

**Universidad Central de Venezuela
Facultad de Ciencias
Escuela de Computación**

Lecturas en Ciencias de la Computación
ISSN 1316-6239

**Proyección sobre el núcleo de las
matrices de acoplamiento de
descomposición de dominios**

Prof. Luis Manuel Hernández R.

RT 2006-01

Centro de Cálculo Científico y Tecnológico de la UCV
CCCT-UCV
Caracas, Febrero, 2006.

Proyección sobre el núcleo de las matrices de acoplamiento de descomposición de dominio.

Luis Manuel Hernández-Ramos **

6 de febrero de 2006

Resumen

El método de proyecciones alternantes oblicuas para problemas de punto de ensilladura (saddle point problems) hace uso del operador de proyección sobre el núcleo de una matriz rectangular. Esta proyección es, en general, costosa de calcular ya que requiere la pseudoinversa de una matriz rectangular. En este artículo se muestra que, en el caso de los problemas de punto de ensilladura resultantes de ciertas discretizaciones de descomposición de dominios, el cálculo de dicha proyección equivale a realizar un promedio ponderado de los valores de los nodos de la interfaz. Es decir; los valores de los nodos fuera de la interfaz no intervienen en el computo. Este resultado se traduce en bajo costo computacional y es válido tanto para el caso conforme, como para el caso no conforme.

AMS: 65F10, 65N22, 65Y05.

Palabras claves: Descomposición de dominios, Precondicionador, Sistemas lineales acoplados.

1. Introducción

Estamos interesados en los sistemas de ecuaciones lineales de la forma,

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix}, \quad (1)$$

*Centro de Cálculo Científico y Tecnológico, Facultad de Ciencias, Universidad Central de Venezuela, Caracas.

** e-mail: luish@kuaimare.ciens.ucv.ve

donde

(H1) B es una matriz $m \times n$, $\text{rang}(B) = m$, $m \leq n$.

(H2) A es una matriz $n \times n$ simétrica positivo definida.

Sistemas de este tipo son llamados «Problemas de punto de ensilladura» (En inglés, Saddle point problems), y aparecen en gran variedad de aplicaciones científicas y de ingeniería tales como: dinámica de fluidos computacional, optimización, y descomposición de dominios. Un excelente resumen sobre los problemas de punto de ensilladura y sus aplicaciones puede encontrarse en el artículo de Benzi, Golub y Liesen [BGL05].

Como la hipótesis **(H2)** implica que la matriz A es invertible, entonces el problema de punto de ensilladura (1) es equivalente a:

$$\begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix} \begin{pmatrix} X \\ \lambda \end{pmatrix} = \begin{pmatrix} F \\ -BA^{-1}F \end{pmatrix}, \quad (2)$$

donde $S = -BA^{-1}B^T$.

Nótese que en el nuevo sistema (2), la matriz

$$\begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix},$$

es triangular por bloques, por lo que resolver dicho sistema equivale a resolver primero el sistema de ecuaciones,

$$S\lambda = BA^{-1}F, \quad (3)$$

y luego resolver el sistema,

$$Ax = f - B^T\lambda.$$

El sistema (3) es llamado sistema condensado, y la matriz S de dimensión $m \times m$, se denomina el complemento de Schur del sistema. El complemento de Schur juega un rol preponderante en la resolución de los sistemas acoplados.

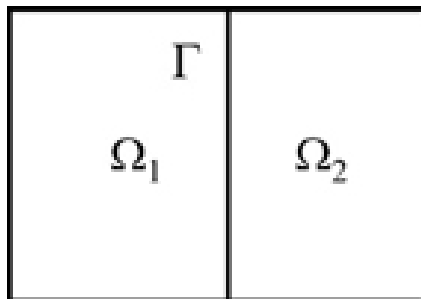


Figura 1: Dominio Rectangular Descompuesto

1.1. Descomposición de Dominios

Particularmente nos interesamos en los problemas de punto de ensilladura provenientes de algunas discretizaciones de descomposición de dominios. A estos sistemas los llamaremos sistemas lineales acoplados, cf. [HR99, HR00, HR02, HR05].

Los métodos de descomposición de dominios fueron desarrollados originalmente para resolver ecuaciones en derivadas parciales en regiones de \mathbb{R}^2 y \mathbb{R}^3 [XZ98, QV99]. En las estrategias de descomposición de dominios para la resolución numérica de ecuaciones en derivadas parciales, se debe resolver un problema definido en un dominio $\bar{\Omega}$, y para facilitar su resolución numérica se descompone dicho dominio en varios sub-dominios Ω_i . Entonces, la solución del problema global en $\bar{\Omega}$ se efectúa a partir de la resolución de sub-problemas locales en cada sub-dominio Ω_i , que luego se adjuntan mediante la imposición de una condición de acoplamiento en las interfaces de los sub-dominios.

Existen varias maneras de imponer estas condiciones de acoplamiento. Por ejemplo, mediante un multiplicador de Lagrange. La discretización del problema de descomposición de dominios conduce a la resolución de sistemas lineales acoplados de punto de ensilladura

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix}. \quad (4)$$

En este caso, la matriz A del sistema es una matriz diagonal por bloques, donde cada bloque A_i proviene de la discretización del problema local en el sub-dominio Ω_i . La matriz B proviene de la discretización de las condiciones de acoplamiento en las interfaces.

Para ejemplificar, supongamos que el dominio Ω es particionado en 2 sub-dominios Ω_1 y Ω_2 , con una interfaz que los separa que denotaremos Γ . Consideraremos dos tipos de condición de acoplamiento según la discretización de

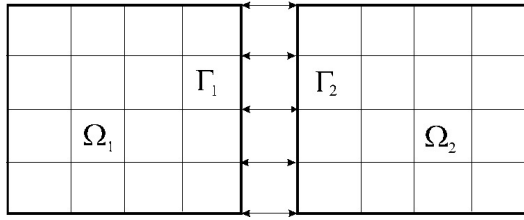


Figura 2: Caso conforme

los problemas locales en cada sub-dominio. En efecto, se pueden utilizar mallas independientes en cada sub-dominio, allí no hay correspondencia entre los puntos de las mallas en la interfaz. Este tipo de mallas se denomina “no conforme”. La otra posibilidad es de utilizar mallas que no son independientes, en el sentido en que ellas se corresponden exactamente en las interfaces. Este tipo de malla se conoce como mallas “conformes”.

1.1.1. Mallas conformes

Cuando las mallas son conformes, cada nodo de la interfaz de un sub-dominio es un nodo de la interfaz de los sub-dominios vecinos. Entonces las condiciones de acoplamiento obligan al vector X a tener el mismo valor en esos nodos. Cuando el sub-dominio se descompone en dos sub-dominios, si se numeran los nodos de la interfaz al final de cada sub-dominio, la matriz B que corresponde a la discretización de la condición de acoplamiento presenta una estructura muy particular de la forma ver ([JUV97]):

$$B = (B_1 \ B_2); B_1 = (0 \ I_\Gamma); B_2 = (0 \ -I_\Gamma). \quad (5)$$

En este caso, la condición $BX = 0$ expresa condiciones del tipo:

$$X_i^{(1)} - X_i^{(2)} = 0 \quad (6)$$

que fuerzan un salto nulo a través de la interfaz.

1.1.2. Mallas no conformes

Cuando las mallas son no conformes, no hay generalmente correspondencia entre los nodos de las interfaces. En este caso se imponen condiciones de acoplamiento de una manera “debil”. Por ejemplo, cuando hay dos sub-dominios con una interfaz Γ , se exige que la solución exacta verifique:

$$\int_{\Gamma} (v_1 - v_2)w dx = 0, \forall v_1 \in V_1, v_2 \in V_2, w \in W, \quad (7)$$

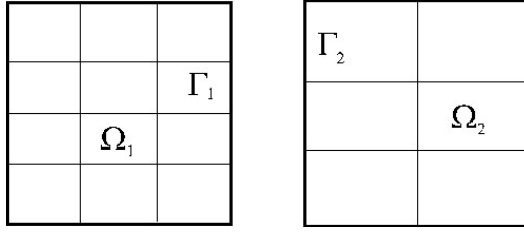


Figura 3: Caso no conforme

donde:

- v_i es la traza sobre Γ de la solución local definida en el sub-dominio Ω_i ,
- V_i es el espacio en el cual se define esta solución local,
- w es una función “test” en un espacio conveniente W . Es frecuente tomar el espacio W como la traza sobre Γ , de menor dimension entre V_i y V_j .

Cuando el problema es discretizado, se eligen los subespacios discretos convenientes $V_i^h \subset V_i$ y $W_i^h \subset W$ y se expresa la condición de acoplamiento como:

$$\int_{\Gamma} (v_1 - v_2) w dx = 0, \quad \forall v_1 \in V_1^h, v_2 \in V_2^h, w \in W^h. \quad (8)$$

Nuevamente, esta condición se expresa bajo forma matricial por $BX = 0$ con $B \in \mathbb{R}^{m \times n}$, y

$$m = \dim W^h = \min(\dim V_1^h, \dim V_2^h) \text{ y } n \leq \dim V_1^h + \dim V_2^h.$$

1.2. El método PAO

El método de proyecciones alternantes oblicuas (PAO) [HR05] es un método desarrollado para resolver problemas de punto de ensilladura; específicamente para problemas de punto de ensilladura provenientes de descomposición de dominios. Este método es una versión del método de los gradientes conjugados preconditionado aplicado al sistema condensado $BA^{-1}B^T\lambda = BA^{-1}F$, y tiene la ventaja que no necesita construir explícitamente la matriz del complemento de Schur (generalmente una matriz densa), ni para las iteraciones, ni para el preconditionador. En el método PAO, el preconditionador para el complemento de Schur es un preconditionador algebraico, basado en

el método de proyecciones alternantes de Von Neumann [NEU50, RE]. El algoritmo 1 muestra el método de proyecciones alternantes oblicuas. La matriz $R = I - P$ denota la matriz de la proyección ortogonal sobre el núcleo de B y P denota la matriz de la proyección ortogonal en el núcleo de B . El cálculo

Algorithm 1 Algoritmo de Proyecciones Alternantes Oblicuas

Elegir $x_0 \in V = \{x : F - Ax \perp \ker B\}$

$$w_0 = RARx_0$$

$$y_0 = w_0$$

for $k := 0, 1, \dots$, **do**

$$d_k = A^{-1}y_k$$

$$\alpha_k = \frac{\langle x_k, w_k \rangle}{\langle d_k, y_k \rangle}$$

$$x_{k+1} = x_k - \alpha_k d_k$$

$$w_{k+1} = RARx_{k+1}$$

$$\beta_{k+1} = \frac{\langle x_{k+1}, w_{k+1} \rangle}{\langle x_k, w_k \rangle}$$

$$y_{k+1} = w_{k+1} + \beta_{k+1}y_k$$

end for

explícito de la matriz R puede requerir el cálculo de la matriz pseudoinversa de B . De esta manera $R = B^+B$, donde $B^+ = B^T(BB^T)^{-1}$ es la matriz pseudoinversa de B . Esto puede resultar costoso para problemas en donde m sea relativamente grande.

2. Matrices de Acoplamiento

El concepto de Matriz de Acoplamiento, generaliza las matrices que expresan las condiciones de acoplamiento dadas en la sección anterior. Se denomina matriz de acoplamiento a una matriz que expresa las condiciones que se imponen en la interfaz entre los sub-dominios para que las soluciones locales se correspondan con la solución global del problema. De allí, se infiere que una matriz de acoplamiento:

1. No debe depender de valores fuera de la interfaz Γ ,
2. Debe acoplar convenientemente las trazas de las soluciones en la interfaz Γ .

Sea $\bar{\Omega}$ el dominio de cálculo; suponemos que $\bar{\Omega}$ es descompuesto en dos sub-dominios separados por una interfaz Γ . Sea $\mathfrak{I}_{\bar{\Omega}}$ el conjunto de índices de nodos

de $\bar{\Omega}$ (es por lo tanto el conjunto de índices de las columnas de B). Se realiza una partición de este conjunto en 2 subconjuntos de índices: $\mathfrak{S}_{\bar{\Omega}} = \mathfrak{S}_{\Omega} \cup \mathfrak{S}_{\Gamma}$ con $\mathfrak{S}_{\Omega} \cap \mathfrak{S}_{\Gamma} = \emptyset$, donde \mathfrak{S}_{Γ} es el conjunto de índices de nodos de la interfaz Γ . Aquí, cada nodo de una interfaz es contado en cada sub-dominio en el cual pertenezca: si pertenece a dos sub-dominios es contado dos veces, si pertenece a tres sub-dominios es contado tres veces, etc.

Sea B la matriz de acoplamiento; entonces $B \in \mathbb{R}^{m \times n}$ con $m \ll n$, es decir; los sub-dominios tienen muchos menos nodos en la interfaz que en el interior. Para que la primera propiedad anterior se verifique, es necesario y suficiente que:

$$B(\vec{e}_i) = 0, \quad \forall i \in \mathfrak{S}_{\Omega}, \quad (9)$$

donde \vec{e}_i es el i -ésimo vector de la base canónica. En el caso conforme, B está dado por (5) (por lo tanto $2m = \text{card}(\mathfrak{S}_{\Gamma})$). En el caso no conforme, la estructura de B depende de la elección del espacio de funciones test W^h . Sabemos, sin embargo que W^h contiene forzosamente las funciones constantes, por lo tanto necesariamente se cumple,

$$B(\vec{1}_{\mathfrak{S}_{\Gamma}}) = 0, \quad (10)$$

donde $\vec{1}_C = \sum_{i \in C} \vec{e}_i$ para todo subconjunto de índices $C \subset \mathfrak{S}_{\bar{\Omega}}$.

Aquí, se tomarán estas dos propiedades para definir una “matriz de acoplamiento”.

Definición 2.1 *Una matriz B se dice de acoplamiento sobre \mathfrak{S}_{Γ} si y solamente si B verifica (9) y (10), i.e.*

1. $\forall i \in \mathfrak{S}_{\Omega} : B(\vec{e}_i) = 0,$
2. $B(\vec{1}_{\mathfrak{S}_{\Gamma}}) = 0.$

Remarquemos aquí que necesariamente toda matriz de acoplamiento B en \mathfrak{S}_{Γ} verifica

$$B(\vec{1}_{\mathfrak{S}_{\bar{\Omega}}}) = 0. \quad (11)$$

3. Proyección sobre el núcleo de una matriz

En esta sección se muestra que efectivamente, el cálculo de la proyección sobre el núcleo de una matriz de acoplamiento, equivale a efectuar un promedio ponderado de los valores de la interfaz.

3.1. Proyección sobre el núcleo por pseudoinversa

Definición 3.1 Sea $B \in \mathbb{R}^{m \times n}$ una matriz cualquiera y B^+ su matriz pseudoinversa. Se define la transformación lineal P , de \mathbb{R}^n en \mathbb{R}^n :

$$\forall y \in \mathbb{R}^n : P(y) = (I - B^+B)y. \quad (12)$$

Denotamos $R = B^+B$. Es fácil mostrar que R es el operador de proyección ortogonal sobre $(\text{Ker } B)^\perp$ [SB80].

Si escribimos,

$$\forall y \in \mathbb{R}^n : P(y) = (I - R)y, \quad (13)$$

entonces necesariamente $P(y) = y - Ry$ es el operador de proyección ortogonal de y sobre $\text{Ker } B$.

Observación:

Si $y \in \text{Ker } B$ entonces $P(y) = (I - B^+B)y = y$.

3.2. Proyección sobre el núcleo de matrices de acoplamiento

Sea $B \in \mathbb{R}^{m \times n}$ una matriz de acoplamiento definida en §2.1, que verifica por tanto las propiedades (9) y (10). Sea P la proyección ortogonal sobre el núcleo de B definida por (12). Entonces (9) implica inmediatamente

$$\forall i \in \mathfrak{S}_\Omega : P(\vec{e}_i) = \vec{e}_i, \quad (14)$$

i.e. P preserva la restricción de los vectores a \mathfrak{S}_Ω . Igualmente (10) implica que

$$P(\vec{1}_{\mathfrak{S}_\Gamma}) = \vec{1}_{\mathfrak{S}_\Gamma}. \quad (15)$$

Esta última propiedad nos da el resultado siguiente.

Téorema 3.1 Sea B una matriz de acoplamiento, y sea P la proyección definida por (12); sea $\vec{y}_{\mathfrak{S}_\Gamma} = \sum_{i \in \mathfrak{S}_\Gamma} y_i \vec{e}_i$. Entonces,

$$P(\vec{y}_{\mathfrak{S}_\Gamma}) = \sum_{i \in \mathfrak{S}_\Gamma} y_i \vec{w}_i, \quad (16)$$

donde $\sum_{i \in \mathfrak{S}_\Gamma} \vec{w}_i = \vec{1}_{\mathfrak{S}_\Gamma}$.

Demostación

Sea $\vec{y}_{\mathfrak{S}_\Gamma} = \sum_{i \in \mathfrak{S}_\Gamma} y_i \vec{e}_i$; entonces $P(\vec{y}_{\mathfrak{S}_\Gamma}) = \sum_{i \in \mathfrak{S}_\Gamma} y_i P(\vec{e}_i)$. Pongamos $\vec{w}_i = P(\vec{e}_i)$; luego:

$$\begin{aligned} \sum_{i \in \mathfrak{S}_\Gamma} \vec{w}_i &= \sum_{i \in \mathfrak{S}_\Gamma} P(\vec{e}_i) \\ &= P\left(\sum_{i \in \mathfrak{S}_\Gamma} \vec{e}_i\right) \\ &= P(\vec{1}_{\mathfrak{S}_\Gamma}) \\ &= \vec{1}_{\mathfrak{S}_\Gamma}. \end{aligned}$$

Corolario 3.1 Sea $\vec{y} = \sum_{i \in \mathfrak{S}_{\bar{\Omega}}} y_i \vec{e}_i$, $\vec{y}_{\mathfrak{S}_\Omega} = \sum_{i \in \mathfrak{S}_\Omega} y_i \vec{e}_i$, entonces:

$$P(\vec{y}) = \vec{y}_{\mathfrak{S}_\Omega} + \sum_{i \in \mathfrak{S}_\Gamma} y_i \vec{w}_i, \quad (17)$$

donde $\sum_{i \in \mathfrak{S}_\Gamma} \vec{w}_i = \vec{1}_{\mathfrak{S}_\Gamma}$.

En consecuencia la proyección sobre el núcleo de una matriz de acoplamiento es computacionalmente económica porque ella modifica solamente las componentes del vector que corresponden a la interfaz, dejando igual los componentes del vector que están fuera de la interfaz. Uno puede entonces esperar encontrar formulas convenientes para calcular P aprovechando la estructura particular de cada problema, como en el ejemplo siguiente:

Ejemplo: Caso conforme.

En el caso conforme B está definida como en (5),

$$B = (B_1 \ B_2); B_1 = (0 \ I_\Gamma); B_2 = (0 \ -I_\Gamma) \quad (18)$$

y B es de rango maximal, por lo tanto

$$B^+ = B^T (BB^T)^{-1} = \frac{1}{2} \begin{pmatrix} 0 \\ I_\Gamma \\ 0 \\ -I_\Gamma \end{pmatrix}, \quad (19)$$

$$R = B^+ B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & I_\Gamma/2 & 0 & -I_\Gamma/2 \\ 0 & 0 & 0 & 0 \\ 0 & -I_\Gamma/2 & 0 & I_\Gamma/2 \end{pmatrix} \quad (20)$$

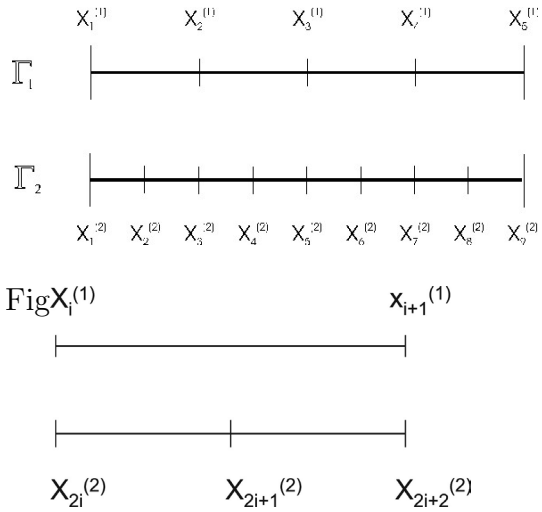


Figura 5: Segmento

y

$$P = I - B^+B = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I_\Gamma/2 & 0 & I_\Gamma/2 \\ 0 & 0 & I & 0 \\ 0 & I_\Gamma/2 & 0 & I_\Gamma/2 \end{pmatrix}. \quad (21)$$

Entonces, los únicos componentes del vector $R(y)$ corresponden a la mitad de los saltos en la interfaz y aquellas componentes del vector $P(y) = (I - R)y$ corresponden al promedio de los valores de cada nodo de la interfaz con el valor de su nodo vecino correspondiente. (cf. [HR99]).

Ejemplo: Caso no conforme. Nodos intercalados.

Supongamos que los nodos de la interfaz Γ_2 están intercalados entre los nodos de la interfaz Γ_1 como muestra la figura 4 (paso h constante en Γ_1 .) Supongamos que en cada interfaz Γ_i se define el espacio V_h^i de las funciones constantes a trozos, $i = 1, 2$. Tomamos $W_h = V_h^1$. En este caso, la matriz B viene dada por:

$$B = \begin{pmatrix} I & 0 & -I/2 & 0 & -I/2 & 0 \end{pmatrix}.$$

Ya que cada $w_i = P(e_i)$, podemos realizar localmente el cálculo de la matriz de proyección P sobre el núcleo de B , remitiéndonos a cualquier segmento de Γ_1 como indica la figura 5. En este caso definimos la matriz B local, que llamaremos \hat{B} :

$$\hat{B} = \begin{pmatrix} 1 & -1/2 & -1/2 \end{pmatrix}.$$

Luego,

$$\hat{B}^+ = \hat{B}^T(\hat{B}\hat{B}^T)^{-1} = \frac{1}{3h} \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix};$$

$$\hat{R} = \hat{B}^+ B = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1/2 & 1/2 \\ -1 & 1/2 & 1/2 \end{pmatrix};$$

y finalmente,

$$\hat{P} = I - R = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1/3 & 5/6 & -1/6 \\ 1/3 & -1/6 & 5/6 \end{pmatrix}.$$

Luego, el efecto de la proyección P sobre el núcleo de B , es un promedio ponderado de únicamente los valores de los nodos que están en el mismo segmento, y los valores de los nodos de los segmentos localmente adyacentes en la interfaz vecina.

4. Conclusiones

En la implementación de dichas proyecciones sobre máquinas paralelas, la comunicación no se hace sino sobre las componentes de los vectores que están en la interfaz. Si las interfaces son de tamaño relativamente pequeño, entonces el costo de las comunicaciones será razonable.

El resultado del teorema 3.1 puede aprovecharse para hacer mas eficiente el cálculo de la proyección sobre el núcleo de B de varias maneras. Por ejemplo, mediante el método de coeficientes indeterminados. Igualmente, dicho resultado puede emplearse para aprovechar ciertas características particulares en la disposición de los nodos del problema a resolver, como por ejemplo: nodos desfasados, nodos intercalados, etc. [HR00]. En el caso conforme se muestra que la proyección se realiza promediando punto a punto los valores de los nodos contiguos en las interfaces.

Referencias

- [BGL05] M. BENZI, G. GOLUB, and J. LIESEN. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

- [ELM02] H.C. ELMAN. Preconditioners for saddle point problems arising in computational fluid dynamics. *Appl. Numer. Math*, 43(4):75–89, 2002.
- [HR99] L.M. HERNANDEZ-RAMOS. Precondicionamiento por pseudoinversa para sistemas lineales acoplados de descomposicion de dominios, 1999.
- [HR00] L.M. HERNANDEZ-RAMOS. Un opérateur de projection pour les systèmes linéaires couplés (thèse pour obtenir le diplôme d'études doctorales), 2000.
- [HR02] L.M. HERNANDEZ-RAMOS. Projections alternées obliques pour les systèmes linéaires couplés. Technical Report R2024, Publications du Laboratoire Jacques-Louis Lions (Paris VI), 2002.
- [HR05] L.M. HERNANDEZ-RAMOS. Alternating oblique projections for coupled linear systems. *Numerical Algorithms*, 38(4):285–303, 2005.
- [JUV97] X. JUVIGNY. *Résolution des Grands Systèmes Linéaires sur des Machines Massivement Parallèles*. PhD thesis, Université Pierre et Marie Curie, Paris VI, 1997.
- [NEU50] J. VON NEUMANN. Functional operators vol. II. The geometry of orthogonal spaces. *Annals of Math. Studies*, 22, 1950. Princeton University Press. This is a reprint of mimeographed lecture notes first distributed in 1933.
- [QV99] A. QUARTERONI and A. VALLI. *Domain Decomposition Methods for Partial Differential Equations*. Oxford University Press, Oxford, 1999.
- [RE] M. RAYDAN and R. ESCALANTE. Alternating projections methods: Theory and application. Livre en Préparation.
- [SB80] J. STOER and R. BULIRSCH. *Introduction to Numerical Analysis*. Springer-Verlag, 1980.
- [XZ98] J. XU and J. ZOU. Some nonoverlapping domain decomposition methods. *SIAM Review*, 40:857–914, 1998.