacentes y las internas están aisladas.

5. Elementos de transición

Las bases libres de nodos se pueden emplear fácilmente para realizar transiciones entre elementos de distinto orden dentro del mismo mallado. Este tipo de estrategias es típica de los llamados **métodos** $h-p$ donde se juega simultáneamente con el tamaño de la malla y los grados polinómicos: la idea es tomar grados más altos en las zonas más regulares de la solución, manteniendo mallas gruesas, y mallados más finos donde la solución presente singularidades.

Vamos a dar una idea somera de cómo cambiar el grado para elementos sobre paralelogramos. Además del grado básico de los polinomios, $p \geq 2$, ahora a cada lado del rectángulo le asociamos un grado polinómico propio:

$$e_j \quad \longmapsto \quad p_j \leq p, \quad j = 1, 2, 3, 4.$$

Se consideran así las siguientes funciones (la construcción es una variante muy simple de la dada en la Sección 3).

- (1) Funciones de vértice (son las mismas de antes):

$$\begin{aligned} \ell_0(\xi) \ell_0(\eta) &= \lambda_2 \lambda_4 \\ \ell_1(\xi) \ell_0(\eta) &= \lambda_1 \lambda_4 \\ \ell_1(\xi) \ell_1(\eta) &= \lambda_1 \lambda_3 \\ \ell_0(\xi) \ell_1(\eta) &= \lambda_2 \lambda_3. \end{aligned}$$

- (2) Funciones de lado: a cada lado le asociamos $p_j - 1$ funciones. La lista completa es

$$\begin{aligned} \ell_0(\xi) \ell_k(\eta) & \quad 2 \leq k \leq p_1 \\ \ell_1(\xi) \ell_k(\eta) & \quad 2 \leq k \leq p_2 \\ \ell_k(\xi) \ell_0(\eta) & \quad 2 \leq k \leq p_3 \\ \ell_k(\xi) \ell_1(\eta) & \quad 2 \leq k \leq p_4. \end{aligned}$$

Si algún $p_j = 1$, no se añade ninguna función a ese lado.

- (3) Funciones internas (son las mismas de antes):

$$\ell_j(\xi) \ell_k(\eta), \quad 2 \leq j, k \leq p.$$

En total tenemos

$$4 + \sum_{j=1}^4 (p_j - 1) + (p - 1)^2 = \sum_{j=1}^4 p_j + (p - 1)^2 \leq 4p + (p - 1)^2 = (p + 1)^2 = \dim \mathbb{Q}_p$$

funciones. El espacio obtenido está contenido en \mathbb{Q}_p . Si $q := \min\{p_1, p_2, p_3, p_4\}$, entonces el espacio contiene a \mathbb{Q}_q . De hecho, se puede escribir que el conjunto de las combinaciones lineales de las funciones dadas antes es el espacio

$$\left\{ p \in \mathbb{Q}_p \mid p|_{e_j} \in \mathbb{P}_{p_j}(e_j), \quad j = 1, \dots, 4 \right\}$$

donde $\mathbb{P}_k(e_j)$ son los polinomios de grado menor o igual que k sobre el lado e_j (esencialmente polinomios de una variable).

Supongamos que tenemos un mallado como el dado en la figura, en el que hemos impuesto el espacio de discretización sobre unos cuantos rectángulos. Para conseguir una transición de unos elementos de un grado a otros de grado superior, la forma en que hemos dado la base es idónea.

Por ejemplo, para el elemento central de la fila inferior, debemos tomar (numeramos lados como en el elemento de referencia) $p_1 = 2$ y $p_2 = 4$. Eso obliga a tomar $p = 4$ como mínimo (y no tiene sentido tomarlo mayor). El lado inferior es libre (podemos tomar $p_3 = 4$ o menor) y el superior no estará decidido hasta que veamos qué viene de ese elemento.

El elemento central de la columna izquierda valen los mismos argumentos. El elemento restante debe tener grado 4 en los lados que tocan al mallado \mathbb{Q}_4 , luego está obligado a tener grado interior igual a cuatro. Los lados restantes son libres y deben ‘acordarse’ con los rectángulos adyacentes.

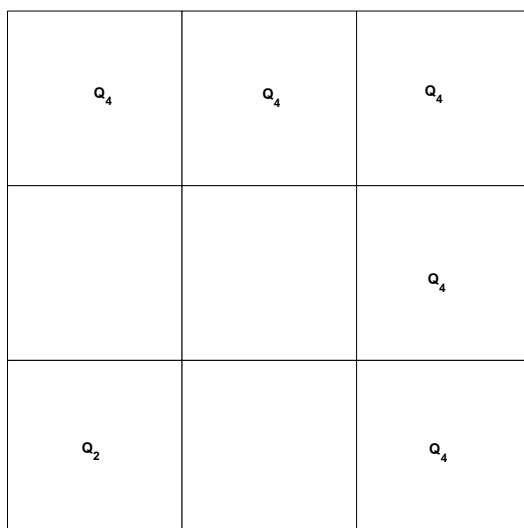


Figura 1.3: Un mallado de paralelogramos con grados impuestos en algunos de ellos.

Una forma más sistemática de hacer este trabajo consiste en asignar directamente grado a todos los elementos y luego modificarlos a la baja sobre los lados adyacentes (el grado debe ser el menor de los de los elementos correspondientes). En la zona de conexión entre elementos de distinto grado habrá unos cuantos que no tendrán el grado completo. Puesto que el ensamblado se hace elemento a elemento, el hecho de que la descripción del espacio sea muy complicada no dificulta en absoluto la implementación que se hace con información fácilmente recopilable.

6. Ejercicios propuestos

1. ¿Cuántas funciones de cada tipo se tienen para la base propuesta de \mathbb{Q}_5 ?

2. En el cubo de referencia $[-1, 1] \times [-1, 1] \times [-1, 1]$, da una construcción general de una base de \mathbb{Q}_p siguiendo las ideas anteriores, pero empleando ahora: funciones de vértice, funciones de arista, funciones de cara y internas. ¿Cuántas funciones de cada tipo se obtienen? **Indicación:** mirando la base de \mathbb{Q}_p en dimensión dos, es fácil notar que hay un patrón en los índices (que valgan 0 ó 1 ó más) para separar los tres tipos de funciones. Se puede repetir esta idea en dimensión tres
3. Haz un recuento de funciones de la base dada de \mathbb{P}_p en la Sección 4 y verifica que su número es igual a

$$\dim \mathbb{P}_p = \frac{(p+1)(p+2)}{2} = \binom{p+2}{2} = 1 + 2 + \dots + p + (p+1)$$

Si se tiene una base de \mathbb{P}_p , ¿qué funciones se añaden para obtener una base de \mathbb{P}_{p+1} ?

4. Completa la asignación de grados del ejemplo de la Sección 5. ¿Qué dimensión tiene el espacio local en cada elemento? ¿Cuántos de los grados de libertad son internos (corresponden a funciones internas) en cada elemento? ¿Cuál es el número total de grados de libertad?
5. Da una construcción de funciones de base para un triángulo escogiendo el grado en los lados.

Lección 2

Elementos de tipo Hermite

Los elementos finitos de tipo Hermite son aquéllos que emplean derivadas como grados de libertad. Su empleo, que resulta simple y bastante natural para problemas unidimensionales, puede llegar a venir forzado por la exigencias del problema. En este capítulo veremos un problema unidimensional (la viga de Euler–Bernoulli) que requiere emplear elementos de Hermite. En el Capítulo 3 veremos su equivalente bidimensional (la placa de Kirchhoff). Independientemente de ello, los elementos de Hermite permiten simplificar la localización de los grados de libertad para elementos de orden relativamente alto. El coste es la dificultad de implementación, especialmente en lo relativo a condiciones de contorno.

1. Elemento de Hermite \mathbb{P}_3 en dimensión uno

Vamos a elegir de nuevo el intervalo $[-1, 1]$ como elemento de referencia. En él consideramos el espacio $\mathbb{P}_3 = \mathbb{P}_3(\xi) := \{a_0 + a_1 \xi + a_2 \xi^2 + a_3 \xi^3 \mid a_i \in \mathbb{R}\}$. Es fácil comprobar que los valores

$$p(-1), \quad p'(-1), \quad p(1), \quad p'(1)$$

determinan unívocamente a $p \in \mathbb{P}_3$. Entonces tenemos cuatro funciones de base, dos de tipo lagrangiano y dos de tipo hermitiano: \widehat{N}_α ($\alpha = 1, \dots, 4$) de forma que

$$\begin{aligned} \widehat{N}_1(-1) &= 1 & \widehat{N}_1(1) &= 0 & \widehat{N}'_1(-1) &= 0 & \widehat{N}'_1(1) &= 0, \\ \widehat{N}_2(-1) &= 0 & \widehat{N}_2(1) &= 1 & \widehat{N}'_2(-1) &= 0 & \widehat{N}'_2(1) &= 0, \\ \widehat{N}_3(-1) &= 0 & \widehat{N}_3(1) &= 0 & \widehat{N}'_3(-1) &= 1 & \widehat{N}'_3(1) &= 0, \\ \widehat{N}_4(-1) &= 0 & \widehat{N}_4(1) &= 0 & \widehat{N}'_4(-1) &= 0 & \widehat{N}'_4(1) &= 1 \end{aligned}$$

Realizamos una partición de $[0, L]$

$$0 = x_0 < x_1 < \dots < x_{n-1} < x_n = L.$$

Consideramos la transformación

$$F_j : [-1, 1] \rightarrow I_j := [x_{j-1}, x_j]$$

dada por

$$x = F_j(\xi) := \frac{x_j + x_{j-1}}{2} + \xi \frac{x_j - x_{j-1}}{2} = x_{j-1} \frac{1 - \xi}{2} + x_j \frac{\xi + 1}{2}.$$

Ahora consideramos cuatro funciones asociadas a I_j . Las dos funciones lagrangianas son elementales

$$N_{1,j} := \widehat{N}_1 \circ F_j^{-1}, \quad N_{2,j} := \widehat{N}_2 \circ F_j^{-1}.$$

Las funciones hermitianas tienen que ser reescaladas: si $h_j = x_j - x_{j-1}$ es la longitud de I_j , definimos

$$N_{3,j} := (h_j/2)\widehat{N}_3 \circ F_j^{-1}, \quad N_{4,j} := (h_j/2)\widehat{N}_4 \circ F_j^{-1}.$$

Este escalado se hace para que se cumplan las siguientes relaciones

$$\begin{aligned} N_{1,j}(x_{j-1}) &= 1 & N_{1,j}(x_j) &= 0 & N'_{1,j}(x_{j-1}) &= 0 & N'_{1,j}(x_j) &= 0, \\ N_{2,j}(x_{j-1}) &= 0 & N_{2,j}(x_j) &= 1 & N'_{2,j}(x_{j-1}) &= 0 & N'_{2,j}(x_j) &= 0 \\ N_{3,j}(x_{j-1}) &= 0 & N_{3,j}(x_j) &= 0 & N'_{3,j}(x_{j-1}) &= 1 & N'_{3,j}(x_j) &= 0 \\ N_{4,j}(x_{j-1}) &= 0 & N_{4,j}(x_j) &= 0 & N'_{4,j}(x_{j-1}) &= 0 & N'_{4,j}(x_j) &= 1. \end{aligned}$$

Del espacio

$$X_h := \left\{ u_h \in \mathcal{C}^1[0, L] \mid u_h|_{I_j} \in \mathbb{P}_3, \quad \forall j \right\}$$

se puede construir una base de forma muy simple. Al nodo x_j le asociamos dos funciones: una de tipo lagrangiano, definida pegando $N_{2,j}$ con $N_{1,j+1}$ y otra de tipo hermitiano, definida pegando $N_{4,j}$ con $N_{3,j+1}$ (si $j = 0$ ó $j = n$, sólo hay que dar una definición en un intervalo). Así, podemos numerar la base como

$$\psi_0, \dots, \psi_n, \quad \varphi_1, \dots, \varphi_n$$

Cada función lagrangiana asociada a un nodo interno cumple

$$\psi_i(x_j) = \delta_{ij}, \quad \psi'_i(x_j) = 0, \quad \forall i, j$$

mientras que las hermitianas cumplen

$$\varphi_i(x_j) = 0, \quad \varphi'_i(x_j) = \delta_{ij}, \quad \forall i, j.$$

El espacio X_h está por tanto determinado por $2(n+1)$ grados de libertad (dos por vértice), frente al espacio de elementos \mathbb{P}_3 continuos, que tiene más grados de libertad, en concreto, $3n + 1$ (2 por intervalo más uno por extremo). A priori, la calidad de ambos espacios es similar (darán el mismo tipo de convergencia), aunque el precio es doble:

- Hay que tener en cuenta que la transformación necesaria para obtener las funciones locales en el elemento I_j es distinta según el tipo de grado de libertad.
- Se pierden los grados de libertad internos (dos por elemento) que tenía el elemento de Lagrange de tipo 3.

Es ilustrativo comparar el empleo de este espacio X_h para calcular las matrices de masa y rigidez asociadas a un problema de orden dos unidimensional

$$\int_0^L u' v' + \int_0^L u v,$$

frente al espacio de elementos \mathbb{P}_3 que solo son continuos. Contando elementos de la base de X_h es fácil ver que este espacio tiene dimensión $2(n+1) = 2n+2$ ya que hay dos funciones por nodo y los nodos son únicamente los extremos de los intervalos. En los elementos \mathbb{P}_3 de tipo Lagrange hay una función por nodo, pero ahora hay dos nodos dentro de cada elemento, luego tenemos

$$(n+1) + 2n = 3n+1 > 2n+2.$$

(Sólo con $n=1$ ambos números coinciden, ya que con un único elemento no hay diferencia entre los espacios). Obviamente los grados de libertad se han perdido al exigirse condiciones de regularidad más estrictas. Sin embargo esta ‘información’ no está perdida en las matrices de elementos finitos. Mirando las figuras se ve que cada bloque 4×4 (correspondiente al ensamblado, que sólo ve el espacio localmente) está más conectado con sus vecinos en el caso de elementos de Hermite. La matriz es más pequeña pero también más ‘rígida’. La razón es simple: las funciones de base de los extremos ven a sus vecinos y no hay funciones internas, mientras que en el caso de elementos de Lagrange tenemos nodos internos y solo los extremos (que no tienen más que una función de base asociada) son compartidos por dos elementos.

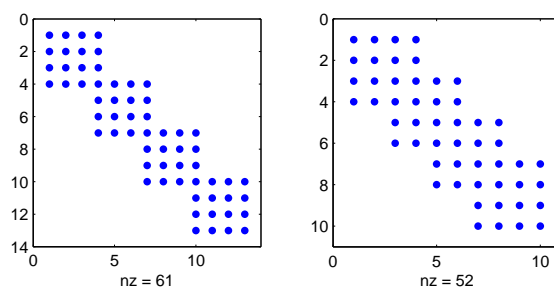


Figura 2.1: Forma de las matrices de elementos finitos para \mathbb{P}_3 Lagrange (izquierda) y Hermite (derecha) con cuatro elementos

2. Viga de Euler–Bernoulli

En esta sección vamos a ver un ejemplo de elemento donde la utilización de elemento de clase \mathcal{C}^1 no es opcional. En su versión más simplificada, la ecuación de la viga de Euler–Bernoulli se escribe como

$$u^{(4)} = f, \quad \text{en } [0, L],$$

con cuatro condiciones de contorno, dos en cada extremo. Tomaremos por ejemplo, las siguientes:

$$u(0) = u'(0) = 0, \quad u''(L) = u'''(L) = 0,$$

es decir, la viga está encastrada en el extremo izquierdo y libre en el derecho.

La formulación variacional de este problema es como sigue. El espacio básico es

$$H^2(0, L) := \left\{ u \in L^2(0, L) \mid u', u'' \in L^2(0, L) \right\}.$$

Entonces, se considera su subespacio,

$$V := \left\{ u \in H^2(0, L) \mid u(0) = u'(0) = 0 \right\}$$

y la formulación variacional

$$\begin{cases} u \in V, \\ \int_0^L u'' v'' = \int_0^L f v, \quad \forall v \in V. \end{cases}$$

El siguiente resultado es fundamental para entender por qué debemos emplear elementos de Hermite en este problema.

Proposición *Sea $u_h : [0, L] \rightarrow \mathbb{R}$ una función polinómica a trozos. Entonces*

$$u_h \in H^2(0, L) \quad \iff \quad u_h \in \mathcal{C}^1([0, L]).$$

El resultado se deduce trivialmente de uno más simple, que es el que obliga a emplear elementos finitos continuos en discretizaciones de problemas de contorno de segundo orden: *sea $u_h : [0, L] \rightarrow \mathbb{R}$ una función polinómica a trozos. Entonces*

$$u_h \in H^1(0, L) \quad \iff \quad u_h \in \mathcal{C}([0, L]).$$

En este caso, podemos tomar

$$X_h := \left\{ u_h \in \mathcal{C}^1[0, L] \mid u_h|_{I_j} \in \mathbb{P}_3, \quad \forall j \right\}$$

como discretización de X_h . Eliminando las funciones de base ψ_0 y φ_0 (la lagrangiana y hermitiana del extremo izquierdo), conseguimos una base de un subespacio de V . Si denotamos a este espacio X_h^0 , entonces

$$\|u - u_h\|_{2,(0,L)} \leq C \inf_{v_h \in X_h^0} \|u - v_h\|_{2,(0,L)} \leq C' h^2,$$

siempre que u sea suficientemente regular.

3. Elemento de Hermite \mathbb{P}_3 en dimensión dos

Los elementos de tipo \mathbb{P}_3 (al igual que los \mathbb{P}_2) requieren una numeración conjunta de lados y vértices que se puede evitar acumulando más grados de libertad en los vértices. Este es el origen de los elementos de tipo Hermite, que involucran valores de derivadas parciales además de valores nodales (lagrangianos) entre las condiciones para determinar las funciones polinómicas en cada triángulo (y por consiguiente para determinar las funciones de base nodales que determinan la base en que se escribe el sistema).

Consideremos los tres vértices de un triángulo P_1^K , P_2^K y P_3^K y su baricentro P_0^K .

Proposición Sea $p \in \mathbb{P}_3$. Entonces p viene unívocamente determinado por

$$p(P_i^K), \quad i = 0, \dots, 3, \quad \nabla p(P_i^K), \quad i = 1, 2, 3$$

Además el valor en el lado que une los vértices P_i^K y P_j^K viene unívocamente determinado por

$$p(P_i^K), \quad p(P_j^K), \quad \nabla p(P_i^K) \cdot v_{ij}^K, \quad \nabla p(P_j^K) \cdot v_{ij}^K,$$

siendo v_{ij}^K el vector que une los citados vértices.

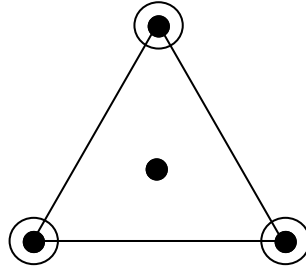


Figura 2.2: Representación gráfica de un triángulo de Hermite (3)

Este resultado permite construir el siguiente espacio de elementos finitos sobre una triangulación de Ω

$$V_h = \left\{ u_h \in \mathcal{C}(\Omega) \left| \begin{array}{l} u_h|_K \in \mathbb{P}_3, \quad \forall K \\ u_h \text{ admite plano tangente en } P_i, \forall i \end{array} \right. \right\}$$

La dimensión de este espacio es

$$3\#\{\text{vértices}\} + \#\{\text{triángulos}\}$$

frente a la construida con los elementos \mathbb{P}_3 lagrangianos, cuya dimensión es

$$\#\{\text{vértices}\} + 2\#\{\text{lados}\} + \#\{\text{triángulos}\}$$

Funciones de base nodales. Podemos considerar en cada triángulo 3 funciones de base por cada vértice y una más por el baricentro. El nodo interior puede ser reducido al ensamblar por condensación estática. Las funciones asociadas a un vértice ocupan los triángulos que rodean al citado vértice. Consideramos así las funciones:

- N_α^K , $\alpha = 0, 1, 2, 3$ de forma que

$$N_\alpha^K(P_\beta^K) = \delta_{\alpha\beta}, \quad \beta = 0, 1, 2, 3,$$

$$\nabla N_\alpha^K(P_\beta^K) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \beta = 1, 2, 3;$$

- $N_{\alpha i}^K$, $\alpha = 1, 2, 3; i = 1, 2$, de forma que

$$N_{\alpha i}^K(P_\beta^K) = 0, \quad \beta = 0, 1, 2, 3,$$

$$\nabla N_{\alpha i}^K(P_\beta^K) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \{1, 2, 3\} \ni \beta \neq \alpha,$$

$$\nabla N_{\alpha 1}^K(P_\alpha^K) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \nabla N_{\alpha 2}^K(P_\alpha^K) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

De la misma forma podemos considerar las funciones de base en el elemento de referencia

- \widehat{N}_α , $\alpha = 0, 1, 2, 3$,
- $\widehat{N}_{\alpha i}$, $\alpha = 1, 2, 3; i = 1, 2$.

Paso al elemento de referencia. Al incluir grados de libertad hermitianos, la relación con las matrices en el elemento de referencia ya no es trivial, puesto que los grados de libertad relacionados con las derivadas ‘se mezclan’. Sea como de costumbre $F_K : \widehat{K} \rightarrow K$, la transformación afín que traslada el triángulo de referencia al triángulo K , de forma que $F_K(\widehat{P}_i) = P_i^K$ (nótese que el baricentro cumple también esta propiedad). Sea

$$B_K = \begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{bmatrix}$$

la matriz jacobiana de la transformación.

Proposición Para $\alpha = 0, 1, 2, 3$

$$\widehat{N}_\alpha \circ F_K^{-1} = N_\alpha^K$$

Para $\alpha = 1, 2, 3$

$$\begin{bmatrix} N_{\alpha,1}^K \\ N_{\alpha,2}^K \end{bmatrix} = B_K \begin{bmatrix} \widehat{N}_{\alpha,1} \circ F_K^{-1} \\ \widehat{N}_{\alpha,2} \circ F_K^{-1} \end{bmatrix}.$$

Demostración. Es lo mismo que probar que para $\alpha = 0, 1, 2, 3$

$$\widehat{N}_\alpha = N_\alpha^K \circ F_K$$

Para $\alpha = 1, 2, 3$

$$\begin{bmatrix} \widehat{N}_{\alpha,1} & \widehat{N}_{\alpha,2} \end{bmatrix} = \begin{bmatrix} N_{\alpha,1}^K \circ F_K & N_{\alpha,2}^K \circ F_K \end{bmatrix} B_K^{-\top}.$$

Las primeras son fáciles de comprobar. Para las segundas, definimos $p_{\alpha,1}^K, p_{\alpha,2}^K$ por las relaciones

$$\begin{bmatrix} p_{\alpha,1}^K & p_{\alpha,2}^K \end{bmatrix} = \begin{bmatrix} \widehat{N}_{\alpha,1} \circ F_K^{-1} & \widehat{N}_{\alpha,2} \circ F_K^{-1} \end{bmatrix} B_K^\top,$$

que son obviamente elementos de \mathbb{P}_3 : cada uno de ellos es combinación lineal de $\widehat{N}_{\alpha,i} \circ F_K^{-1}$ ($i = 1, 2$) y, por tanto, se anula en los cuatro nodos. Tomando el gradiente a ambos lados de esta expresión se obtiene

$$B_K^\top \begin{bmatrix} \nabla p_{\alpha,1}^K \circ F_K & \nabla p_{\alpha,2}^K \circ F_K \end{bmatrix} = \begin{bmatrix} \widehat{\nabla} \widehat{N}_{\alpha,1} & \widehat{\nabla} \widehat{N}_{\alpha,2} \end{bmatrix} B_K^\top.$$

Si evaluamos esta expresión en \widehat{P}_β (con $\beta \neq \alpha$) obtenemos

$$B_K^\top \begin{bmatrix} \nabla p_{\alpha,1}^K(P_\beta^K) & \nabla p_{\alpha,2}^K(P_\beta^K) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

mientras que evaluando en \widehat{P}_α ,

$$B_K^\top \begin{bmatrix} \nabla p_{\alpha,1}^K(P_\alpha^K) & \nabla p_{\alpha,2}^K(P_\alpha^K) \end{bmatrix} = B_K^\top,$$

luego

$$\begin{bmatrix} \nabla p_{\alpha,1}^K(P_\alpha^K) & \nabla p_{\alpha,2}^K(P_\alpha^K) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Esto demuestra que $p_{\alpha,1}^K$ y $p_{\alpha,2}^K$ cumplen las seis condiciones de los gradientes, luego son las funciones de base correspondientes. \square

Ensamblado por paso al elemento de referencia La dificultad esencial de este tipo de elementos consiste en que las funciones de base sobre el elemento no son exactamente las trasladadas del elemento de referencia. Vamos a reenumerar los grados de libertad y las funciones de base del 0 al 9 de la siguiente forma:

- $\alpha = 0$ corresponde al nodo interior (baricéntrico);
- $\alpha = 1, 2, 3$ corresponden a los tres vértices;
- $\alpha = 4, 5$ corresponden al gradiente en el primer vértice;
- $\alpha = 6, 7$ corresponden al gradiente en el segundo vértice;
- $\alpha = 8, 9$ corresponden al gradiente en el tercer vértice.

En el fondo se están manejando tres bases:

- base en el elemento de referencia: \widehat{N}_α
- base en el elemento K : N_α^K

- base de referencia trasladada al elemento K : $\tilde{N}_\alpha^K = \hat{N}_\alpha \circ F_K^{-1}$

Las dos últimas bases están relacionadas por el siguiente cambio matricial

$$\begin{bmatrix} N_0^K \\ \vdots \\ N_9^K \end{bmatrix} = \begin{bmatrix} I_4 & & & \\ & B_K & & \\ & & B_K & \\ & & & B_K \end{bmatrix} \begin{bmatrix} \tilde{N}_0^K \\ \vdots \\ \tilde{N}_9^K \end{bmatrix}.$$

El proceso comienza por construir las cuatro matrices típicas (esta vez son 10×10) en el elemento de referencia

$$\hat{\mathbf{K}}_0, \quad \hat{\mathbf{K}}_{xx}, \quad \hat{\mathbf{K}}_{xy}, \quad \hat{\mathbf{K}}_{yy}.$$

Seguidamente se construye la matriz por cambio de variable al elemento K

$$\mathbf{A}_K = |\det B_K| \left(\hat{\mathbf{K}}_0 + c_{11}^K \hat{\mathbf{K}}_{xx} + c_{22}^K \hat{\mathbf{K}}_{yy} + c_{12}^K (\hat{\mathbf{K}}_{xy} + \hat{\mathbf{K}}_{xy}^\top) \right)$$

Esta matriz 10×10 es, en el fondo,

$$\int_K \nabla \tilde{N}_\beta^K \cdot \nabla \tilde{N}_\alpha^K + \int_K \tilde{N}_\beta^K \tilde{N}_\alpha^K,$$

luego corresponde a una reordenación de la matriz

$$a_K(\tilde{N}_\beta^K, \tilde{N}_\alpha^K).$$

Puesto que lo que queremos es $a_K(N_\beta^K, N_\alpha^K)$, basta con hacer un cambio de coordenadas en la forma bilineal

$$\mathbf{M}_K = \begin{bmatrix} I_4 & & & \\ & B_K & & \\ & & B_K & \\ & & & B_K \end{bmatrix} \mathbf{A}_K \begin{bmatrix} I_4 & & & \\ & B_K^\top & & \\ & & B_K^\top & \\ & & & B_K^\top \end{bmatrix}.$$

Esta operación es simplemente una manipulación de tres pares de filas y tres pares de columnas. Por último, puede realizarse una condensación estática del nodo interior (que aquí hemos puesto en la posición 0), equivalente a hacer ceros en la matriz \mathbf{M}_K , hasta obtener una matriz

$$\left[\begin{array}{c|ccc} m_{00}^K & m_{01}^K & \dots & m_{09}^K \\ \hline 0 & & & \\ \vdots & & \mathbf{M}_*^K & \\ 0 & & & \end{array} \right],$$

cuyo bloque \mathbf{M}_*^K (de tamaño 9×9) es el que realmente ensamblamos.

Dificultades A la hora de tratar las condiciones de contorno surgen algunas dificultades con este tipo de elementos finitos:

- Los nodos lagrangianos situados sobre la frontera Dirichlet son bastante simples de tratar, se asigna su valor y se ignoran los tests.
- Los nodos de tipo Hermite situados sobre la frontera Dirichlet son más complicados; para conseguir que una función del espacio V_h se anule sobre un lado Dirichlet hace falta que una combinación lineal de los grados de libertad hermitianos se anule (pero no han de anularse los dos). Esto dificulta la imposición de las condiciones de contorno y la eliminación de los grados de libertad que no hay que emplear como test.
- Sobre un lado Neumann, hay seis funciones de base que no se cancelan, lo cual vuelve a complicar el paso al elemento de referencia; el tipo de operaciones que hay que realizar es muy similar al que hemos hecho para las matrices de masa y rigidez.

Convergencia Realizando con un cierto cuidado el ajuste de las condiciones de contorno tipo Dirichlet para no perder convergencia, los elementos de Hermite \mathbb{P}_3 tienen el mismo orden de convergencia que los Lagrange, esto es, $\mathcal{O}(h^3)$.

4. Ejercicios propuestos

1. Calcula las funciones de base locales del elemento de referencia Hermite \mathbb{P}_3 en dimensión uno. (**Nota.** Una vez calculadas las del vértice izquierdo, las otras se obtienen fácilmente empleando el cambio de variables $\xi \leftrightarrow -\xi$. ¿Cómo?)
2. Haz un dibujo de las funciones de base ψ_i y φ_i del espacio X_h definido en la Sección 1.
3. Comprueba que en dimensión uno, todo polinomio de \mathbb{P}_4 está unívocamente determinado por los siguientes grados de libertad:

$$p(-1), \quad p'(-1), \quad p(0), \quad p(1), \quad p'(1).$$

Construyendo así un espacio de elementos finitos de clase \mathcal{C}^1 , ¿cuántos grados de libertad se obtienen?

4. Considera el elemento de Hermite (3) en el triángulo de referencia. Demuestra que si $p \in \mathbb{P}_3(\xi, \eta)$ se anula en los dos vértices del lado inferior y si la derivada respecto de ξ se anula en los dos vértices, entonces la función p se anula en el lado, pero que la derivada respecto de η no tiene por qué anularse. Deduce que el elemento de Hermite (3) sirve para definir un espacio de elementos finitos continuos pero no de clase \mathcal{C}^1 .
5. Explica cómo se pueden construir las funciones de base del elemento de Hermite (3) en el triángulo de referencia.

Lección 3

Elementos de clase \mathcal{C}^1

1. Elemento de Argyris

Mientras que la construcción de elementos de clase \mathcal{C}^1 es relativamente sencilla en dimensión uno, veremos en esta lección las dificultades nada menores que plantean estos elementos en dimensión dos. La utilización de este tipo de elementos es requisito de conformidad si se discretizan ecuaciones de placas en su formulación directa en la variable de desplazamiento. Comenzamos esta lección viendo el elemento \mathcal{C}^1 más conocido sobre triángulos.

Sea K un triángulo. Consideramos los siguientes grados de libertad locales:

- valor de la función en los vértices
- valor del gradiente en los vértices
- valor de las segundas derivadas en los vértices
- valor de la derivada normal en los puntos medios de los lados.

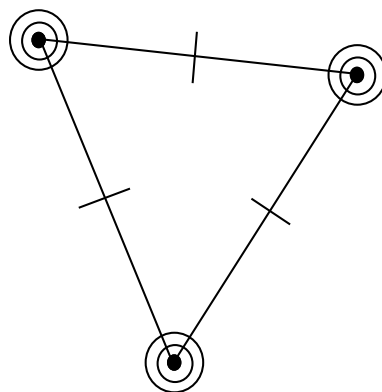


Figura 3.1: Representación esquemática del elemento de Argyris. Los dos círculos concéntricos hacen referencia a exigir todas las primeras y segundas derivadas

Así, en total se tienen

$$3 \times (1 + 2 + 3) + 3 = 21$$

grados de libertad. El espacio considerado es

$$\mathbb{P}_5 := \left\{ \sum_{i+j \leq 5} a_{ij} \xi^i \eta^j \mid a_{ij} \in \mathbb{R} \right\},$$

que tiene dimensión 21.

Proposición *Una función de \mathbb{P}_5 está determinada por los 21 grados de libertad locales. Además, la restricción a un lado de un elemento de \mathbb{P}_5 es un polinomio de grado cinco en una variable. Su valor está determinado por los grados de libertad del lado. Igualmente, las derivadas parciales (primeras) restringidas a los lados son polinomios de grado cuatro en una variable, determinados por los grados de libertad del lado.*

Demostración. La primera se deduce de la segunda. Vamos a ver simplemente esta última, que resulta reveladora de cómo funciona el elemento de Argyris en los lados. Sean $p \in \mathbb{P}_5$ y denotemos \vec{x}_1, \vec{x}_2 a dos vértices de K . Entonces es inmediato que

$$\phi(t) = p((1-t)\vec{x}_1 + t\vec{x}_2)$$

es un polinomio de grado menor o igual que cinco en la variable t . Los grados de libertad en el triángulo determinan

$$\begin{aligned} \phi(0), \quad \phi'(0), \quad \phi''(0), \\ \phi(1), \quad \phi'(1), \quad \phi''(1), \end{aligned}$$

y por tanto $\phi(t)$. Nótese que

$$\phi'(t) = \nabla p((1-t)\vec{x}_1 + t\vec{x}_2) \cdot (\vec{x}_2 - \vec{x}_1)$$

luego para ver la propiedad correspondiente al gradiente, basta estudiar la derivada normal. Si \vec{n} es el vector normal al lado que estamos estudiando, entonces

$$\psi(t) = \nabla p((1-t)\vec{x}_1 + t\vec{x}_2) \cdot \vec{n} = \partial_n p((1-t)\vec{x}_1 + t\vec{x}_2)$$

es un polinomio de grado menor o igual que cuatro. Los datos en los lados nos dan

$$\psi(0), \quad \psi'(0), \quad \psi(1/2), \quad \psi(1), \quad \psi'(1),$$

luego determinan $\psi(t)$. □

Del resultado anterior se deduce que: si todos los grados de libertad asociados a un lado del triángulo (los de los vértices y del punto medio) se anulan, entonces tanto el polinomio como su gradiente se anulan en ese lado.

Seguidamente consideramos una triangulación en las condiciones habituales. En cada vértice de la triangulación fijamos el valor de la función, su gradiente y sus segundas derivadas. En los puntos medios de los lados, la derivada normal. Si la función es un polinomio de grado menor o igual que cinco en cada triángulo y los grados de libertad comunes a varios triángulos coinciden, entonces, la función resultante cumple:

- $u_h \in \mathcal{C}^1(\Omega)$

- u_h es de clase \mathcal{C}^2 en los vértices.

Por tanto, el espacio discreto resultante de pegar elementos de Argyris sobre una triangulación es

$$V_h = \left\{ u_h \in \mathcal{C}^1(\Omega) \left| \begin{array}{l} u_h|_K \in \mathbb{P}_5, \forall K \\ u_h \in \mathcal{C}^2 \text{ en los vértices} \end{array} \right. \right\}$$

Dado $u_h \in V_h$ y L un lado exterior, entonces los 13 grados de libertad asociados al lado L se pueden reorganizar de la siguiente forma:

- valor en los dos extremos,
- valor de la derivada tangencial en los dos extremos,
- valor de la derivada normal en los extremos y en el punto medio,
- valor de las segundas derivadas tangenciales en los extremos,
- valor de las segundas derivadas cruzadas tangencial-normal en los extremos,
- valor de las segundas derivadas normales en los extremos.

Notemos las siguientes propiedades (que se deducen de la proposición antes demostrada):

- $u_h|_L = 0$ si y sólo si los 6 grados de libertad tangenciales (puros) se anulan.
- $\partial_n u_h|_L = 0$ si y sólo si los 3 grados de libertad normales y los dos tangenciales-normales se anulan.
- $u_h|_L = \partial_n u_h|_L = 0$ si y sólo si se anulan los 11 grados de libertad conjuntos.

Así, los grados de libertad locales $\partial_{nn}^2(P)$ no se ven involucrados en las condiciones de contorno.

Funciones de base nodales. Cada vértice se emplea seis veces y da lugar, por tanto, a seis funciones de base nodales cuyo soporte es el macrotriángulo que rodea al vértice. Cada punto medio se emplea una única vez. Así

$$\dim V_h = 6 \times \#\{\text{nodos}\} + \#\{\text{lados}\}$$

Para el problema de placa que veremos en la siguiente sección, hace falta saber trabajar con el espacio:

$$V_h^0 = V_h \cap \left\{ u \in H^2(\Omega) \left| \begin{array}{l} u = 0, \quad \text{en } F_1 \\ \partial_n u = 0, \quad \text{en } F_2 \end{array} \right. \right\},$$

siendo F_1 y F_2 dos partes no necesariamente disjuntas de la frontera. La construcción de una base nodal para el espacio V_h^0 no es simple. No se trata de eliminar funciones de base nodales, ya que éstas se refieren a derivadas primeras y segundas en coordenadas cartesianas globales (salvo las de los lados) y las restricciones que hay que imponer están en coordenadas tangenciales-normales referidas a cada lado.

2. Problema de placa de Kirchhoff

Consideremos un dominio poligonal del plano Ω y su frontera Γ . La ecuación de placa de Kirchhoff está asociada al operador bilaplaciano

$$\Delta^2 u = \Delta \Delta u = f, \quad \text{en } \Omega$$

La forma bilineal asociada al problema viene impuesta por la forma de la energía potencial del modelo de placa. Es la siguiente:

$$\begin{aligned} a(u, v) &= \nu \int_{\Omega} \Delta u \Delta v + (1 - \nu) \int_{\Omega} \mathbf{H} u : \mathbf{H} v \\ &= \int_{\Omega} \Delta u \Delta v + (1 - \nu) \int_{\Omega} (2\partial_{xy} u \partial_{xy} v - \partial_{xx} u \partial_{yy} v - \partial_{yy} u \partial_{xx} v) \end{aligned}$$

siendo $\nu \in (0, 1/2)$ el módulo de Poisson y $\mathbf{H} u$ la matriz hessiana de u (matriz de las segundas derivadas).

Denotemos ∂_n y ∂_σ a las derivadas normal exterior y tangencial en la frontera (la derivada tangencial se hace en sentido positivo, contrario a las agujas del reloj). Se darán dos condiciones de contorno en cada trozo de frontera:

$$\begin{aligned} &u \\ &\partial_n u \\ &\nu \Delta u + (1 - \nu) \partial_{nn}^2 u = \mathcal{B}_2 u \quad (\text{fuerza transversal}) \\ &\partial_n(\Delta u) + (1 - \nu) \partial_{n\sigma\sigma}^3 u = \mathcal{B}_1 u \quad (\text{momento de flexión}) \end{aligned}$$

Esta es la forma de las condiciones de contorno en dominios con frontera poligonal. Cuando el dominio tiene parte de la frontera curva, las condiciones tienen un aspecto menos simple. No se admiten todos los pares de condiciones de contorno. Los pares válidos son:

- Placa encastrada

$$u|_{\Gamma}, \quad \partial_n u|_{\Gamma}.$$

- Placa libre

$$\mathcal{B}_1 u|_{\Gamma}, \quad \mathcal{B}_2 u|_{\Gamma}.$$

- Placa simplemente apoyada

$$u|_{\Gamma}, \quad \mathcal{B}_2 u|_{\Gamma}.$$

- Placa rodante

$$\partial_n u|_{\Gamma}, \quad \mathcal{B}_1 u|_{\Gamma}.$$

Las condiciones de tipo 1 ó 3 en algún trozo del dominio dan lugar a solución única. Si no las hay se deben fijar puntos, ya que hay unicidad salvo desplazamientos verticales planos.

Podemos entonces dividir la frontera en $\Gamma_1, \Gamma_2, \Gamma_3$ y Γ_4 en función de qué par de condiciones de contorno se aplica en cada trozo. Así, incluso contando con que alguna de esas zonas sean vacías, tenemos como condiciones de contorno:

$$\begin{aligned} u &= g_1, & \text{en } F_1 &:= \Gamma_1 \cup \Gamma_3, \\ \partial_n u &= g_2, & \text{en } F_2 &:= \Gamma_1 \cup \Gamma_4, \\ \mathcal{B}_2 u &= g_3, & \text{en } F_3 &:= \Gamma_2 \cup \Gamma_3 = \Gamma \setminus F_2, \\ \mathcal{B}_1 u &= g_4, & \text{en } F_4 &:= \Gamma_2 \cup \Gamma_4 = \Gamma \setminus F_1. \end{aligned}$$

No se pueden imponer la fuerza de torsión y la derivada normal en la misma zona de frontera. Tampoco se pueden imponer simultáneamente el desplazamiento y la flexión. Las dos condiciones primeras son esenciales (se imponen fuera de la forma bilineal) y las dos últimas son naturales (aparecen incorporadas en la formulación variacional).

El espacio natural para el problema de placa viene impuesto por la forma bilineal antes escrita:

$$H^2(\Omega) = \left\{ u \in L^2(\Omega) \mid \partial^\alpha u \in L^2(\Omega), |\alpha| \leq 2 \right\}.$$

Además, se considera el espacio donde se imponen las condiciones esenciales de forma homogénea:

$$H_D^2(\Omega) := \left\{ u \in H^2(\Omega) \mid \begin{array}{l} u = 0, \quad \text{en } F_1 \\ \partial_n u = 0, \quad \text{en } F_2 \end{array} \right\}$$

Todas las condiciones de tipo restricción de u o de su derivada normal se imponen fuera de la formulación variacional.

La **fórmula de Rayleigh–Green** es la equivalente en el bilaplaciano a la fórmula de Green para el laplaciano:

$$a(u, v) = \int_{\Omega} \Delta^2 u v + \int_{\Gamma} (\mathcal{B}_2 u \partial_n v - \mathcal{B}_1 u v).$$

A partir de esta fórmula se deduce fácilmente la formulación variacional del problema:

$$\left[\begin{array}{l} u \in H^2(\Omega), \\ u = g_1, \quad \text{en } F_1, \\ \partial_n u = g_2, \quad \text{en } F_2, \\ a(u, v) = \int_{\Omega} f v - \int_{\Gamma \setminus F_1} g_4 v + \int_{\Gamma \setminus F_2} g_3 \partial_n v, \quad \forall v \in H_D^2(\Omega). \end{array} \right.$$

El punto de partida para la discretización es una triangulación \mathcal{T}_h en las condiciones habituales. Si se emplean funciones polinómicas en cada elemento, surge el siguiente resultado:

Proposición *Si u_h es una función polinómica a trozos, entonces*

$$u_h \in H^2(\Omega) \iff u_h \in \mathcal{C}^1(\Omega).$$

Por tanto, se puede emplear el elemento de Argyris para este tipo de discretizaciones. Notemos que en la forma bilineal se emplean segundas derivadas de las funciones de base, mientras que en el término independiente se emplean el valor de las funciones, más su restricción a la frontera y su derivada normal en la frontera. Si

$$V_h = \left\{ u_h \in \mathcal{C}^1(\Omega) \left| \begin{array}{l} u_h|_K \in \mathbb{P}_5, \forall K \\ u_h \in \mathcal{C}^2 \text{ en los vértices} \end{array} \right. \right\},$$

$$V_h^0 = V_h \cap \left\{ u \in H^2(\Omega) \left| \begin{array}{l} u = 0, \quad \text{en } F_1 \\ \partial_n u = 0, \quad \text{en } F_2 \end{array} \right. \right\},$$

la discretización del problema es

$$\left[\begin{array}{l} u_h \in V_h, \\ u_h = g_{1,h}, \quad \text{en } F_1, \\ \partial_n u_h = g_{2,h}, \quad \text{en } F_2, \\ a(u_h, v_h) = \int_{\Omega} f v_h - \int_{\Gamma \setminus F_1} g_4 v_h + \int_{\Gamma \setminus F_2} g_3 \partial_n v_h, \quad \forall v_h \in V_h^0. \end{array} \right.$$

Los datos Dirichlet, impuestos de forma directa como en este caso, deben ser aproximados primero. En particular $g_{1,h}$ debe ser un polinomio de grado cinco en cada lado de la triangulación que esté apoyado sobre F_1 y debe ser globalmente \mathcal{C}_1 . DE la misma forma $g_{2,h}$ debe ser un polinomio de grado menor o igual que cuatro sobre cada lado que esté en F_2 y globalmente continuo.

Puesto que el problema es elíptico se tiene el Lema de Céa aplicado a esta situación particular:

$$\|u - u_h\|_{2,\Omega} \leq C \inf \left\{ \|u - v_h\|_{2,\Omega} \left| v_h \in V_h, \begin{array}{l} v_h = g_{1,h}, \quad \text{en } F_1, \\ \partial_n v_h = g_{2,h}, \quad \text{en } F_2 \end{array} \right. \right\}.$$

Combinando el lema de Céa con la teoría de interpolación se obtiene el resultado

$$\|u - u_h\|_{2,\Omega} \leq C(u)h^4$$

en las hipótesis habituales de regularidad suficiente de la solución y de no aplanamiento de los triángulos. Los inconvenientes principales del elemento de Argyris son los siguientes:

- Es difícil dar valores en el contorno. Habitualmente hay que imponer las condiciones de tipo esencial en forma débil.
- Es difícil encontrar y evaluar las funciones de base locales para hacer el ensamblado.
- Relación entre las funciones de base locales y las globales y tipos de funciones de base.
- Es muy difícil pasar al elemento de referencia.

3. Elementos de Bell y BFS

Sea K un triángulo y sean L_1, L_2 y L_3 sus lados. Definimos el espacio

$$\mathbb{P}(K) = \left\{ v \in \mathbb{P}_5 \mid \partial_n v|_{L_i} \in \mathbb{P}_3(t), \forall i \right\}$$

que tiene dimensión 18 (la restricción al lado de la definición es la habitual parametrización lineal del lado).

Como grados de libertad locales se toman los 18 de los vértices del elemento de Argyris:

- valor en los vértices,
- gradiente en los vértices,
- segundas derivadas en los vértices.

El elemento resultante, llamado **triángulo de Bell** es un elemento de clase \mathcal{C}^1 , algo más simple que el de Argyris, pero que presenta básicamente las mismas dificultades de implementación. A pesar de su mayor simplicidad aparente, la ganancia es limitada y se pierde un orden de convergencia.

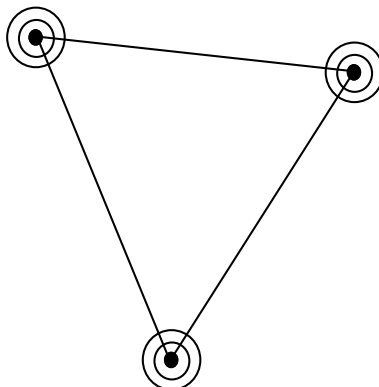


Figura 3.2: Representación esquemática del elemento de Bell

La imposibilidad de simplificar la situación bajando el número de grados de libertad, pero manteniendo la exigencia de continuidad \mathcal{C}^1 viene reflejada en el siguiente resultado: *un elemento triangular de clase \mathcal{C}^1 necesita las dos derivadas en los vértices.*

Para mallados de rectángulos orientados en los ejes cartesianos, se puede hacer lo siguiente. En todos los rectángulos, como el elemento de referencia, tomamos \mathbb{Q}_3 como espacio discreto. Concentramos cuatro grados de libertad por vértice:

- valor de la función,
- valor del gradiente,
- valor de la derivada segunda cruzada $\partial_{\xi\eta}^2$.

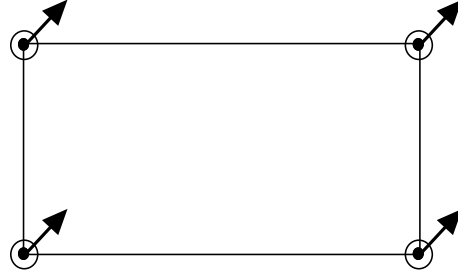


Figura 3.3: Representación esquemática del elemento BFS. La flecha diagonal en cada vértice hace referencia a la derivada segunda cruzada

Es bastante fácil demostrar que el elemento resultante es unisolvante (los grados de libertad determinan unívocamente el valor) y que es de clase \mathcal{C}^1 . El elemento resultante se denomina **elemento BFS (Bogner–Fox–Schmidt)**.

Ante la dificultad de emplear elementos de clase \mathcal{C}^1 , algunas veces conviene buscar otro tipo de técnicas de elementos finitos.

- Utilizar elementos no \mathcal{C}^1 ó no \mathcal{C}^0 incluso, esto es, elementos no conformes. En tal caso hay que estudiar el error de consistencia que se produce, algo que no ocurre con las formulaciones habituales.
- Cambiar el tipo de formulación y pasar a emplear formulaciones mixtas.

4. Ejercicios propuestos

1. Vamos a estudiar en este ejercicio el elemento BFS sobre el cuadrado de referencia $[-1, 1] \times [-1, 1]$.
 - a) Sea $q \in \mathbb{Q}_3$. Demuestra que si los ocho grados de libertad asociados al lado $\xi = -1$ (cuatro para $(-1, -1)$ y otros cuatro para $(-1, 1)$) se anulan, entonces, $q = 0$ en ese lado e igualmente $\partial_\xi q = 0$ en ese lado.
 - b) Sea $q \in \mathbb{Q}_3$ y supongamos que los dieciséis grados de libertad asociados a q en el cuadrado BFS se anulan. Escribiendo

$$q(\xi, \eta) = p_0(\eta) + p_1(\eta) (\xi - 1) + p_2(\eta) (\xi - 1)^2 + p_3(\eta) (\xi - 1)^3$$

(cada p_i es un polinomio de grado menor o igual que tres), demuestra que por anularse q y $\partial_\xi q$ en $\xi = 1$, necesariamente, $p_0(\eta) = p_1(\eta) = 0$. Seguidamente, mirando lo que ocurre en $\xi = -1$, deducir que $p_2(\eta) = p_3(\eta) = 0$, luego el elemento es unisolvante (los grados de libertad determinan unívocamente al polinomio).

2. Sobre el cuadrado de referencia $[-1, 1] \times [-1, 1]$ vamos a dar una estrategia para calcular las funciones de base del elemento BFS.

- a) Razonando como en el ejercicio anterior, demuestra que las cuatro funciones de base asociadas al vértice $(-1, -1)$ se pueden escribir en la forma

$$(\xi - 1)^2 (\eta - 1)^2 p(\xi, \eta), \quad p \in \mathbb{Q}_1.$$

- b) Escribiendo ahora el polinomio restante en la forma

$$p(\xi, \eta) = a_0 + a_1 (\xi + 1) + a_2 (\eta + 1) + a_3 (\xi + 1) (\eta + 1),$$

da una forma simple de calcular los coeficientes resultantes para las cuatro funciones de base asociadas al vértice $(-1, -1)$.

- c) Para los vértices restantes, con los cambios de variables

$$\begin{array}{ccc} \xi \leftrightarrow -\xi & \xi \leftrightarrow \xi & \xi \leftrightarrow -\xi \\ \eta \leftrightarrow \eta & \eta \leftrightarrow -\eta & \eta \leftrightarrow -\eta \end{array}$$

se pueden emplear para calcular las demás funciones de base. ¿Cómo?

3. Consultando en bibliografía (o en la web) o pensándolo tú mismo, averigua si es posible emplear elementos BFS sobre mallados generales de paralelogramos (la dificultad está en el grado de libertad que lleva segundas derivadas en el elemento de referencia).

Lección 4

Elementos no conformes

Se llaman elementos finitos no conformes a aquéllos que no satisfacen las hipótesis necesarias de continuidad que les permiten pertenecer al espacio donde se haya la solución. Por ejemplo, emplear elementos finitos no continuos cuando la solución se busca en $H^1(\Omega)$ (problemas de orden dos) o emplear elementos que no sean de clase C^1 cuando trabajamos en $H^2(\Omega)$ (problema de placa de Kirchhoff, viga de Euler–Bernouilli) es emplear elementos no conformes. En los elementos no conformes, además del error de aproximación hay un término de error de consistencia. Por este último término se puede afirmar que *no todo vale*. Espacios con buenas propiedades de aproximación no son válidos para resolver determinadas ecuaciones simplemente porque el error de consistencia no tiende a cero. Este error no existía en ninguno de los tipos de elementos que nos hemos encontrado hasta el momento, razón por la cual solo teníamos que preocuparnos del orden aproximación.

1. Error de las aproximaciones no conformes

Vamos a estudiar la idea de aproximación no conforme sobre un problema bastante simple, para analizar las nuevas fuentes de error de este tipo de discretizaciones.

Comenzamos con un problema de Dirichlet homogéneo para la ecuación de Poisson (Laplace no homogénea) sobre un dominio poligonal en el plano Ω , cuya frontera denotamos Γ :

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ u = 0, & \text{en } \Gamma. \end{cases}$$

La formulación variacional de este problema se establece en el espacio

$$V^0 := \left\{ u \in H^1(\Omega) \mid u = 0, \text{ en } \Gamma \right\}.$$

La formulación es simplemente

$$\begin{cases} u \in V^0, \\ \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v, \quad \forall v \in V^0. \end{cases}$$

Todos los métodos clásicos de elementos finitos se pueden describir de una u otra manera como la elección de una triangulación (una partición de Ω siguiendo determinadas reglas) y la construcción de un espacio V_h , de funciones polinómicas a trozos y continuas. A partir de allí, se construye el subespacio

$$V_h^0 = \left\{ u_h \in V_h \mid u_h = 0, \quad \text{en } \Gamma \right\}$$

y se escribe el problema variacional discreto

$$\begin{cases} u_h \in V_h^0, \\ \int_{\Omega} \nabla u_h \cdot \nabla v_h = \int_{\Omega} f v_h, \quad \forall v_h \in V_h^0. \end{cases}$$

La razón de exigir que las funciones discretas sean continuas es conseguir que

$$V_h \subset H^1(\Omega)$$

y, por tanto,

$$V_h^0 \subset V^0.$$

En ese caso, todo el error del método es un error de aproximación:

$$\|u - u_h\|_{1,\Omega} \leq C \inf \left\{ \|u - v_h\|_{1,\Omega} \mid v_h \in V_h^0 \right\},$$

es decir, el error del método depende exclusivamente de la capacidad del espacio V_h^0 (espacio en el que buscamos la solución) para aproximar a la solución, independientemente de que se cumpla una ecuación o no.

En cierto modo, el requisito de continuidad de los elementos finitos se nota también en la forma bilineal

$$\int_{\Omega} \nabla u_h \cdot \nabla v_h.$$

Matricialmente, esta forma bilineal se convierte en la matriz de rigidez. Si las funciones u_h, v_h no fueran continuas en las interfaces entre triángulos producirían algo así como deltas de Dirac en los lados y habría que incorporar esas distribuciones singulares al problema de alguna manera.

En cambio, si observamos la forma bilineal como la implementamos al ensamblar,

$$\sum_K \int_K \nabla u_h \cdot \nabla v_h,$$

no parece haber ningún inconveniente a la hora de emplear elementos que no sean continuos. La idea de las discretizaciones no conformes es la de escoger como V_h , un espacio de funciones polinómicas a trozos **no continuas**.

La utilización de elementos no conformes viene motivada por bastantes situaciones, entre ellas destacan:

- la dificultad de encontrar elementos de clase \mathcal{C}^1 simples para problemas planteados en $H^2(\Omega)$ (veremos un ejemplo al final de este capítulo),

- la necesidad de eliminar el efecto de *locking* producido por el exceso de rigidez de algunos elementos finitos en situaciones especiales.

Dada una triangulación \mathcal{T}_h , consideramos la forma bilineal

$$a_h(u, v) = \sum_K \int_K \nabla u \cdot \nabla v.$$

Si $u, v \in H^1(\Omega)$, entonces

$$a_h(u, v) = a(u, v)$$

(podemos pegar las integrales sobre los triángulos K para obtener una integral sobre Ω , ya que los gradientes existen como funciones en todo el dominio). Consideraremos como norma de energía

$$\|u\|_h = \left(\sum_K \int_K |\nabla u|^2 \right)^{1/2} = a_h(u, u)^{1/2}.$$

El error del método se mirará en el sentido siguiente

$$\|u - u_h\|_h = \left(\sum_K \int_K |\nabla u - \nabla u_h|^2 \right)^{1/2},$$

ya que u_h no va a estar en $H^1(\Omega)$ y no tendrá sentido hablar de $\|u - u_h\|_{1,\Omega}$.

Tomamos $V_h^0 \not\subset V^0$ un espacio que aproxime en cierto modo a las funciones de V^0 y proponemos

$$\left[\begin{array}{l} u_h \in V_h^0, \\ \sum_K \int_K \nabla u_h \cdot \nabla v_h = \int_{\Omega} f v_h, \quad \forall v_h \in V_h^0. \end{array} \right.$$

El estudio del error de este tipo de discretizaciones se deduce de un resultado muy simple y muy genérico, conocido a menudo como **segundo lema de Strang**. Vamos a quitar el superíndice nulo (procedente de las condiciones de contorno homogéneas) para simplificar las notaciones.

Proposición *Consideremos dos problemas variacionales, uno exacto y otro aproximado*

$$\left[\begin{array}{l} u \in V, \\ a(u, v) = \ell(v), \quad \forall v \in V, \end{array} \right. \quad \left[\begin{array}{l} u_h \in V_h, \\ a_h(u_h, v_h) = \ell_h(v_h), \quad \forall v_h \in V_h, \end{array} \right.$$

en las siguientes condiciones:

- V_h es de dimensión finita.
- $a_h(\cdot, \cdot)$ está definida en $V_h + V$, es simétrica y definida positiva.
- $a_h(\cdot, \cdot)$ es una extensión de la forma bilineal original, esto es,

$$a_h(u, v) = a(u, v), \quad \forall u, v \in V.$$

- ℓ_h es una extensión a $V_h + V$ de ℓ , es decir,

$$\ell_h(v) = \ell(v), \quad \forall v \in V.$$

Se define la norma de energía discreta

$$\|u\|_h := \sqrt{a_h(u_h, u_h)}, \quad u_h \in V_h + V.$$

Entonces

$$\|u - u_h\|_h \leq \inf_{v_h \in V_h} \|u - v_h\|_h + \sup_{0 \neq w_h \in V_h} \frac{|a_h(u - u_h, w_h)|}{\|w_h\|_h}.$$

Demostración. Consideramos el problema intermedio

$$\begin{cases} \tilde{u}_h \in V_h, \\ a_h(\tilde{u}_h, v_h) = a_h(u, v_h), \quad \forall v_h \in V_h. \end{cases}$$

La solución de este problema cumple

$$\|u - \tilde{u}_h\|_h = \inf_{v_h \in V_h} \|u - v_h\|_h,$$

ya que está definida por las ecuaciones que dan la mejor aproximación de u en V_h empleando la norma de energía discreta. Entonces

$$\|u - u_h\|_h \leq \|u - \tilde{u}_h\|_h + \|u_h - \tilde{u}_h\|_h.$$

Finalmente, se tiene que

$$\|u_h - \tilde{u}_h\|_h = \sup \frac{|a_h(u_h - \tilde{u}_h, w_h)|}{\|w_h\|_h} = \sup \frac{|\ell_h(w_h) - a_h(u, w_h)|}{\|w_h\|_h} = \sup \frac{|a_h(u_h - u, w_h)|}{\|w_h\|_h},$$

luego se obtiene la cota del enunciado. □

El primer término del error es un **error de aproximación**, similar al que nos encontramos en aproximaciones conformes. El segundo es nuevo: es un **error de consistencia**, que mide hasta qué punto la solución exacta no es solución de las ecuaciones discretas. La razón de la existencia de este término es que

$$a_h(u, v) = \ell(v), \quad \forall v \in V \quad \not\Rightarrow \quad a_h(u, v_h) = \ell_h(v_h), \quad \forall v_h \in V_h,$$

ya que

$$V_h \not\subset V.$$

2. Elementos de Crouzeix–Raviart

Vamos a regresar al problema con el que comenzábamos el capítulo. Añadimos al problema un término de reacción:

$$\begin{cases} -\Delta u + u = f, & \text{en } \Omega, \\ u = 0, & \text{en } \Gamma. \end{cases}$$

Denotaremos

$$V^0 := \left\{ u \in H^1(\Omega) \mid u = 0, \text{ en } \Gamma \right\}.$$

Las formas bilineales son:

$$\begin{aligned} a(u, v) &= \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} u v, \\ a_h(u, v) &= \sum_K \left(\int_K \nabla u \cdot \nabla v + \int_K u v \right), \end{aligned}$$

donde \mathcal{T}_h es una triangulación de Ω y la suma \sum_K se realiza sobre los elementos de la triangulación. Notemos que el término de masa/reacción se da partido a trozos en la definición de a_h , aunque al no haber derivadas en él es irrelevante si lo consideramos así o no.

El elemento que explicaremos se define sobre cualquier triángulo como sigue: el espacio discreto es \mathbb{P}_1 y los grados de libertad son los valores de la función en los puntos medios de los lados. Es obvio que un polinomio de grado menor o igual que uno está unívocamente determinado por el valor en los tres puntos medios de los lados, pero en el valor sobre un lado influyen los tres grados de libertad.

Pegando elementos finitos de este tipo, se obtiene el espacio discreto

$$V_h = \left\{ v_h : \Omega \rightarrow \mathbb{R} \mid \begin{array}{l} v_h|_K \in \mathbb{P}_1, \forall K \\ v_h \text{ continua en puntos medios de lados} \end{array} \right\}.$$

Notemos que

$$V_h \not\subset H^1(\Omega),$$

ya que las funciones de V_h no son continuas. La condición de Dirichlet se impone anulando los valores sobre los puntos medios de los lados frontera:

$$V_h^0 = \left\{ v_h \in V_h \mid v_h(P_i) = 0, \forall P_i \in \Gamma \right\}$$

Las funciones de V_h^0 no se anulan sobre la frontera.

Proposición *La forma bilineal discreta*

$$a_h(u, v) = \sum_K \left(\int_K \nabla u \cdot \nabla v + \int_K u v \right)$$

es definida positiva en $V^0 + V_h^0$.

Demostración. Supongamos que

$$w = w_1 + w_2 \in H^1(\Omega) + V_h$$

cumple

$$a_h(w, w) = 0.$$

Por tanto

$$\sum_K \int_K |\nabla w|^2 = 0.$$

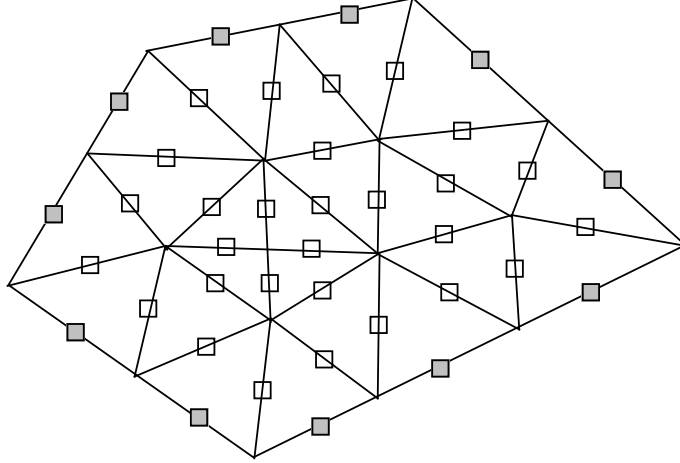


Figura 4.1: Una triangulación y los nodos correspondientes a elementos Crouzeix–Raviart, indicando los nodos Dirichlet

de donde se deduce que w es constante en cada triángulo. Esto implica que w es polinómica a trozos, luego al mirar

$$w_1 = w_2 - w \in H^1(\Omega),$$

obligatoriamente w_1 debe ser continua (¿por qué?). Así w es constante a trozos y continua en los puntos medios de los lados y, por tanto, es constante.

Si $w_1 = 0$ en Γ y w_2 se anula sobre los puntos medios de los lados exteriores (esto es lo que significa que $w \in V^0 + V_h^0$), entonces es fácil ver primero que $w_1 = 0$ en todo Ω y, de allí, $w_2 = 0$. \square

Del término de reacción/masa de la forma bilineal se deduce directamente que $a_h(w, w) = 0$ implica $w = 0$. Damos esta demostración más complicada para enfatizar que bastaría el operador $-\Delta$ con parte de condición Dirichlet para obtener esta propiedad.

El esquema discreto

$$\begin{cases} u_h \in V_h^0, \\ a_h(u_h, v_h) = \ell(v_h), \quad \forall v_h \in V_h^0, \end{cases}$$

encaja en la teoría de la Sección 1. Hay, por tanto, dos términos de error. El error de aproximación es simple, ya que se trata de aproximación polinómica a trozos y no hay ninguna diferencia con que la discretización sea conforme o no. Si $u \in H^2(\Omega)$

$$\inf_{v_h \in V_h} \|u - v_h\|_h \leq C(u) h$$

en condiciones habituales sobre no deformación de la malla.

Para el término de consistencia notemos que

$$\begin{aligned}
a_h(u, w_h) - \ell_h(w_h) &= \sum_K \left(\int_K \nabla u \cdot \nabla w_h + \int_K u w_h - \int_K f w_h \right) \\
&= \sum_K \left(\int_K \nabla u \cdot \nabla w_h + \int_K u w_h + \int_K (\Delta u - u) w_h \right) \\
&= \sum_K \int_{\partial K} (\partial_n u) w_h = \sum_e \int_e \partial_n u [w_h]_e
\end{aligned}$$

donde e son los lados de los triángulos y $[w_h]$ mide el salto de w_h a través de e en la dirección normal. Se supone que tomamos $w_h \equiv 0$ en el exterior de Ω para definir los saltos en lados frontera.

Nótese que $[w_h]_e(P) = 0$ si P es el punto medio de e . Así, para cualesquiera $\xi_1, \xi_2 \in \mathbb{R}$

$$\int_e \partial_n u [w_h]_e = \int_e (\partial_n u - \xi_1) [w_h - \xi_2]_e,$$

y por un argumento estándar se demuestra que

$$\left| \int_e \partial_n u [w_h]_e \right| \leq Ch |u|_{2, \tilde{K}} \left(|w_h|_{1, K_1} + |w_h|_{1, K_2} \right),$$

con las notaciones de la gráfica ($\tilde{K} = K_1 \cup K_2$) Uniendo todo se llega a que

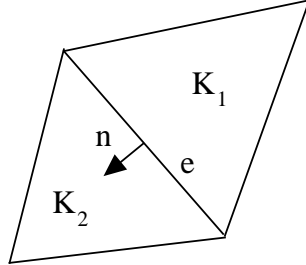


Figura 4.2: *Dos triángulos contiguos y una única normal orientada*

$$\|u - u_h\|_h = \mathcal{O}(h).$$

El elemento finito anterior recibe comúnmente el nombre de **elemento de Crouzeix–Raviart** aunque se debe a Gilbert Strang. Su utilización, junto con constantes a trozos para presiones en el problema de Stokes, constituye el primer par no conforme de elementos de Crouzeix–Raviart, que existen para órdenes superiores.

Funciones de base. Se asocia una función de base nodal a cada punto medio. Los soportes son

- dos triángulos, para nodos interiores,

- un triángulo, para nodos exteriores Neumann.

El número total de grados de libertad es igual al número de lados de la triangulación. El nivel de adyacencia es muy bajo. Dos nodos son adyacentes si son lados del mismo triángulo.

Para las condiciones de Dirichlet no homogéneas, se anclan los nodos Dirichlet y se eliminan los tests correspondientes. Si un triángulo tiene un lado Neumann, las tres funciones de base aportan a la integral sobre el lado ya que **no se anulan** en el lado. Esto implica que el ensamblado Neumann se realizará no por lados como en los elementos lagrangianos comunes, sino por triángulos exteriores.

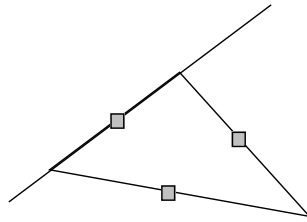


Figura 4.3: *Los tres nodos de un triángulo apoyado en la frontera Neumann son relevantes*

Para el ensamblado, se puede pasar al elemento de referencia puesto que son elementos lagrangianos. Los requisitos para el ensamblado son los siguientes:

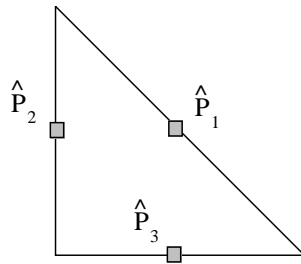


Figura 4.4: *Triángulo de referencia*

- Las habituales listas de coordenadas de vértices de la triangulación, de triángulos (con numeraciones de sus vértices) y de lados Dirichlet y Neumann.
- Una numeración de los lados con referencia local a los extremos que los componen y una lista de cuáles son los lados de cada triángulo (respetando un orden local similar al tomado por sus vértices). Estas dos listas son relativamente simples de preprocesar a partir de la información de las listas habituales.
- De la lista anterior se deduce la lista de lados Dirichlet.
- Para el ensamblado de condiciones Neumann hay que conocer qué triángulos tienen lados en la frontera Neumann o cuáles son estos lados.

3. El elemento de Morley para el bilaplaciano

La forma bilineal del problema de placa de Kirchhoff es

$$a(u, v) = \nu \int_{\Omega} \Delta u \Delta v + (1 - \nu) \int_{\Omega} \text{Hu} : \text{Hv}.$$

Esta forma se puede partir según una triangulación y definirse así

$$a_h(u, v) = \sum_K \left(\nu \int_K \Delta u \Delta v + (1 - \nu) \int_K \text{Hu} : \text{Hv} \right).$$

El **elemento de Morley** es polinómico \mathbb{P}_2 sobre cada triángulo. Los grados de libertad locales son

- valor en los vértices,
- derivada normal en los puntos medios.

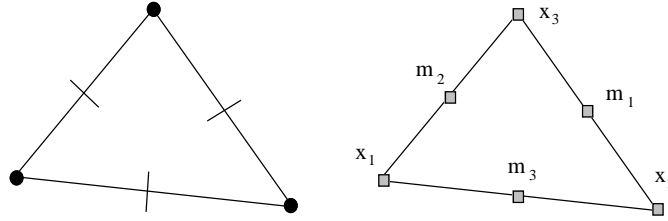


Figura 4.5: Representación esquemática del elemento de Morley y numeración local de sus nodos

Proposición *Los seis grados de libertad locales dados determinan unívocamente un polinomio \mathbb{P}_2 en un triángulo.*

Demostración: Denotemos, como en la figura, \mathbf{x}_α a los vértices y \mathbf{m}_α a los puntos medios de los lados. Si $p \in \mathbb{P}_2$ y $p(\mathbf{x}_\alpha) = 0$, $\alpha = 1, 2, 3$, entonces

$$\partial_\tau p(\mathbf{m}_\alpha) = 0, \quad \alpha = 1, 2, 3.$$

Así

$$\left. \begin{array}{l} \partial_\tau p(\mathbf{m}_\alpha) = 0 \\ \partial_n p(\mathbf{m}_\alpha) = 0 \end{array} \right\} \Rightarrow \nabla p(\mathbf{m}_\alpha) = 0.$$

Como $\nabla p \in (\mathbb{P}_1)^2$ y se anula en tres puntos no alineados, obligatoriamente $\nabla p = 0$, luego p es constante. Como $p(\mathbf{x}_\alpha) = 0$, necesariamente $p = 0$. □

No obstante, si coinciden los tres grados de libertad referidos a un lado (valor en los vértices más derivada normal en el punto medio) de dos triángulos contiguos, el resultado no es una función continua. Esto implica que el elemento de Morley no es de clase \mathcal{C}^0 . El espacio discreto obtenido es

$$V_h = \left\{ u_h : \Omega \rightarrow \mathbb{R} \mid u_h|_K \in \mathbb{P}_2, \forall K, \quad \left. \begin{array}{l} u_h \text{ continua en vértices y con} \\ \text{derivada normal continua en puntos medios} \end{array} \right\}.$$

Nótese que $\nabla u_h(\mathbf{m}_i)$ está bien definido en todos los puntos medios ya que fijamos la derivada normal y la tangencial viene dada por la coincidencia de valores en los vértices.

Hay dos tipos de funciones de base:

- asociadas a vértices,
- asociadas a puntos medios.

Unos comentarios finales:

- Para añadir restricciones Dirichlet (nótese que las condiciones de Dirichlet son dos) se exigen $u_h(P) = 0$ en los vértices Dirichlet y $\partial_n u_h(P) = 0$ en los puntos medios Dirichlet. Esto equivale a construir V_h^0 eliminando las funciones de base nodales asociadas a todos los nodos Dirichlet.
- Al ser los grados de libertad de tipo Hermite asociados a los lados y no a los vértices, cuando se restringe una función de base nodal a un triángulo, sale una de las funciones de base del elemento triangular.
- En cambio, al pasar al elemento de referencia no se preservan las funciones de base, sino que se mezclan. Esto complica el ensamblado como en todo tipo de elementos de Hermite.
- La forma bilineal es definida positiva en $H_0^2(\Omega) + V_h^0$ (la demostración de esto es paralela a la realizada con el elemento de Crouzeix–Raviart).
- El análisis de convergencia es muy similar al del elemento de Crouzeix–Raviart

$$\|u - u_h\|_h = \mathcal{O}(h).$$

- El tipo de información que se emplea para el ensamblado es como en el caso de elementos \mathbb{P}_2 .

4. Ejercicios propuestos

1. Calcular las funciones de base en el elemento de referencia para el elemento de Crouzeix–Raviart.
2. Averiguar cómo es el elemento de Crouzeix–Raviart de orden dos.
3. Con la forma bilineal de la placa de Kirchhoff y el espacio V_h de los elementos de Morley, demostrar que si

$$a_h(w, w) = 0, \quad w \in H^2(\Omega) + V_h$$

entonces, necesariamente w es globalmente un polinomio de grado uno sobre todo Ω . Además si descomponemos $w = w_1 + w_2$, con $w_1 \in H^2(\Omega)$ y $w_2 \in V_h$, obligatoriamente w_1 es un polinomio de grado menor o igual que dos sobre todo Ω .

Lección 5

Formulaciones mixtas y condiciones inf–sup

Qué son los elementos finitos mixtos no se entiende hasta que uno no se encuentra con formulaciones mixtas o con un problema cuya estructura natural es mixta. Esto puede provocar inicialmente una cierta confusión: uno puede hacer una formulación mixta de la ecuación de Laplace, pero también puede apreciar la estructura mixta del problema de Stokes. Al final el nombre va a ser lo de menos y en lo que habrá que concentrarse es que en los problemas mixtos va a haber más de una incógnita y que alguna de las ecuaciones se pueden entender de forma generalizada como restricciones o alguna de las incógnitas como un multiplicador. En este capítulo llevaremos en paralelo el estudio de los problemas con forma mixta y de sus discretizaciones de tipo Galerkin, observando cómo estas ideas se van reflejando en un ejemplo característico: la formulación mixta del laplaciano (con dos incógnitas, el potencial y el gradiente), que corresponde también a un problema de flujo de Darcy linealizado.

1. Problemas con estructura mixta

Un **problema mixto** tiene la siguiente forma general:

- Hay dos espacios de Hilbert, X y M . A veces se dice que las variables en X son las variables primales, y las de M los multiplicadores (ya veremos por qué más adelante).
- Hay una forma bilineal continua en X , frecuentemente simétrica (aunque esto no sea imprescindible)

$$a : X \times X \rightarrow \mathbb{R}.$$

- Hay una forma bilineal continua

$$b : X \times M \rightarrow \mathbb{R}.$$

El problema consiste en hallar p y u :

$$\begin{cases} p \in X, & u \in M \\ a(p, q) + b(q, u) = \ell(q), & \forall q \in X, \\ b(p, v) = \chi(v), & \forall v \in M. \end{cases}$$

Los datos son dos aplicaciones lineales y continuas

$$\ell : X \rightarrow \mathbb{R}, \quad \chi : M \rightarrow \mathbb{R}.$$

La segunda ecuación se considera una **restricción**. Esto no es un sistema triangular donde primero se resuelve una ecuación y luego se sustituye en la otra ecuación.

Un ejemplo muy simple lo constituye la llamada **formulación mixta** para el problema de Laplace. Sea Ω un dominio del plano o del espacio y Γ su frontera. En lugar de plantear el problema

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ u = g_0, & \text{en } \Gamma, \end{cases}$$

mediante una formulación variacional tradicional, añadimos una incógnita al problema

$$\mathbf{p} = \nabla u.$$

Esta función debe cumplir la restricción

$$\operatorname{div} \mathbf{p} = -f, \quad \text{en } \Omega.$$

Por el teorema de la divergencia

$$\int_{\Omega} u(\operatorname{div} \mathbf{q}) + \int_{\Omega} \nabla u \cdot \mathbf{q} = \int_{\Gamma} u(\mathbf{q} \cdot \mathbf{n}),$$

luego podemos escribir que

$$\int_{\Omega} \mathbf{p} \cdot \mathbf{q} + \int_{\Omega} u(\operatorname{div} \mathbf{q}) = \int_{\Gamma} g_0(\mathbf{q} \cdot \mathbf{n}), \quad \forall \mathbf{q}.$$

La restricción sobre la divergencia de \mathbf{p} se escribe como

$$\int_{\Omega} (\operatorname{div} \mathbf{p}) v = - \int_{\Omega} f v, \quad \forall v.$$

El problema encaja con la estructura mixta si tomamos las formas bilineales

$$\begin{aligned} a(\mathbf{p}, \mathbf{q}) &:= \int_{\Omega} \mathbf{p} \cdot \mathbf{q}, \\ b(\mathbf{p}, v) &:= \int_{\Omega} (\operatorname{div} \mathbf{p}) v \end{aligned}$$

y los datos

$$\begin{aligned}\ell(\mathbf{q}) &:= \int_{\Gamma} g_0(\mathbf{q} \cdot \mathbf{n}) \\ \chi(v) &:= - \int_{\Omega} f v.\end{aligned}$$

Los espacios para las variables deben ser suficientemente exigentes como para que las formas bilineales correspondientes sean continuas. En concreto es razonable tomar

$$X := \mathbf{H}(\text{div}, \Omega) := \left\{ \mathbf{p} : \Omega \rightarrow \mathbb{R}^d \left| \begin{array}{l} p_i \in L^2(\Omega), \quad \forall i \\ \text{div } \mathbf{p} \in L^2(\Omega) \end{array} \right. \right\}$$

con la norma

$$\|\mathbf{p}\|_{\text{div}, \Omega} := \left[\int_{\Omega} |\mathbf{p}|^2 + \int_{\Omega} |\text{div } \mathbf{p}|^2 \right]^{1/2}$$

y

$$M = L^2(\Omega).$$

La condición de Dirichlet en esta formulación queda incorporada como un término independiente. Pasa de ser esencial (en la formulación variacional clásica) a ser natural en la formulación mixta. La condición de Neumann pasaría a ser esencial.

Nota. Las ecuaciones anteriores aparecen también en los llamados flujos de Darcy, un modelo lineal simple para flujo en medios porosos. \square

Antes de entrar en resultados sobre resolubilidad del problema y sobre su discretización vamos a notar un resultado bastante simple.

Proposición *Supongamos que la forma bilineal a es simétrica y semidefinida positiva, es decir*

$$a(p, q) = a(q, p), \quad \forall p, q \in X$$

y

$$a(p, p) \geq 0, \quad \forall p \in X.$$

Entonces el problema

$$\left[\begin{array}{ll} p \in X, & u \in M \\ a(p, q) + b(q, u) &= \ell(q), \quad \forall q \in X, \\ b(p, v) &= \chi(v), \quad \forall v \in M. \end{array} \right.$$

es equivalente al problema de minimización con restricciones

$$\frac{1}{2}a(p, p) - \ell(p) = \text{mín!}, \quad \begin{array}{l} p \in X, \\ b(p, v) = \chi(v), \quad \forall v \in M. \end{array}$$

Idea de la demostración. En el problema de minimización se añade un multiplicador de Lagrange $u \in M$ para hacerse cargo de las restricciones. Los mínimos del problema son necesariamente puntos críticos del lagrangiano

$$\mathcal{L}(p, u) := \frac{1}{2}a(p, p) - \ell(p) + b(p, u) - \chi(u).$$

Seguidamente se demuestra que los puntos críticos del Llagrangiano son puntos de silla y que se calculan con las ecuaciones de la formulación mixta. \square

Aplicado a nuestro ejemplo, nos damos cuenta de que estamos minimizando

$$\frac{1}{2} \int_{\Omega} |\mathbf{p}|^2 - \int_{\Gamma} g_0(\mathbf{p} \cdot \mathbf{n}) = \text{mín!}, \quad \text{div } \mathbf{p} = -f.$$

2. Discretizaciones de problemas mixtos

Para aproximar el problema

$$\begin{cases} p \in X, & u \in M, \\ a(p, q) + b(q, u) = \ell(q), & \forall q \in X, \\ b(p, v) = \chi(v), & \forall v \in M, \end{cases}$$

tomamos pares de espacios de dimensión finita

$$X_h \subset X, \quad M_h \subset M$$

y consideramos un método de Galerkin

$$\begin{cases} p_h \in X_h, & u_h \in M_h, \\ a(p_h, q_h) + b(q_h, u_h) = \ell(q_h), & \forall q_h \in X_h, \\ b(p_h, v_h) = \chi(v_h), & \forall v_h \in M_h. \end{cases}$$

Este problema es equivalente a un sistema de ecuaciones que vamos a desarrollar seguidamente.

Tomamos bases de ambos espacios discretos

$$\begin{aligned} \{\varphi_1, \dots, \varphi_N\}, & \quad \text{base de } X_h, \\ \{\mu_1, \dots, \mu_K\}, & \quad \text{base de } M_h \end{aligned}$$

y descomponemos las incógnitas en la base.

$$p_h = \sum_{j=1}^N p_j \varphi_j, \quad u_h = \sum_{j=1}^K u_j \mu_j.$$

Entonces, el problema es equivalente a

$$\begin{cases} a\left(\sum_j p_j \varphi_j, \varphi_i\right) + b\left(\varphi_i, \sum_j u_j \mu_j\right) = \ell(\varphi_i), & i = 1, \dots, N \\ b\left(\sum_j p_j \varphi_j, \mu_i\right) = \chi(\mu_i), & i = 1, \dots, K \end{cases}$$

Así se llega al sistema con $N + K$ ecuaciones e incógnitas

$$\begin{cases} \sum_{j=1}^N a(\varphi_j, \varphi_i) p_j + \sum_{j=1}^K b(\varphi_i, \mu_j) u_j = \ell(\varphi_i), & i = 1, \dots, N, \\ \sum_{j=1}^N b(\varphi_j, \mu_i) p_j = \chi(\mu_i), & i = 1, \dots, K. \end{cases}$$

Podemos apreciar la estructura en bloques del sistema agrupando ecuaciones e incógnitas como sigue

$$\begin{aligned} \mathbf{A} &= (a_{ij}), \quad a_{ij} = a(\varphi_j, \varphi_i), & (N \times N) \\ \mathbf{B} &= (b_{ij}), \quad b_{ij} = b(\varphi_j, \mu_i), & (K \times N) \\ \mathbf{p} &= (p_1, \dots, p_N)^\top, & \mathbf{u} = (u_1, \dots, u_K)^\top \\ \mathbf{a} &= (\ell(\varphi_1), \dots, \ell(\varphi_N))^\top, & \mathbf{b} = (\chi(\mu_1), \dots, \chi(\mu_K))^\top \end{aligned}$$

el sistema se estructura

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}.$$

Podemos regresar aquí al problema de minimización con restricciones y ver qué ha ocurrido. Para esto hay que suponer que la forma bilineal a es simétrica y semidefinida positiva.

Puesto que el problema

$$\begin{cases} p_h \in X_h, & u_h \in M_h, \\ a(p_h, q_h) + b(q_h, u_h) = \ell(q_h), & \forall q_h \in X_h, \\ b(p_h, v_h) = \chi(v_h), & \forall v_h \in M_h, \end{cases}$$

es un caso particular de problema mixto establecido en espacios de dimensión finita, el problema es equivalente al problema de minimización con restricciones

$$\begin{aligned} \frac{1}{2}a(p_h, p_h) - \ell(p_h) &= \text{mín!}, & p_h \in X_h, \\ b(p_h, v_h) &= \chi(v_h), & \forall v_h \in M_h. \end{aligned}$$

Es obvio aquí que no sólo hemos restringido el ámbito de búsqueda del mínimo (de X a X_h), sino que hemos modificado la restricción. Aunque p_h cumpla

$$b(p_h, v_h) = \chi(v_h), \quad \forall v_h \in M_h$$

no hemos de esperar que

$$b(p_h, v) = \chi(v), \quad \forall v \in M.$$

La variable discreta $u_h \in M_h$ es el multiplicador de Lagrange discreto.

Si tomamos el punto de vista del sistema, nos encontramos con que

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$$

equivale a

$$\frac{1}{2}\mathbf{p} \cdot (\mathbf{A} \mathbf{p}) - \mathbf{a} \cdot \mathbf{p} = \text{mín!}, \quad \mathbf{p} \in \mathbb{R}^N,$$

$$\mathbf{B} \mathbf{p} = \mathbf{b}.$$

Esto es un problema de minimización de un funcional cuadrático con restricciones lineales. Las incógnitas $\mathbf{u} \in \mathbb{R}^K$ corresponden a los multiplicadores de Lagrange que hay que incorporar para tratar las K restricciones.

3. Aspectos teóricos: problema exacto

En los problemas mixtos la demostración de existencia y unicidad cobra un valor adicional sobre los casos tradicionales. La razón es que la estabilidad de los métodos numéricos no se hereda automáticamente de la formulación como ocurre en los problemas elípticos. Vamos, por tanto, primero a observar en qué condiciones se puede asegurar la existencia y unicidad de solución y luego intentar repetir los argumentos en el caso discreto.

Lo primero que hay que notar es que si ha de haber solución del problema, necesariamente para todo χ debe haber algún elemento de X que cumpla la restricción (segunda ecuación de la formulación)

$$b(p_\chi, v) = \chi(v), \quad \forall v \in M.$$

Si se tiene un elemento p_χ que cumple la restricción todos los demás son $p_\chi + X_0$, donde

$$X_0 = \left\{ p \in X \mid b(p, v) = 0, \quad \forall v \in M \right\}.$$

El primer resultado importante es una forma múltiple de garantizar que existen elementos que cumplen la restricción, sea cual sea χ (es, por tanto, un resultado de suprayectividad).

Proposición *Las siguientes propiedades son equivalentes:*

- Para todo $\chi : M \rightarrow \mathbb{R}$ lineal y continuo, existe p_χ tal que

$$b(p_\chi, v) = \chi(v), \quad \forall v \in M.$$

- Existe $\beta > 0$ tal que

$$\sup_{0 \neq p \in X} \frac{b(p, u)}{\|p\|_X} \geq \beta \|u\|_M, \quad \forall u \in M.$$

- Existe $\beta > 0$ que cumple: si $u \neq 0$, se puede encontrar $0 \neq p \in X$ de forma que

$$b(p, u) \geq \beta \|p\|_X \|u\|_M.$$

- Se cumple la siguiente cota:

$$\inf_{0 \neq u \in M} \left[\sup_{0 \neq p \in X} \frac{b(p, u)}{\|p\|_X \|u\|_M} \right] > 0.$$

La última condición se suele denominar condición ínfimo–supremo o simplemente condición inf–sup. En el caso discreto, que nos encontraremos más adelante, suele ser llamada condición de Babuška–Brezzi o condición inf–sup discreta.

Proposición *Supongamos que existen $\alpha, \beta > 0$ tal que*

$$\sup_{0 \neq p \in X} \frac{b(p, u)}{\|p\|_X} \geq \beta \|u\|_M, \quad \forall u \in M$$

y

$$a(p, p) \geq \alpha \|p\|_X^2, \quad \forall p \in X_0.$$

Entonces el problema

$$\begin{cases} p \in X, & u \in M, \\ a(p, q) + b(q, u) = \ell(q), & \forall q \in X, \\ b(p, v) = \chi(v), & \forall v \in M, \end{cases}$$

tiene solución única, que es función continua de los datos. La cota de continuidad depende de forma creciente de $1/\alpha$ y $1/\beta$.

Enfaticemos una vez más que la primera condición es necesaria. La segunda condición del resultado puede ser relajada. En todo caso, vamos a examinar qué hay que hacer para verificar las hipótesis del resultado.

- Probar la condición inf–sup para b . En general, lo más práctico, si no se puede aplicar la primera forma de escribirla, es intentar la tercera.
- Identificar el conjunto

$$X_0 = \left\{ p \in X \mid b(p, v) = 0, \quad \forall v \in M \right\}.$$

- Verificar la elipticidad de a sobre X_0 .

Vamos a intentar mostrar cómo se aplican los pasos al ejemplo estándar de este tema. La condición inf–sup para

$$b(\mathbf{p}, u) = \int_{\Omega} (\operatorname{div} \mathbf{p}) u$$

es la parte más compleja. Hay que demostrar que

$$\sup_{0 \neq \mathbf{p} \in \mathbf{H}(\operatorname{div}, \Omega)} \frac{1}{\|\mathbf{p}\|_{\operatorname{div}, \Omega}} \int_{\Omega} (\operatorname{div} \mathbf{p}) u \geq \beta \|u\|_{0, \Omega}, \quad \forall u \in L^2(\Omega),$$

para algún $\beta > 0$.

Demostración. Empleamos la tercera forma de escribir la condición inf-sup para demostrarla. Sea $0 \neq u \in L^2(\Omega)$. Resolvemos el problema

$$\begin{cases} -\Delta\phi = -u, & \text{en } \Omega, \\ \phi = 0, & \text{en } \Gamma, \end{cases}$$

y definimos $\mathbf{p} = \nabla\phi$ de forma que $\text{div } \mathbf{p} = u$. Notemos que el problema que permite calcular ϕ en función de u permite asegurar que

$$\|\phi\|_{1,\Omega} \leq C_\Omega \|u\|_{0,\Omega}.$$

Así

$$\begin{aligned} \|\mathbf{p}\|_{\text{div},\Omega} &= \left[\int_\Omega |\mathbf{p}|^2 + \int_\Omega |\text{div } \mathbf{p}|^2 \right]^{1/2} \\ &= \left[\int_\Omega |\nabla\phi|^2 + \int_\Omega |u|^2 \right]^{1/2} \leq \sqrt{C_\Omega^2 + 1} \|u\|_{0,\Omega}. \end{aligned}$$

Llegados aquí ya está todo probado, puesto que, por un lado podemos asegurar que $\mathbf{p} \neq 0$ (¿por qué?) y además

$$b(\mathbf{p}, u) = \int_\Omega (\text{div } \mathbf{p}) u = \int_\Omega |u|^2 \geq (C_\Omega^2 + 1)^{-1/2} \|\mathbf{p}\|_{\text{div},\Omega} \|u\|_{0,\Omega},$$

lo cual es equivalente a la condición inf-sup. □

Seguidamente, es fácil notar que

$$X_0 = \left\{ \mathbf{p} \in \mathbf{H}(\text{div}, \Omega) \mid \text{div } \mathbf{p} = 0 \right\}.$$

La elipticidad es inmediata ya que si $\mathbf{p} \in X_0$, entonces

$$a(\mathbf{p}, \mathbf{p}) = \int_\Omega |\mathbf{p}|^2 = \|\mathbf{p}\|_{\text{div},\Omega}^2.$$

4. Aspectos teóricos: problema discreto

El principal inconveniente de las formulaciones mixtas es que, aunque se cumplan las hipótesis que garantizan que el problema está bien planteado, no hay ninguna garantía de que las hipótesis se heredan en el caso discreto.

El primer inconveniente es el siguiente. Si miramos la cantidad

$$\beta_h := \inf_{0 \neq u_h \in M_h} \left[\sup_{0 \neq p_h \in X_h} \frac{b(p_h, u_h)}{\|p_h\|_X \|u_h\|_M} \right]$$

no hay ninguna garantía de que $\beta_h > 0$. Si $\beta_h = 0$, entonces la restricción discreta

$$b(p_h, v_h) = \chi(v_h), \quad \forall v_h \in M_h,$$

podría no cumplirse. Esto viene a ser lo mismo que notar que el bloque de ecuaciones

$$\mathbf{B} \mathbf{p} = \mathbf{b}$$

porque a \mathbf{B} le falta rango por columnas. Visto así, es evidente que necesariamente

$$\dim X_h = N \geq K = \dim M_h,$$

puesto que si el rango debe ser máximo por columnas, debe haber tantas (o más) columnas como filas. Este es un aspecto de los métodos mixtos que puede chocar algo con la intuición: las restricciones no deben ser demasiado discretizadas.

Incluso si $\beta_h > 0$, es necesario que β_h no tienda a cero conforme se refinan las discretizaciones. Si esto ocurre, las cotas de estabilidad se estropearán. Vamos a repetir los argumentos del problema exacto sobre el problema discreto. Necesitaremos del espacio de elementos de X_h que cumplen la restricción discreta

$$X_h^0 = \left\{ p_h \in X_h \mid b(p_h, v_h) = 0, \forall v_h \in M_h \right\}.$$

Proposición *Supongamos que existen $\alpha^*, \beta^* > 0$ tal que*

$$\sup_{0 \neq p_h \in X_h} \frac{b(p_h, u_h)}{\|p_h\|_X} \geq \beta^* \|u_h\|_M, \quad \forall u_h \in M_h$$

y

$$a(p_h, p_h) \geq \alpha^* \|p_h\|_X^2, \quad \forall p_h \in X_h^0.$$

Entonces el problema

$$\begin{cases} p_h \in X_h, & u_h \in M_h, \\ a(p_h, q_h) + b(q_h, u_h) = \ell(q_h), & \forall q_h \in X_h, \\ b(p_h, v_h) = \chi(v_h), & \forall v_h \in M_h, \end{cases}$$

tiene solución única y se tiene la estimación casi-óptima del error

$$\|p - p_h\|_X + \|u - u_h\|_M \leq C \left[\inf_{q_h \in X_h} \|p - q_h\|_X + \inf_{v_h \in M_h} \|u - v_h\|_M \right].$$

5. Forma matricial de la condición inf-sup

Dados dos espacios de dimensión finita $M_h \subset M$ y $X_h \subset X$, vamos a intentar estimar la constante β_h

$$\beta_h := \inf_{0 \neq u_h \in M_h} \left[\sup_{0 \neq p_h \in X_h} \frac{b(p_h, u_h)}{\|u_h\|_M \|p_h\|_X} \right]$$

esto es, estimar la mejor β_h de forma que se tenga la desigualdad

$$\sup_{0 \neq p_h \in X_h} \frac{b(p_h, u_h)}{\|p_h\|_X} \geq \beta_h \|u_h\|_M, \quad \forall u_h \in M_h.$$

El interés numérico es comprobar, para un par de sucesiones de espacios, que β_h se mantiene apartado de cero.

Tomamos bases de los espacios X_h y M_h :

$$\begin{aligned} \varphi_1, \dots, \varphi_N, & \quad \text{base de } X_h, \\ \mu_1, \dots, \mu_K, & \quad \text{base de } M_h. \end{aligned}$$

Construimos las siguientes matrices:

$$\begin{aligned} \mathbf{B} & : b_{ij} = b(\varphi_j, \mu_i), \\ \mathbf{R} & : r_{ij} = (\varphi_i, \varphi_j)_X, \\ \mathbf{M} & : m_{ij} = (\mu_i, \mu_j)_M. \end{aligned}$$

Las matrices \mathbf{R} y \mathbf{M} son simétricas y definidas positivas. La matriz \mathbf{B} es

$$K \times N = \dim M_h \times \dim X_h.$$

Relación de las matrices con las normas y la forma bilineal. Sean $\mathbf{p} \in \mathbb{R}^N$ y $\mathbf{u} \in \mathbb{R}^K$. Entonces, si

$$p_h = \sum_j p_j \varphi_j, \quad u_h = \sum_j u_j \mu_j$$

tenemos que

$$b(p_h, u_h) = \mathbf{u} \cdot \mathbf{B} \mathbf{p}.$$

Además

$$\begin{aligned} \|u_h\|_M & = (\mathbf{u} \cdot \mathbf{M} \mathbf{u})^{1/2} = |\mathbf{M}^{1/2} \mathbf{u}|, \\ \|p_h\|_X & = (\mathbf{p} \cdot \mathbf{R} \mathbf{p})^{1/2} = |\mathbf{R}^{1/2} \mathbf{p}|. \end{aligned}$$

La desigualdad

$$\sup_{0 \neq p_h \in X_h} \frac{b(p_h, u_h)}{\|p_h\|_X} \geq \beta_h \|u_h\|_M, \quad \forall u_h \in M_h.$$

es entonces equivalente a que

$$\sup_{0 \neq \mathbf{p} \in \mathbb{R}^N} \frac{\mathbf{u} \cdot \mathbf{B} \mathbf{p}}{|\mathbf{R}^{1/2} \mathbf{p}|} \geq \beta_h |\mathbf{M}^{1/2} \mathbf{u}|, \quad \forall \mathbf{u} \in \mathbb{R}^K.$$

De esta desigualdad se deduce trivialmente que $\mathbf{B}^\top \mathbf{u} = \mathbf{0}$ debe implicar que $\mathbf{u} = \mathbf{0}$, luego el rango de \mathbf{B} debe ser K . Así, necesariamente

$$\dim M_h \leq \dim X_h.$$

Interpretación como valores propios/singulares. Como

$$\sup_{0 \neq \mathbf{p} \in \mathbb{R}^N} \frac{\mathbf{u} \cdot \mathbf{B} \mathbf{p}}{|\mathbf{R}^{1/2} \mathbf{p}|} = \sup_{0 \neq \mathbf{p} \in \mathbb{R}^N} \frac{(\mathbf{B}^\top \mathbf{u}) \cdot \mathbf{p}}{|\mathbf{R}^{1/2} \mathbf{p}|} = \sup_{0 \neq \mathbf{v} \in \mathbb{R}^N} \frac{(\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{u}) \cdot \mathbf{v}}{|\mathbf{v}|},$$

la condición se reescribe como

$$|\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{M}^{-1/2} \mathbf{n}| = \sup_{0 \neq \mathbf{v} \in \mathbb{R}^N} \frac{(\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{M}^{-1/2} \mathbf{n}) \cdot \mathbf{v}}{|\mathbf{v}|} \geq \beta_h |\mathbf{n}|, \quad \forall \mathbf{n} \in \mathbb{R}^K,$$

es decir

$$\inf_{\mathbf{0} \neq \mathbf{n} \in \mathbb{R}^K} \frac{|\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{M}^{-1/2} \mathbf{n}|}{|\mathbf{n}|} \geq \beta_h.$$

La cantidad de la izquierda es el menor valor singular de la matriz $\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{M}^{-1/2}$, es decir, la raíz cuadrada del mínimo valor propio de

$$(\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{M}^{-1/2})^\top (\mathbf{R}^{-1/2} \mathbf{B}^\top \mathbf{M}^{-1/2}) = \mathbf{M}^{-1/2} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^\top \mathbf{M}^{-1/2}.$$

El menor valor propio (son todos positivos si $\text{rango}(\mathbf{B}) = K$), es el menor valor λ , tal que existe $\mathbf{u} \neq \mathbf{0}$ de forma que

$$\mathbf{B} \mathbf{R}^{-1} \mathbf{B}^\top \mathbf{u} = \lambda \mathbf{M} \mathbf{u}.$$

6. Verificación numérica

Método de potencias. Si $\mathbf{p} \neq \mathbf{0}$ cumple que

$$\mathbf{A} \mathbf{p} = \lambda \mathbf{p}, \quad |\mathbf{p}| = 1,$$

entonces $\mathbf{p}^\top \mathbf{A} \mathbf{p} = \lambda$. El método de potencias para una matriz cuadrada \mathbf{A} se basa en que al definir una sucesión

$$\mathbf{p}_{n+1} = \mathbf{A} \mathbf{p}_n$$

la sucesión de vectores normalizados $(1/|\mathbf{p}_n|) \mathbf{p}_n$ converge a un vector propio asociado al valor propio dominante (al de mayor módulo) con probabilidad uno (si el valor propio dominante es negativo, la subsucesión par y la impar convergen por separado). Si aplicamos esta misma idea a la matriz $\mathbf{A} = \mathbf{B}^{-1}$ (supuesto que la inversa exista, claro), tenemos que al normalizar la sucesión

$$\mathbf{p}_{n+1} = \mathbf{B}^{-1} \mathbf{p}_n$$

convergemos al mayor valor propio de \mathbf{B}^{-1} , que es el inverso del menor valor propio de \mathbf{B} . Esto se debe a que

$$\mathbf{B}^{-1} \mathbf{p} = \lambda_{\text{máx}}(\mathbf{B}^{-1}) \mathbf{p}$$

implica que

$$\mathbf{B} \mathbf{p} = (1/\lambda_{\text{máx}}(\mathbf{B}^{-1})) \mathbf{p} = \lambda_{\text{mín}}(\mathbf{B}) \mathbf{p}.$$

Nótese que cada iteración del método requiere la resolución de un sistema de ecuaciones con \mathbf{B} como matriz.

Valores propios generalizados. Si queremos calcular el mínimo λ tal que

$$\mathbf{C} \mathbf{p} = \lambda \mathbf{M} \mathbf{p}$$

siendo \mathbf{M} simétrica definida positiva, podemos intentar aplicar el método anterior a la matriz $\mathbf{B} = \mathbf{M}^{-1} \mathbf{C}$. Las iteraciones para generar la sucesión de vectores son de la forma

$$\mathbf{p}_{n+1} = \mathbf{C}^{-1} \mathbf{M} \mathbf{p}_n.$$

En este caso, si $\mathbf{C}\mathbf{p} = \lambda\mathbf{M}\mathbf{p}$ y $\mathbf{p}^\top\mathbf{M}\mathbf{p} = 1$, es cuando tenemos que

$$\mathbf{p}^\top\mathbf{C}\mathbf{p} = \lambda.$$

El algoritmo resultante, va realizando normalizaciones progresivas de la sucesión de vectores, aprovechando que lo importante es su dirección, no su magnitud. Las normalizaciones se realizan para exigir que en la sucesión, todos los vectores cumplan que $\mathbf{x}_n^\top\mathbf{M}\mathbf{x}_n = 1$.

Algoritmo. La entrada del algoritmo es \mathbf{x}_0 tal que $\mathbf{x}_0^\top\mathbf{M}\mathbf{x}_0 = 1$.

```

for  $n \geq 0$ 
   $\mathbf{y}_n := \mathbf{M}\mathbf{x}_n$ 
  resolver  $\mathbf{C}\mathbf{z}_{n+1} := \mathbf{y}_n$ 
   $\xi_n := (\mathbf{z}_{n+1}^\top\mathbf{M}\mathbf{z}_{n+1})^{1/2}$ 
   $\lambda_{n+1} := (\mathbf{z}_{n+1}^\top\mathbf{y}_n)/\xi_n^2$ 
   $\mathbf{x}_{n+1} = (1/\xi_n)\mathbf{z}_{n+1}$ 
end-for

```

Uno de los dos productos por \mathbf{M} que hay que hacer en cada iteración se puede ahorrar reescribiendo el algoritmo de forma ligeramente distinta.

El caso de la condición inf-sup. Para la condición inf-sup $\mathbf{C} = \mathbf{B}^\top\mathbf{R}^{-1}\mathbf{B}$. Resolver sistemas con \mathbf{C} como matriz puede llegar a ser costoso si no se quiere calcular \mathbf{R}^{-1} . Si se emplea el método del gradiente conjugado, en cada iteración hay que resolver un sistema con \mathbf{R} como matriz, lo cual aconseja realizar una descomposición de Cholesky de \mathbf{R} si esto es factible.

7. Ejercicios propuestos

1. En un problema mixto, si a es elíptica en X_0 , demuestra que si $X_h^0 \subset X_0$, entonces la estabilidad numérica se limita a la condición inf-sup.
2. En un problema mixto donde

$$a(p, p) \geq \alpha\|p\|_X^2, \quad \forall p \in X,$$

demuestra que tanto a nivel continuo como discreto, las condiciones inf-sup son necesarias y suficientes para asegurar resolubilidad y estabilidad.

3. **Criterio de Fortin.** Supongamos que se cumple la condición inf-sup para b

$$\sup_{0 \neq p \in X} \frac{b(p, u)}{\|p\|_X} \geq \beta\|u\|_M, \quad \forall u \in M$$

y que existe $\pi_h : X \rightarrow X_h$ de forma que

$$\|\pi_h p\|_X \leq C\|p\|_X, \quad \forall p \in X,$$

y que para todo p

$$b(\pi_h p - p, v_h) = 0, \quad \forall v_h \in M_h.$$

Demuestra que se tiene la condición inf-sup discreta para el par de espacios X_h, M_h con constante $C\beta$ (emplear la tercera caracterización de la condición).

4. Dar una formulación mixta del problema

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ \partial_\nu u = g_1, & \text{en } \Gamma, \end{cases}$$

empleando u y $\mathbf{p} = \nabla u \in \mathbf{H}(\text{div}, \Omega)$ como incógnitas. La condición $\mathbf{p} \cdot \mathbf{n} = g_1$ se impondrá de forma esencial. ¿Cuál es la condición inf-sup que hay que verificar?

Lección 6

Elementos que preservan la divergencia

El capítulo precedente nos planteó el problema de encontrar espacios de elementos finitos dentro del espacio $\mathbf{H}(\text{div}, \Omega)$. Como veremos enseguida, la condición de conformidad en este espacio es menos exigente que la de espacios de tipo H^1 , lo cual permite emplear un tipo de elementos finitos algo menos ‘rígidos’. Hay que entender que la elección de espacios discretos dentro de $\mathbf{H}(\text{div}, \Omega)$ está muchas veces condicionada por la necesidad de encontrar otro espacio con el que cumpla algún tipo de condición inf-sup. Esta lección va a servir para conocer los órdenes más bajos de una familia muy empleada de elementos finitos que preservan la divergencia, los elementos de Raviart–Thomas.

1. Condición de preservación de la divergencia

Consideremos el espacio

$$\mathbf{H}(\text{div}, \Omega) := \left\{ \mathbf{p} : \Omega \rightarrow \mathbb{R}^2 \mid p_1, p_2, \text{div } \mathbf{p} \in L^2(\Omega) \right\}.$$

En el espacio $\mathbf{H}(\text{div}, \Omega)$ se puede definir la traza normal: dado un fragmento $\Gamma_0 \subset \Gamma$ de la frontera de Ω , tiene sentido hablar de

$$\mathbf{p} \cdot \mathbf{n} \quad \text{en } \Gamma_0,$$

siendo \mathbf{n} el vector normal exterior. El teorema de la divergencia se escribe para $\mathbf{p} \in \mathbf{H}(\text{div}, \Omega)$ y $u \in H^1(\Omega)$ como

$$\int_{\Omega} \mathbf{p} \cdot \nabla u + \int_{\Omega} (\text{div } \mathbf{p}) u = \int_{\Gamma} (\mathbf{p} \cdot \mathbf{n}) u.$$

El resultado principal, que afecta a cómo se debe sustituir $\mathbf{H}(\text{div}, \Omega)$ por un subespacio discreto suyo es el siguiente:

Proposición *Sea \mathcal{T}_h una triangulación de Ω en las condiciones habituales. Sea $\mathbf{p}_h : \Omega \rightarrow \mathbb{R}^2$ una función cuya restricción a cada elemento es polinómica en ambas componentes.*

Entonces

$$\mathbf{p}_h \in \mathbf{H}(\text{div}, \Omega) \iff \begin{array}{l} \mathbf{p}_h \cdot \mathbf{n} \text{ es continua} \\ \text{en los lados de los elementos} \end{array}$$

Esto implica que la condición de conformidad en $\mathbf{H}(\text{div}, \Omega)$ no exige la continuidad de los elementos finitos. Basta tener la continuidad de la componente normal en las interfases entre elementos.

Si una función polinómica a trozos ha de pertenecer al espacio $\mathbf{H}(\text{div}, \Omega)$, entonces, en cada lado L debe cumplirse que

$$\mathbf{p}_h^+ \cdot \mathbf{n} = \mathbf{p}_h^- \cdot \mathbf{n},$$

siendo \mathbf{p}_h^+ y \mathbf{p}_h^- los valores de \mathbf{p}_h en los triángulos que rodean a L .

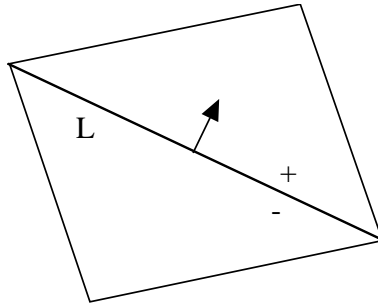


Figura 6.1: Dos triángulos contiguos definen una única normal

2. Elementos de Raviart–Thomas

Consideramos el siguiente conjunto de funciones polinómicas:

$$\mathbb{D}_1 = \left\{ \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \alpha_0 \begin{bmatrix} x \\ y \end{bmatrix} \mid \alpha_0, \alpha_1, \alpha_2 \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} \alpha_1 + \alpha_0 x \\ \alpha_2 + \alpha_0 y \end{bmatrix} \mid \alpha_0, \alpha_1, \alpha_2 \in \mathbb{R} \right\}$$

Notemos que

$$\dim \mathbb{D}_1 = 3, \quad (\mathbb{P}_0)^2 \subset \mathbb{D}_1 \subset (\mathbb{P}_1)^2$$

y que si $\mathbf{p} \in \mathbb{D}_1$, entonces su divergencia es constante:

$$\nabla \cdot \mathbf{p} = 2\alpha_0 \in \mathbb{P}_0$$

Si $\mathbf{p} \in \mathbb{D}_1$, L es un lado y \mathbf{n} es la normal sobre el lado L , entonces

$$\mathbf{p}|_L \cdot \vec{n} \in \mathbb{P}_0$$

(nótese que los puntos de L cumplen $\mathbf{x} \cdot \mathbf{n} \equiv \text{constante}$).

Grados de libertad locales. Si K es un triángulo y $\mathcal{E}(K)$ es el conjunto de los lados de K , consideramos los flujos normales

$$\phi_L(\mathbf{p}) = \text{long } L \vec{p}|_L \cdot \mathbf{n}_L = \int_L \mathbf{p} \cdot \mathbf{n}, \quad L \in \mathcal{E}(K).$$

Proposición *Los flujos normales determinan unívocamente el valor de $\mathbf{p} \in \mathbb{D}_1$.*

Demostración. Si $\phi_L(\mathbf{p}) = 0$ en los tres lados de un triángulo, entonces

$$\int_K \underbrace{\nabla \cdot \mathbf{p}}_{= 2\alpha_0} = \int_{\partial K} \mathbf{p} \cdot \mathbf{n} = 0,$$

luego $\mathbf{p} \equiv (\alpha_1, \alpha_2)$. Así, como

$$\mathbf{p} \cdot \mathbf{n}_j = 0, \quad j = 1, 2, 3$$

obligatoriamente $\mathbf{p} = 0$. □

Definimos el espacio de Raviart–Thomas:

$$X_h = \left\{ \mathbf{p}_h : \Omega \rightarrow \mathbb{R}^2 \mid \begin{array}{l} \mathbf{p}_h|_K \in \mathbb{D}_1, \quad \forall K \\ \text{flujo normal continuo en todos los lados} \end{array} \right\}.$$

El valor del flujo normal en todos los lados de la triangulación determina \mathbf{p}_h , luego

$$\dim X_h = \#\{\text{lados}\}.$$

Para trabajar con estos elementos, en cada lado de la triangulación hay que fijar una dirección normal. Una opción simple es tomar en el sentido contrario a las agujas del reloj desde el vértice de menor número global.

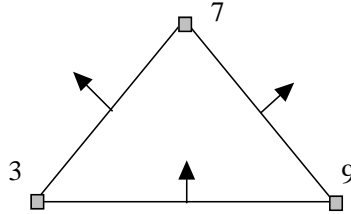


Figura 6.2: La numeración global determina los sentidos de las normales

Tomando una numeración de los lados, podemos definir para el lado L_i una función $\varphi_i \in X_h$ tal que

$$\phi_{L_j}(\varphi_i) = \delta_{ij}$$

Los soportes ocupan dos triángulos contiguos (uno si es lado exterior). Estas funciones definen una base de X_h .

Las transformaciones afines entre triángulos no preservan el conjunto \mathbb{D}_1 . Hay que hacer unos pequeños ajustes para obtener la base en un elemento cualquiera a partir de la base del elemento de referencia. Si $F_K : \hat{K} \rightarrow K$

$$F_K(\hat{\mathbf{x}}) = B_K \hat{\mathbf{x}} + \mathbf{b}_K$$

es un cambio de variable afín desde el elemento de referencia, es fácil comprobar que si $\widehat{\mathbf{q}} \in \mathbb{D}_1$, entonces

$$\mathbf{q} = \frac{1}{|\det B_K|} B_K \widehat{\mathbf{q}} \circ F_K^{-1} \in \mathbb{D}_1.$$

Así, para todo v

$$\begin{aligned} \int_{\partial K} (\mathbf{q} \cdot \mathbf{n}) v &= \int_K (\operatorname{div} \mathbf{q}) v + \int_K \mathbf{q} \cdot \nabla v \\ &= \int_{\widehat{K}} (\operatorname{div} \widehat{\mathbf{q}}) (v \circ F_K) + \int_{\widehat{K}} \widehat{\mathbf{q}} \cdot \nabla (v \circ F_K) \\ &= \int_{\partial \widehat{K}} (\widehat{\mathbf{q}} \cdot \widehat{\mathbf{n}}) (v \circ F_K). \end{aligned}$$

De allí se deduce que si \widehat{L} es un lado de \widehat{K} y L es su imagen por F_K , entonces

$$\int_{\widehat{L}} \widehat{\mathbf{q}} \cdot \widehat{\mathbf{n}} = \int_L (\mathbf{q} \cdot \mathbf{n}).$$

Por tanto, si $\widehat{\mathbf{N}}_1, \widehat{\mathbf{N}}_2, \widehat{\mathbf{N}}_3$ es la base en el elemento de referencia

$$\mathbf{N}_\alpha = \frac{1}{|\det B_K|} B_K \widehat{\mathbf{N}}_\alpha \circ F_K^{-1}, \quad \alpha = 1, 2, 3,$$

es la base en el elemento K (orientada hacia el exterior). Esta transformación de funciones vectoriales se denomina **transformación de Piola**.

3. Aplicación

Retomemos el ejemplo del capítulo anterior donde aparecía el espacio $\mathbf{H}(\operatorname{div}, \Omega)$. Se trataba de un problema mixto sobre los espacios

$$X = \mathbf{H}(\operatorname{div}, \Omega), \quad M = L^2(\Omega),$$

determinado por las formas bilineales

$$\begin{aligned} a(\mathbf{p}, \mathbf{q}) &= \int_{\Omega} \mathbf{p} \cdot \mathbf{q}, \\ b(\mathbf{p}, u) &= \int_{\Omega} (\nabla \cdot \mathbf{p}) u, \end{aligned}$$

y con datos

$$\begin{aligned} \ell(\mathbf{q}) &= \int_{\Gamma} u_0 \mathbf{q} \cdot \mathbf{n}, \\ \chi(v) &= - \int_{\Omega} f v. \end{aligned}$$

El problema tiene la forma

$$\begin{cases} \mathbf{p} \in X, & u \in M, \\ a(\mathbf{p}, \mathbf{q}) + b(\mathbf{q}, u) = \ell(\mathbf{q}), & \forall \mathbf{q} \in X, \\ b(\mathbf{p}, v) = \chi(v), & \forall v \in M. \end{cases}$$

y corresponde a una formulación del problema

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ u = u_0, & \text{en } \Gamma, \end{cases}$$

empleando las variables u y $\mathbf{p} = \nabla u$.

Como espacios discretos tomamos X_h el espacio de Raviart–Thomas explicado en la sección precedente y M_h el espacio de las funciones constantes a trozos sobre los triángulos

$$M_h = \left\{ u_h : \Omega \rightarrow \mathbb{R} \mid u_h|_K \in \mathbb{P}_0, \quad \forall K \right\}.$$

El método discreto es entonces:

$$\begin{cases} \mathbf{p}_h \in X_h, & u_h \in M_h \\ a(\mathbf{p}_h, \mathbf{q}_h) + b(\mathbf{q}_h, u_h) = \ell(\mathbf{q}_h), & \forall \mathbf{q}_h \in X_h, \\ b(\mathbf{p}_h, v_h) = \chi(v_h) & \forall v_h \in M_h. \end{cases}$$

El espacio M_h tiene una base obvia

$$\mu_i = \begin{cases} 1, & \text{en el triángulo } K_i, \\ 0, & \text{en los demás.} \end{cases}$$

Así

$$\dim X_h = \#\{\text{lados}\}, \quad \dim M_h = \#\{\text{triángulos}\}.$$

Nótese que X_h no es un producto de espacios, es decir, no se discretizan por separado las componentes de \mathbf{p} , sino conjuntamente. Para X_h tomamos como base φ_i

$$\text{long } L_j (\varphi_i|_{L_j} \cdot \mathbf{n}_j) = \delta_{ij}.$$

Numeramos los lados y los triángulos

$$N = \#\{\text{lados}\}, \quad M = \#\{\text{triángulos}\}$$

y definimos

$$\begin{aligned} a_{ij} &= a(\varphi_j, \varphi_i) = \int_{\Omega} \varphi_j \cdot \varphi_i, & i, j &= 1, \dots, N \\ b_{ij} &= b(\varphi_j, \mu_i) = \int_{K_i} \nabla \cdot \varphi_j, & i &= 1, \dots, M \\ & & j &= 1, \dots, N \end{aligned}$$

Sobre la matriz \mathbf{A} es fácil notar que $a_{ij} = 0$ si L_i y L_j no son lados de un mismo triángulo. Igualmente, $b_{ij} = 0$ si L_j no es lado de K_i . Si L_j es lado de K_i

$$b_{ij} = \int_{K_i} \nabla \cdot \varphi_j = \int_{\partial K_i} \varphi_i \cdot \mathbf{n} = \pm 1$$

donde el signo depende de la dirección tomada en el flujo de φ_j . Con las notaciones de la figura

$$b_{ij} = 1, \quad b_{\ell j} = -1.$$

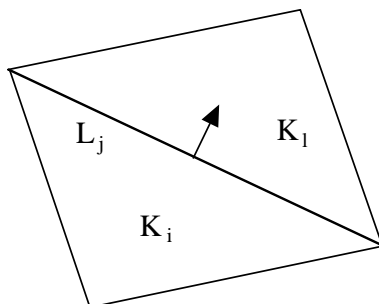


Figura 6.3: Dos triángulos contiguos y el lado común (sus numeraciones son independientes)

Ensamblado. La matriz \mathbf{B} se ensambla conociendo cuáles son los lados de cada triángulo. Se hace fila a fila (es decir, triángulo a triángulo). Por ejemplo, siguiendo la figura

$$b_{43} = -1, \quad b_{46} = -1, \quad b_{48} = 1.$$

Los requisitos para este ensamblado son:

- una lista de triángulos con sus lados,
- una lista de lados con sus extremos y con los *signos*, esto es, sabiendo si el triángulo queda o no a izquierda.

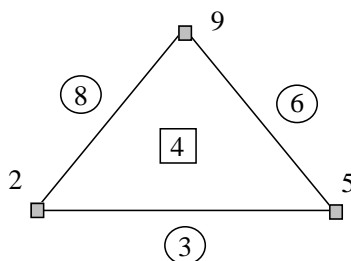


Figura 6.4: Numeraciones de triángulos, lado y vértices para la construcción de B

Para ensamblar \mathbf{A} , descomponemos como de costumbre

$$a_{ij} = \sum_K \left[\int_K \varphi_j \cdot \varphi_i \right].$$

Las funciones de base locales del triángulo K , todas con flujo exterior, son las que cumplen

$$\text{long } L_\alpha (\mathbf{N}_\beta^K \cdot \mathbf{n}_\alpha) = \delta_{\alpha\beta}, \quad \alpha, \beta = 1, 2, 3.$$

Entonces

$$\varphi_j|_K = \begin{cases} \pm \mathbf{N}_\alpha^K, & \text{si } L_j \text{ es el lado local } L_\alpha^K, \\ 0, & \text{si } L_j \text{ no es lado de } K. \end{cases}$$

Así el proceso se limita a calcular la matriz 3×3

$$\mathbf{M}_K = \left[\int_K \mathbf{N}_\alpha^K \cdot \mathbf{N}_\beta^K \right]$$

y ensamblarla contando con los cambios de signo. Cada flujo que apunte hacia el interior requiere del cambio de signo de la fila y la columna correspondiente (el elemento diagonal no cambia, por tanto, de signo). Por ejemplo, si la normal del primer lado apunta hacia el interior y las demás hacia el exterior, la transformación por realizar es

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \longmapsto \begin{bmatrix} m_{11} & -m_{12} & -m_{13} \\ -m_{21} & m_{22} & m_{23} \\ -m_{31} & m_{32} & m_{33} \end{bmatrix}.$$

Recordemos que si $\widehat{\mathbf{N}}_1, \widehat{\mathbf{N}}_2, \widehat{\mathbf{N}}_3$ es la base en el elemento de referencia, entonces

$$\mathbf{N}_\alpha = \frac{1}{|\det B_K|} B_K \widehat{\mathbf{N}}_\alpha \circ F_K^{-1}, \quad \alpha = 1, 2, 3,$$

es la base en el elemento K (orientada hacia el exterior). Para ensamblar se puede hacer lo siguiente:

$$\begin{aligned} \int_K \mathbf{N}_\alpha^K \cdot \mathbf{N}_\beta^K &= \frac{1}{|\det B_K|} \int_{\widehat{K}} (B_K^\top B_K) \widehat{\mathbf{N}}_\alpha \cdot \widehat{\mathbf{N}}_\beta \\ &= \frac{1}{|\det B_K|} \left[c_{11}^K \int_{\widehat{K}} \widehat{N}_\alpha^1 \widehat{N}_\beta^1 + c_{12}^K \int_{\widehat{K}} \widehat{N}_\alpha^1 \widehat{N}_\beta^2 + c_{12}^K \int_{\widehat{K}} \widehat{N}_\alpha^2 \widehat{N}_\beta^1 + c_{22}^K \int_{\widehat{K}} \widehat{N}_\alpha^2 \widehat{N}_\beta^2 \right] \end{aligned}$$

siendo

$$C_K = \begin{bmatrix} c_{11}^K & c_{12}^K \\ c_{21}^K & c_{22}^K \end{bmatrix} = B_K^\top B_K, \quad \widehat{\mathbf{N}}_\alpha = (\widehat{N}_\alpha^1, \widehat{N}_\alpha^2).$$

Todo se reduce al cálculo de tres matrices 3×3 en el elemento de referencia:

$$\widehat{\mathbf{M}}_{\ell m} = \int_{\widehat{K}} \widehat{N}_\alpha^\ell \widehat{N}_\beta^m, \quad \ell, m = 1, 2.$$

Resolución del sistema. Si \mathbf{A} es definida positiva (como en nuestro caso), el sistema

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$$

se puede transformar en

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ 0 & -\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} - \mathbf{B}\mathbf{A}^{-1}\mathbf{a} \end{bmatrix}$$

resolver el sistema inferior con un iterativo (cada iteración requiere resolver un sistema con \mathbf{A} como matriz) y después resolver la primera variable.

4. Ideas sobre el análisis

Para verificar que esta discretización es estable, primero hay que comprobar la condición ínfimo-supremo

$$\sup_{\mathbf{0} \neq \mathbf{q}_h \in X_h} \frac{b(\mathbf{q}_h, u_h)}{\|\mathbf{q}_h\|_{\text{div}, \Omega}} \geq \beta^* \|u_h\|_{L^2(\Omega)}, \quad \forall u_h \in M_h,$$

con $\beta^* > 0$ independiente del mallado. El criterio de Fortin dice que si existe $\pi_h : X \rightarrow X_h$ tal que

$$\begin{aligned} \|\pi_h \mathbf{q}\|_{\text{div}, \Omega} &\leq C \|\mathbf{q}\|_{\text{div}, \Omega}, \\ b(\pi_h \mathbf{q}, v_h) &= b(\mathbf{q}, v_h), \quad \forall v_h \in M_h, \end{aligned}$$

entonces se cumple la condición anterior ínfimo supremo. Así pues, basta con construir π_h en las condiciones anteriores. La segunda condición implica que

$$\int_K \nabla \cdot (\pi_h \mathbf{q}) = \int_K \nabla \cdot \mathbf{q}, \quad \forall K,$$

es decir,

$$\int_{\partial K} (\pi_h \mathbf{q}) \cdot \mathbf{n} = \int_{\partial K} \mathbf{q} \cdot \mathbf{n}, \quad \forall K.$$

Para ello básicamente se toman los flujos lado a lado

$$\int_L \mathbf{q} \cdot \mathbf{n}$$

y se construye $\pi_h \mathbf{q}$ interpolándolos.

Lo segundo que hay que hacer es identificar el conjunto

$$X_h^0 = \left\{ \mathbf{p}_h \in X_h \mid b(\mathbf{p}_h, v_h) = 0, \quad \forall v_h \in M_h \right\}.$$

Sea $\mathbf{p}_h \in X_h$. Entonces

$$\begin{aligned} \mathbf{p}_h \in X_h^0 &\iff b(\mathbf{p}_h, v_h) = 0, & \forall v_h \in M_h \\ &\iff \int_K \underbrace{\nabla \cdot \mathbf{p}_h}_{= \text{cte}} = 0, & \forall K \\ &\iff \nabla \cdot \mathbf{p}_h = 0, \end{aligned}$$

luego en este caso

$$X_h^0 \subset X_0.$$

Así, la propiedad

$$a(\mathbf{q}_h, \mathbf{q}_h) \geq \alpha^* \|\mathbf{q}_h\|_{\text{div}, \Omega}^2, \quad \forall \mathbf{q}_h \in X_h^0,$$

se hereda del caso continuo.

En consecuencia, el método discreto da un sistema con solución única y se tiene la estimación

$$\|\mathbf{p} - \mathbf{p}_h\|_{\text{div}, \Omega} + \|u - u_h\|_{0, \Omega} \leq C \left[\inf_{\mathbf{q}_h \in X_h} \|\mathbf{p} - \mathbf{q}_h\|_{\text{div}, \Omega} + \inf_{v_h \in M_h} \|u - v_h\|_{0, \Omega} \right] = \mathcal{O}(h).$$

en las condiciones habituales.

5. Otras condiciones de contorno

Cuando el problema original tiene condiciones mixtas (parte Dirichlet, parte Neumann),

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ u|_{\Gamma_D} = u_0, & \partial_n u|_{\Gamma_N} = g, \end{cases}$$

la formulación mixta es algo más complicada. Tomamos $\mathbf{p} = \nabla u$ como incógnita adicional y llegamos a

$$\begin{cases} \mathbf{p} - \nabla u = 0, & \text{en } \Omega, \\ \nabla \cdot \mathbf{p} = -f, & \text{en } \Omega, \\ u|_{\Gamma_D} = u_0, & \mathbf{p} \cdot \mathbf{n}|_{\Gamma_N} = g. \end{cases}$$

Además de $X = \mathbf{H}(\text{div}, \Omega)$ y $M = L^2(\Omega)$, hay que considerar el espacio:

$$\tilde{X} = \left\{ \mathbf{q} \in \mathbf{H}(\text{div}, \Omega) \mid \mathbf{q} \cdot \mathbf{n}|_{\Gamma_N} = 0 \right\}$$

La formulación mixta es entonces

$$\begin{cases} \mathbf{p} \in X, u \in M, \\ \mathbf{p} \cdot \mathbf{n} = g, & \text{en } \Gamma_N \\ \int_{\Omega} \mathbf{p} \cdot \mathbf{q} + \int_{\Omega} (\nabla \cdot \mathbf{q}) u = \int_{\Gamma_D} u_0 (\mathbf{q} \cdot \mathbf{n}), & \forall \mathbf{q} \in \tilde{X}, \\ \int_{\Omega} (\nabla \cdot \mathbf{p}) v = - \int_{\Omega} f v, & \forall v \in M. \end{cases}$$

Las hipótesis sobre el problema mixto, tanto a nivel exacto como de discretización se realizan sobre el par formado por los espacios \tilde{X} y M .

Para discretizar, empleamos el espacio de Raviart–Thomas anterior X_h y

$$\tilde{X}_h = \left\{ \mathbf{p}_h \in X_h \mid \mathbf{p}_h \cdot \mathbf{n}|_L = 0, \quad \forall L \text{ lado Neumann} \right\} \subset \tilde{X}.$$

M_h vuelve a ser el espacio de las funciones constantes a trozos sobre triángulos. Además, hemos de considerar los valores g_i , obtenidos promediando g en el lado Neumann L_i (o evaluando g en el punto medio de ese lado).

$$\begin{cases} \mathbf{p}_h \in X_h, & u_h \in M_h, \\ \mathbf{p}_h \cdot \mathbf{n}|_{L_i} = g_i, & \forall L_i \text{ lado Neumann,} \\ a(\mathbf{p}_h, \mathbf{q}_h) + b(\mathbf{q}_h, u_h) = \ell(\mathbf{q}_h), & \forall \mathbf{q}_h \in \tilde{X}_h, \\ b(\mathbf{p}_h, v_h) = \chi(v_h) & \forall v_h \in M_h \end{cases}$$

A la hora de imponer el primer bloque de condiciones de contorno, lo que se hace es ensamblar el sistema completo sin incorporar las condiciones de Neumann

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & 0 \end{bmatrix}.$$

Luego se ignoran las filas del primer bloque correspondientes a lados Neumann y se dan valores a las incógnitas correspondientes, quedando un sistema con idéntica estructura.

Las incógnitas:

$$\mathbf{p}_h = \sum_j p_j \boldsymbol{\varphi}_j, \quad u_h = \sum_j u_j \mu_j$$

donde p_j es el flujo total normal en el lado L_j (dirección prefijada) y u_j es un valor constante por triángulo. Si L_i es lado Dirichlet

$$\int_{\Gamma_D} u_0 \boldsymbol{\varphi}_i \cdot \mathbf{n} = \frac{1}{\text{long } L_i} \int_{L_i} u_0$$

luego el ensamblado de las condiciones Dirichlet (que ahora son naturales) es inmediato con estos elementos finitos.

6. Elemento de Raviart–Thomas de orden dos

En cada triángulo consideramos ahora el espacio \mathbb{D}_2 de polinomios de la forma

$$\underbrace{\begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + x \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} + y \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix}}_{\mathbf{p}_1} + \underbrace{(\gamma_1 x + \gamma_2 y)}_{p_1^0} \begin{bmatrix} x \\ y \end{bmatrix}$$

de forma que

$$(\mathbb{P}_1)^2 \subset \mathbb{D}_2 \subset (\mathbb{P}_2)^2, \quad \dim \mathbb{D}_2 = 8.$$

Descomponiendo $\mathbf{p} \in \mathbb{D}_2$ en la forma

$$\mathbf{p} = \mathbf{p}_1 + p_1^0 \mathbf{x}, \quad \begin{aligned} \mathbf{p}_1 &\in (\mathbb{P}_1)^2 \\ p_1^0(x, y) &= \gamma_1 x + \gamma_2 y \end{aligned}$$

se tienen:

- La divergencia de \mathbf{p} es un polinomio de grado menor o igual que uno:

$$\nabla \cdot \mathbf{p} = \nabla \cdot \mathbf{p}_1 + 3p_1^0 = (\alpha_1 + \beta_2) + 3(\gamma_1 x + \gamma_2 y)$$

- $\nabla \cdot \mathbf{p} = 0$ equivale a

$$\mathbf{p} = \mathbf{p}_1, \quad \nabla \cdot \mathbf{p}_1 = 0.$$

- En cada lado $\mathbf{p} \cdot \mathbf{n}|_L \in \mathbb{P}_1(t)$.

Proposición Sea K un triángulo. Los siguientes grados de libertad

$$\int_K \mathbf{p} \in \mathbb{R}^2, \quad \mathbf{p} \cdot \mathbf{n}|_L, \quad L \in \mathcal{E}(K)$$

determinan $\mathbf{p} \in \mathbb{D}_2$.

En el problema modelo que estamos estudiando un buen par es

$$\begin{aligned} X_h &= \left\{ \mathbf{q}_h \in \mathbf{H}(\text{div}, \Omega) \mid \mathbf{q}_h|_K \in \mathbb{D}_2, \forall K \right\} \\ M_h &= \left\{ v_h : \Omega \rightarrow \mathbb{R} \mid v_h|_K \in \mathbb{P}_1, \forall K \right\}. \end{aligned}$$

De esta forma

$$\begin{aligned} \dim X_h &= 2\# \{\text{triángulos}\} + 2\# \{\text{lados}\} \\ \dim M_h &= 3\# \{\text{triángulos}\} \end{aligned}$$

y el orden de convergencia es $\mathcal{O}(h^2)$.

7. Ejercicios propuestos

1. Si $\mathbf{p}_h \in X_h$

$$\mathbf{p} = \mathbf{p}_h|_K = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \alpha_0 \begin{bmatrix} x \\ y \end{bmatrix},$$

demuestra que entonces

$$\alpha_0 = \frac{1}{2 \text{área } K} \sum_{L \in \mathcal{E}(K)} \phi_L(\mathbf{p}).$$

2. Calcula las funciones de base del elemento de Raviart–Thomas de orden uno sobre el de referencia.
3. Demuestra que los grados de libertad del elemento Raviart–Thomas de orden dos determinan unívocamente el valor del elemento. Para ello, haz los siguientes pasos.

a) Si los grados de libertad se anulan, de la identidad

$$\int_K (\nabla \cdot \mathbf{p})^2 = - \int_{\partial K} (\nabla \cdot \mathbf{p}) (\mathbf{p} \cdot \mathbf{n})$$

deduce que $\mathbf{p} \in (\mathbb{P}_1)^2$.

b) Terminan de deducir que $\mathbf{p} = \mathbf{0}$.

4. Para el elemento de Raviart–Thomas de orden dos, ¿cómo se prueba que los dos grados de libertad de un lado determinan el flujo normal en el lado?
5. Define el elemento Raviart–Thomas de grado mínimo sobre un tetraedro.
6. Averigua cuál es el elemento más simple que preserva la divergencia sobre un mallado de paralelogramos.
7. Para resolver sistemas con la estructura general

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$$

donde \mathbf{A} es simétrica definida positiva, hay un método iterativo llamado **algoritmo de Uzawa**. Averigua en qué consiste.

Lección 7

Formulaciones mixtas en elasticidad

En este capítulo voy a dar unas breves pinceladas sobre formulaciones mixtas de las ecuaciones de la elasticidad lineal. Este es un tema en el que se sigue trabajando y donde la última respuesta está por ser alcanzada. La formulación primal, en la variable desplazamiento, del problema de la elasticidad lineal se establece por medio de las ecuaciones de Navier o de Lamé. Estas ecuaciones tienen un acusado paralelismo con la ecuación de Laplace. Por contra, las formulaciones mixtas de estas ecuaciones presentan numerosas dificultades adicionales por culpa de que las nuevas variables introducidas (tensores de tensiones o de deformaciones) forman matrices simétricas. En este capítulo sigo bastante de cerca el Capítulo VI.3 del libro de Braess que menciono en la bibliografía.

1. Preliminares

Recordemos brevemente que las magnitudes en elasticidad lineal son tres, todas ellas definidas sobre la configuración de referencia $\Omega \subset \mathbb{R}^d$ ($d = 2$ ó 3), esto es, sobre la posición del sólido sin deformar:

- el campo de desplazamientos $\mathbf{u} : \Omega \rightarrow \mathbb{R}^d$,
- el tensor de deformaciones $\boldsymbol{\varepsilon} : \Omega \rightarrow \mathbb{R}_{\text{sim}}^{d \times d}$
- el tensor de tensiones $\boldsymbol{\sigma} : \Omega \rightarrow \mathbb{R}_{\text{sim}}^{d \times d}$.

Bajo la hipótesis de pequeñas deformaciones (primera hipótesis de la elasticidad lineal), las deformaciones se deducen de los desplazamientos por la regla

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u}) := \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^\top).$$

Para materiales con comportamiento lineal (segunda hipótesis de la elasticidad lineal), las tensiones son función lineal de las deformaciones. En el caso de un sólido lineal homogéneo e isótropo, esta relación viene dada por la Ley de Hooke. En el caso tridimensional, esta relación es:

$$\begin{aligned} \boldsymbol{\sigma} &= \frac{E}{1 + \nu} \left(\boldsymbol{\varepsilon} + \frac{\nu}{1 - 2\nu} (\text{tr} \boldsymbol{\varepsilon}) I_3 \right) \\ &= \frac{E}{1 + \nu} \boldsymbol{\varepsilon} + \frac{E \nu}{(1 + \nu)(1 - 2\nu)} (\text{tr} \boldsymbol{\varepsilon}) I_3 = 2\mu \boldsymbol{\varepsilon} + \lambda (\text{tr} \boldsymbol{\varepsilon}) I_3. \end{aligned}$$

A veces puede resultar más cómodo desplegar las matrices (tensores) de tensiones y deformaciones en forma de vectores. Para el caso tridimensional sólo hay seis componentes, ya que los tensores son simétricos:

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_{23} \end{bmatrix} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu & & & \\ \nu & 1-\nu & \nu & & & \\ \nu & \nu & 1-\nu & & & \\ & & & 1-2\nu & & \\ & & & & 1-2\nu & \\ & & & & & 1-2\nu \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{23} \end{bmatrix}.$$

Escribiendo

$$E_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

se puede simplificar la escritura de la matriz anterior a poner

$$\mathcal{C} := \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} (1-2\nu)I_3 + \nu E_3 & \\ & (1-2\nu)I_3 \end{bmatrix} = \begin{bmatrix} 2\mu I_3 + \lambda E_3 & \\ & 2\mu I_3 \end{bmatrix}.$$

Las ecuaciones de equilibrio son

$$\operatorname{div} \boldsymbol{\sigma} + \mathbf{f} = \mathbf{0}, \quad \text{en } \Omega.$$

La frontera Γ se divide en dos partes:

- la frontera de desplazamientos dados (frontera Dirichlet) Γ_D ,
- la frontera de tracciones dadas (frontera Neumann), Γ_N , que es su complementario.

Las condiciones de contorno son

$$\mathbf{u} = \mathbf{u}_0, \quad \text{en } \Gamma_D, \quad \boldsymbol{\sigma} \mathbf{n} = \mathbf{t}_0, \quad \text{en } \Gamma_N,$$

donde \mathbf{n} es el campo de vectores normales exteriores sobre la frontera y $\mathbf{u}_0, \mathbf{t}_0$ son datos.

La formulación primal se realiza en la variable desplazamiento

$$\begin{cases} \mathbf{u} \in \mathbf{H}^1(\Omega), \\ \mathbf{u} = \mathbf{u}_0, & \text{en } \Gamma_D, \\ \int_{\Omega} \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_N} \mathbf{t}_0 \cdot \mathbf{v}, & \forall \mathbf{v} \in \mathbf{H}_{\Gamma_D}^1(\Omega), \end{cases}$$

siendo

$$\mathbf{H}_{\Gamma_D}^1(\Omega) := \{\mathbf{v} \in \mathbf{H}^1(\Omega) \mid \mathbf{v} = \mathbf{0}, \quad \text{en } \Gamma_D\}.$$

Desde el punto de vista de esta formulación, sólo el problema de tracciones puras (el problema de Neumann, cuando no hay frontera de desplazamientos dados) presenta problemas, ya que no tiene solución única, puesto que los movimientos rígidos infinitesimales

$$\mathbf{u}_{\text{rig}}(\mathbf{x}) = \mathbf{a} + \mathbf{b} \times \mathbf{x}, \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^d,$$

tienen deformación nula, luego no son tenidos en cuenta por el tensor de tensiones ni por la condición de contorno de tracción normal dada.

2. Formulación de Hellinger–Reissner (I)

La primera de las formulaciones de Hellinger–Reissner calcula simultáneamente los desplazamientos y las tensiones, buscando estas últimas sobre un espacio de tipo H^1 . Recordemos que la hipótesis de conformidad de un espacio de tipo H^1 es la continuidad. Las tensiones se buscan en el espacio

$$\mathbb{L}_{\text{sim}}^2(\Omega) := \{ \boldsymbol{\tau} : \Omega \rightarrow \mathbb{R}_{\text{sim}}^{3 \times 3} \mid \tau_{ij} \in L^2(\Omega), \quad \forall i, j \},$$

mientras que los desplazamientos se buscan en $\mathbf{H}_{\Gamma_D}^1(\Omega)$.

La formulación parte simplemente de la expresión

$$\mathcal{C}^{-1} \boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\mathbf{u})$$

(\mathcal{C}^{-1} es la inversa de la matriz \mathcal{C} si vemos $\boldsymbol{\sigma}$ y $\boldsymbol{\varepsilon}$ como vectores desplegados, o bien, la expresión inversa de la ley de Hooke). Esta expresión debe ser testada únicamente por matrices simétricas, ya que ambos lados de la ecuación son simétricos. La segunda ecuación de la formulación es simplemente la formulación primal, pero tomando $\boldsymbol{\sigma}$ directamente como incógnita.

Las ecuaciones son:

$$\left[\begin{array}{l} \boldsymbol{\sigma} \in \mathbb{L}_{\text{sim}}^2(\Omega), \quad \mathbf{u} \in \mathbf{H}^1(\Omega), \\ \mathbf{u} = \mathbf{u}_0, \quad \text{en } \Gamma_D, \\ \int_{\Omega} \mathcal{C}^{-1} \boldsymbol{\sigma} : \boldsymbol{\tau} - \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\tau} = 0, \quad \forall \boldsymbol{\tau} \in \mathbb{L}_{\text{sim}}^2(\Omega), \\ - \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{v}) = - \int_{\Omega} \mathbf{f} \cdot \mathbf{v} - \int_{\Gamma_N} \mathbf{t}_0 \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{H}_{\Gamma_D}^1(\Omega). \end{array} \right.$$

La forma bilineal diagonal es elíptica en todo el espacio

$$a(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \geq C \|\boldsymbol{\sigma}\|_{0,\Omega}^2, \quad \forall \boldsymbol{\sigma} \in \mathbb{L}_{\text{sim}}^2(\Omega),$$

por lo que podemos despreocuparnos de ella a nivel continuo y discreto. La forma bilineal

$$b(\boldsymbol{\sigma}, \mathbf{v}) := \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{v})$$

(el signo es irrelevante) cumple la condición inf–sup siempre que haya una parte de condición de contorno de desplazamientos dados (como en la formulación primal). Esto se debe a que, por la desigualdad de Korn, se tiene

$$\|\mathbf{v}\|_{1,\Omega}^2 \leq C \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{H}_{\Gamma_D}^1(\Omega).$$

Entonces, dado \mathbf{v} , basta tomar $\boldsymbol{\sigma} = \boldsymbol{\varepsilon}(\mathbf{v}) \in \mathbb{L}_{\text{sim}}^2(\Omega)$, de modo que

$$\|\mathbf{v}\|_{1,\Omega}^2 \leq C \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{v}) = b(\boldsymbol{\sigma}, \mathbf{v}).$$

Como por elección de $\boldsymbol{\sigma}$

$$\|\boldsymbol{\sigma}\|_{0,\Omega} \leq C' \|\mathbf{v}\|_{1,\Omega},$$

la condición inf–sup se deduce fácilmente.

Requisitos para la discretización. El espacio que empleemos para discretizar la variable $\boldsymbol{\sigma}$, digamos $\mathbb{L}_{\text{sim},h}^2$ puede contener elementos discontinuos. En cambio el espacio necesario para discretizar la variable \mathbf{u} , \mathbf{H}_h^1 , debe ser de elementos finitos continuos. La condición inf–sup discreta que hay que verificar es

$$\sup_{\mathbf{0} \neq \boldsymbol{\sigma}_h \in \mathbb{L}_{\text{sim},h}^2} \frac{1}{\|\boldsymbol{\sigma}_h\|_{0,\Omega}} \int_{\Omega} \boldsymbol{\sigma}_h : \boldsymbol{\varepsilon}(\mathbf{v}_h) \geq \beta \|\mathbf{v}_h\|_{1,\Omega}, \quad \forall \mathbf{v}_h \in \mathbf{H}_h^1, \quad \mathbf{v}_h = \mathbf{0}, \quad \text{en } \Gamma_D.$$

Viendo el análisis realizado en continuo es fácil percatarse de que basta con que $\boldsymbol{\varepsilon}(\mathbf{v}_h) \in \mathbb{L}_{\text{sim},h}^2$ para todo $\mathbf{v}_h \in \mathbf{H}_h^1$, para que se tenga la condición ínfimo–supremo. Por tanto, podemos escoger simplemente:

- un espacio de elementos \mathbb{P}_k continuos sobre una triangulación para cada componente de \mathbf{u}_h
- un espacio de elementos \mathbb{P}_{k-1} discontinuos sobre la misma triangulación para las componentes matriciales de $\boldsymbol{\sigma}_h$ (exigiendo la simetría del tensor).

El inconveniente principal que se le puede poner a esta formulación es la discontinuidad de la aproximación del tensor de tensiones obtenido y el gran número de grados de libertad que requiere.

Por otra parte, si se emplean elementos discontinuos para $\boldsymbol{\sigma}_h$ el bloque diagonal es diagonal por bloques, ya que no hay conexión entre los elementos. Esto permite trasladar fácilmente la resolución a las variables desplazamiento.

3. Formulación de Hellinger–Reissner (II)

La segunda formulación de Hellinger–Reissner realiza integración por partes (aplica el teorema de la divergencia) en la fórmula

$$\mathcal{C}^{-1} \boldsymbol{\sigma} - \boldsymbol{\varepsilon}(\mathbf{u}).$$

Formalmente, tenemos

$$\int_{\Omega} \mathcal{C}^{-1} \boldsymbol{\sigma} : \boldsymbol{\tau} + \int_{\Omega} (\operatorname{div} \boldsymbol{\tau}) \cdot \mathbf{u} = \int_{\Gamma_D} (\boldsymbol{\tau} \mathbf{n}) \cdot \mathbf{u}_0, \quad \boldsymbol{\tau} \mathbf{n} = \mathbf{0}, \quad \text{en } \Gamma_N.$$

Como ocurre en la formulación mixta del laplaciano (flujo de Darcy), la condición de Dirichlet (desplazamientos) se convierte en natural y la condición de Neumann (tracciones normales) pasa a ser esencial. La segunda ecuación será simplemente un test de la ecuación de equilibrio

$$\operatorname{div} \boldsymbol{\sigma} = -\mathbf{f}.$$

El espacio para buscar las tensiones es

$$\mathbb{H}_{\text{sim}}(\operatorname{div}, \Omega) := \{ \boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}, \Omega) \mid \boldsymbol{\tau} = \boldsymbol{\tau}^{\top} \},$$

mientras que los desplazamientos se buscan en $\mathbf{L}^2(\Omega)$. La formulación es

$$\left[\begin{array}{l} \boldsymbol{\sigma} \in \mathbb{H}_{\text{sim}}(\operatorname{div}, \Omega), \quad \mathbf{u} \in \mathbf{L}^2(\Omega), \\ \boldsymbol{\sigma} \mathbf{n} = \mathbf{t}_0, \quad \text{en } \Gamma_N, \\ \int_{\Omega} \mathcal{C}^{-1} \boldsymbol{\sigma} : \boldsymbol{\tau} + \int_{\Omega} (\operatorname{div} \boldsymbol{\tau}) \cdot \mathbf{u} = \int_{\Gamma_D} (\boldsymbol{\tau} \mathbf{n}) \cdot \mathbf{u}_0, \quad \forall \boldsymbol{\tau} \in \mathbb{H}_{\text{sim}, \Gamma_N}(\operatorname{div}, \Omega), \\ \int_{\Omega} (\operatorname{div} \boldsymbol{\sigma}) \cdot \mathbf{v} = - \int_{\Omega} \mathbf{f} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{L}^2(\Omega), \end{array} \right.$$

siendo

$$\mathbb{H}_{\text{sim}, \Gamma_N}(\operatorname{div}, \Omega) := \{ \boldsymbol{\tau} \in \mathbb{H}_{\text{sim}}(\operatorname{div}, \Omega) \mid \boldsymbol{\tau} \mathbf{n} = \mathbf{0}, \quad \text{en } \Gamma_N \}.$$

En el estudio de este problema *hay que separar el caso del problema de Dirichlet*, que introduce alguna dificultad adicional.

Estudio teórico de la formulación. Como de costumbre, comenzamos examinando la condición inf-sup para la forma bilineal

$$\int_{\Omega} (\operatorname{div} \boldsymbol{\sigma}) \cdot \mathbf{v}.$$

Dada $\mathbf{v} \in \mathbf{L}^2(\Omega)$, resolvemos el problema

$$\left[\begin{array}{l} \mathbf{u} \in \mathbf{H}^1(\Omega), \\ \mathbf{u} = \mathbf{0}, \quad \text{en } \Gamma_D, \\ (\mathcal{C} \boldsymbol{\varepsilon}(\mathbf{u})) \mathbf{n} = \mathbf{0}, \quad \text{en } \Gamma_N, \\ \operatorname{div} \mathcal{C} \boldsymbol{\varepsilon}(\mathbf{u}) = \mathbf{v}, \quad \text{en } \Omega, \end{array} \right.$$

y tomamos $\boldsymbol{\sigma} = \mathcal{C} \boldsymbol{\varepsilon}(\mathbf{u})$ de manera que

$$\|\boldsymbol{\sigma}\|_{0, \Omega} \leq C \|\mathbf{u}\|_{1, \Omega} \leq C' \|\mathbf{v}\|_{0, \Omega}$$

y

$$\int_{\Omega} (\operatorname{div} \boldsymbol{\sigma}) \cdot \mathbf{v} = \int_{\Omega} |\mathbf{v}|^2,$$

lo cual basta para mostrar la condición inf-sup. El núcleo de esta forma bilineal es

$$\{\boldsymbol{\tau} \in \mathbb{H}_{\text{sim}, \Gamma_N}(\operatorname{div}, \Omega) \mid \operatorname{div} \boldsymbol{\tau} = \mathbf{0}\} \subset \{\boldsymbol{\tau} \in \mathbb{H}(\operatorname{div}, \Omega) \mid \operatorname{div} \boldsymbol{\tau} = \mathbf{0}\},$$

y sobre el segundo espacio la forma bilineal

$$\int_{\Omega} \mathcal{C}^{-1} \boldsymbol{\sigma} : \boldsymbol{\tau}$$

es trivialmente elíptica.

Desde el punto de vista discreto los desplazamientos se pueden buscar en un espacio de elementos discontinuos. Las columnas de la matriz de tensiones deben pertenecer a un espacio que respete la conformidad de la divergencia, pero la matriz resultante debe ser simétrica, lo cual introduce la dificultad de que no podemos emplear elementos como los de Raviart–Thomas para las columnas de forma independiente, ya que en estos elementos las componentes son distintas y no se va a cumplir la condición de simetría.

4. Formulaciones de Hu–Washizu

Al igual que en el caso de la formulación de Hellinger–Reissner hay dos opciones, según se transfieran propiedades de derivabilidad al tensor de tensiones (que pasa a estar en $\mathbf{H}(\operatorname{div})$, mientras que los desplazamientos pueden ser discontinuos) o no. En esta formulación son incógnitas los tres campos: deformaciones, desplazamientos y tensiones. En las dos opciones que encontraremos, $\boldsymbol{\sigma}$ hará el rol de ‘multiplicador’.

Formulación de Hu–Washizu (I). Las ecuaciones en forma fuerte son:

- la ecuación de estado (Ley de Hooke en el caso homogéneo e isótropo)

$$\mathcal{C} \boldsymbol{\varepsilon} = \boldsymbol{\sigma},$$

que sólo tiene que ser testada por matrices simétricas,

- la ecuación de equilibrio

$$\operatorname{div} \boldsymbol{\sigma} + \mathbf{f}$$

puesta en forma débil (como en la forma primal), de manera que incorpora la condición de tracciones dadas

$$\int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_N} \mathbf{t}_0 \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{H}_{\Gamma_D}^1(\Omega),$$

- la ecuación que liga deformaciones a la parte simétrica del gradiente

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u}).$$

La condición de Dirichlet (desplazamientos) se mantiene como esencial y se impone aparte. Notemos el abuso de notación que supone llamar $\boldsymbol{\varepsilon}$ a la incógnita campo de deformaciones y al operador

$$\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^\top).$$

Las ecuaciones son

$$\left[\begin{array}{l} \boldsymbol{\varepsilon} \in \mathbb{L}_{\text{sim}}^2(\Omega), \quad \mathbf{u} \in \mathbf{H}^1(\Omega), \quad \boldsymbol{\sigma} \in \mathbb{L}_{\text{sim}}^2(\Omega), \\ \mathbf{u} = \mathbf{u}_0, \quad \text{en } \Gamma_D, \\ \int_{\Omega} \boldsymbol{\varepsilon} : \mathcal{C} \boldsymbol{\eta} \quad - \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\eta} = 0, \quad \forall \boldsymbol{\eta} \in \mathbb{L}_{\text{sim}}^2(\Omega), \\ \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_N} \mathbf{t}_0 \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{H}_{\Gamma_D}^1(\Omega), \\ - \int_{\Omega} \boldsymbol{\varepsilon} : \boldsymbol{\tau} + \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\tau} = 0, \quad \forall \boldsymbol{\tau} \in \mathbb{L}_{\text{sim}}^2(\Omega). \end{array} \right.$$

Vamos rápidamente a realizar una serie de observaciones elementales.

- En principio la utilización de ambas variables $\boldsymbol{\sigma}$ y $\boldsymbol{\varepsilon}$ puede parecer una reiteración innecesaria de información. Cuando el material está regido por la Ley de Hooke, realmente es así. Dependiendo de la elección del espacio para discretizar estas dos variables (si es el mismo y tiene determinado tipo de estructura muy elemental), tendremos simplemente $\boldsymbol{\sigma}_h = \mathcal{C} \boldsymbol{\varepsilon}_h$ y nos podemos ahorrar la segunda variable.
- En cambio, cuando el material es más complejo (no es isótropo o no es homogéneo), obtener la ley inversa que permite calcular $\boldsymbol{\varepsilon}$ en función de $\boldsymbol{\sigma}$ no es práctico. Esa función es realizada a nivel discreto por la primera ecuación.
- El requisito de simetría para las variables matriciales es innecesario. Cualquier solución de las ecuaciones tiene a $\boldsymbol{\varepsilon}$ simétrico (ver la tercera ecuación) y, por consiguiente, también a $\boldsymbol{\sigma}$ (ver la primera ecuación). Sin embargo, exigir la simetría en esta formulación reduce el número de grados de libertad.

Aspectos teóricos sobre la formulación. La forma bilineal que impone la restricción es ahora

$$b((\boldsymbol{\varepsilon}, \mathbf{u}), \boldsymbol{\tau}) = - \int_{\Omega} \boldsymbol{\varepsilon} : \boldsymbol{\tau} + \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\tau}.$$

Para demostrar la condición inf-sup basta hacer lo siguiente: dado $\boldsymbol{\tau}$ tomar $\boldsymbol{\varepsilon} = -\boldsymbol{\tau}$ y $\mathbf{u} = \mathbf{0}$, ya que entonces

$$b((-\boldsymbol{\tau}, \mathbf{0}), \boldsymbol{\tau}) = \int_{\Omega} \boldsymbol{\tau} : \boldsymbol{\tau} = \|\boldsymbol{\tau}\|_{0,\Omega}^2$$

y

$$\|\boldsymbol{\tau}\|_{0,\Omega} = \|\boldsymbol{\varepsilon}\|_{0,\Omega} = \|\boldsymbol{\varepsilon}\|_{0,\Omega} + \|\mathbf{u}\|_{1,\Omega}.$$

El núcleo de esta restricción es el conjunto

$$\{(\boldsymbol{\varepsilon}, \mathbf{u}) \in \mathbb{L}_{\text{sim}}^2(\Omega) \times \mathbf{H}_{\Gamma_D}^1(\Omega) \mid \boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u})\},$$

sobre el que se tiene elipticidad de la forma bilineal diagonal, ya que si $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u})$, entonces

$$a((\boldsymbol{\varepsilon}, \mathbf{u}), (\boldsymbol{\varepsilon}, \mathbf{u})) = \int_{\Omega} \boldsymbol{\varepsilon} : \mathcal{C}\boldsymbol{\varepsilon} = \frac{1}{2} \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{u}) : \mathcal{C}\boldsymbol{\varepsilon}(\mathbf{u}) + \frac{1}{2} \int_{\Omega} \boldsymbol{\varepsilon} : \mathcal{C}\boldsymbol{\varepsilon}$$

y el resultado se sigue de la desigualdad de Korn.

Formulación de Hu–Washizu (II). La diferencia principal con la anterior formulación estriba en que en la ecuación de equilibrio no se aplica el teorema de la divergencia, de manera que sí hay que aplicarlo en la relación entre desplazamientos y deformaciones. La condición de Dirichlet pasa a ser natural y la de Neumann esencial. Las ecuaciones se pueden escribir

$$\left[\begin{array}{l} \boldsymbol{\varepsilon} \in \mathbb{L}_{(\text{sim})}^2(\Omega), \quad \mathbf{u} \in \mathbf{L}^2(\Omega), \quad \boldsymbol{\sigma} \in \mathbb{H}_{\text{sim}, \Gamma_N}(\text{div}, \Omega), \\ \boldsymbol{\sigma} \mathbf{n} = \mathbf{t}_0, \quad \text{en } \Gamma_N, \\ \int_{\Omega} \boldsymbol{\varepsilon} : \mathcal{C}\boldsymbol{\eta} \quad - \int_{\Omega} \boldsymbol{\sigma} : \boldsymbol{\eta} = 0, \quad \forall \boldsymbol{\eta} \in \mathbb{L}_{(\text{sim})}^2(\Omega), \\ \int_{\Omega} \text{div} \boldsymbol{\sigma} \cdot \mathbf{v} = - \int_{\Omega} \mathbf{f} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{L}^2(\Omega), \\ \int_{\Omega} \boldsymbol{\varepsilon} : \boldsymbol{\tau} \quad + \int_{\Omega} \mathbf{u} \cdot \text{div} \boldsymbol{\tau} = \int_{\Gamma_D} \mathbf{u}_0 \cdot \boldsymbol{\tau} \mathbf{n}, \quad \forall \boldsymbol{\tau} \in \mathbb{H}_{\text{sim}, \Gamma_N}(\Omega). \end{array} \right.$$

(En principio parece ser que los signos no encajan para una formulación mixta, pero basta cambiar de signo la segunda y tercera ecuaciones para que sí lo hagan).

5. Formulación–elemento PEERS

6. Ejercicios propuestos

1. Propón espacios discretos para la formulación de Hu–Washizu (I), de manera que se tenga estabilidad del esquema numérico.

Lección 8

Estimación a posteriori del error

El objetivo genérico de la **estimación a posteriori** del error es la búsqueda de una cantidad computable económicamente a partir de la solución numérica obtenida y que sirva para dar una cota superior del error, sin alejarse demasiado del error cometido. En esta lección final doy unas pinceladas de unos cuantos métodos de estimación a posteriori del error e intentaremos ver qué se puede esperar de un estimador. Este capítulo sigue bastante de cerca el primer capítulo del libro de Verfürth citado en la bibliografía.

1. Objetivos y terminología

Si la solución de un problema mediante elementos finitos es u_h , u es la solución exacta y la norma natural del problema (aquella en la que hemos obtenido estimaciones a priori) es $\|\cdot\|$, entonces el estimador a posteriori del error es cualquier magnitud

$$\text{Est}_h := \text{Est}(u_h, \text{datos})$$

tal que existan dos constantes, independientes de la solución y de los datos, de forma que

$$C_1 \text{Est}_h \leq \|u - u_h\| \leq C_2 \text{Est}_h.$$

La segunda desigualdad es la más importante: se denomina **fiabilidad** (reliability) del estimador e indica que el estimador acota superiormente al error. La primera desigualdad (el estimador acota inferiormente al error) se denomina **eficiencia** del estimador e implica la imposibilidad de que el estimador dé por mala una solución bien calculada. Que las constantes C_1 y C_2 sean próximas a la unidad es una medida de que el estimador se ajusta al error.

A veces no es posible obtener una desigualdad como la de fiabilidad y es suficiente con una desigualdad del tipo

$$\|u - u_h\| \leq C_2 \text{Est}_h + \text{Osc}(h)$$

donde el término $\text{Osc}(h)$ debe tender a cero cuando h tiende a cero (triangulaciones más finas) y se puede estimar dependiendo únicamente de los datos del problema, no de la solución. Este término suele estar referido a la variación (oscilación) de los datos, como veremos claramente en el caso de los estimadores residuales.

Una exigencia de tipo práctica sobre el estimador (que, insisto, debe depender únicamente de la solución numérica y de los datos) es que su cómputo sea económico, es decir, el tiempo necesario para calcularlo sea mucho menor que el de resolver el problema. Además, la cota de fiabilidad no debe emplear ningún tipo de regularidad de los datos, ya que los estimadores se emplean precisamente para evaluar el error en previsión de que las soluciones sean poco regulares y, sobre todo, sin saber si lo son o no.

Finalmente, una última exigencia que se aplica a la mayor parte de los estimadores es que sean **locales**. Esto ocurrirá de varias maneras distintas, pero se puede condensar en la siguiente idea. Se espera que el estimador se construya a partir de una lista de nodos, elementos o lados de manera que

$$\text{Est}_h^2 = \sum_j \text{Est}_{j,h}^2,$$

donde $\text{Est}_{j,h}$ emplee únicamente información (sobre la solución y los datos) cercana a elemento, nodo o lado j . El estimador es local si podemos acotar

$$C_3 \text{Est}_{j,h} \leq \|u - u_h\|_j$$

donde $\|u - u_h\|_j$ es el error en una zona que rodea al elemento, nodo o lado j . Si esta cota se hace de manera que

$$\sum_j \|\cdot\|_j^2 \leq C_4 \|\cdot\|^2,$$

se tiene que el carácter local del estimador implica su eficiencia. La ventaja de este carácter local es que sabemos que si un sumando del error concentra gran parte del error, entonces estamos seguros de que el error procede de esa zona geométrica y no de otra. ¡Esto no es tan obvio como parece! Puede ocurrir que un estimador se calcule con cantidades en una zona, pero que el error esté provocado por algo que ocurre en otro sitio.

Un estimador a posteriori del error se debe complementar con tres procesos más:

- Un proceso de comparación con una tolerancia. Si

$$\text{Est}_h \leq \text{Tol},$$

la solución se da por buena. La tolerancia debe ser relativa a alguna magnitud relacionada con la solución, por supuesto.

- Un proceso de marcado, que decide qué términos del error son los más grandes. Por ejemplo, se pueden reenumerar los sumandos locales

$$\text{Est}_{1,h} \leq \text{Est}_{2,h} \leq \dots \leq \text{Est}_{N-1,h} \leq \text{Est}_{N,h}$$

y luego tomar, o bien, un número fijo de los más grandes, o bien, el mínimo número posible de términos de forma que

$$\eta \text{Est}_h^2 \leq \sum_{j=k}^N \text{Est}_{j,h}^2$$

con $\eta < 1$. Esto quiere decir, que se buscan el mínimo número de términos que producen un porcentaje fijado del error total. Decididos qué términos producen la mayor parte del error, se marcan los elementos correspondientes.

- Un proceso de refinamiento del mallado, que busca un nuevo mallado subdividiendo los elementos marcados y, posiblemente, algunos cercanos para evitar la aparición de nodos colgantes (hanging nodes).

Habitualmente estos tres pasos deben ser repetidos hasta que la tolerancia deseada es alcanzada. Para evitar la degeneración de las triangulaciones (ver sección final) a veces el mallado nuevo se guarda con memoria del mallado del que procede, puesto que cada remallado puede requerir deshacer parte del mallado más reciente.

Algunos procedimientos de adaptatividad de mallados emplean engrosamiento de mallados (coarsening) con objeto de reducir el número de grados de libertad en zonas donde es notorio que hay muy poco error. Esto es un tema todavía poco desarrollado.

2. Notaciones

Dada una triangulación \mathcal{T}_h y un elemento genérico $K \in \mathcal{T}_h$, denotamos

- $\mathcal{E}(K)$ = lados de K
- $\mathcal{N}(K)$ =vértices (nodos) de K
- \mathcal{E}_h =conjunto de todos los lados de la triangulación
- \mathcal{N}_h =conjunto de todos los nodos de la triangulación
- $\mathcal{N}(e)$ =vértices de $e \in \mathcal{E}_h$

Consideramos además las siguiente particiones de los conjuntos de lados (interiores, Dirichlet, Neumann)

$$\mathcal{E}_h = \mathcal{E}_{h,\Omega} \cup \mathcal{E}_{h,D} \cup \mathcal{E}_{h,N}$$

y nodos (interiores, Dirichlet y Neumann)

$$\mathcal{N}_h = \mathcal{N}_{h,\Omega} \cup \mathcal{N}_{h,D} \cup \mathcal{N}_{h,N}$$

(un nodo Neumann es un nodo de lado Neumann que no es de lado Dirichlet).

A cada lado le asignamos un vector normal \mathbf{n}_e obligatoriamente exterior en los lados Neumann. Denotaremos $[\varphi]_e$ el salto de una función a lo largo de un lado.

3. Estimadores residuales

Sea Ω un dominio poligonal del plano. Los datos del problema serán $f : \Omega \rightarrow \mathbb{R}$ y $g : \Gamma_N \rightarrow \mathbb{R}$. Consideramos el problema

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ u|_{\Gamma_D} = 0, & u|_{\Gamma_N} = g. \end{cases}$$

En forma variacional, en el espacio $V = \{v \in H^1(\Omega) \mid v|_{\Gamma_D} = 0\}$, se tiene

$$\begin{cases} u \in V, \\ \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v + \int_{\Gamma_N} g v, \quad \forall v \in V. \end{cases}$$

Se puede poner condición de contorno Dirichlet no homogénea sin demasiadas dificultades adicionales.

Dada una triangulación \mathcal{T}_h aplicamos el método de elementos finitos \mathbb{P}_1 , esto es, tomando

$$V_h = \{v_h \in \mathcal{C}(\bar{\Omega}) \mid v_h|_{\Gamma_D} = 0, v_h|_K \in \mathbb{P}_1, \forall K\},$$

consideramos el problema

$$\begin{cases} u_h \in V_h, \\ \int_{\Omega} \nabla u_h \cdot \nabla v_h = \int_{\Omega} f v_h + \int_{\Gamma_N} g v_h, \quad \forall v_h \in V_h. \end{cases}$$

Dada \tilde{u} , definimos el **residuo** $R(\cdot; \tilde{u}) : V \rightarrow \mathbb{R}$

$$R(v; \tilde{u}) = \int_{\Omega} f v + \int_{\Gamma_N} g v - \int_{\Omega} \nabla \tilde{u} \cdot \nabla v.$$

Por definición

$$R(v; u) = 0, \quad \forall v \in V \quad \iff \quad u \text{ es la solución.}$$

Si u_h es la solución discreta

$$R(v; u_h) = \int_{\Omega} \nabla(u - u_h) \cdot \nabla v$$

y

$$R(v_h; u_h) = 0, \quad \forall v_h \in V_h.$$

Además

$$\|u - u_h\|_{1,\Omega} \approx_h \sup_{0 \neq w} \frac{R(w; u_h)}{\|w\|_{1,\Omega}} =: \mathcal{R}_h.$$

El objetivo es, por tanto, estimar \mathcal{R}_h .

Proposición

$$\mathcal{R}_h^2 \leq \sum_K h_K^2 \int_K |f|^2 + \sum_{e \in \mathcal{E}_{h,N}} h_e \int_e |g - \partial_n u_h|^2 + \sum_{e \in \mathcal{E}_{h,\Omega}} h_e \int_e [\partial_n u_h]^2$$

Cada uno de los tres términos en el lado derecho de la expresión anterior mide un efecto distinto:

- Como en cada triángulo $\Delta u_h = 0$ (puesto que u_h es lineal a trozos),

$$\int_K |f|^2 = \int_K |\Delta u - \Delta u_h|^2$$

mide el incumplimiento de la ecuación diferencial.

- El término

$$\int_e |g - \partial_n u_h|^2$$

mide el incumplimiento de la condición de Neumann.

- Finalmente, el término

$$\int_e [\partial_n u_h]^2$$

mide la falta de regularidad de la solución numérica, que provoca saltos en las derivadas normales.

Nótese también que, fijado un triángulo K ,

$$h_K^2 \int_K |f|^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_{h,N}} h_e \int_e |g - \partial_n u_h|^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_{h,\Omega}} h_e \int_e [\partial_n u_h]^2$$

son los sumandos referidos a K y sus lados. Si promediamos

$$f_K = \frac{1}{\text{área } K} \int_K f, \quad g_e = \frac{1}{h_e} \int_e g,$$

definimos la siguiente aproximación de la cantidad anterior

$$\eta_{R,K}^2 = h_K^2 \int_K |f_K|^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_{h,N}} h_e \int_e |g_e - \partial_n u_h|^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_{h,\Omega}} h_e \int_e |[\partial_n u_h]|^2$$

que ya es computable. (Todas las funciones que aparecen son constantes). Como estimador del error definimos

$$\text{Est}_h = \left[\sum_K \eta_{R,K}^2 \right]^{1/2}$$

Notemos que el trabajo de cómputo del estimador incluye dos fases: una que se realiza únicamente sobre los datos, promediándolos sobre los elementos; otra que emplea estos promedios y la solución numérica para construir el estimador.

La **fiabilidad** del estimador queda reflejada en la siguiente desigualdad:

$$\mathcal{R}_h^2 \leq C \text{Est}_h^2 + \sum_K h_K^2 \int_K |f - f_K|^2 + \sum_e h_e \int_e |g - g_e|^2.$$

Como ya se había indicado, en el lado derecho de la desigualdad, hay términos de orden alto además del estimador. Estos términos reflejan únicamente cómo los promedios de los datos aproximan a los datos, de ahí su nombre de términos de oscilación. Notemos que

estos términos indican que si la malla no resuelve las oscilaciones fuertes de los datos, es difícil pensar que la solución numérica no hereda este problema de aproximación.

Se tiene además una cota de **estimación local**. Para ello se introduce el macrotriángulo ω_K , formado por la unión de triángulos apoyados en K (tienen lados comunes), es decir,

$$\omega_K = \cup\{K' \mid \mathcal{E}(K') \cap \mathcal{E}(K) \neq \emptyset\}.$$

Entonces, se tiene que

$$\eta_{R,K}^2 \leq C\|u - u_h\|_{1,\omega_K}^2 + \sum_{K' \subset \omega_K} h_{K'}^2 \int_{K'} |f - f_{K'}|^2 + \sum_{e \in \mathcal{E}(K) \cap \mathcal{E}_{h,N}} \int_e |g - g_e|^2.$$

Esto indica que la parte del estimador relacionada con el triángulo K acota inferiormente al error en el macrotriángulo circundante, sumado a términos de oscilación de los datos, también en los elementos circundantes.

Si se supone una cierta **regularidad en los datos** f y g (regularidad en los datos, ¡no en las soluciones!), se demuestra fácilmente que

$$\int_K |f - f_K| = \mathcal{O}(h_K^2), \quad \int_e |g - g_e|^2 = \mathcal{O}(h_e^2),$$

luego

$$\|u - u_h\|_{1,\Omega} \leq C\text{Est}_h + \mathcal{O}(h).$$

Por tanto, conforme la malla se refina y los términos de oscilación se van resolviendo, el estimador es cada vez más fiable. Por otro lado,

$$\eta_{R,K} \leq C\|u - u_h\|_{1,\omega_K} + \underbrace{\mathcal{O}(h_K^2) + \mathcal{O}(h_e^{3/2})}_{\text{locales}},$$

lo cual da una eficiencia local, dependiente únicamente de parámetros de la malla en la zona circundante al triángulo K .

4. Estimadores locales

Funciones burbuja. Si $\lambda_{i,K}$ ($i = 1, 2, 3$) son las funciones coordenadas baricéntricas de un triángulo K , entonces definimos

$$b_K = \begin{cases} 27\lambda_{1,K}\lambda_{2,K}\lambda_{3,K}, & \text{en } K, \\ 0 & \text{en otro caso.} \end{cases}$$

Si e es un lado delimitado por los triángulos K_1 y K_2 y numeramos los vértices comunes de la misma forma, definimos

$$b_e := \begin{cases} 4\lambda_{1,K_1}\lambda_{2,K_1}, & \text{en } K_1, \\ 4\lambda_{1,K_2}\lambda_{2,K_2}, & \text{en } K_2, \\ 0 & \text{en otro caso.} \end{cases}$$

Se tiene que

$$0 \leq b_K \leq 1, \quad 0 \leq b_e \leq 1$$

y ambas toman valor máximo unidad.

Funciones asociadas a un punto. Dado un nodo no Dirichlet de la triangulación $x \in \mathcal{N}_{h,\Omega} \cap \mathcal{N}_{h,N}$ y su macrotriángulo asociado ω_x , formado por los triángulos que tienen a x como vértice

$$\omega_x = \cup\{K \mid x \in \mathcal{N}(K)\}.$$

Entonces consideramos las funciones

- b_K con $x \in \mathcal{N}(K)$, esto es, $K \subset \omega_x$
- b_e con $x \in \mathcal{N}(e)$ y $e \in \mathcal{E}_{h,\Omega}$
- b_e con $e \in \partial\omega_x \cap \Gamma_N$

Se define el espacio \mathcal{V}_x de las combinaciones lineales de las funciones burbuja anteriores. Todas las funciones de \mathcal{V}_x tienen soporte en ω_x .

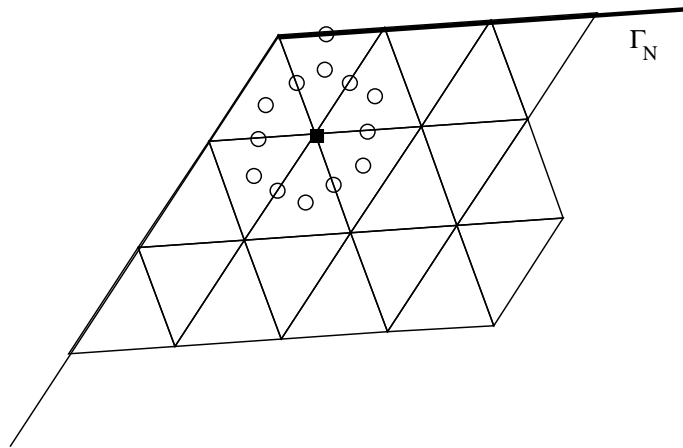


Figura 8.1: Funciones burbuja asociadas a un nodo

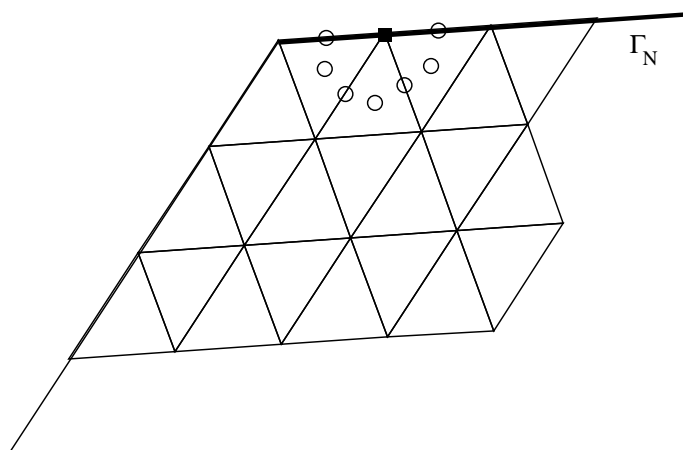


Figura 8.2: Funciones burbuja asociadas a un nodo (otra situación)

El estimador local asociado al nodo x se compone de dos fases. Primero se resuelve el siguiente problema local

$$\begin{cases} v_x \in \mathcal{V}_x \\ \int_{\omega_x} \nabla v_x \cdot \nabla w = \sum_{K \subset \omega_x} \int_K f_K w + \sum_e \int_e g_e w - \int_{\omega_x} \nabla u_h \cdot \nabla w, \quad \forall w \in \mathcal{V}_x. \end{cases}$$

Se define entonces el estimador

$$\eta_{D,x}^2 := \int_{\Omega} |\nabla v_x|^2 = \int_{\omega_x} |\nabla v_x|^2.$$

Hay una interpretación simple de este estimador. Si $w \in \mathcal{V}_x$, entonces

$$\int_{\Omega} \nabla(u - u_h) \cdot \nabla w = R(w; u_h) \approx \int_{\omega_x} \nabla v_x \cdot \nabla w = \int_{\Omega} \nabla v_x \cdot \nabla w,$$

luego estamos haciendo una estimación local del residuo. De hecho, el problema que define v_x es un método de Galerkin aproximado en el espacio local \mathcal{V}_x para

$$\begin{cases} -\Delta \varphi = f, & \text{en } \omega_x \\ \varphi_{\partial \omega_x \setminus \Gamma_N} = u_h, & \partial_n \varphi|_{\Gamma_N \cap \partial \omega_x} = g. \end{cases}$$

El estudio de este estimador se realiza por comparación con el estimador residual precedente: para todo x

$$\eta_{D,x}^2 \leq C \sum_{K \subset \omega_x} \eta_{R,K}^2$$

y para todo K

$$\eta_{R,K}^2 \leq C \sum_{x \in \mathcal{N}(K)} \eta_{D,x}^2,$$

donde en la suma última hay que eliminar los vértices Dirichlet, en los que no se realiza estimación. La primera propiedad nos permite asegurar que $\eta_{D,x}$ es un buen estimador local. De hecho, se puede obtener la siguiente cota, algo más fina que la obtenida sumando las cotas de error para los estimadores residuales:

$$\eta_{D,x}^2 \leq C \|u - u_h\|_{1,\omega_x}^2 + \sum_{K \subset \omega_x} h_K^2 \int_K |f - f_K|^2 + \sum_{e \subset \Gamma_N \cap \partial \omega_x} \int_e |g - g_e|^2$$

La segunda nos permite emplear

$$\sum_x \eta_{D,x}^2 = \sum_x \int_{\omega_x} |\nabla v_x|^2$$

como estimador global del error, aunque como tal es excesivamente complicado. La idea de este tipo de estimadores es localizar zonas donde ya presuponemos que puede haber mucho error, no construir un estimador completo del error.

5. Estimadores jerárquicos

En los dos casos que expondremos a continuación

$$V_h \subset W_h \subset X$$

W_h es un subespacio más fino ligado a la misma triangulación. La idea es evaluar el residuo de la solución numérica sobre este subespacio más grande. La parte de V_h no produce residuo y, por tanto, solo hay que averiguar el residuo de la parte añadida.

Refinamiento por lados

El espacio W_h se compondrá de las funciones de V_h más una función por cada lado no Dirichlet. A cada lado asignamos un función $\varphi_e \in W_h$ de forma que

$$\text{sop } \varphi_e \subset \omega_e, \quad 0 \leq \varphi_e \leq 1, \quad Ch_e \leq \int_e \varphi_e$$

$$\int_{\omega_e} |\nabla \varphi_e|^2 \leq C h_e^{-2} \int_{\omega_e} |\varphi_e|^2.$$

El estimador es entonces

$$\begin{aligned} r_e &= \int_{\Omega} f \varphi_e + \int_{\Gamma_N} g \varphi_e - \int_{\Omega} \nabla u_h \cdot \nabla \varphi_e \\ &= R(\varphi_e; u_h) = \int_{\Omega} (\nabla u - \nabla u_h) \cdot \nabla \varphi_e. \end{aligned}$$

Si e es un lado interior

$$r_e = \int_{\omega_e} f \varphi_e - \int_e [\partial_n u_h] \varphi_e$$

mientras que si e es un lado Neumann (de un triángulo K)

$$r_e = \int_K f \varphi_e + \int_e g \varphi_e.$$

Como estimador global podemos emplear

$$\text{Est}_h^2 := \sum_{e \notin \mathcal{E}_{h,D}} r_e^2.$$

Se tiene una desigualdad de estimación local

$$|r_e| \leq C \|u - u_h\|_{1, \omega_e},$$

del que se deduce la eficiencia del estimador. Además se tiene un resultado de eficiencia para datos suficientemente regulares

$$\|u - u_h\|_{1, \Omega} \leq C \text{Est}_h + \mathcal{O}(h)$$

Ejemplos.

- $W_h = \{u_h \in \mathcal{C}(\Omega) \mid u_h|_K \in \mathbb{P}_2, \forall K\}$ surge de tomar $\varphi_e = b_e$ para cada lado.
- Si $\mathcal{T}_{h/2}$ es la partición uniforme a mitad de la triangulación, entonces $W_h = V_{h/2}$ surge de tomar como φ_e la función de base nodal del punto medio.

En ambos casos hemos construido el espacio dando una base de un suplementario de V_h .

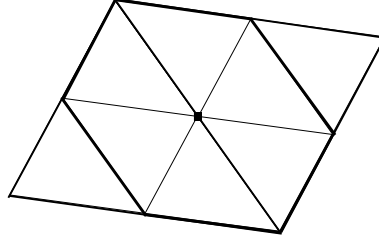


Figura 8.3: Refinamiento de un par de triángulos y (en trazo más grueso) soporte de la función añadida para la estimación del error

Refinamiento por lados y triángulos

El espacio W_h se compondrá de las funciones de V_h más una función por cada lado no Dirichlet y otra por triángulo. A cada lado asignamos una función $\varphi_e \in W_h$ (como antes) de forma que

$$\text{sop } \varphi_e \subset \omega_e, \quad 0 \leq \varphi_e \leq 1, \quad Ch_e \leq \int_e \varphi_e$$

$$\int_{\omega_e} |\nabla \varphi_e|^2 \leq C h_e^{-2} \int_{\omega_e} |\varphi_e|^2.$$

A cada triángulo asociamos $\varphi_K \in W_h$ tal que

$$\text{sop } \varphi_K \subset K, \quad 0 \leq \varphi_K \leq 1, \quad Ch_K^2 \leq \int_K \varphi_K,$$

$$\int_K |\nabla \varphi_K|^2 \leq C h_K^{-2} \int_K |\varphi_K|^2.$$

El estimador para cada triángulo es entonces

$$r_K := \int_{\Omega} f \varphi_K + \int_{\Gamma_N} g \varphi_K - \int_{\Omega} \nabla u_h \cdot \nabla \varphi_K = R(\varphi_K; u_h) = \int_K f \varphi_K.$$

El estimador de cada lado es el mismo de antes. Como estimador global podemos emplear

$$\text{Est}_h^2 = \sum_{e \notin \mathcal{E}_{h,D}} r_e^2 + \sum_K r_K^2.$$

Las propiedades son similares a las anteriores.

Ejemplos.

- $W_h = \{u_h \in \mathcal{C}(\Omega) \mid u_h|_K \in \mathbb{P}_3, \forall K\}$ surge de tomar $\varphi_e = b_e$ para cada lado y $\varphi_K = b_K$ para cada triángulo.
- Si $\mathcal{T}_{h/4}$ es la partición uniforme a cuartos de la triangulación (cada triángulo produce dieciséis), entonces $W_h = V_{h/4}$ surge de tomar:

φ_e la función de base nodal del punto medio en el espacio $V_{h/2}$

φ_K la función que vale uno en el triángulo central y cero en la frontera exterior.

Con ellas no se llega a cubrir el espacio entre V_h y W_h .

- Si $\tilde{\mathcal{T}}_h$ es la triangulación obtenida por subdivisión en seis de cada triángulo, $W_h = \tilde{V}_h$ surge de tomar φ_e la función nodal del punto medio y φ_K la baricéntrica.

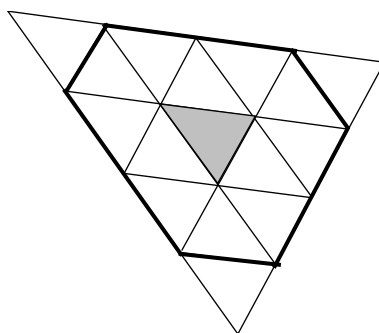


Figura 8.4: Refinamiento cuádruple de un triángulo y soporte de la función interior añadida para la estimación del error (se indica la zona en la que la función es constante y en trazo grueso la frontera del soporte)

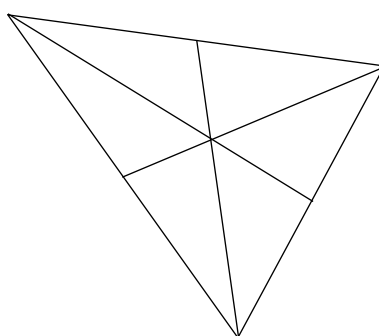


Figura 8.5: Otro tipo de partición de un triángulo para estimación del error

6. Promediado del gradiente (Zienkiewicz–Zhu)

Vamos a simplificar el problema a un problema sólo con condiciones Dirichlet homogéneas,

$$\begin{cases} -\Delta u = f, & \text{en } \Omega, \\ u|_{\Gamma} = 0, \end{cases}$$

que resolveremos empleando elementos lineales sobre una triangulación. La idea de base del método de Zienkiewicz–Zhu consiste en calcular un gradiente continuo discreto empleando promedios locales del gradiente constante a trozos de la solución numérica. Seguidamente se emplea la diferencia entre el gradiente discreto y el corregido como estimación del error.

Si u_h es la solución por elementos finitos \mathbb{P}_1 ($u_h \approx u$), intentaremos definir

$$\gamma u_h \approx \nabla u_h$$

tal que

$$\int_{\Omega} |\nabla u - \gamma u_h|^2 \leq \beta \int_{\Omega} |\nabla u - \nabla u_h|^2, \quad \text{con } \beta < 1.$$

En tal caso

$$\int_{\Omega} |\gamma u_h - \nabla u_h|^2 \approx \int_{\Omega} |\nabla u - \nabla u_h|^2 = \text{error}.$$

Notemos que ∇u_h es constante en cada triángulo.

Consideramos los siguientes espacios de funciones vectoriales:

- $\vec{W}_h = \{\vec{w}_h : \Omega \rightarrow \mathbb{R}^2 \mid \vec{w}_h|_K \in (\mathbb{P}_1)^2, \forall K\}$.
- $\vec{V}_h = \{\vec{v}_h \in \vec{W}_h \mid \vec{v}_h \in \mathcal{C}(\Omega)\}$.

Si $v_h \in V_h$, entonces $\nabla u_h \in \vec{W}_h$. Definimos además

$$(\vec{v}_h, \vec{w}_h)_h = \sum_K \frac{\text{área } K}{3} \left[\sum_{x \in \mathcal{N}(K)} \vec{v}_h|_K(x) \cdot \vec{w}_h|_K(x) \right] \approx \int_{\Omega} \vec{v}_h \cdot \vec{w}_h.$$

- Si \vec{v}_h ó \vec{w}_h es constante a trozos

$$(\vec{v}_h, \vec{w}_h)_h = \int_{\Omega} \vec{v}_h \cdot \vec{w}_h$$

- En \vec{W}_h

$$(\vec{v}_h, \vec{v}_h)_h^{1/2}$$

define una norma equivalente (independientemente de h) a la de energía.

- Si \vec{v}_h y \vec{w}_h son continuas

$$(\vec{v}_h, \vec{w}_h)_h = \frac{1}{3} \sum_x \text{área } \omega_x \vec{v}_h(x) \cdot \vec{w}_h(x).$$

Para estimar el error, podemos pensar en resolver el problema

$$\begin{cases} \gamma u_h \in \vec{V}_h, \\ (\gamma u_h, \vec{v}_h)_h = (\nabla u_h, \vec{v}_h)_h, \quad \forall \vec{v}_h \in \vec{V}_h. \end{cases}$$

La solución explícita en un nodo x (esto ya da la solución en todo el dominio) es

$$\gamma u_h(x) = \sum_{K \subset \omega_x} \frac{\text{área } K}{\text{área } \omega_x} \nabla u_h|_K.$$

A partir de allí se definen para cada triángulo

$$\eta_{Z,K}^2 = \int_K |\gamma u_h - \nabla u_h|^2$$

y

$$\eta_{Z,h}^2 = \sum_K \eta_{Z,K}^2.$$

la ventaja de este estimador es que su coste computacional es muy bajo. Se han elaborado muchas variantes de este estimador, que es uno de los más empleados en la práctica para adaptatividad.

7. Estrategias de refinamiento para triángulos

8. Ejercicios propuestos

1. Supongamos que un estimador organizado por triángulos cumple la desigualdad local

$$\eta_K \leq C \|u - u_h\|_{1,\omega_K}.$$

Demuestra que el estimador es eficiente.

2. El estimador local por nodos cumple

$$\eta_{R,K}^2 \leq C \sum_{x \in \mathcal{N}(K)} \eta_{D,x}^2.$$

Deduce una cota de fiabilidad (con términos de oscilación) para el estimador

$$\sum_x \int_{\omega_x} |\nabla v_x|^2.$$

3. Dado un triángulo K consideramos las siguientes funciones burbuja

- $b_{K'}$ donde $K' \subset \omega_K$ (esto es, tiene un lado común con K),
- b_e , donde $e \in \mathcal{E}(K)$ es un lado no Dirichlet,

- b_e , donde e es un lado Neumann de ω_K ,

y el espacio \mathcal{V}_K generado por las funciones anteriores. Define un estimador local por triángulos siguiendo las mismas ideas del estimador por puntos. ¿Cuál es el dominio donde está el soporte de todas las funciones de \mathcal{V}_K ? Dibuja las posibilidades de elección de funciones burbujas para triángulos con todos los lados interiores, con un lado Dirichlet o con un lado Neumann (habitualmente los malladores devuelven triangulaciones con un único lado exterior).

Lección 9

Referencias bibliográficas

Los siguientes textos contienen aspectos teóricos sobre el método de los elementos finitos en un lenguaje matemático. El libro de Ciarlet es la referencia obligada del método en la comunidad matemática. El primero de ellos incide más en detalles computacionales (métodos iterativos, etc) y en aplicaciones a mecánica de sólido.

- Braess, Dietrich. *Finite elements. Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, Cambridge, 1997.
- Brenner, Susanne C.; Scott, L. Ridgway. *The mathematical theory of finite element methods*. Texts in Applied Mathematics, 15. Springer-Verlag, New York, 1994.
- Ciarlet, Philippe G. *The finite element method for elliptic problems*. Studies in Mathematics and its Applications, Vol. 4. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978
- Ženšek, Alexander. *Nonlinear elliptic and evolution problems and their finite element approximations*. Computational Mathematics and Applications. Academic Press, Inc., London, 1990.

El texto

- Brezzi, Franco; Fortin, Michel. *Mixed and hybrid finite element methods*. Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

es la referencia obligada sobre variantes de las formulaciones y métodos mixtos.

El siguiente texto

- Joly, Pascal *Mise en œuvre de la méthode des éléments finis*. Mathématiques & Applications, 2. Ellipses, Paris, 1990.

contiene explicaciones detalladas sobre cómo almacenar de forma eficiente las matrices de los elementos finitos y una exposición cuidada del ensamblado de los elementos clásicos. Una buena referencia para estrategias de implementación de elementos de alto orden es:

- Solin, Pavel; Segeth, Karel; Dolezel, Ivo. *Higher-order Finite Element Methods*. Chapman Hall/CRC Press, 2003.

Los textos

- Ainsworth, Mark; Oden, John Tinsley *A posteriori error estimation in finite element analysis*. Wiley-Interscience, 2000.
- Verfürth, Rüdiger. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, 1996.

realizan exposiciones amplias sobre distintos métodos de estimación a posteriori del error.

Índice general

1. Elementos lagrangianos de orden alto	2
1. Funciones de Lobatto	2
2. Bases libres de nodos	3
3. Construcción de bases de \mathbb{Q}_p	5
4. Construcción de bases en triángulos	8
5. Elementos de transición	11
6. Ejercicios propuestos	12
2. Elementos de tipo Hermite	14
1. Elemento de Hermite \mathbb{P}_3 en dimensión uno	14
2. Viga de Euler–Bernoulli	16
3. Elemento de Hermite \mathbb{P}_3 en dimensión dos	18
4. Ejercicios propuestos	22
3. Elementos de clase \mathcal{C}^1	23
1. Elemento de Argyris	23
2. Problema de placa de Kirchhoff	26
3. Elementos de Bell y BFS	29
4. Ejercicios propuestos	30
4. Elementos no conformes	32
1. Error de las aproximaciones no conformes	32
2. Elementos de Crouzeix–Raviart	35
3. El elemento de Morley para el bilaplaciano	40
4. Ejercicios propuestos	41
5. Formulaciones mixtas y condiciones inf–sup	42
1. Problemas con estructura mixta	42
2. Discretizaciones de problemas mixtos	45
3. Aspectos teóricos: problema exacto	47
4. Aspectos teóricos: problema discreto	49
5. Forma matricial de la condición inf–sup	50
6. Verificación numérica	52
7. Ejercicios propuestos	53

6. Elementos que preservan la divergencia	55
1. Condición de preservación de la divergencia	55
2. Elementos de Raviart–Thomas	56
3. Aplicación	58
4. Ideas sobre el análisis	62
5. Otras condiciones de contorno	63
6. Elemento de Raviart–Thomas de orden dos	64
7. Ejercicios propuestos	65
7. Formulaciones mixtas en elasticidad	67
1. Preliminares	67
2. Formulación de Hellinger–Reissner (I)	69
3. Formulación de Hellinger–Reissner (II)	70
4. Formulaciones de Hu–Washizu	72
5. Formulación–elemento PEERS	74
6. Ejercicios propuestos	74
8. Estimación a posteriori del error	75
1. Objetivos y terminología	75
2. Notaciones	77
3. Estimadores residuales	77
4. Estimadores locales	80
5. Estimadores jerárquicos	83
6. Promediado del gradiente (Zienkiewicz–Zhu)	86
7. Estrategias de refinamiento para triángulos	87
8. Ejercicios propuestos	87
9. Referencias bibliográficas	89