



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA

Generalización del método de proyecciones alternantes oblicuas al caso de proyección sobre una variedad lineal

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela por la **Br. Olga C. Domínguez R.** para optar al título de Licenciada en Matemática.

Tutor: Luis M. Hernández R.

Caracas, Venezuela

Abril de 2009

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado “**Generalización del método de proyecciones alternantes oblicuas al caso de proyección sobre una variedad lineal**”, presentado por la **Br. Olga C. Dominguez R.**, titular de la Cédula de Identidad **14.166.918**, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de **Licenciada en Matemática**.

Luis M. Hernández R.
Tutor

Francisco Tovar
Jurado

Marcos Raydan
Jurado

Dedicatoria

Dedicada a mi Padre, que a pesar de no haber contado con su presencia física durante este largo trayecto universitario, muchas veces amargo, pero otras veces dulce, estuvo día tras día presente en mi mente y corazón. Que Dios lo tenga en la gloria

Y a ti Madre, con todo el amor del mundo.

Los amo profundamente.

Agradecimiento

Gracias especiales y preferentes a mi tutor Luis Manuel Hernández por toda esa inquebrantable paciencia y atención prestada durante el desarrollo de este trabajo. Gracias por apoyarme y hasta los muchos empujones necesario para terminar este trabajo.

Gracias al profesor Marcos Raydan también por su invaluable apoyo y colaboración que de forma directa e indirecta contribuyeron en la realización de esta tesis.

Gracias a la profesora Marisela Domínguez, por brindarme desinteresadamente esa parte humana que en muchas ocasiones pareciera, por involuntarios motivos, ser olvidada por los profesores. Gracias a Marisela por el apoyo dado mediante sus palabras de aliento y sabios consejos en mis flaquezas y al profesor Ramón Bruzual por haber sido partícipe de ese hecho.

Gracias a mi amada familia, mi madre Mirta Rivero, mi hermana Mary y mi refrescante sobrino Christian por acompañarme siempre, obsequiandome su cariño y alegrías y haber tenido paciencia y confianza para verme alcanzar esta meta. Gracias a mi abuelo Manuel Rivero, por su querencia y su apoyo económico tan oportuno para asistir a talleres que ayudaron en mi formación como profesional en esta área.

Gracias a mi pequeño gran mundo donde están todas esas personas que durante mi carrera me acompañaron y constantemente me apoyaron, me ayudaron a levantar en mis caídas, me hicieron reír, llorar, y sobre todo me hicieron este largo camino mucho más abordable y más amenos. A ti Ricardo e Isabel que permanecieron a mi lado en gran parte de mi carrera y me regalaron profusos momentos de alegría, comprensión y cariño. Gracias por sus consejos y sobre todo a ti Ricardo por alentarme y hacerme entrar en razón en todos esos lapsos críticos y en los que no veía luz con algunas materias de la carrera.

Gracias Hugo y Daniel L., mis amigos más antiguos de la facultad, a Carolina y Daniel H. por ser excelentes compañeros de estudio, a Gustavo, Ely y Marcos, por siempre hacerme reír y pasarla bien en todo el tiempo que compartí a su lado. Gracias a German, Marco, Hairol

y Marcelo por todos esos momentos de entretenimiento y distracción que hemos vivido y seguiremos compartiendo dentro y fuera de la universidad y sobre todo a ti Carlos Javier que siempre estuviste ahí presto a gozar mis alegrías y compartir mis pesares sin importar su dimensión y tener siempre las palabras oportunas para elevar el ánimo. Gracias por todos esos momentos gratos en que compartimos dichas y desdichas.

Gracias Juan Miguel, mi “ser humano favorito”, por darme todo ese afecto incondicional, por cada minuto de bienestar que me has regalado y por tu constante disposición para ayudarme y colaborar no sólo con la finalización de éste trabajo, sino con las demás labores de mi cotidianidad.

“Gracias a la vida” como dice la canción porque me ha dado salud a mi y a mi familia para poder culminar esta etapa y continuar con las que vienen, por darme los recursos necesarios que me permitieron cubrir mis necesidades económicas de mi entorno, gracias por todas esas oportunidades laborales que fluyeron a mi y que conjuntamente con mi formación académica, han hecho posible lo que hoy en día soy. Gracias a ese gran Dios por poner en mi camino a toda esa gente maravillosa que supieron confiar en mi y valorarme por todos aquellos aspectos realmente importantes en la vida de un ser humano y que además me dieron a conocer la verdadera felicidad.

Índice general

Introducción	1
Capítulo 1. Sistemas de punto de ensilladura	3
1. Introducción a los problemas de punto de ensilladura	3
2. Propiedades de las matrices de punto de ensilladura	5
3. Resolución numérica de los sistemas de punto de ensilladura	9
4. Esquemas estacionarios	13
5. Métodos no estacionarios y de subespacios de Krylov	16
6. Precondicionamiento	21
Capítulo 2. Método de proyecciones alternantes (MAP)	25
1. Definición del método MAP	27
2. Velocidad de convergencia	29
3. Método de proyecciones oblicuas alternadas aplicado a sistemas acoplados	32
Capítulo 3. Generalización del método de proyecciones oblicuas alternadas (PAO)	37
1. Nociones básicas	37
2. Definición del problema	37
3. Experimentación numérica	41
4. Resultados	44
Conclusiones	62
Bibliografía	63

Introducción

Los sistemas de punto de ensilladura son sistemas lineales por bloque 2×2 que pueden presentar diversas características según las propiedades que tengan las matrices por bloque que lo definen. Estos sistemas escritos en su forma mas general presentan la siguiente estructura:

$$(0.1) \quad \begin{pmatrix} A & B_1^T \\ B_2 & -C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

donde

$$A \in R^{n \times n}, \quad B_1, B_2 \in R^{m \times n}, \quad C \in R^{m \times m} \text{ con } n \geq m.$$

En vista que cualquier sistema en un principio podría ser visto con una estructura como la anteriormente planteada, se consideran como sistema de punto de ensilladura aquellos cuyos bloques de matrices cumplan una o más de las siguientes características:

(P1): A es una matriz simétrica: $A = A^T$.

(P2): La parte simétrica de A , $H \equiv \frac{1}{2}(A + A^T)$, es positiva semidefinida.

(P3): $B_1 = B_2 = B$

(P4): C es simétrica ($C = C^T$) y positiva semidefinida.

(P5): $C = 0$

Uno de los tipos de problemas de punto de ensilladura que más ocurren en aplicaciones prácticas es cuando todas las hipótesis dadas anteriormente se cumplen, y se tiene por lo tanto un sistema lineal simétrico. Para este caso, recientemente, L.M. Hernández-Ramos [20] propuso un método bastante eficiente, denominado método de proyecciones alternantes oblicuas. Este método fue propuesto para los problemas simétricos de punto de ensilladura (0.1) donde la componente g del vector del lado derecho es igual a cero. En dicho artículo, se prueba que la solución del problema de punto de ensilladura con $g = 0$ es el único punto en la intersección de una variedad lineal y un subespacio; y se propone un método

basado en proyectar alternadamente entre ambos para generar una sucesión que converge a la solución, utilizando una proyección oblicua para la variedad lineal y una proyección ortogonal en el subespacio. Sin embargo, en muchas aplicaciones prácticas es necesario resolver este problema cuando $g \neq 0$. Para ello, en este trabajo proponemos una generalización al método PAO adaptado a este caso particular.

En [20] se proponen dos versiones del método PAO. En dicho artículo se prueba que el método de proyecciones alternantes oblicuas equivale al método de mínimo descenso preconditionado aplicado al sistema del complemento de Schur (sistema condensado) $BA^{-1}B^T y = BA^{-1}f$. Tomando este preconditionador, se propone luego una nueva versión basada en el método de los gradientes conjugados preconditionado. Posteriormente, [8], propone una versión del método de Barzilai Borwein para estos sistemas condensados. Sin embargo, Dai et al. en este artículo no realiza pruebas preconditionando el método de Barzilai-Borwein con el preconditionador propuesto por L.M. Hernández-Ramos [20]. Como segunda parte de este trabajo, se realizarán pruebas numéricas de la generalización del método PAO y GC-PAO al caso donde $g \neq 0$ y se comparará su rendimiento con los métodos de Uzawa, Barzilai-Borwein e igualmente con una versión del método de Barzilai-Borwein que utiliza el preconditionador propuesto en [20].

Sistemas de punto de ensilladura

En este capítulo, se tocarán los diferentes aspectos ligados a la caracterización y resolución de problemas de puntos de ensilladura. Para ello, nos basaremos principalmente en los artículos de M. Benzi, G. Golub y J. Liesen [3], un pequeño resumen realizado por M. Rozložnik [29], así como también otros artículos citados en la bibliografía, principalmente el artículo de Luis M. Hernández [20].

1. Introducción a los problemas de punto de ensilladura

Desde hace algunos años ha existido mucho interés en la resolución numérica de grandes sistemas lineales llamados puntos de ensilladura (Saddle Point Problems), motivado a numerosas aplicaciones en las ramas técnicas y científicas que conducen a este tipo de problemas. Por ejemplo, entre las aplicaciones de este tipo de problemas se encuentran la resolución de formulaciones mixtas de problemas de mecánica de fluidos, problemas de descomposición de dominios y para algoritmos de optimización lineal y no lineal. En este trabajo, nos interesaremos particularmente en el caso donde el sistema de punto de ensilladura $Au = b$ tiene la siguiente estructura particular por bloques.

$$(1.1) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

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 positiva semidefinida en cuyo caso se trata de un sistema lineal simétrico.

Este sistema proviene de condiciones de optimización de primer orden de problemas de programación cuadrática:

$$(1.2) \quad \min J(x) = \frac{1}{2}x^T Ax - f^T x$$

$$(1.3) \quad \text{sujeto a: } Bx = g.$$

En este caso, la variable y de (1.1) representa el vector de multiplicadores de Lagrange. Toda solución (x_*, y_*) de (1.1) es un punto de ensilladura para el lagrangeano

$$(1.4) \quad \mathcal{L}(x, y) = \frac{1}{2}x^T Ax - f^T x + (Bx - g)^T y,$$

de donde proviene el nombre “problemas de punto de ensilladura” dado a (1.1). Recordemos que un punto de ensilladura es un punto $(x_*, y_*) \in \mathfrak{R}^{n \times m}$ que verifica

$$(1.5) \quad \mathcal{L}(x_*, y) \leq \mathcal{L}(x_*, y_*) \leq \mathcal{L}(x, y_*), \quad \forall x \in \mathfrak{R}^n, \forall y \in \mathfrak{R}^m,$$

o de una manera equivalente,

$$(1.6) \quad \min_x \max_y \mathcal{L}(x, y) = \mathcal{L}(x_*, y_*) = \max_y \min_x \mathcal{L}(x, y).$$

Entre las principales aplicaciones en las que se alcanza sistemas de puntos de ensilladura se pueden nombrar los problemas de dinámica de fluidos computacionales (Turek 1999, Wesseling 2001), estimación de mínimos cuadrados con restricciones (Bjorck 1996, Golub y Van Loan 1996), optimización con restricciones (Wright 1992, Wright 1997), aproximación por elementos finitos para EDP elípticas (Quarteroni y Valli 1994). En problemas como dinámica de fluidos o elasticidad lineal, el sistema de punto de ensilladura resulta de la discretización de sistemas de ecuaciones parciales con restricciones donde frecuentemente las restricciones representan leyes básicas de conservación. Debido a que las ecuaciones de punto de ensilladuras pueden ser derivadas de condiciones de equilibrio para sistemas físicos, estas ecuaciones también son llamadas ecuaciones de equilibrio. Ver Strang (1986, 1988). En literaturas sobre optimización, se hace referencia a este tipo de sistemas como “sistemas KKT” proveniente de los problemas de las condiciones de optimización de primer orden de Karush-Kuhn-Tucker. Sistemas como (1.1) también surgen de problemas de descomposición de dominio donde los multiplicadores de Lagrange son usados para asegurar la continuidad en las fronteras. Ver Chan y Mathew (1994), Farhat y Roux (1991), Hu, Shi y Yu (2004), L. M. Hernández-Ramos (2005).

2. Propiedades de las matrices de punto de ensilladura

En esta sección, se establecerán algunas propiedades de las matrices de punto de ensilladura

$$(1.7) \quad \mathcal{A} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix},$$

como la existencia de distintas factorizaciones, existencia de la inversa, propiedades espectrales y condicionamiento.

2.1. Factorización por bloques y complemento de Schur. Si la matriz A es invertible, la matriz de punto de ensilladura \mathcal{A} admite la siguiente factorización triangular por bloques:

$$(1.8) \quad \mathcal{A} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix},$$

donde $S = -BA^{-1}B^T$ es el complemento de Schur de A en \mathcal{A} . Numerosas propiedades importantes de la matriz de punto de ensilladura \mathcal{A} son derivadas de (1.8).

Por ejemplo, las siguientes factorizaciones LU equivalentes de la matriz \mathcal{A} son muy utilizadas:

$$(1.9) \quad \mathcal{A} = \begin{pmatrix} A & 0 \\ B & S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix}$$

y

$$(1.10) \quad \mathcal{A} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}.$$

2.2. Condición para la existencia de la inversa de la matriz de punto de ensilladura. Sea A una matriz no singular. La matriz \mathcal{A} es una matriz no singular si y sólo si, el complemento de Schur S es también no singular. Desafortunadamente, en general, no se puede decir mucho sobre la matriz de complemento de Schur $S = -BA^{-1}B^T$. Por lo tanto es necesario introducir algunas hipótesis sobre A y sobre B para realizar este análisis.

2.2.1. *Caso simétrico.* Comenzaremos por analizar el caso donde la matriz A es simétrica positiva definida. En este caso la matriz del complemento de Schur $S = -BA^{-1}B^T$ es una matriz simétrica negativa semidefinida. Luego es evidente que S y en consecuencia \mathcal{A} son matrices invertibles si, y sólo si, la matriz B es de rango completo ($Rang(B) = m$). En este caso la matriz de complemento de Schur S es simétrica negativa definida y los problemas de punto de ensilladura (1.1- 1.4) tienen solución única x_* .

En el caso en que la matriz A no sea positiva definida, para que la matriz \mathcal{A} sea invertible es suficiente que A sea positiva definida sobre $\ker B$, es decir que $x^T Ax \neq 0$ con $x \in \ker B$. De otra forma, si A es simétrica positiva semidefinida, el siguiente resultado es conocido [2]:

TEOREMA 1.1. Supongamos que A es simétrica positiva semidefinida, y B de rango completo ($Rang(B) = m$). Luego, una condición necesaria y suficiente para que la matriz \mathcal{A} sea no singular es que: $\ker A \cap \ker B = \{0\}$.

Ver demostración en [3].

2.2.2. *Caso general:* Para el caso no simétrico, es más difícil encontrar las condiciones suficientes para que \mathcal{A} sea invertible en función del kernel de A y de B . El siguiente teorema establece las condiciones suficientes para la existencia de una solución única en función de la parte simétrica de la matriz A , $H = \frac{1}{2}(A + A^T)$. [3]:

TEOREMA 1.2. Supongamos que la parte simétrica de A , $H = \frac{1}{2}(A + A^T)$ es positiva semidefinida y que B es de rango completo ($Rang(B) = m$). Luego,

- (1) $\ker(H) \cap \ker(B) = \{0\} \Rightarrow \mathcal{A}$ es invertible.
- (2) \mathcal{A} es invertible $\Rightarrow \ker A \cap \ker B = \{0\}$.

Ver demostración en [3].

2.3. Inversa de una matriz de punto de ensilladura. Sea A una matriz no singular. Luego la matriz \mathcal{A} es invertible si, y sólo si, la matriz del complemento de Schur $S = -BA^{-1}B^T$ es también no singular. En este caso, se tiene la siguiente expresión para la inversa de \mathcal{A}

$$(1.11) \quad \mathcal{A}^{-1} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B^T S^{-1}BA^{-1} & -A^{-1}B^T S^{-1} \\ -S^{-1}BA^{-1} & S^{-1} \end{pmatrix}.$$

y la expresión explícita de la solución (x_*, y_*) del problema (1.1) para el caso en que $g = 0$ sería

$$(1.12) \quad \begin{pmatrix} x_* \\ y_* \end{pmatrix} = \begin{pmatrix} (I + A^{-1}B^T S^{-1}B)A^{-1}f \\ -S^{-1}BA^{-1}f \end{pmatrix}.$$

La matriz $\Pi = -A^{-1}B^T S^{-1}B$ es una matriz de proyección. En efecto,

- $\Pi^2 = \Pi$
- $A\Pi w \in (\ker B)^\perp$
- $w - \Pi w \in \ker B$.

En este caso, Π representa una proyección oblicua sobre la imagen de $A^{-1}B^T$ y ortogonal a $\ker B^\perp$ [19, 20].

La primera componente del vector solución x_* se puede escribir como

$$(1.13) \quad x_* = (I - \Pi)x_c,$$

donde $x_c = A^{-1}f$ es la solución del problema sin restricciones (1.2).

Como A es una matriz positiva definida, entonces la matriz $I - \Pi$ se convierte en la matriz de proyección A -ortogonal sobre $\ker B$. En este caso, A^{-1} es también una matriz positiva definida, y la segunda componente del vector solución y_* es la solución del problema de mínimos cuadrados generalizados $B^T u \approx f$ con respecto a la norma inducida por A^{-1} $\|v\|_{A^{-1}} = (\langle v, v \rangle_{A^{-1}})^{1/2}$ [?]:

$$(1.14) \quad \|f - B^T y_*\|_{A^{-1}} = \min_u \|f - B^T u\|_{A^{-1}}.$$

Es posible dar una expresión para la inversa de \mathcal{A} donde la matriz A no es necesariamente invertible. Denotemos por $Z \in R^{n \times (n-m)}$ una matriz cuyas columnas forman una base ortonormal de $\ker B$. Si la parte simétrica H de A es positiva semidefinida, entonces la matriz $(n-m) \times (n-m)$ dada por $Z^T A Z$ es invertible dado que su parte simétrica $Z^T H Z$ es positiva definida. Si llamamos $W = Z(Z^T A Z)^{-1} Z^T$, tenemos la expresión siguiente para la inversa de \mathcal{A} :

$$(1.15) \quad \mathcal{A}^{-1} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}^{-1} = \begin{pmatrix} W & (I - WA)B^+ \\ (B^+)^T(I - AW) & -(B^+)^T(A - AWA)B^+ \end{pmatrix}$$

donde $B^+ = B^T(BB^T)^{-1}$ es la matriz pseudo-inversa de B .

Observación: $B^T(BB^T)^{-1}B = I - ZZ^T$ es la matriz de la proyección ortogonal sobre

$(\ker B)^\perp$. Más adelante esta matriz se denotará como $R = B^+B$ [18].

Esta expresión explícita de \mathcal{A}^{-1} tiene una utilidad práctica computacional muy limitada, debido a esto su interés es principalmente teórico.

2.4. Propiedades espectrales de matrices de punto de ensilladura. En esta sección, se enunciarán algunos resultados sobre los autovalores de las matrices de punto de ensilladura. Estos resultados son interesantes a la hora de resolver los problemas a través de métodos iterativos.

TEOREMA 1.3. Sea A una matriz simétrica definida positiva y B una matriz de rango completo. Sean μ_1 y μ_n el más grande y el más pequeño valor propio de A respectivamente, y sean σ_1 y σ_m el valor singular más grande y más pequeño de B . Si denotamos por $\sigma(\mathcal{A})$ el espectro de \mathcal{A} . Entonces

$$(1.16) \quad \sigma(\mathcal{A}) \subset I^- \cup I^+$$

donde

$$(1.17) \quad I^- = \left[\frac{1}{2} \left(\mu_n - \sqrt{\mu_n^2 + 4\sigma_1^2} \right), \frac{1}{2} \left(\mu_1 - \sqrt{\mu_1^2 + 4\sigma_m^2} \right) \right]$$

y

$$(1.18) \quad I^+ = \left[\mu_n, \frac{1}{2} \left(\mu_1 + \sqrt{\mu_1^2 + 4\sigma_1^2} \right) \right]$$

2.5. Condicionamiento. Los problemas de punto de ensilladura provenientes de casos prácticos suelen ser mal condicionados, por lo tanto en esos casos se puede aprovechar los beneficios de la forma de la matriz \mathcal{A} para disminuir el efecto de su mal condicionamiento. En el caso donde el bloque A es simétrico positivo definido y B es de rango completo, la matriz \mathcal{A} es también simétrica y el número de condición viene dado por :

$$(1.19) \quad \kappa(\mathcal{A}) = \frac{\max |\lambda(\mathcal{A})|}{\min |\lambda(\mathcal{A})|}$$

Del teorema (1.3), podemos deducir que el condicionamiento de \mathcal{A} crece a medida que $\mu_n = \lambda_{\min}(A)$ o cuando $\sigma_{\min}(B)$ tiende a cero (suponiendo $\lambda_{\max}(A)$ y $\sigma_{\max}(B)$ constantes). Este incremento del número de condición de \mathcal{A} deteriora la velocidad de convergencia de la mayoría de los métodos iterativos, como por ejemplo para los métodos de Krylov.

En el uso de métodos directos basados en factorización triangular para la solución de estos

sistemas se ha visto (ver Bjorck, 1996) que multiplicar por un escalar positivo a la matriz bloque A mejora el comportamiento de estos métodos, procedimiento que puede ser visto como un preconditionamiento. Las técnicas del Lagrangiano aumentado constituyen una buena aproximación para tratar los problemas con mal condicionamiento (c.f. [16]). Aquí se supone $A = A^T$ (eventualmente singular) y B de rango completo. La idea es reemplazar el sistema de punto de ensilladura (1.1) por el sistema equivalente,

$$(1.20) \quad \begin{pmatrix} A + B^T W B & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f + B^T W g \\ g \end{pmatrix}.$$

La matriz $W_{m \times m}$ debe ser una matriz simétrica positiva semidefinida seleccionada convenientemente. La escogencia más simple es tomar $W = \gamma I$ ($\gamma > 0$). En este caso, si la matriz A es positiva definida, entonces el bloque (1, 1) de (1.20) es no singular y positivo definido. El objetivo en la elección de W , es volver el nuevo sistema (1.20) más fácil de resolver con métodos iterativos que el sistema original. Cuando se utiliza $W = \gamma I$, la elección $\gamma = \|A\|_2 / \|B\|_2^2$ proporciona frecuentemente buenos resultados en casos prácticos en el sentido de que el número de condición de la matriz (1.20) es de una manera aproximada minimizado.

3. Resolución numérica de los sistemas de punto de ensilladura

Además de la clasificación actual de los métodos numéricos de resolución de sistemas lineales en métodos directos y métodos iterativos, los métodos para los problemas de punto de ensilladura se clasifican también en: métodos acoplados y métodos desacoplados [3]. Los métodos desacoplados permiten calcular los vectores desconocidos x y y de forma separada, es decir, se calcula primero x para luego calcular a partir de éste el vector y , o viceversa. Esta forma de resolución encuentra la solución de dos sistemas lineales de menor tamaño, conocidos como sistemas reducidos, en este caso de tamaño $m \times m$, una para hallar x y otro para y . Los métodos desacoplados pueden ser directos, iterativos o bien una combinación de ambos. Por ejemplo, uno de los sistemas reducidos puede ser resuelto por un método directo y el otro por uno iterativo según convenga tomando en cuenta las propiedades de las matrices. Uno de los métodos desacoplados más representativos es el método de reducción al complemento de Schur.

A diferencia de los métodos desacoplados, los métodos acoplados resuelven el sistema (1.1) de una forma global. Los vectores desconocidos x y y son calculados simultáneamente, sin

utilizar de manera explícita los sistemas reducidos. Entre los métodos acoplados se encuentran los métodos directos basados en una factorización de la matriz global \mathcal{A} , y los métodos iterativos de Krylov aplicados al sistema global (1.1), utilizando algún preconditionador. Generalmente los preconditionadores utilizados para los métodos acoplados se basan en los preconditionadores para los métodos desacoplados.

3.1. Reducción al complemento de Schur. Se considera el problema (1.1), donde

$$(1.21) \quad Ax + B^T y = f, \quad Bx = g.$$

Si se supone A y \mathcal{A} de (1.1) no singulares, entonces la matriz del complemento de Schur $S = -BA^{-1}B^T$ es también no singular. Si se multiplica ambos lados de la primera ecuación (1.21) por BA^{-1} se obtiene

$$(1.22) \quad Bx + BA^{-1}B^T y = BA^{-1}f,$$

si se utiliza $Bx = g$ se tiene

$$(1.23) \quad BA^{-1}B^T y = BA^{-1}f - g,$$

que es un sistema reducido de orden m para la variable y . Este sistema se denomina sistema condensado, donde la matriz $BA^{-1}B^T$ es el negativo del complemento de Schur $S = -BA^{-1}B^T$.

Si se calcula y_* de la expresión (1.23), el vector x_* se puede luego obtener de la resolución de

$$(1.24) \quad Ax_* = f - B^T y_*,$$

que es otro sistema reducido de orden n para x , el cual utiliza el bloque (1,1) de \mathcal{A} (la matriz A). Esta es justamente la aplicación de la eliminación de Gauss por bloques aplicada al bloque (1,1) de la matriz \mathcal{A} . Esto quiere decir que si se utiliza la factorización LU por bloque (1.10), se obtiene el sistema transformado:

$$(1.25) \quad \begin{pmatrix} I & 0 \\ -BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} I & 0 \\ -BA^{-1} & I \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix},$$

que es equivalente a

$$(1.26) \quad \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ -BA^{-1}f + g \end{pmatrix}.$$

La resolución de este sistema triangular por bloques, por sustitución regresiva, equivale a resolver los sistemas (1.23) y luego (1.24), para obtener x y y . Estos sistemas pueden ser resueltos por métodos directos o por métodos iterativos. En el caso particular donde A y $-S$ son positivas definidas, se puede utilizar métodos muy eficaces como la factorización de Cholesky o bien el método de gradiente conjugado.

Esta forma de resolución se vuelve atractiva si el orden m del sistema reducido es pequeño, y si los sistemas lineales que utilizan la matriz A pueden ser resueltos eficazmente. Un inconveniente de esta aproximación viene del hecho que la matriz A debe ser no singular y principalmente por el hecho que la matriz de complemento de Schur $S = -BA^{-1}B^T$ puede ser densa y por lo tanto ser muy costoso calcular una factorización. Desafortunadamente, son muchos los casos en los que la matriz del complemento de Schur S son densas. Se pueden tener también problemas de inestabilidades numéricas a la hora de construir la matriz S , sobre todo cuando la matriz A está mal condicionada.

En el caso en donde A es positiva semidefinida y singular, los métodos de reducción al complemento de Schur pueden ser aplicados con técnicas del lagrangiano aumentado que reemplaza al sistema original (1.1) por el sistema equivalente (1.20), pero en cuyo caso el bloque (1, 1) es no singular.

Si la matriz S es muy costosa de calcular, la reducción al sistema condensado puede ser obtenida a través de la resolución de (1.23) por un método iterativo que utilice sólo los productos matriz - vector $Sy = -BA^{-1}B^Ty$ que no requieren de una construcción explícita de S .

En efecto, para los métodos iterativos, el vector residual asociado a un y cualquiera está dado por

$$(1.27) \quad r_y = BA^{-1}f - Sy = BA^{-1}(f - B^Ty) - g.$$

Si se define $x = A^{-1}(f - B^Ty)$ como en (1.24), entonces el vector residual se convierte en $r_y = Bx - g$, el cual utiliza sólo un producto matriz vector con la matriz B . Esta observación a inspirado métodos tales como el método de Uzawa. En el caso en el que el complemento de

Schur esté bien condicionado, esta es una opción a tomar en consideración. De lo contrario, cuando el sistema del complemento de Schur está mal condicionado, se necesita emplear técnicas de preconditionamiento. El preconditionamiento del sistema condensado (1.23) no es algo trivial, cuando la matriz S no está dada explícitamente.

4. Esquemas estacionarios

En esta sección se abordaran los métodos iterativos llamados “estacionarios”. Entre los métodos estacionarios se encuentran los primeros métodos utilizados con éxito para resolver problemas de punto de ensilladura, y que aún continúan siendo sujetos de estudio para utilizarlos con preconditionadores de métodos de tipo Krylov y también para el desarrollo de versiones inexactas [?].

4.1. Método de Uzawa. Los métodos de tipo Uzawa fueron propuestos originalmente por los matemáticos economistas Arrow, Hurwicz y Uzawa. Estos métodos consisten en un esquema estacionario con iteraciones simultáneas para las variables x y y que pueden ser expresados en términos de factorización de la matriz \mathcal{A} . A este método también se le puede dar igualmente una interpretación en términos de métodos iterativos clásicos aplicados a los sistemas condensados

$$(1.28) \quad \begin{cases} BA^{-1}B^T y = BA^{-1}f \\ Ax = f - B^T y, \end{cases}$$

Los métodos de tipo Uzawa son clasificados como métodos acoplados pero algunas veces también se han considerado como métodos desacoplados.

Este método a tenido mucha popularidad para resolver problemas de dinámica de fluidos, especialmente para resolver los problemas estacionarios de Stokes.

Supongamos A invertible. Sean los vectores iniciales x_0 y y_0 , el método de Uzawa está dado por la siguiente iteración:

$$(1.29) \quad \begin{cases} Ax_{k+1} = f - B^T y_k \\ y_{k+1} = y_k + \omega(Bx_{k+1} - g), \end{cases}$$

donde ω es un parámetro de relajación. Podemos ver que este esquema iterativo puede ser escrito en términos de una partición de la matriz $\mathcal{A} = \mathcal{P} - \mathcal{Q}$ (ver Golub y Overton, 1988) y de la iteración de punto fijo

$$(1.30) \quad \mathcal{P}u_{k+1} = \mathcal{Q}u_k + b,$$

donde

$$(1.31) \quad \mathcal{P} = \begin{pmatrix} A & 0 \\ B & -\frac{1}{\omega}I \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} 0 & -B^T \\ 0 & -\frac{1}{\omega}I \end{pmatrix}, \quad \text{y} \quad u_k = \begin{pmatrix} x_k \\ y_k \end{pmatrix}.$$

Aquí la matriz de iteración está dada por

$$(1.32) \quad \mathcal{T} = \mathcal{P}^{-1}\mathcal{Q} = \begin{pmatrix} 0 & -A^{-1}B^T \\ 0 & I - \omega BA^{-1}B^T \end{pmatrix},$$

y todos los valores propios de \mathcal{T} son reales (y al menos n de ellos son iguales a cero).

Por otro lado, si se utiliza la primera ecuación de (1.29) para eliminar x_{k+1} de la segunda ecuación, se obtiene

$$(1.33) \quad y_{k+1} = y_k + \omega(BA^{-1}f - BA^{-1}B^T y_k - g),$$

así el método de Uzawa se convierte en el método estacionario de Richardson aplicado al sistema condensado

$$(1.34) \quad BA^{-1}B^T y = BA^{-1}f - g.$$

Si denotamos como el menor y el mayor valor propio de $BA^{-1}B^T$ como λ_{min} y λ_{max} respectivamente, decimos que el método de Richardson converge para todo ω que cumpla

$$(1.35) \quad 0 < \omega < \frac{2}{\lambda_{max}}.$$

Además, el radio espectral de la matriz de iteración $I - \omega BA^{-1}B^T$ es mínimo para el valor

$$(1.36) \quad \omega_* = \frac{2}{\lambda_{min} + \lambda_{max}}.$$

El método de Uzawa funciona bien para ciertas discretizaciones de problemas de Stokes estacionarios, donde la matriz de complemento de Schur es espectralmente equivalente a la matriz identidad. En este caso, los valores propios del complemento de Schur están acotados superior e inferiormente por constantes positivas y la velocidad de convergencia no dependen del tamaño de la malla h . Este no es el caso para otros tipos de problemas, como por ejemplo los problemas de generalización provenientes de solución de problemas de Stokes no estacionarios empleando métodos implícitos. La convergencia de Uzawa para este caso es más lenta, en particular para los problemas de Oseen con un número de Reynolds muy grande (para una viscosidad pequeña). Se puede mejorar la convergencia con la ayuda de un preconditionador adecuado, pero conseguir un buen preconditionador usualmente no es una tarea fácil.

La aplicación del método de Uzawa requiere resolver un sistema lineal con la matriz A en cada iteración. Pero este sistema puede ser resuelto bien sea por un método directo o más

frecuentemente por una iteración interna. En el caso en el que se utilice un solucionador inexacto, se emplea una aproximación más simple para la resolución de dicho sistema en cada iteración [?].

El método de Arrow-Hurwicz es una alternativa menos costosa que el método de Uzawa, pues en algunos casos resolver el sistema lineal con A puede ser muy costoso. Se considera la función objetivo

$$(1.37) \quad \phi(x) = \frac{1}{2}x^T Ax - x^T(f - B^T y_k).$$

Cuando en el método de Uzawa se calcula $Ax_{k+1} = f - B^T y_k$, el iterado x_{k+1} minimiza la función objetivo (1.37). Para derivar un método menos costoso, se puede tomar un solo paso en la dirección del gradiente negativo de $\phi(x)$, con un paso fijo α . El método resultante está dado por,

$$(1.38) \quad \begin{cases} x_{k+1} = x_k + \alpha(f - Ax_k - B^T y_k) \\ y_{k+1} = y_k + \omega(Bx_{k+1} - g), \end{cases}$$

Al igual que el método de Uzawa, el método de Arrow-Hurwicz también puede ser visto como una iteración de punto fijo, dada por la partición $\mathbf{A} = \mathcal{P} - \mathcal{Q}$ donde

$$(1.39) \quad \mathcal{P} = \begin{pmatrix} \frac{1}{\alpha}I & 0 \\ B & -\frac{1}{\omega}I \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} \frac{1}{\alpha}I - A & -B^T \\ 0 & -\frac{1}{\omega}I \end{pmatrix},$$

La convergencia de este algoritmo depende de dos parámetros de relajación, α y ω . La velocidad de convergencia del método de Arrow-Hurwicz es usualmente lenta. Para mejorarla, gran cantidad de variantes se han propuesto, entre los cuales se puede encontrar la versión preconditionada

$$(1.40) \quad \begin{cases} Ax_{k+1} = x_k + \alpha Q_A^{-1}(f - Ax_k - B^T y_k) \\ y_{k+1} = y_k + \omega Q_B^{-1}(Bx_{k+1} - g), \end{cases}$$

donde las matrices Q_A y Q_B son matrices de preconditionamiento seleccionadas convenientemente. Se puede ver también esta versión preconditionada del método de Arrow-Hurwicz como una versión inexacta del método de Uzawa.

5. Métodos no estacionarios y de subespacios de Krylov

Entre los métodos no estacionarios, los métodos basados en los subespacios de Krylov se encuentran entre los más populares.

Sea el sistema lineal

$$(1.41) \quad Ax = b,$$

donde A es una matriz no singular y b un vector. Los métodos basados en subespacios de Krylov generan una serie de aproximaciones de la forma

$$(1.42) \quad x_n = x_0 + K_n(A, r_0),$$

donde $r_0 = b - Ax_0$ es el vector residual inicial, $K_n(A, r_0)$ es el n -ésimo subespacio de Krylov asociado con A y r_0 definido por

$$(1.43) \quad K_n(A, r_0) = \text{span}\{r_0, Ar_0, \dots, A^{n-1}r_0\}.$$

A partir de (1.42) se puede deducir que el error $x - x_n$ de una aproximación x_n está dado por la expresión $x - x_n = P_n(A)(x - x_0)$ y $r_n = P_n(A)(x - x_0)$ donde P_n es un polinomio de grado a lo sumo n que satisface $P_n(0) = 1$. La clase entera de estos polinomios se denota como Π_n . Entre los métodos basados en los subespacios de Krylov más conocidos se encuentra el método de gradiente conjugado (GC), introducido por Hestenes y Stiefel en 1952. Este método fue propuesto para resolver sistemas $Ax = b$ para el caso en que la matriz A es simétrica positiva definida. La solución aproximada general obtenida por GC verifica

$$(1.44) \quad \|x - x_n\|_A = \min_{u \in x_0 + K_n(A, r_0)} \|x - u\|_A.$$

Si utilizamos $x - x_n = P_n(A)(x - x_0)$ y una expresión de $x - x_0$ en términos de una base ortonormal de vectores propios de A , se obtiene una cota del error de la forma,

$$(1.45) \quad \|x - x_n\|_A \leq \min_{P \in \Pi_n} \max_{\lambda_i \in \sigma(A)} |P(\lambda_i)| \|x - x_0\|_A,$$

donde $\sigma(A)$ es el especto de la matriz A (conjunto de valores propios).

La velocidad de convergencia del método GC depende fuertemente de la distribución de los valores propios de A . También podemos encontrar una cota menos ajustada pero aún más

práctica, como por ejemplo la cota dada por la expresión

$$(1.46) \quad \frac{\|x - x_n\|_A}{\|x - x_0\|_A} \leq 2 \left(\frac{\sqrt{k(A) - 1}}{\sqrt{k(A) + 1}} \right)^n$$

donde $k(A)$ es el número de condición de la matriz A .

Cuando la matriz A no es simétrica positiva definida, se utilizan frecuentemente otros métodos de tipo Krylov menos eficientes que el GC como por ejemplo MINRES, GMRES, etc.

En el caso simétrico pero indefinido, se usa frecuentemente el método del mínimo residual (MINRES). Este método calcula una solución aproximada que minimiza la norma del vector residual sobre una serie de subespacios de Krylov

$$(1.47) \quad \|b - Ax_n\| = \min_{u \in x_0 + K_n(A, r_0)} \|b - Au\|$$

de manera análoga el método de gradiente conjugado. De igual forma se puede obtener una cota de la norma del vector residual de la forma

$$(1.48) \quad \|r_n\| \leq \min_{P \in \Pi_n} \max_{\lambda_i \in \sigma(A)} |P(\lambda_i)| \|r_0\|.$$

En los problemas en que A no sea positiva definida, es decir, que dicha matriz presenta valores propios negativos puede ser más difícil obtener una expresión práctica para acotar la velocidad de convergencia en función de los valores propios de la matriz, como en el caso del método GC.

Aún menos claro es el caso en que la matriz A no es simétrica. Para ello existen muchos métodos iterativos, pero no son usados del todo en la práctica. El método más analizado es el GMRES, creado por Saad y Schultz, que está basado sobre una minimización de la norma del vector residual.

5.1. Métodos de Krylov para problemas de punto de ensilladura. Para los problemas de punto de ensilladura, los métodos de Krylov pueden ser utilizados globalmente sobre los problemas $\mathcal{A}u = b$ como un método acoplado, o bien para resolver el problema de los sistemas reducidos a través de un método desacoplado.

Dentro de los métodos acoplados, los métodos de Krylov constituyen una aproximación de la solución de la forma

$$(1.49) \quad u_n = u_0 + K_n(\mathcal{A}, r_0)$$

donde \mathcal{A} es la matriz global del sistema (1.1). Entre los métodos desacoplados, es frecuente utilizar los métodos de Krylov para resolver el sistema condensado

$$(1.50) \quad BA^{-1}B^T y = BA^{-1}f - g.$$

PROPOSICIÓN 1.1. Si $A \in \mathfrak{R}^{n \times n}$ es una matriz simétrica definida positiva, entonces A^{-1} también lo es.

DEMOSTRACIÓN. Si A es una matriz simétrica entonces $(A^{-1})^T = A^{-1}$. Si consideramos $x \in \mathfrak{R}^{n \times 1}$ con $n \neq 0$ entonces

$$(1.51) \quad x^T A^{-1} x = x^T A^{-1} A A^{-1} x = (A^{-1} x)^T A (A^{-1} x) > 0$$

pues A es simétrica definida positiva, además $A^{-1} x \neq 0$. Se concluye así que A^{-1} es una matriz simétrica definida positiva. □

PROPOSICIÓN 1.2. Si $A \in \mathfrak{R}^{n \times n}$ es una matriz simétrica definida positiva y $B \in \mathfrak{R}^{m \times n}$ (con $n \geq m$) es de rango completo, entonces la matriz $BA^{-1}B^T$ es simétrica positivo definida.

DEMOSTRACIÓN. Es fácil ver que S es simétrica. Sea $x \in \mathfrak{R}^{n \times 1}$ con $n \neq 0$ entonces

$$(1.52) \quad x^T (BA^{-1}B^T) x = (B^T x)^T A^{-1} (A^T x) > 0$$

pues como se vio en la proposición anterior A^{-1} es simétrica definida positiva, además $A^T x \neq 0$ pues la matriz B es de rango completo lo que permite concluir que $BA^{-1}B^T$ es una matriz simétrica definida positiva. □

En vista de que el método de Uzawa equivale en efecto, al método de Richardson aplicado a dicho sistema condensado (1.50) y sabiendo que la matriz del complemento de Schur S es definida positiva podemos también desarrollar versiones Krylov del método de Uzawa. Para ello nos beneficiaremos del hecho que el residuo del sistema condensado para un iterado y_k cualquiera, como se vio anteriormente, está dado por

$$(1.53) \quad BA^{-1}f - BA^{-1}B^T y_k - g = Bx_k - g.$$

Nos aprovecharemos de esto para escribir una versión gradiente conjugado del método de Uzawa (GC-Uzawa).

Algoritmo GC-Uzawa

(1) **Inicialización:**

- Seleccionar $y_0 \in \mathfrak{R}^m$
- Resolver $Ax_0 = f - B^T y_0$
- $r_0 = Bx_0 - g$
- $d_0 = r_0$

(2) **Iteraciones: Para** $k = 0, \dots$ hasta convergencia

- Resolver $Az_k = B^T d_k$
- $\alpha_k = \langle r_k, r_k \rangle / \langle d_k, Bz_k \rangle$
- $y_{k+1} = y_k + \alpha_k d_k$
- $r_{k+1} = r_k - \alpha_k Bz_k$
- $x_{k+1} = x_k - \alpha_k z_k$
- $\beta_{k+1} = \langle r_{k+1}, r_{k+1} \rangle / \langle r_k, r_k \rangle$
- $d_{k+1} = r_{k+1} + \beta_{k+1} d_k$

(3) **fin**

5.2. Método de Barzilai-Borwein aplicado a sistemas de punto de ensilladura.

Este método, es un método de tipo gradiente en el que Barzilai y Borwein propusieron un nuevo tamaño de paso para la dirección de búsqueda que mejora notablemente el comportamiento de este tipo de métodos, ante el mal condicionamiento que muy frecuentemente presentan los problemas de punto de ensilladura.

Barzilai y Borwein proponen un tamaño de paso α_k de la siguiente forma

$$(1.54) \quad \alpha_k = \frac{\langle s_{k-1}, s_{k-1} \rangle}{\langle s_{k-1}, y_{k-1} \rangle}$$

donde $s_{k-1} = x_k - x_{k-1}$ y $y_{k-1} = g_k - g_{k-1}$.

Para mayor detalle ver [8].

Barzilai y Borwein en 1988 realizan un análisis de convergencia para el caso de funciones cuadráticas en dos variables, pero años después Raydan en [26] prueba la convergencia global al caso de funciones cuadráticas estrictamente convexas para más de dos variables.

En la práctica computacional se ha visto que el método de Barzilai y Borwein es muy eficiente, razón por la cual se ha considerado a modo de comparación con el comportamiento del método del gradiente conjugado y el método GC- Uzawa aplicados a este problema.

El método de Barzilai-Borwein será aplicado al sistema condensado

$$(1.55) \quad BA^{-1}B^T y = BA^{-1}f - g$$

El algoritmo de Barzilai - Borwein adaptado a la resolución de problemas de punto de ensilladura viene dado por

Barzilai Borwein aplicado a sistemas de punto de ensilladura

(1) Inicialización:

- Seleccionar $x_0 \in V$
- Calcular g_0 inicial
- Calcular α_0 inicial

(2) Iteraciones: Para $k = 0, \dots$ hasta convergencia

- $s_k = \frac{-1}{\alpha_k} g_k$
- $x_{k+1} = x_k + s_k$
- $y_{k-1} = g_k - g_{k-1}$
- $\alpha_k = \langle s_{k-1}, s_{k-1} \rangle / \langle s_{k-1}, y_{k-1} \rangle$

(3) fin

Cabe destacar que este método, a diferencia de los demás métodos clásicos de tipo gradiente, es un método que no es de descenso monótono pues su dirección de búsqueda $-\alpha_k g_k$ no se trata de una sucesión monótona decreciente. Este método en teoría no garantiza el descenso de la función a minimizar $f(x)$ sin embargo, para el caso en que la función es una cuadrática, como lo es en este caso, en [28] se probó la convergencia global, es decir que converge a un punto estacionario independientemente del iterado inicial.

6. Precondicionamiento

La convergencia y la confiabilidad de los métodos iterativos son a menudo mejoradas con la utilización de preconditionadores. Con el preconditionamiento, las direcciones de búsqueda de los métodos iterativos son modificadas por información proveniente de otro método directo o iterativo del operador del sistema. Este método debe ser efectuado de una manera simple y con un bajo costo computacional.

En muchos casos preconditionar equivale a transformar el sistema original $Au = b$, en un sistema equivalente $\tilde{A}u = \tilde{b}$ con mejores características espectrales para su resolución con métodos iterativos.

Un preconditionador es una matriz M ó (M^{-1}) que realiza esta transformación (ver por ejemplo [31]) premultiplicando o postmultiplicando el sistema original.

De manera general, existen dos formas de construir preconditionadores. Una de ellas está basada en técnicas puramente algebraicas, tales como la factorización incompleta, las inversas aproximadas o los métodos algebraicos multinivel. Estos preconditionadores necesitan poco conocimiento del problema y pueden ser aplicados de cierta manera como una caja negra.

La segunda forma consiste en desarrollar un preconditionador “a la medida” del problema particular a resolver. A diferencia de la forma de preconditionamiento mencionada anteriormente, para encontrar un preconditionador bueno por este camino se requiere el conocimiento muy específico de las características del problema, tales como el origen del problema, su geometría, el método de discretización empleado, etc. Por supuesto, mientras más información se tenga del problema, mejor será la calidad del preconditionador empleado.

6.1. Precondicionamiento para problemas de punto de ensilladura. Para los problemas de punto de ensilladura, la construcción de buenos preconditionadores requieren explorar la estructura por bloques del problema. Para los métodos acoplados, la matriz del sistema global de punto de ensilladura

$$(1.56) \quad \mathcal{A} = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$$

es una matriz no definida, y esto representa un problema para la búsqueda de un preconditionador del sistema.

En el caso donde la matriz \mathcal{A} es simétrica no definida, no está del todo claro que propiedades debe tener la matriz de preconditionamiento \mathcal{M} . Si \mathcal{M} es simétrica positiva definida el sistema preconditionado puede ser escrito como

$$(1.57) \quad \mathcal{M}^{-1/2} \mathcal{A} \mathcal{M}^{-T/2} (\mathcal{M}^{T/2} x) = \mathcal{M}^{-1/2} b$$

En este caso, se puede utilizar un método iterativo para resolver sistemas simétricos no definidos, como por ejemplo MINRES.

Si el preconditionador \mathcal{M} es no definido, $\mathcal{M}^{1/2}$ no existe y el preconditionamiento del sistema simétrico $\mathcal{A}u = b$ se convierte en un sistema preconditionado no simétrico

$$(1.58) \quad \mathcal{A} \mathcal{M}^{-1} \hat{u} = b, \quad \hat{u} = \mathcal{M}u.$$

El mismo problema se presenta si la matriz de punto de ensilladura \mathcal{A} o la matriz de preconditionadora \mathcal{M} son no simétricas. La matriz del sistema preconditionado $\mathcal{A} \mathcal{M}^{-1}$ es en general no simétrica, y se debe utilizar un método iterativo o simétrico, como por ejemplo GMRES.

6.2. Precondicionadores por bloques. En esta sección, consideraremos los preconditionadores diagonales por bloques y triangulares por bloques. La eficacia de estos esquemas de preconditionamiento depende de la calidad de aproximación de la matriz A y del complemento de Schur S disponible. Es por este motivo que la construcción de estos preconditionadores está ligada a la construcción de preconditionadores para los métodos desacoplados anteriormente descritos.

6.2.1. *Precondicionadores diagonales por bloques.* Los preconditionadores diagonales por bloques tienen la forma

$$(1.59) \quad \mathcal{M} = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix}$$

Normalmente, el bloque M_1 es un preconditionador para la matriz A y el bloque M_2 es un preconditionador para el negativo del complemento de Schur $-S = BA^{-1}B^T$ o para una aproximación $BM_1^{-1}B^T$. Cuando A es simétrica positiva definida, el bloque M_1 generalmente también es simétrico positivo definido. El bloque M_2 es normalmente positivo definido o negativo definido. En el caso donde el bloque M_2 es positiva definida, el sistema preconditionado se mantiene simétrico no definido.

Se puede mostrar (ver [22]) que si $M_1 = A$ y $M_2 = \pm BA^{-1}B^T$, es decir

$$(1.60) \quad \mathcal{M} = \begin{pmatrix} A & 0 \\ 0 & \pm S \end{pmatrix}$$

entonces la matriz del sistema preconditionado $\mathcal{A}\mathcal{M}^{-1}$ tiene solamente tres valores propios diferentes no nulos $\{1, \frac{1 \pm \sqrt{5}}{2}\}$, y en ese caso, los métodos de Krylov terminan en a lo sumo tres iteraciones. Por supuesto, en la práctica no se tiene las inversas de las matrices A y S exactamente. Sin embargo, con buenos preconditionadores de A y de S se espera un acercamiento de los valores propios del sistema preconditionado alrededor de $\{1, \frac{1 \pm \sqrt{5}}{2}\}$, y en consecuencia una velocidad de convergencia de los métodos de Krylov razonables.

6.2.2. *Precondicionadores triangulares por bloques.* La segunda gran clase de preconditionadores tienen la forma triangular por bloques

$$(1.61) \quad \mathcal{M} = \begin{pmatrix} M_1 & B^T \\ 0 & M_2 \end{pmatrix}$$

donde M_1 y M_2 son preconditionadores para A y para $BA^{-1}B^T$ respectivamente. Hay que remarcar que estos preconditionadores son no simétricos.

La versión exacta de estos preconditionadores es

$$(1.62) \quad \mathcal{M} = \begin{pmatrix} A & B^T \\ 0 & \pm BA^{-1}B^T \end{pmatrix}$$

que dan lugar a un sistema preconditionado

$$(1.63) \quad \mathcal{A}\mathcal{M}^{-1} = \begin{pmatrix} I & 0 \\ BA^{-1} & \pm I \end{pmatrix}$$

que tiene como máximo dos valores propios diferentes de cero $\{\pm 1\}$.

6.3. Aproximación del Complemento de Schur. La mayoría de los preconditionadores para los problemas de punto de ensilladura tienen la necesidad de construir un preconditionador o una inversa aproximada del complemento de Schur asociado. La construcción de estos preconditionadores es un proceso delicado, que depende en general de las características del problema a tratar. De hecho, una buena aproximación es fundamental para el desempeño de los métodos a utilizar.

La matriz del complemento de Schur $S = -BA^{-1}B^T$ es una matriz usualmente densa, que se evita construir explícitamente en los métodos iterativos. En la sección 5.1, se muestra que para los métodos de tipo Krylov, es posible construir el vector residual del sistema condensado asociado al sistema de punto de ensilladura, sin tener la necesidad de construir dicha matriz. La dificultad se presenta cuando se desea dar una aproximación a su inversa y más aún cuando se busca encontrar un preconditionador “paralelizable” y algebraico. En la literatura, existen gran cantidad de aproximaciones para preconditionar el complemento de Schur para los diversos problemas de punto de ensilladura. La mayoría depende fuertemente del problema a tratar.

6.3.1. *El Preconditionador $(B^+)^T AB^+$.* En años recientes, la aproximación a la inversa de S dada por $S^{-1} \approx (B^+)^T AB^+$, y sus variantes han sido utilizadas satisfactoriamente en los problemas de punto de ensilladura. Este preconditionador fue propuesto de manera independiente por H.C. Elman [12], para los problemas generados a partir de discretizaciones de la ecuación de Navier - Stokes, por O. Widlund y A. Klawonn [23], para el método FETI y también, de manera puramente algebraica por L. Hernández-Ramos en [20]

El preconditionador $(B^+)^T AB^+$ tiene una interpretación algebraica simple: Si la matriz B es invertible, la inversa del complemento de Schur viene dada por $(B^{-1})^T AB^{-1}$. Pero como en general la matriz B es una matriz rectangular no invertible (de rango completo), se propone construir una aproximación $S^{-1} \approx (B^+)^T AB^+$ que reemplaza su inversa por la matriz pseudoinversa denotada por $B^+ = B^T(BB^T)^{-1}$.

CAPÍTULO 2

Método de proyecciones alternantes (MAP)

A lo largo de este capítulo se hablará del método de proyecciones alternantes y sus variantes. Para su desarrollo, se consideró gran parte de la teoría y los resultados propuestos por R. Escalante y M. Raydan en el libro [27] que presenta un buen desarrollo de este tema y la teoría necesaria para abarcar los aspectos básicos para el estudio del método de proyecciones alternantes. Consideremos las siguientes definiciones:

DEFINICIÓN 2.1. Un espacio de Hilbert es un espacio vectorial con producto interno que es completo con respecto a la norma dada por el producto interno.

DEFINICIÓN 2.2. Dado un espacio de Hilbert H y una variedad lineal M de H el ortogonal de M es

$$(2.1) \quad M^\perp = \{z \in H \mid \langle z, x \rangle = 0 \ \forall x \in M\}$$

PROPOSICIÓN 2.1. Sean H un espacio de Hilbert y M una variedad lineal de H entonces M^\perp es un subespacio (cerrado) de H .

TEOREMA 2.1. Sean H un espacio de Hilbert y M un subespacio (cerrado) de H entonces para cada $x \in H$ existe un único $y \in M$ y un único $z \in M^\perp$ tal que $x = y + z$. Además

$$(2.2) \quad \|x\|^2 = \|y\|^2 + \|z\|^2; \langle x, y \rangle = \text{dist}(x; M) :$$

El vector y es llamado la proyección ortogonal de x sobre M y lo denotaremos por $P_M(x)$. Para la demostración ver [5].

DEFINICIÓN 2.3. Sea H un espacio de Hilbert y sea $X \subseteq H$. Un operador $T : X \rightarrow Y$ en H es una función que tiene uno o más valores Tx en H correspondiente a cada elemento x de S . Al conjunto S se le conoce como **dominio** del operador T y al conjunto $R(T) = \{Tx/x \in S\}$ se le conoce como **rango** de T

DEFINICIÓN 2.4. Sea $T : X \rightarrow Y$ un operador en un espacio de Hilbert H , con X e Y espacios normados. Se dice que T es acotado si existe un número M tal que $\|Tx\| \leq M\|x\| \quad \forall x \in X$. Si T es acotado, se define su norma como $\|T\| = \sup\{\frac{\|Tx\|}{\|x\|} : x \in X, x \neq 0\}$

DEFINICIÓN 2.5. Sea $T : X \rightarrow Y$ un operador en un espacio de Hilbert H , con X e Y espacios normados. El conjunto de todos los elementos (x, Tx) con $x \in D(T)$ y $Tx \in R(T)$ se conoce como gráfico de T .

Para el desarrollo de este trabajo se tiene particular interés en los operadores de proyección por lo tanto consideraremos solo este tipo de operadores para la teoría que desarrollaremos a continuación.

Recordemos ahora las definiciones que caracterizan a los operadores de proyección.

DEFINICIÓN 2.6. Sea $T : X \rightarrow Y$ un operador en un espacio de Hilbert H , se dice que T es un operador lineal si $G(T)$ es subespacio lineal de $H \times H$.

DEFINICIÓN 2.7. Un operador $T : X \rightarrow Y$ sobre un espacio de Hilbert H , es inyectivo si existe un único elemento Tx asociado a x para todo $x \in D(T)$.

La definición anterior mantiene que si T es un operador inyectivo y ocurre que $(x, y) \in G(T)$ y $(x, y') \in G(T)$ entonces $y = y' = Tx$.

Sea S un subespacio cerrado de H . Si $x_0 \in S$ es un vector tal que $x - x_0 \in S^\perp$, dicho vector se denomina la proyección ortogonal de x en S . La operación que denota la proyección de x en S se denota $P_S(x) = x_0$.

TEOREMA 2.2. Un operador P definido en un espacio de Hilbert H es un operador de proyección si y sólo si:

- (1) P es inyectivo, lineal y acotado con $D(P) = H$.
- (2) $\langle Px, y \rangle = \langle x, Py \rangle \quad \forall x, y \in H$ (P es autoadjunto) y
- (3) $P^2 = P$ (P es idempotente).

PROPOSICIÓN 2.2. Sea X un espacio vectorial y P una proyección en X . Sea $R(P)$ el rango de P y $\ker(P)$ su núcleo, entonces:

- $x \in R(P)$ si y sólo si $P(x) = x$.

- $\ker(P) = R(I - P)$.
- $I - P$ es una proyección.
- $R(P) = \ker(I - P)$.
- $X = R(P) + \ker(P)$.
- $X = R(P) + R(I - P)$.

1. Definición del método MAP

El método de proyecciones alternantes (MAP), en su formulación más simple, fue propuesto originalmente por John von Neumann [35] quien trató el problema de encontrar la proyección de un punto de ensilladura dados en un espacio de Hilbert, sobre la intersección de dos subespacios cerrados.

Tiempo después, Cheney et Goldstein [7] realizaron la extensión del análisis de un esquema propuesto por Von Neumann en el caso de dos conjuntos convexos cerrados. En particular, ellos establecieron la convergencia del método sobre hipótesis débiles.

Supongamos que se tienen dos conjuntos convexos M y $N \in R^n$ y denotemos P_M y P_N las proyecciones sobre dichos conjuntos respectivamente. El algoritmo comienza con un valor inicial cualquiera $x_0 \in M$ y a partir de éste se realizan proyecciones sobre M y N de forma alternada de la siguiente forma:

$$x_k = P_N P_M x_{k-1} \text{ para } k = 1, 2, \dots$$

Sea H un espacio de Hilbert y sea M un subespacio cerrado H , la proyección ortogonal sobre M se denota por P_M . En particular, P_M es lineal, autoadjunta ($P_M^* = P_M$), idempotente ($P_M^2 = P_M$), y $P_M(x)$ es la mejor aproximación (o el punto más cercano) de x a M :

$$(2.3) \quad \|x - P_M(x)\| = d(x, M),$$

donde $d(x, M) = \inf\{\|x - y\| \mid y \in M\}$.

P_M se denomina la proyección ortogonal sobre M por causa de la propiedad que la define:

$$(2.4) \quad \langle x - P_M(x), y \rangle = 0 \quad \forall y \in M$$

Esto quiere decir que, $x - P_M(x)$ es ortogonal a M . Con la notación

$$(2.5) \quad M^\perp := \{y \in H \mid \langle x, y \rangle = 0 \quad \forall x \in M\},$$

esto se escribe $x - P_M(x) \in M^\perp$.

El teorema que sigue a continuación, es fundamental en la idea principal del método MAP.

TEOREMA 2.3. Sean M y N dos subespacios cerrados de H . Luego, $P_M P_N = P_N P_M$ si, y sólo si, $P_M P_N = P_{M \cap N}$

Esto quiere decir que, P_M y P_N conmutan si y sólo si, su composición es también una proyección ortogonal.

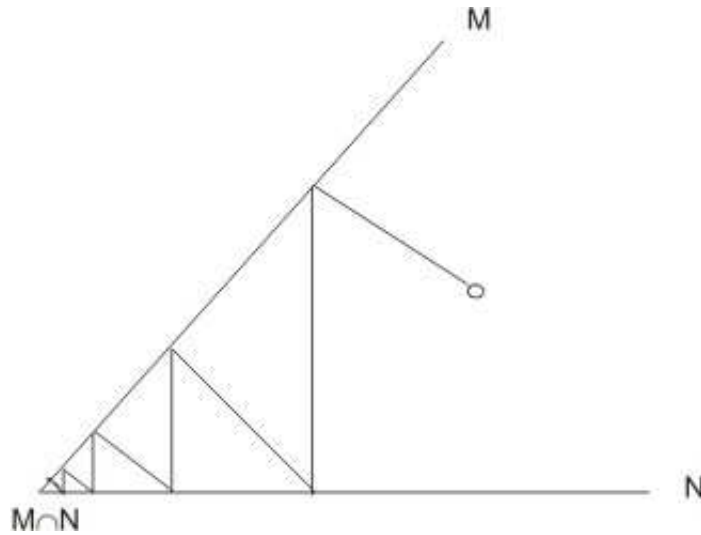


FIGURA 2.1. Método MAP

1.1. El método de Von Neumann. Von Neumann [35] se interesó por el caso en que P_M y P_N no conmutan. Él probó el siguiente resultado:

TEOREMA 2.4.

$$(2.6) \quad \lim_{n \rightarrow \infty} (P_M P_N)^n x = P_{M \cap N} x$$

La ecuación 2.6 dio origen a un algoritmo conocido como el método de proyecciones alternantes (MAP):

Algoritmo MAP

(1) **inicialización:**

- $x_0 = x \in H$

(2) **iteración:** Para $k = 0, 1, \dots$, **Hacer**

$$\bullet x_n = P_M P_N x_{n-1}$$

(3) **Fin**

Geoméricamente el método MAP busca encontrar la mejor aproximación x en $A \cap B \subset H$ proyectando x sobre A , luego proyectando el resultado sobre B , y se continua proyectando el resultado de forma alternada sobre A y sobre B . La sucesión generada converge a $P_{A \cap B}$. La utilidad práctica del algoritmo MAP es que muchas veces es más fácil calcular las proyecciones sobre A ó B individualmente que calcular directamente la proyección sobre $A \cap B$.

1.1.1. *Una formulación equivalente.* Si se cambia M por M^\perp y N por N^\perp en el teorema 2.4, y si utilizamos el hecho de que $P_{M^\perp} = I - P_M, P_{N^\perp} = I - P_N$, y $M^\perp \cap N^\perp = \overline{(N + M)}^\perp$, se obtiene la reformulación siguiente del teorema 2.4:

TEOREMA 2.5. Para cada $x \in H$,

$$(2.7) \quad \lim_{n \rightarrow \infty} [(I - P_M)(I - P_N)]^n x = I - P_{\overline{(M+N)}} x$$

1.1.2. *Extensión de Halperin.* La extensión del teorema (2.4) a más de dos subespacios es posible. Esta fue desarrollada por Halperin [17].

Sean A_1, A_2, \dots, A_k , k subespacios cerrados de H .

TEOREMA 2.6.

$$(2.8) \quad \lim_{n \rightarrow \infty} (P_{A_k} P_{A_{k-1}} \dots P_{A_1})^n x = P_{\bigcap_1^k A_i} x$$

2. Velocidad de convergencia

La velocidad de convergencia del método MAP depende del ángulo entre los subespacios concernientes.

Recordemos que si $x \in H$ y $y \in H$, el ángulo θ entre los vectores x y y está definido como:

$$(2.9) \quad \cos \theta = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

2.1. Ángulo entre dos subespacios. La siguiente definición fue introducida originalmente por Friedrichs [14], ésta es la definición más común en la literatura del método MAP para trabajar con el ángulo entre dos subespacios.

DEFINICIÓN 2.8. El ángulo $\theta(M, N)$ entre dos subespacios cerrados M y N es un ángulo comprendido entre $[0, \pi/2]$ cuyo coseno $c(M, N)$ está dado por:

$$(2.10) \quad c(M, N) = \sup\{\langle x, y \rangle \mid x \in M \cap (M \cap N)^\perp, \|x\| \leq 1, y \in N \cap (M \cap N)^\perp, \|y\| \leq 1\}.$$

Otros autores definen el ángulo $\theta(M, N)$ sin tomar en consideración el factor $(M \cap N)^\perp$ que aparece en la expresión precedente. La siguiente definición se debe a Dixmier [10]:

DEFINICIÓN 2.9. El ángulo mínimo θ_0 entre dos entornos cerrados M y N de H , está definido como el ángulo comprendido entre $[0, \pi/2]$ cuyo coseno $c_0(M, N)$ está definido por:

$$(2.11) \quad c_0(M, N) = \sup\{\langle x, y \rangle \mid x \in M, \|x\| \leq 1, y \in N, \|y\| \leq 1\}$$

Observación:

Las siguientes propiedades son consecuencia de las definiciones:

- (1) $0 \leq c(M, N) \leq c_0(M, N) \leq 1$
- (2) $c(M, N) = c(N, M)$ y $c_0(M, N) = c_0(N, M)$
- (3) $c(M, N) = c_0(M \cap (M \cap N)^\perp, N \cap (M \cap N)^\perp)$
- (4) $|\langle x, y \rangle| \leq c_0(M, N) \|x\| \|y\| \quad \forall x \in M, y \in N$ (Desigualdad de Schwarz)

El lema a continuación posee numerosas propiedades,

LEMA 2.1 (Kayalar, Weinert [?] et Deutsch [9]). Se puede verificar que:

- (1) $c(M, N) = c_0(M, N \cap (M \cap N)^\perp) = c_0(M \cap (M \cap N)^\perp, N)$
- (2) $c(M, N) = \|P_M P_N\|$
- (3) $c_0(M, N) = \|P_M P_N - P_{M \cap N}\| = \|P_M P_N P_{(M \cap N)^\perp}\|$

2.2. Velocidad de convergencia del método MAP. De acuerdo con el teorema 2.6,

$(P_r \dots P_2 P_1)^n x$ converge a $P_M x$ para todo $x \in H$ (donde $M = \bigcap_{i=1}^r M_i$ y $P_i = P_{M_i}$).

Podemos notar que, para cada $i = 1 \dots r$, en $P_i P_M = P_M$, y también

$$\begin{aligned} P_i P_{M^\perp} &= P_i (I - P_M) \\ &= P_i - P_i P_M \\ &= P_i - P_M P_i \\ &= (I - P_M) P_i \\ &= P_{M^\perp} P_i. \end{aligned}$$

Entonces, podemos deducir que para cada $x \in H$:

$$\begin{aligned} \|(P_r, \dots, P_2 P_1)^n x - P_M x\| &\leq \|(P_r, \dots, P_2 P_1)^n - P_M\| \|x\| \\ &= \|(P_r, \dots, P_2 P_1 P_{M^\perp})^n\| \|x\| \\ &\leq \|(P_r, \dots, P_2 P_1 P_{M^\perp})\|^n \|x\| \end{aligned}$$

Luego, la velocidad de convergencia del algoritmo MAP está determinado por la norma del operador $(P_r, \dots, P_2 P_1 P_{M^\perp})$. En particular, para $r = 2$, y a partir del lema 2.1 podemos deducir que:

$$(2.12) \quad \|(P_2 P_1)^n - P_M\| \leq \|P_2 P_1 P_{M^\perp}\| = c(M_1, M_2)^n.$$

Por lo tanto, en el caso de dos subespacios, esta cota no es la mejor. Aronszajn [1] dio la siguiente cota:

Para todo $x \in H$ y $n \geq 1, n \in \mathbb{Z}$.

$$(2.13) \quad \|(P_2 P_1)^n x - P_M x\| \leq c(M_1, M_2)^{2n-1} \|x\|.$$

Mas tarde, en 1988, Kayalar et Weinert encontraron que,

TEOREMA 2.7 (Kayalar y Weinert, 1988 [?]).

$$(2.14) \quad \|(P_2 P_1)^n - P_M\| = c(M_1, M_2)^{2n-1}, \quad (n = 1, 2, \dots)$$

2.2.1. *Velocidad de convergencia en el caso de múltiples subespacios.* Para el caso de múltiples subespacios ($r \geq 2$), Smith, Solmon y Wagner [32] probó una cota para la velocidad de convergencia en función de los ángulos de los subespacios concernientes.

TEOREMA 2.8 (Smith, Solmon y Wagner [32]). Para cada $i = 1, \dots, r$ sean M_i subespacios cerrados de H . Luego para cada $x \in H$ y para algún entero $n \geq 1$ se sigue que:

$$(2.15) \quad \|(P_{M_r} P_{M_{r-1}} \dots P_{M_1})^n x - P_{\cap_{i=1}^r M_i} x\| \leq c^n \|x - P_{\cap_{i=1}^r M_i} x\|$$

donde

$$(2.16) \quad c = [1 - \prod_{i=1}^{r-1} \sin^2 \theta_i]^{1/2},$$

aquí θ_i es el ángulo M_i y $N_i = \cap_{j=i+1}^r M_j$. Para la prueba ver [27]

2.2.2. *Teorema de Convergencia.* Sea $\{x_k\}$ la sucesión generada por el método acelerado y $x \in H$ el punto inicial, entonces, $\forall n \in \mathcal{N}$:

$$(2.17) \quad \|x_k - P_M x\| \leq \|x - P_M x\| \left(\prod_{i=0}^{k-1} f_i \right) c^k,$$

donde $0 \leq f_i \leq 1$ verifica

$$(2.18) \quad \|x_{i+1} - P_M x\| = f_i \|Qx_i - P_M x\|.$$

Ver demostración en [27].

3. Método de proyecciones oblicuas alternadas aplicado a sistemas acoplados

Recientemente Luis M. Hernández en [19] propuso un método de proyecciones alternadas aplicado a sistemas de punto de ensilladura (1.1) para el caso en que la matriz A es simétrica definida positiva, B es de rango completo y la función g es igual a cero.

En dicho trabajo se aplicó el método de proyecciones alternadas para resolver problemas de punto de ensilladura con la siguiente estructura:

$$(2.19) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix},$$

donde

(H1): $B \in \mathfrak{R}^{m \times n}$ es una matriz $rank(B) = m, m \leq n$,

(H2): $A \in \mathfrak{R}^{n \times n}$ es una matriz semidefinida positiva.

Mediante la siguiente proposición se demostró que bajo las hipótesis dadas sobre las matrices bloques A y B se garantiza una solución única del sistema (2.19).

PROPOSICIÓN 2.3. Bajo las hipótesis (H1,H2) el sistema (2.19) tiene solución única $(x, \lambda)^T$ que satisface:

$$(2.20) \quad \begin{cases} f - Ax \perp \ker B, \\ x \in \ker B. \end{cases}$$

Demostración

Bajo las hipótesis (H1,H2) el sistema:

$$(2.21) \quad \begin{pmatrix} A & B^T \\ 0 & BA^{-1}B^T \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} f \\ BA^{-1}f \end{pmatrix},$$

es equivalente al sistema (2.19) y $BA^{-1}B^T$ es una matriz definida positiva. Como consecuencia (2.21) tiene solución única. Esta solución x verifica $f - Ax = B^T$ con $B^T \lambda \perp \ker B$. \square

Luego, el método propuesto en [19] consiste en hallar la solución (x_*, y_*) del sistema proyectando de manera alternada primero ortogonalmente sobre el $\ker B$ y luego de forma oblicua sobre una variedad lineal V definida como $V = \{x \in X / f - Ax \in (\ker B)^\perp\}$. Usando el hecho que $V \cap \ker B = \{x_*\}$ en donde x_* es la primera componente del vector solución (x_*, y_*) se propone entonces el siguiente esquema iterativo

DEFINICIÓN 2.10 (Método PAO).

$$(2.22) \quad \begin{cases} x_0 \in V \\ x_{k+1} = x_k + \alpha_k d_k, \end{cases}$$

con

$$(2.23) \quad d_k = \tilde{Q}Px_k - x_k,$$

donde

- d_k es la dirección de descenso, que es una aproximación de $e_k = x - x_k$,
- α_k es el tamaño del paso,
- P es la matriz de proyección ortogonal sobre $\ker B$,
- \tilde{Q} es la matriz de una proyección sobre la variedad lineal V .

Luego a partir de este valor se calcula y_* usando el hecho que $B^T y_k = f - Ax_k$.

Para reducir los costos computacionales en [19] se propuso utilizar una proyección oblicua \tilde{Q} sobre V definida por,

$$(2.24) \quad \forall y \in \mathfrak{R}^n, \tilde{Q}y = A^{-1}Pf + Qy,$$

donde

$$(2.25) \quad \begin{cases} R = I - P, \text{ la proyección ortogonal sobre } (\ker B)^\perp, \\ Q = A^{-1}RA, \end{cases}$$

entonces

$$(2.26) \quad \tilde{Q}y = A^{-1}Pf + A^{-1}(I - P)Ay$$

$$(2.27) \quad = A^{-1}P(f - Ay) + y.$$

dando lugar al siguiente algoritmo.

Algoritmo de proyecciones oblicuas alternadas PAO

(1) **inicialización:**

- $x_0 = x \in V$

(2) **iteración: Para** $k = 0, 1, \dots$, **Hacer**

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

- $\alpha_k = \frac{\langle -x_k, d_k \rangle_A}{\langle d_k, d_k \rangle_A} = -\frac{\langle Ad_k, x_k \rangle}{\langle Ad_k, d_k \rangle}$

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

(3) **Fin**

En ([19]) se puede ver la demostación de las siguientes propiedades que posee el método PAO.

LEMA 2.2. Bajo las hipótesis (H1,H2), el método PAO satisface:

Para todo $k \geq 0$,

(1) $d_k \in \ker B^{\perp A}$,

(2) $x_k \in V$,

(3) $\langle Ad_k, d_k \rangle = 0 \Leftrightarrow x_k = x$

(4) $d_k = \tilde{Q}Px_k - x_k$, luego el método de GC-Uzawa preconditionado es equivalente al método PAO.

LEMA 2.3. Bajo las hipótesis (H1,H2), el método PAO satisface:

Para todo $k \geq 0$,

(1) $\langle x, d_k \rangle_A = 0$,

(2) $\langle e_k, d_k \rangle_A = \langle -x_k, d_k \rangle_A$, con $e_k = x - x_k$,

(3) $\langle e_{k+1}, d_k \rangle_A = 0$,

(4) $\langle e_k, d_k \rangle_A = \|Rx_k\|_A^2$.

LEMA 2.4. Bajo las hipótesis (H1,H2), el método PAO satisface:

Para todo $k \geq 0$,

(1) $\|e_{k+1}\|_A^2 = \|e_k\|_A^2 - \frac{\|Rx_k\|_A^4}{\|d_k\|_A^2} = \|e_k\|_A^2 - \alpha_k^2 \|d_k\|_A^2$,

(2) $\|e_{k+1}\|_A \leq \|e_k\|_A$,

(3) $\|e_{k+1}\|_A = \|e_k\|_A$ iff $x = x_k$.

LEMA 2.5. Bajo las hipótesis (H1,H2), el método PAO satisface:

Para todo $k \geq 0$,

- (1) $\|e_{k+1}\|_A = \|e_k\|_A \sqrt{1 - \rho_k^2}$, donde $\rho_k = \frac{\|Rx_k\|_A^2}{\|d_k\|_A \|e_k\|_A}$,
- (2) $\rho_k \geq \frac{\|Rx_k\|_A}{\|e_k\|_A} \frac{1}{\|Q\|_A}$, donde $Q = A^{-1}RA$ es la proyección oblicua sobre $(\ker B)^{\perp A}$.
- (3) $\frac{\|Rx_k\|_A}{\|e_k\|_A} \geq 1$,
- (4) $\|e_{k+1}\|_A \leq \|e_k\|_A \sqrt{1 - \rho^2}$, donde $\rho = \frac{1}{\|Q\|_A} > 0$. Luego el método converge.

Otro método propuesto en [19] es la versión Gradiente Conjugado PAO (GC-PAO) que proviene de considerar el método de Uzawa preconditionado con $Q_B^{-1} = (B^+)^T AB^+$. Este método adaptado al problema de punto de ensilladura es el siguiente:

$$(2.28) \quad \begin{cases} x_{k+1} = A^{-1}(f - B^T y_{k+1}) \\ y_{k+1} = y_k + \alpha Q_B^{-1} B x_k, \end{cases}$$

Sustituyendo y_{k+1} en $x_{k+1} = A^{-1}(f - B^T y_{k+1})$ se obtiene el siguiente esquema iterativo para hallar el vector solución x_* del sistema

$$(2.29) \quad x_{k+1} = x_k - \alpha A^{-1} B^T Q_B^{-1} B x_k$$

Si usamos el preconditionador $Q_B^{-1} = (B^+)^T AB^+$, en la iteración anterior se obtiene que la dirección de descenso coincide con la propuesta en el método PAO $d_k = -A^{-1}RARx_k$ lo que los hace métodos equivalentes.

Para el caso en que la matriz A es simétrica definida positiva se propone entonces la versión gradiente conjugado PAO aplicado al sistema del complemento de Schur (o sistema condensado) $BA^{-1}B^T y = BA^{-1}f$ con este preconditionador.

Algoritmo GC-PAO

(1) **Inicialización:**

- Seleccionar $x_0 \in V$
- $w_0 = RARx_0$
- $z_0 = w_0$

(2) **Iteraciones: Para $k = 0, \dots$**

- $d_k = A^{-1}z_k$
- $\alpha_k = \langle x_k, w_k \rangle / \langle d_k, z_k \rangle$
- $x_{k+1} = x_k - \alpha_k d_k$

- $w_{k+1} = RARx_{k+1}$
- $\beta_{k+1} = \langle x_{k+1}, w_{k+1} \rangle / \langle x_k, w_k \rangle$
- $z_{k+1} = w_{k+1} + \beta_{k+1}z_k$

(3) fin

CAPÍTULO 3

Generalización del método de proyecciones oblicuas alternadas (PAO)

A lo largo de este capítulo se desarrollará la base teórica requerida para generalizar el método de proyecciones oblicuas alternadas al caso de variedades lineales para resolver los sistemas de la forma (1.1).

1. Nociones básicas

DEFINICIÓN 3.1. Sea V un espacio vectorial. Una variedad lineal $M \subseteq V$ es un conjunto de la forma $M = \{s + p/s \in S\}$, donde S es un subespacio de V y $p \in V$.

DEFINICIÓN 3.2. Un conjunto C en un espacio vectorial se dice que es un conjunto convexo si $(1 - \alpha)x + \alpha y \in C$ donde $x, y \in C$ y $0 < \alpha < 1$.

2. Definición del problema

Estamos interesados en resolver los sistemas lineales por bloques llamados sistemas de punto de ensilladura (Saddle Point Problems) como se definió en (1.1) a través del método de proyecciones oblicuas alternantes (PAO) propuesto por L.M. Hernández-Ramos para el caso en que la función g del vector $(f \ g)^T$ es distinta de cero.

Sea el siguiente sistema lineal por bloques

$$(3.1) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

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 semidefinida positiva.

Escribamos el sistema anterior en su forma equivalente

$$(3.2) \quad \begin{cases} Ax + B^T y = f \\ Bx = g \end{cases}$$

Como se vio en el capítulo 1, bajo las hipótesis dadas sobre las matrices A y B , la matriz \mathcal{A} del sistema (3.1) es invertible por lo tanto el sistema $\mathcal{A}u = b$ tiene solución única (x_*, y_*) . Luego, si definimos las siguientes variedades lineales

$$V = \{x \in X / f - Ax \in (\ker B)^\perp\} \text{ y}$$

$$W = \{x \in X / Bx = g\}$$

entonces el vector x_* es tal que

$$\{x_*\} = V \cap W,$$

donde W también se puede escribir como $W = \ker B + c$, con c una solución particular del sistema $Bx = g$.

En este trabajo se propone, de manera similar a [20], utilizar proyecciones alternadas, en este caso sobre las variedades lineales V y W para hallar la solución aproximada (x_*, y_*) del sistema (3.1).

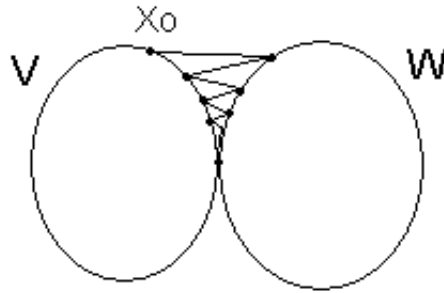


FIGURA 3.1. Método PAO

Se considera de nuevo la proyección oblicua \tilde{Q} sobre la variedad lineal V como se definió en [20]

$$(3.3) \quad \tilde{Q}(y) = A^{-1}P(f - Ay) + y$$

Luego, para calcular la proyección ortogonal sobre W que llamaremos P_w recurriremos a la siguiente proposición de [27].

PROPOSICIÓN 3.1. Sea W una variedad lineal definida por $W = S + c$ con S un conjunto convexo de X y c perteneciente a un espacio vectorial. Sea $P_W(x)$ la proyección ortogonal del vector x en la variedad lineal W , entonces

$$(3.4) \quad P_W(x) = P_{S+c}(x) = P_S(x - c) + c$$

Si denotamos $x = P_W(x_0)$ como la proyección de $x_0 = A^{-1}f$ sobre $W = S + c$ con $S = \ker B$ y c una solución particular de $Bx = g$ y usando la proposición anterior podemos verificar que x_0 es la proyección A-ortogonal sobre $\ker B$.

$$\begin{aligned} x_0 - P_W(x_0) &= x_0 - (P_S(x_0 - c) + c) \\ &= (x_0 - c) - P_S(x_0 - c) \perp S \end{aligned}$$

2.1. Método PAO generalizado. Para hallar la solución partiremos de un valor inicial $x_0 = A^{-1}f$ de la variedad lineal V que es la solución del problema sin restricciones (1.2). Luego, el método PAO se define de la siguiente forma.

$$(3.5) \quad \begin{cases} x_0 \in V \\ x_{k+1} = x_k + \alpha_k d_k \end{cases}$$

en donde $d_k = \tilde{Q}P_W x_k - x_k$, considerando nuevamente los operadores P y Q como

$$(3.6) \quad \begin{cases} R = I - P, \text{ la proyección ortogonal sobre } (\ker B)^\perp, \\ Q = A^{-1}RA, \end{cases}$$

entonces

$$(3.7) \quad \tilde{Q}y = A^{-1}PF + A^{-1}(I - P)Ay$$

$$(3.8) \quad = A^{-1}P(F - Ay) + y.$$

LEMA 3.1. En el método PAO, para el caso en que $g \neq 0$, la dirección de descenso viene dada por $d_k = -A^{-1}RAR(x_k - c)$.

DEMOSTRACIÓN.

$$\begin{aligned}
\tilde{Q}P_W x_k - x_k &= A^{-1}PF + A^{-1}RAP_W x_k - x_k \\
A(\tilde{Q}P_W x_k - x_k) &= PF + RAP_W x_k - Ax_k \\
&= PF + RA(P(x_k - c) + c) - Ax_k \\
&= PF + RAP(x_k - c) + RAc - Ax_k \\
&= PF + RA(x_k - c) + RAR(x_k - c) + RAc - Ax_k \\
&= PF + RAx_k - RAR(x_k - c) - Ax_k \\
&= PF - PAx_k - RAR(x_k - c) \\
&= P(F - Ax_k) - RAR(x_k - c) \\
A(\tilde{Q}P_W x_k - x_k) &= -RAR(x_k - c) \\
d_k &= -A^{-1}RAR(x_k - c)
\end{aligned}$$

□

Teniendo este resultado se puede escribir entonces la versión generalizada del algoritmo PAO al caso de proyectar sobre dos variedades lineales.

Algoritmo PAO

El algoritmo adaptado para resolver el sistema (3.2) por proyecciones oblicuas alternadas es el siguiente.

(1) **inicialización:**

- $x_0 = x \in V$
- Calcular $c = B^+g$ (solución particular de $Bx = g$) ;

(2) **iteración: Para $k = 0, 1, \dots$, Hacer**

- $d_k = -A^{-1}RAR(x_k - c)$
- $\alpha_k = \frac{\langle -x_k, d_k \rangle_A}{\langle d_k, d_k \rangle_A} = -\frac{\langle Ad_k, x_k \rangle}{\langle Ad_k, d_k \rangle}$
- $x_{k+1} = x_k + \alpha_k d_k$

(3) **Fin**

Viendo este resultado nos damos cuenta que la solución al problema (3.1) se puede hallar de manera equivalente resolviendo el siguiente problema de punto de ensilladura

$$(3.9) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x - c \\ y \end{pmatrix} = \begin{pmatrix} f - Ac \\ 0 \end{pmatrix},$$

Nótese que este problema tiene la forma de los sistemas resueltos en [20], por lo tanto puede utilizarse este mismo algoritmo para hallar $x - c$ y luego sumando el valor de $c \in W$ a la solución del sistema (3.1) para hallar la solución x_* del sistema original. Este resultado se puede interpretar igualmente a partir del resultado que se desprende en [3],[20] de que en la solución al problema (1.1).

$$(3.10) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x_{*1} \\ y_{*1} \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix},$$

el vector x_{*1} es la proyección A-ortogonal de $x_0 = A^{-1}f$ sobre $KerB$ y que en el sistema

$$(3.11) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x_{*2} \\ y_{*2} \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

el vector x_{*2} es la proyección A-ortogonal del $A^{-1}f$ sobre W

de esta manera, usando de nuevo la proposición (3.1) tenemos que $x_{*2} = P_W(A^{-1}f) = P_{kerB}(A^{-1}f - c) + c$ y así resolviendo un problema del tipo (3.10) resolvemos (3.11)

3. Experimentación numérica

Para la experimentación numérica de este trabajo se implementaron, además de la versión gradiente conjugado del método Uzawa (GCUzw), los algoritmos del método de Uzawa y el método de Barzilai-Borwein (BB) para evaluar y comparar su comportamiento con sistemas de punto de ensilladura con las hipótesis dadas en (1.1) con sistemas con distintas dimensiones. También se realizaron experimentos con cada uno de estos métodos usando el preconditionador $(B^+)^T AB^+$ aplicado al sistema condensado (sistema del complemento de Schur).

Las matrices empleadas en los experimentos provienen de problemas de optimización con restricciones y fueron tomadas de una librería con matrices para pruebas CUTeR (The Constrained and Unconstrained Testing Environment). El entorno CUTeR es un conjunto de subrutinas de Fortran que crea problemas de prueba en el contexto de optimización lineal y no lineal.

Todos los problemas seleccionados para la experimentación con matrices del CUTeR son problemas de la forma

$$(3.12) \quad \min J(x) = \frac{1}{2}x^T Ax - f^T x$$

$$(3.13) \quad \text{sujeto a: } Bx = g.$$

donde $A \in \mathbb{R}^{n \times n}$ es una matriz simétrica y $B \in \mathbb{R}^{m \times n}$ la matriz de restricciones es de rango completo. A la matriz bloque A proveniente del CUTeR se le aplicó una transformación para garantizar que se trataran de matrices definidas positivas. Dicha transformación consiste en $A_{(i,i)} + 0,01$ con $i = 1, \dots, n$, es decir, sumar el escalar 0,01 a los elementos de la diagonal.

También se realizaron experimentos con matrices provenientes de problemas de resultan de aplicar el método de descomposición de dominio a las ecuaciones en derivadas parciales de Poisson

$$(3.14) \quad \begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

donde Ω representa dominios rectangulares acotados de \mathbb{R}^2 , $\partial\Omega$ representa la frontera y f una función dada.

En esta discretización, se obtiene un sistema de punto de ensilladura de la forma

$$(3.15) \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix},$$

donde

- $B \in \mathbb{R}^{m \times n}$ es una matriz $\text{rank}(B) = m, m \leq n$,
- $A \in \mathbb{R}^{n \times n}$ es una matriz simétrica definida positiva.

Experimento	Matrices CUTer	n	m	cond(A)	cond(S)
1	GOULDQP3	699	344	601	423,95
2	KASIP	1021	1001	101	541,103
3	AUG2DCQP	3280	1600	1	989,27
4	AUG3DCQP	3873	1000	1	66,45
5	HUESMOD	5000	2	1	64,8
6	CVXQP1S	100	50	$1,1769 \times 10^5$	$7,0158 \times 10^6$
7	CVXQP2S	100	25	$1,1769 \times 10^5$	$1,8961 \times 10^5$
8	CVXQP3S	100	75	$1,1544 \times 10^5$	$1,2053 \times 10^8$
9	QGROW15	645	300	38901	7316,9
10	QGROW22	946	440	53601	7245,3

CUADRO 1. Dimensiones de las matrices CUTer

Experimento	n	m	cond(A)	cond(S)
1	400	39	294.2512	$8,6897 \times 10^{12}$
2	576	47	409.0419	$2,0664 \times 10^{15}$
3	784	55	542.6497	$6,9041 \times 10^{17}$
4	900	59	616.5139	$5,9601 \times 10^{17}$
5	1156	67	778.3686	$2,2222 \times 10^{18}$

CUADRO 2. Dimensiones de las matrices de Descomposición de Dominio

Estas matrices fueron generadas mediante un código facilitado por el Prof L. Hernández. Para este caso se generó adicionalmete una función g , pues dichos códigos generan problemas del tipo (2.19) en donde la función g era igual a cero.

En el cuadro 1 se especifican las dimensiones y condicionamiento de las matrices CUTer empleadas en cada experimento y en el cuadro 2 las características de las matrices de descomposición de dominio.

4. Resultados

En los Cuadro 3 y 4 se presentan los resultados obtenidos en un procesador Pentium 4(R) de 1.8 GHz, 1.5 GB RAM con funciones implementadas en MATLAB para los distintos métodos aplicados a las matrices CUTer y matrices provenientes de descomposición de dominio respectivamente, indicándose la cantidad de iteraciones requeridas por cada uno para su convergencia con una tolerancia de 10^{-6} y una cantidad máxima de iteraciones establecida igual a 1000. Adicionalmente en el Cuadro 4 se muestran los tiempos empleados para la convergencia en cada uno de los métodos calculados a través de la función de Matlab `cputime`.

El criterio de parada utilizado fue

$$(3.16) \quad \frac{\|r_k\|}{\|r_0\|} < tol$$

donde r_k representa el error en la iteración k , y r_0 el error inicial calculado a partir del valor inicial $x_0 = A^{-1}f$.

Experimento	Uzw	Uzw PRE	GC	GC PRE	BB	BB PRE
1	369	126	33	28	104	36
2	136	23	6	8	72	13
3	1000	2	106	2	171	2
4	219	2	33	2	62	2
5	24	2	2	2	8	2
6	1000*	100*	151	36	1000*	208
7	1000*	1000*	47	29	1000*	381
8	1000*	1000*	410	37	1000*	268
9	1000*	1000*	118	23	505*	432*
10	1000*	1000*	145	26	394*	452

CUADRO 3. Iteraciones requeridas para la convergencia con las matrices CUTer. (*)Se alcanzó la cantidad máxima de iteraciones.

En las siguientes gráficas se registra la evolución de la medida del error en norma cuadrática con el paso de las iteraciones en cada uno de los métodos empleados.

Experimento	Uzw	Uzw PRE	GC	GC PRE	BB	BB PRE
1	2.41	4.81	0.92	1.23	0.54	1.04
2	0.67	38.56	0.31	19.46	0.25	22.59
3	2.25	0.76	4.87	1.45	0.32	1.03
4	0.53	1.79	1.01	3.37	0.14	2.43
5	0.09	0.26	0.03	0.45	0.03	0.34
6	1.85	5.57	1.35	0.29	1.89	1.12
7	1.84	4.07	0.39	0.18	1.96	1.54
8	1.96	7.85	5.03	0.4	0.84	2.12
9	7.14	–	3.81	8.51	3.39	–
10	10.40	–	5.04	12.45	3.39	–

CUADRO 4. Tiempo de convergencia (Matrices CUTer)

Experimento	Uzw	Uzw PRE	GC	GC PRE	BB	BB PRE
1	740	232	49	39	122	54
2	335	54	8	9	104	21
3	800	1	122	1	418	1
4	403	1	39	1	82	1
5	41	1	1	1	10	1

CUADRO 5. Problemas de descomposición de dominio. Iteraciones requeridas para la convergencia.

A través de estas gráficas se verificó el comportamiento esperado de los métodos iterativos seleccionados para la experimentación en donde se verificó que la versión preconditionada mejoró notablemente la convergencia en cada caso. El método de gradiente conjugado preconditionado con $(B^+)^T AB^+$ convergió con mayor rapidez que el resto de los métodos. En algunos casos se observó que los métodos preconditionados requieren de un alto costo computacional lo que podría provenir de las iteraciones internas requeridas para resolver los sistemas que involucran a la matriz de proyección $R = B^+B$.

Experimento	Uzw	Uzw PRE	GC	GC PRE	BB	BB PRE
1	2.73	0.37	0.34	0.12	0.35	0.18
2	4.71	0.57	0.53	0.18	0.53	0.32
3	7.81	1.06	0.59	0.26	1.25	0.45
4	9.26	1.23	0.79	0.31	1.12	0.57
5	12.06	1.67	0.93	0.48	1.81	0.82

CUADRO 6. Problemas de descomposición de dominio. Tiempo requerido para la convergencia.

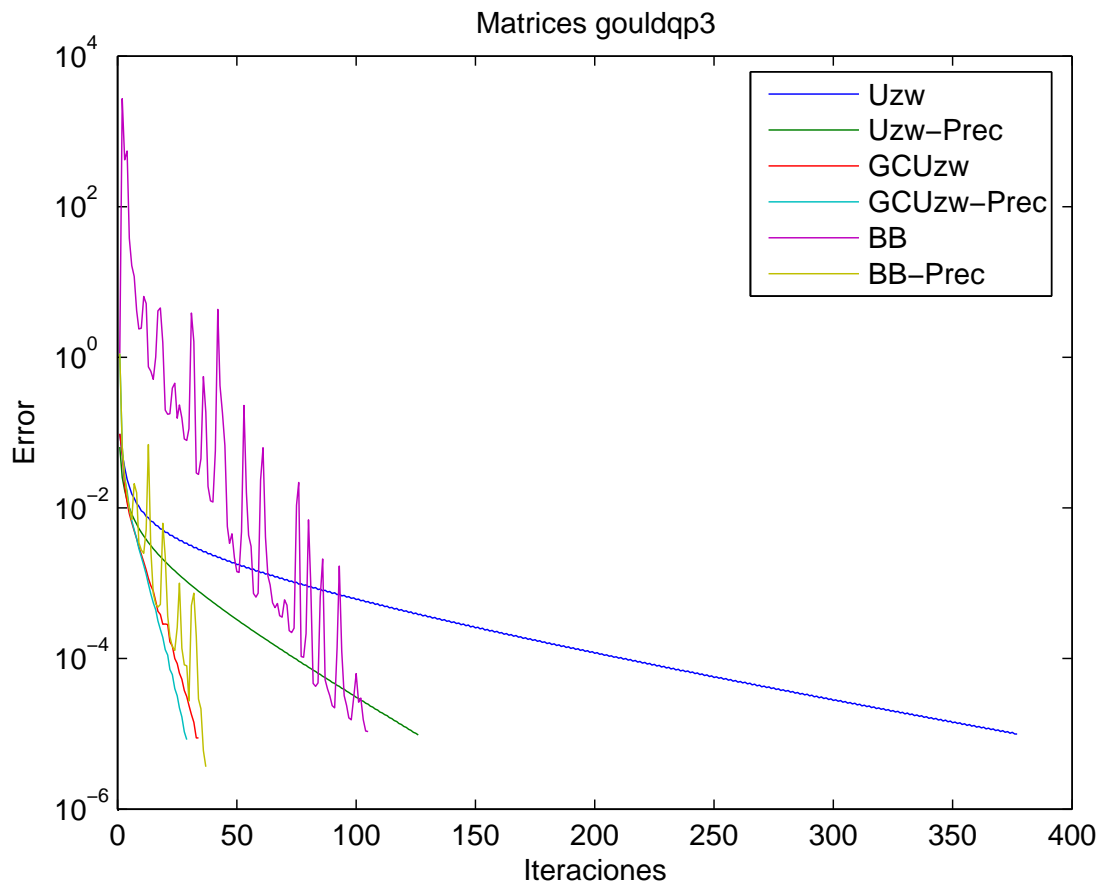


FIGURA 3.2. Experimento 1

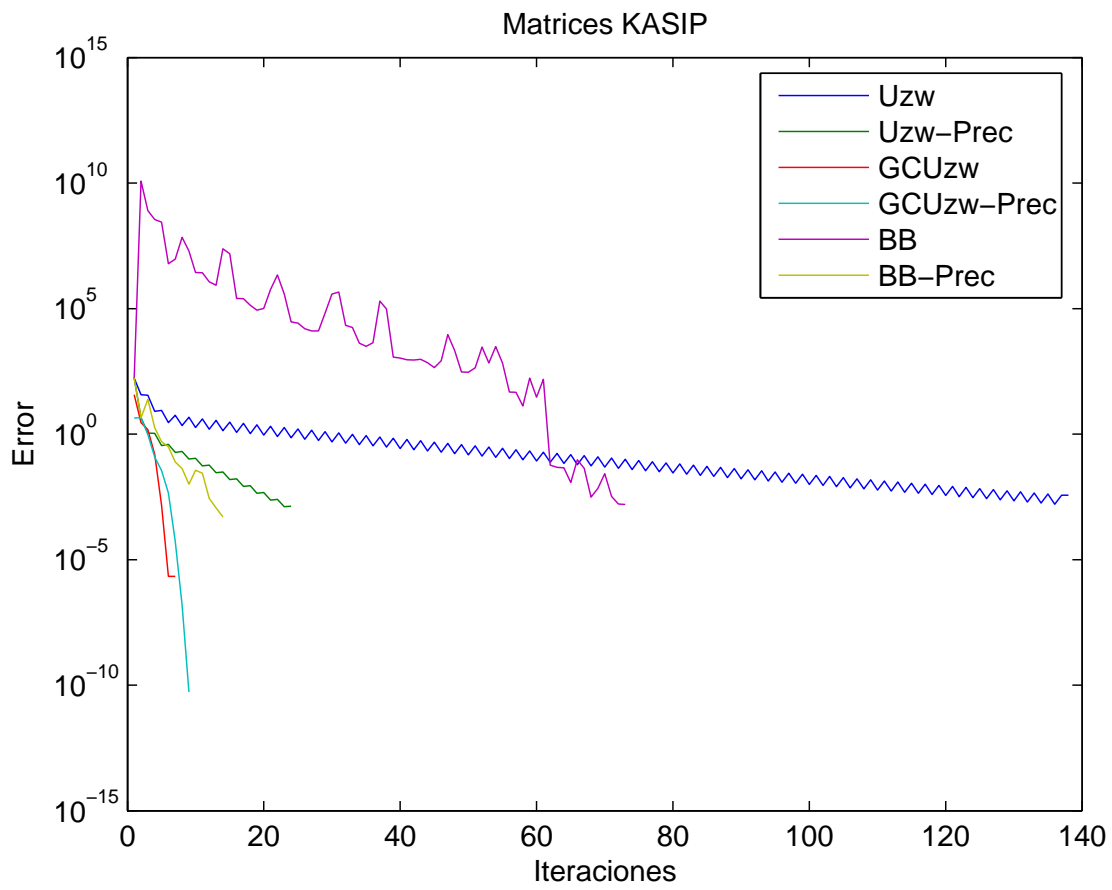


FIGURA 3.3. Experimento 2

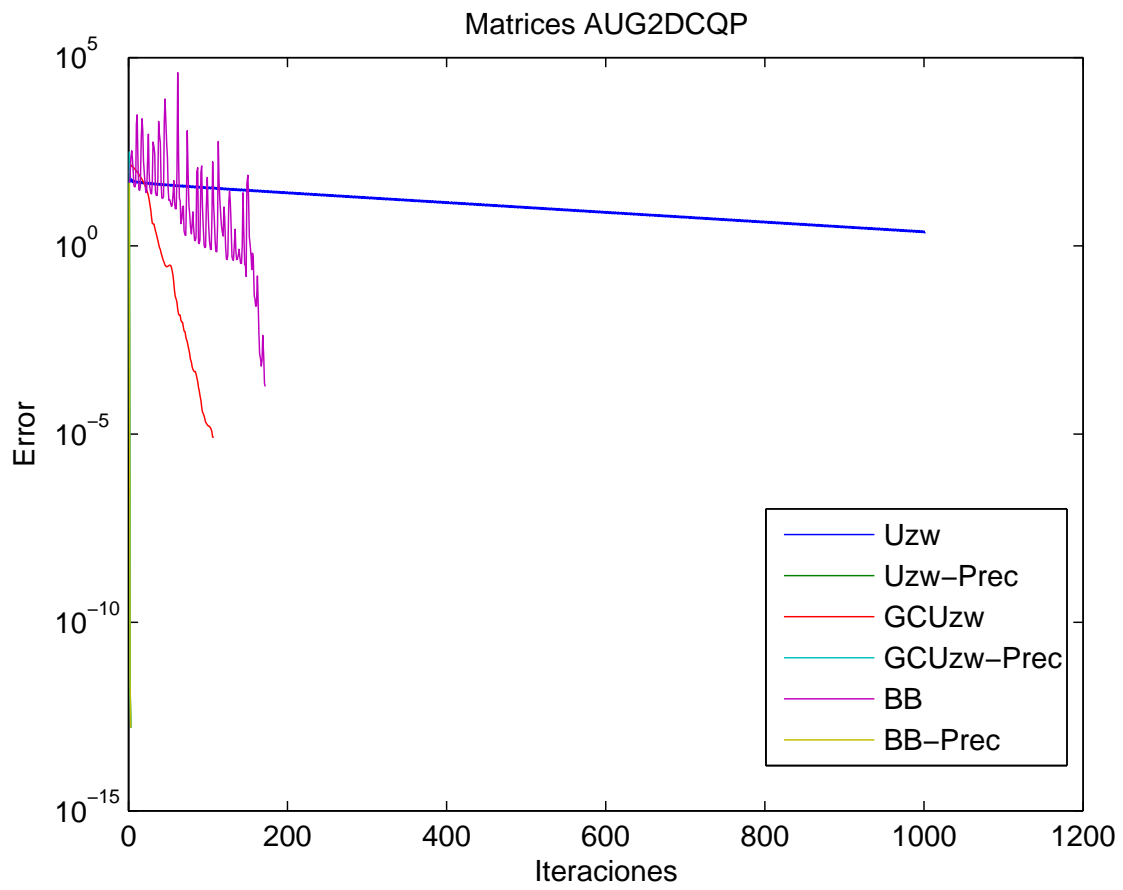


FIGURA 3.4. Experimento 3

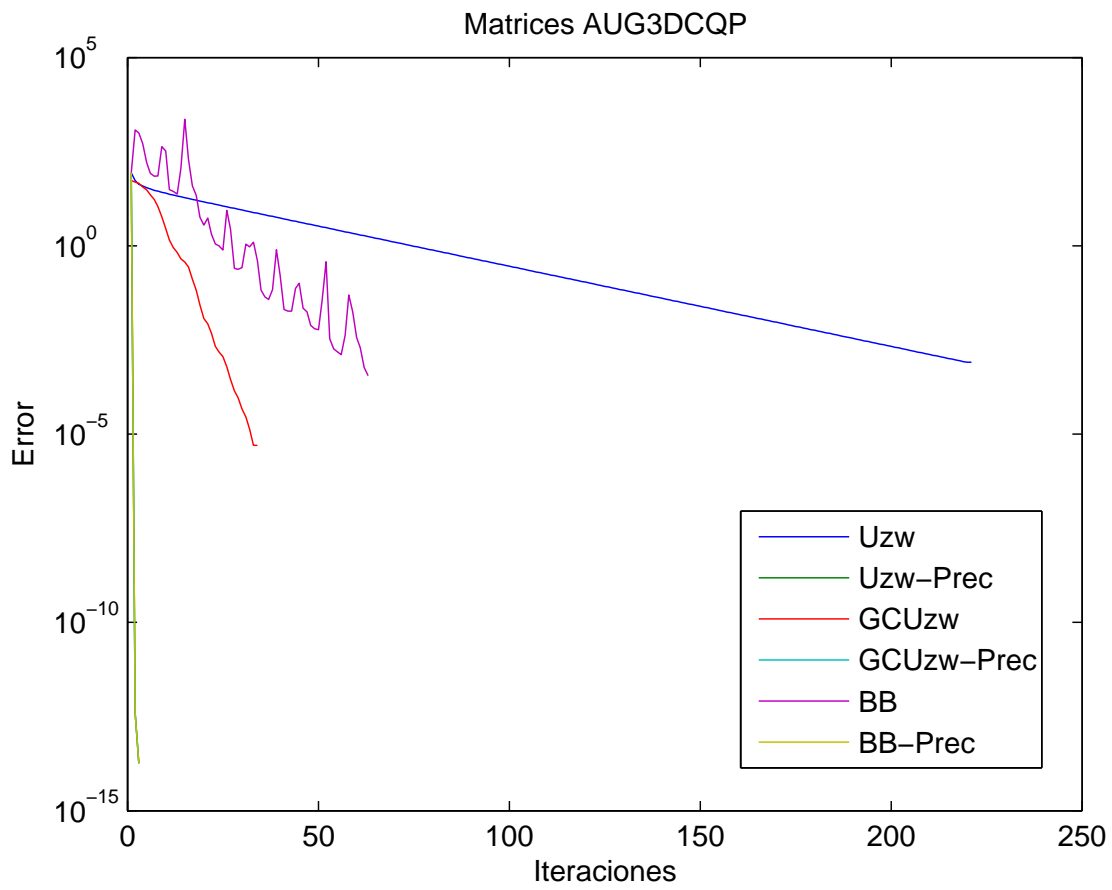


FIGURA 3.5. Experimento 4

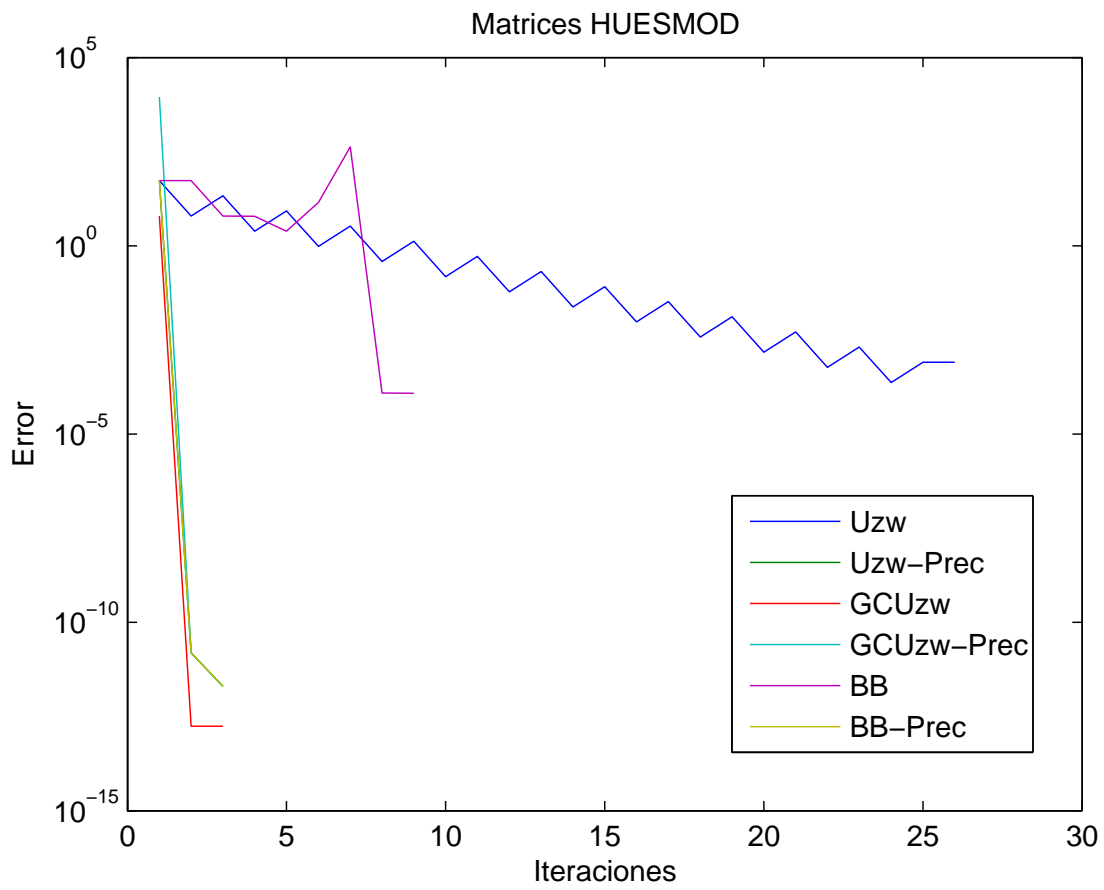


FIGURA 3.6. Experimento 5

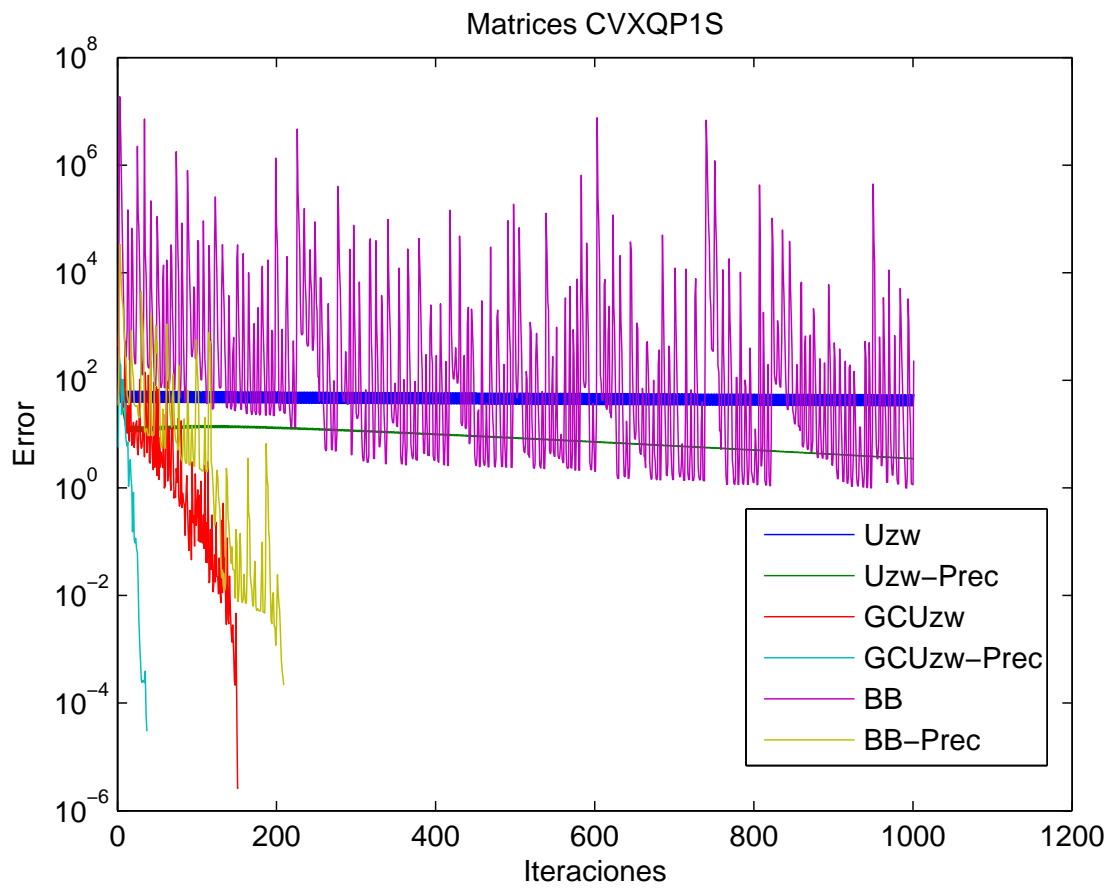


FIGURA 3.7. Experimento 6

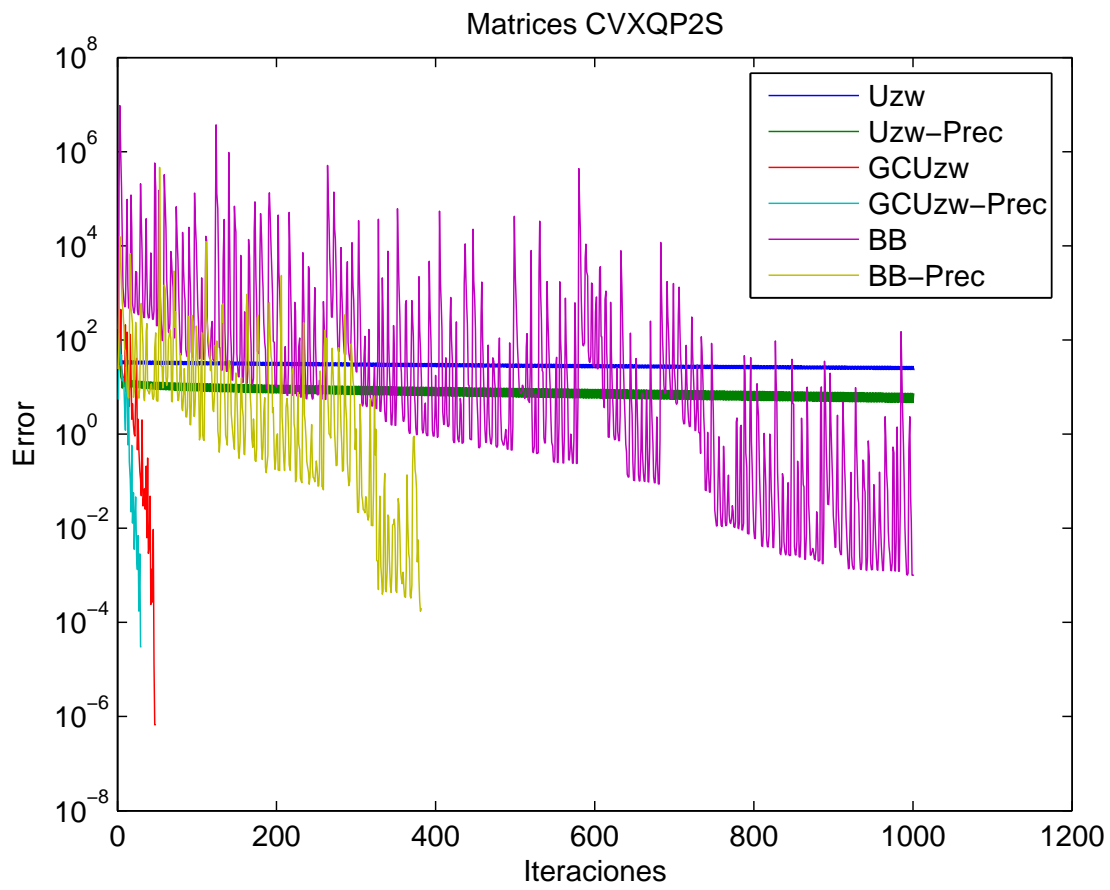


FIGURA 3.8. Experimento 7

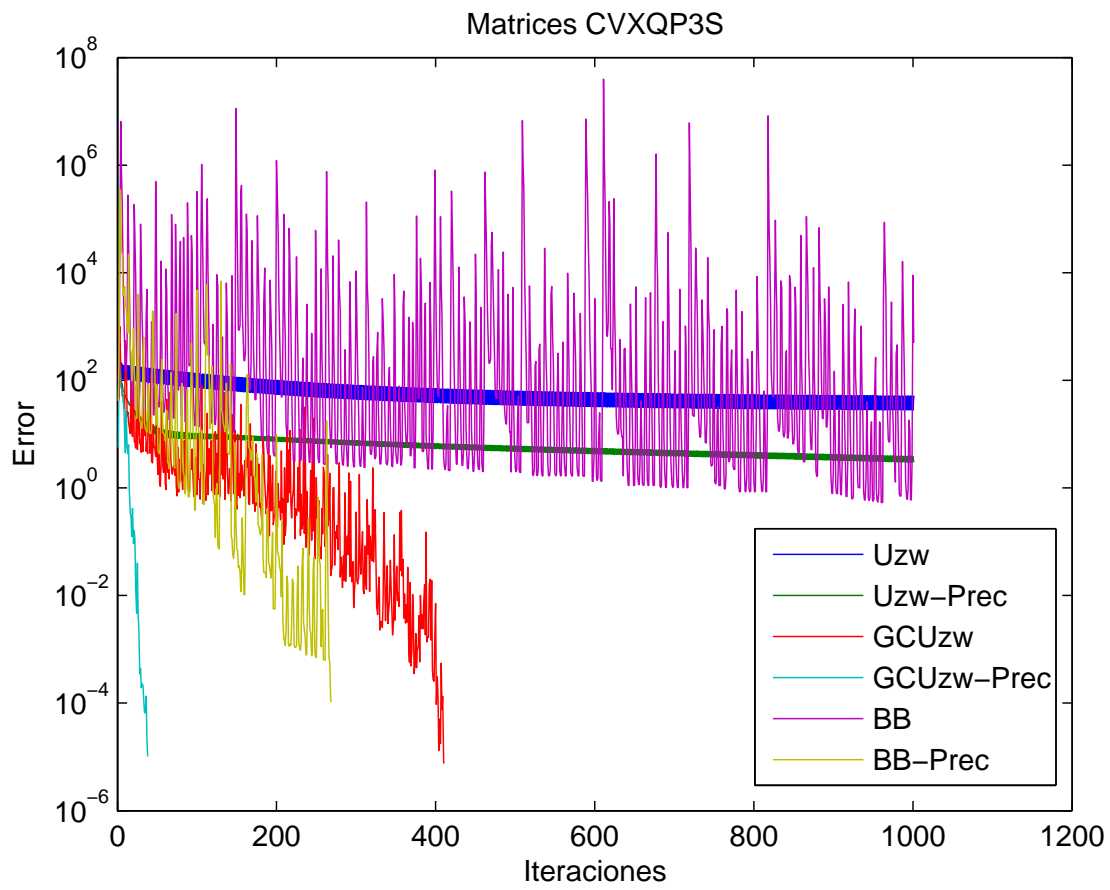


FIGURA 3.9. Experimento 8

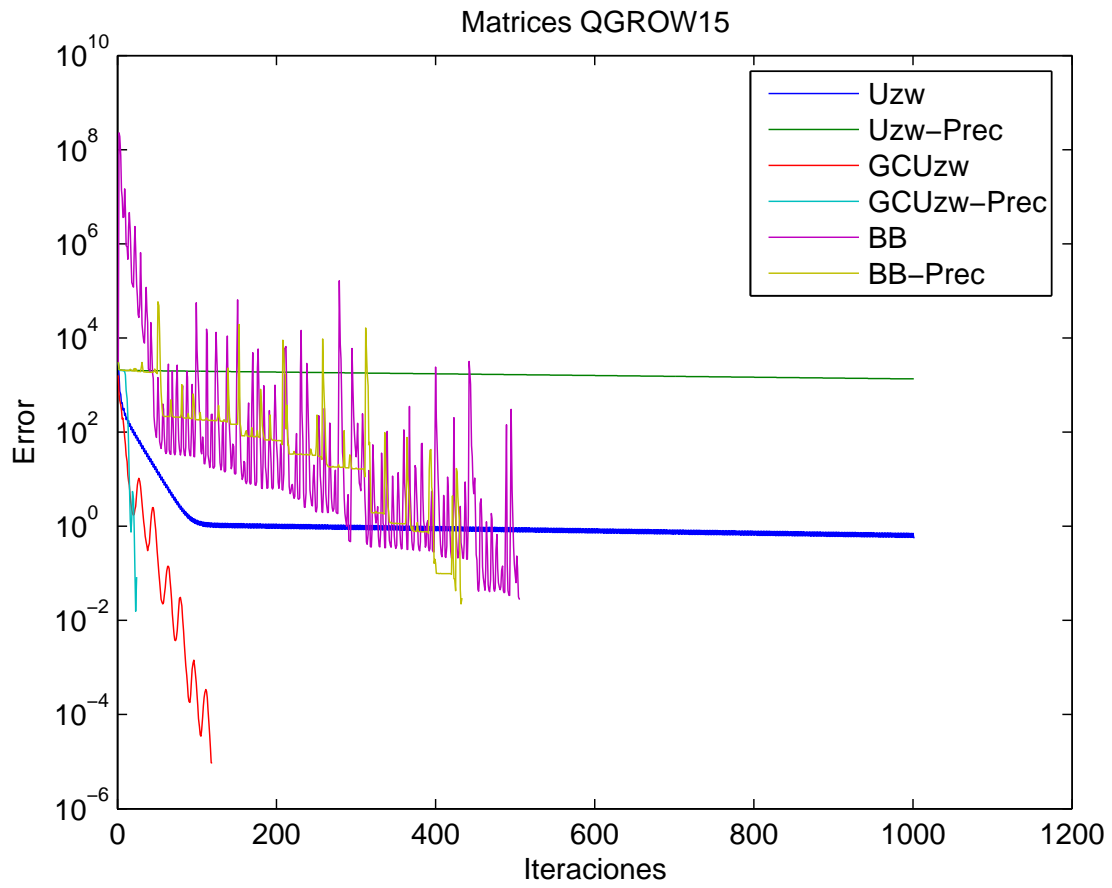


FIGURA 3.10. Experimento 9

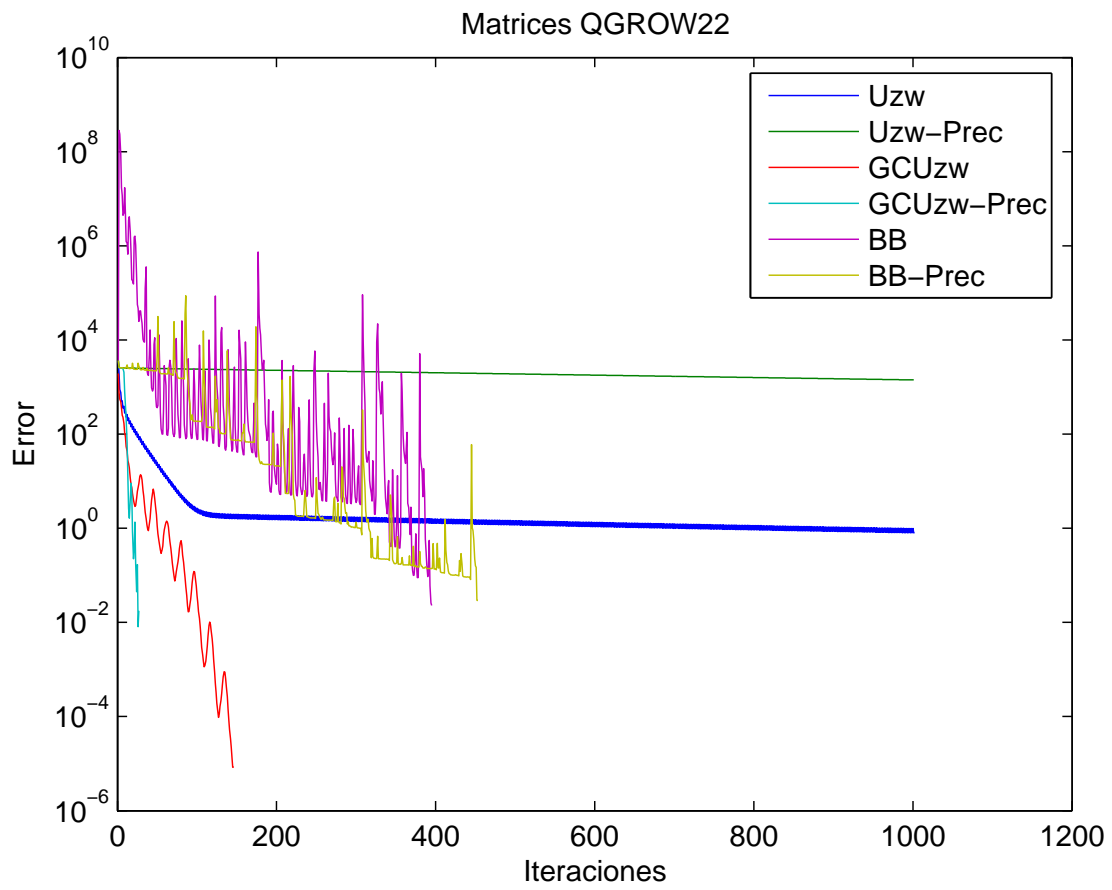


FIGURA 3.11. Experimento 10

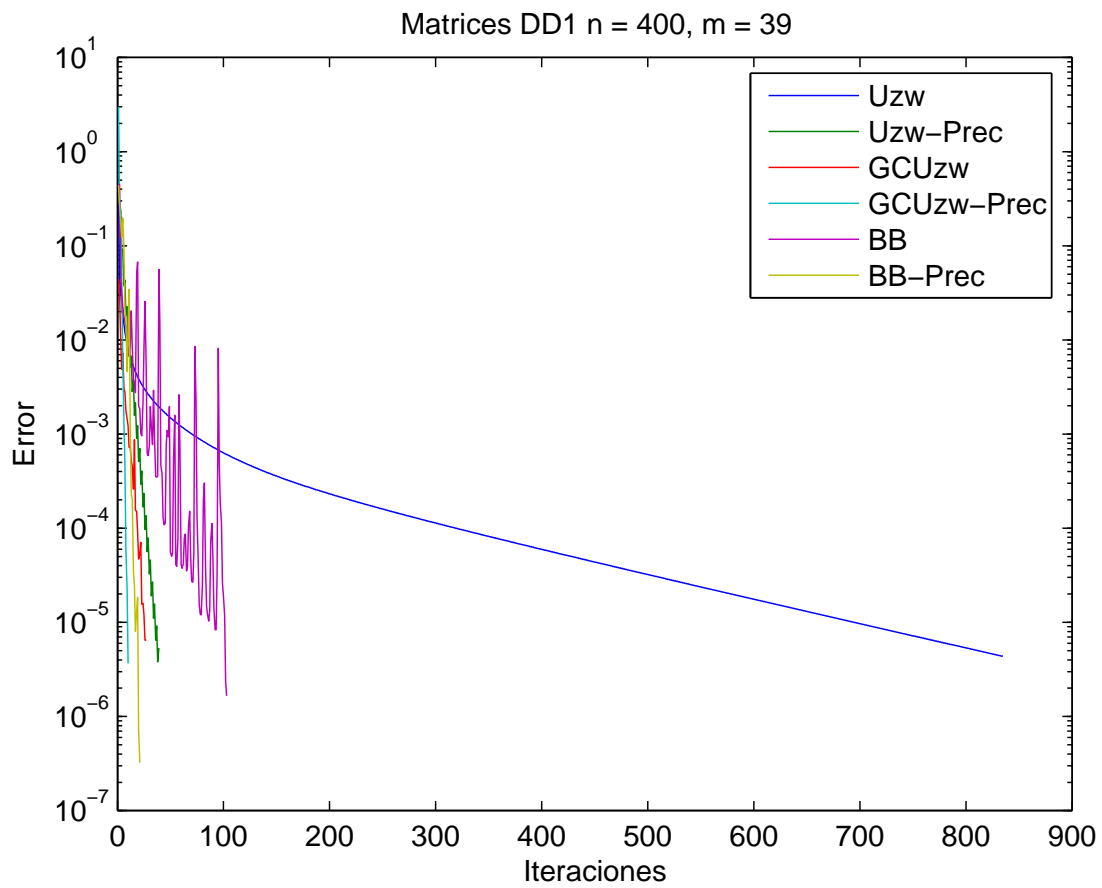


FIGURA 3.12. Experimento 11

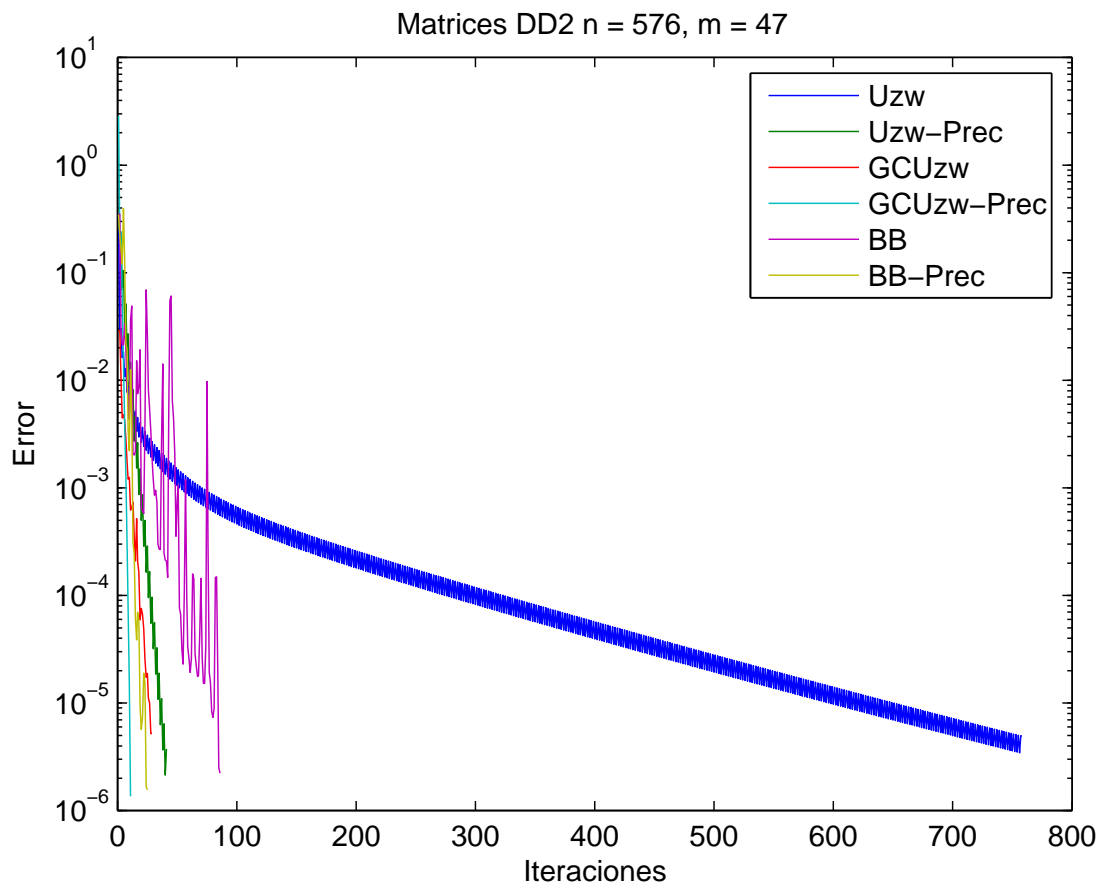


FIGURA 3.13. Experimento 12

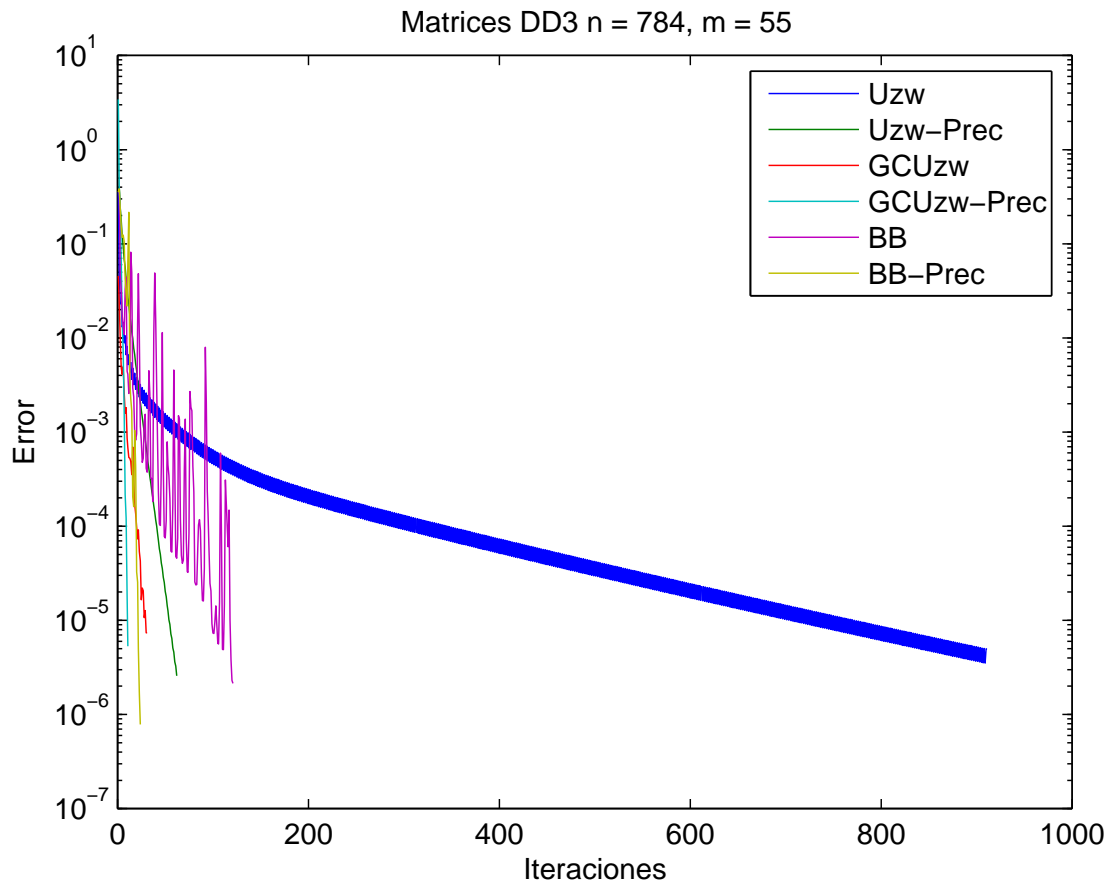


FIGURA 3.14. Experimento 13

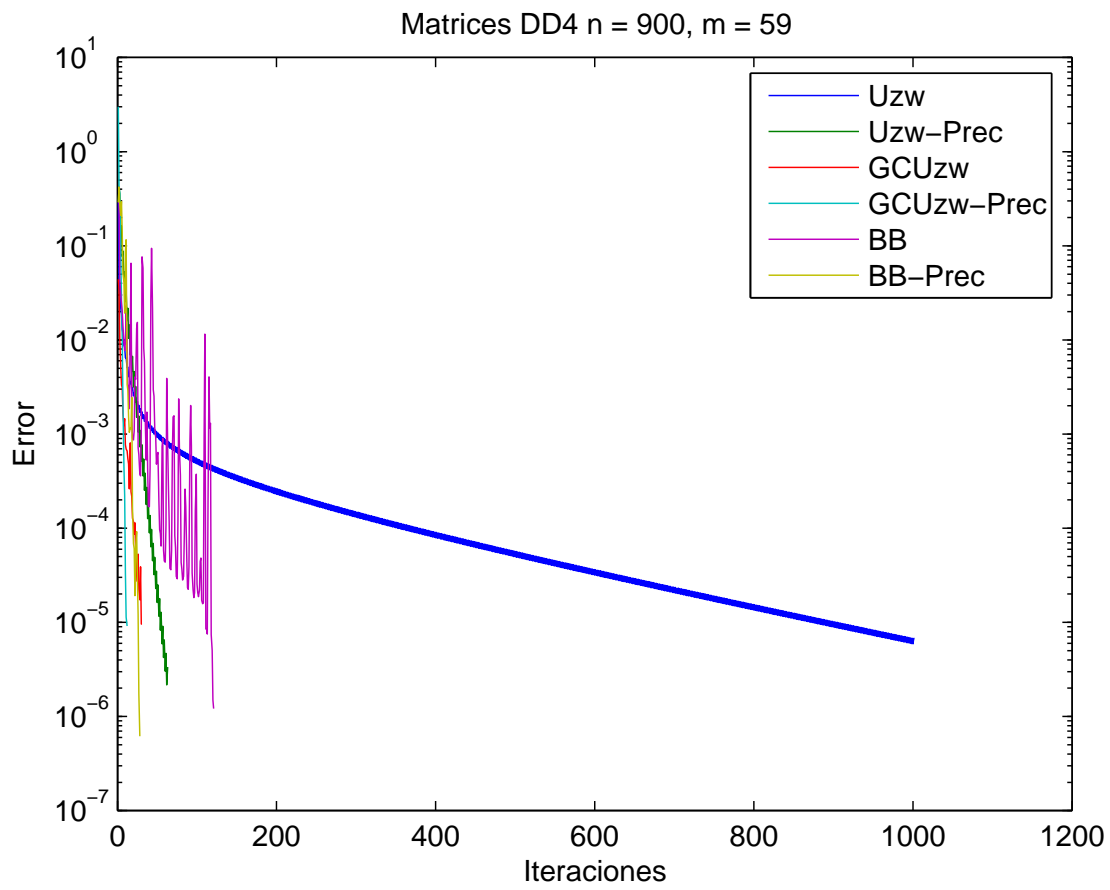


FIGURA 3.15. Experimento 14

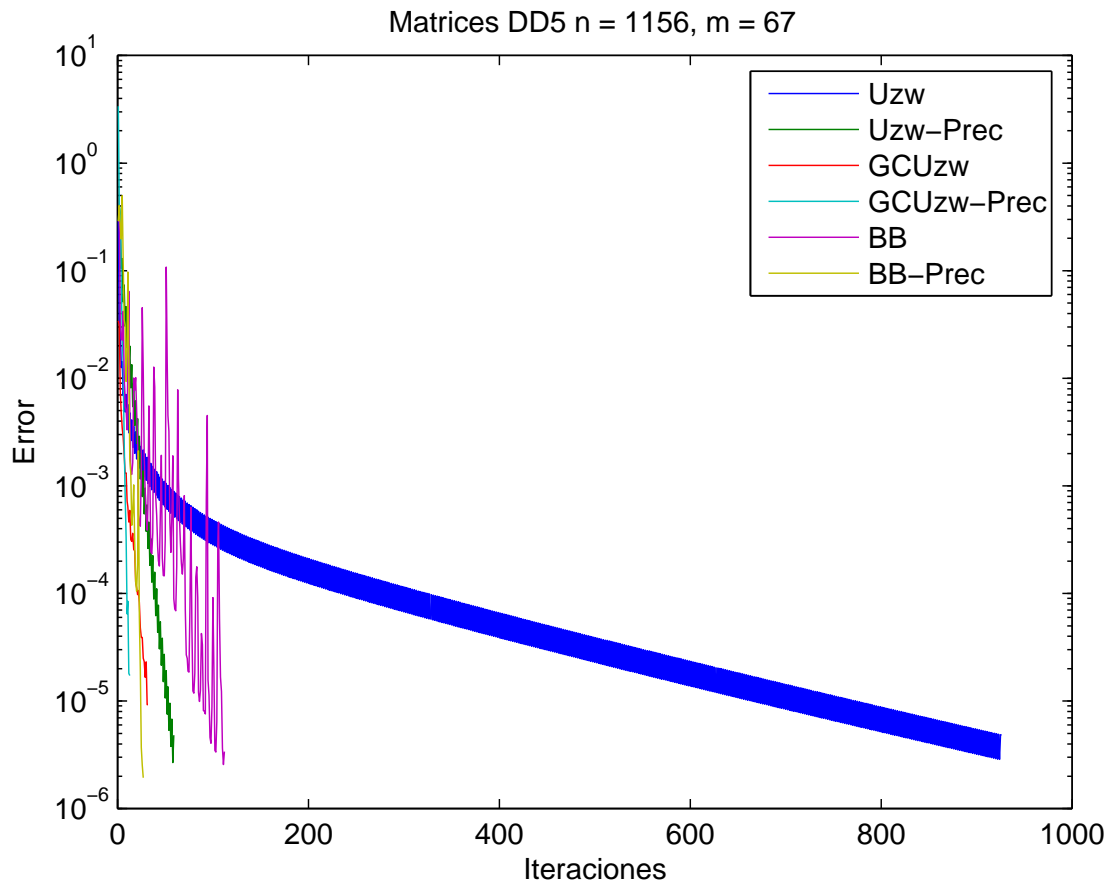


FIGURA 3.16. Experimento 15

Conclusiones

Durante el desarrollo de este trabajo se demostró que efectivamente se podía encontrar la solución aproximada a los sistemas lineales de puntos de ensilladura donde la matriz A es una matriz simétrica definida positiva y la matriz B de rango completo y en la cuál el vector g del lado derecho del sistema era distinto al vector cero aplicando el método de proyecciones oblicuas alternadas sobre las variedades lineales $V = \{x \in X / f - Ax \in (ker B)^\perp\}$ y $W = \{x \in X / Bx = g\}$ lo que amplía la aplicabilidad de este método sobre sistemas de punto de ensilladura más generales.

Los resultados experimentales, empleando 3 métodos iterativos distintos confirman que efectivamente hallar la solución al sistema (3.1) es equivalente a hallar la solución del sistema (3.9) en cuyo caso se verificó que la parte x, x_{*2} era el vector solución x_* trasladada un término c con c una solución particular de $Bx = g$.

Para futuros trabajos que enriquezcan esta línea de investigación se recomienda

- Mejorar la eficiencia de los códigos empleados en los métodos precondicionado para reducir el número de iteraciones internas requeridas para la convergencia de dichos métodos.
- Realizar pruebas adicionales empleando las técnicas del lagrangeano aumentado para mejorar el número de condición de la matriz por bloques \mathcal{A} y comparar el rendimiento en los métodos iterativos empleados.
- Ampliar el estudio de este método considerando los casos en que la matriz del bloque $(1, 1)$ no es positiva definida o al caso en que la matriz C de 0.1 es distinta de cero.

Bibliografía

- [1] N. ARONSZAJN. Theory of reproducing kernels. Trans. Amer. Math. Soc., vol 68, pp. 337-404, 1959.
- [2] M. BENZI y G. GOLUB. A preconditioner for generalized saddle point problems. SIAM J. Matrix Anal. Appl, 2004.
- [3] M. BENZI, G. GOLUB, y J. LIESEN. Numerical solution of saddle point problems. Acta Numérica, 2005.
- [4] A. BJORK, Numerical Methods for Least Squares Problems, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.
- [5] R. BRUZUAL, M. DOMINGUEZ, Espacios de Hilbert. Caracas, 2005.
- [6] T. CHAN y T. P. MATHEW, Domain decomposition algorithms, Acta Numerica , pp. 61-143,1994.
- [7] W. CHENEY y A. GOLDSTEIN. Proximity maps for convex sets. Proc. Amer. Math. Soc, vol 10, pp. 448-450, 1959.
- [8] Y. DAI y L. LIAO. R-linear convergence of the Barzilai and Borwein gradient method. IMA Journal of Numerical Analysis, 2002.
- [9] F. DEUTSCH. Rate of convergence of the method of alternating projections. ISNM, vol 72, pp. 96-107, 1984.
- [10] J. DIXMIER. Étude sur les variétés et les opérateurs de Julia avec quelques applications. Bull. Soc. Math. France, vol 77, pp. 11-101, 1949.
- [11] L.T. DOS SANTOS. A parallel subgradient projections method for the convex feasibility problem. J. Comput. Appl. Math., vol. 1, pp. 307-320, 1987.
- [12] H. ELMAN. Preconditioning for the steady-state navier-stokes equations with low viscosity. SIAM J. Sci. Comput, 1999.
- [13] C. FARHAT y F.-X. ROUX, A method of finite element tearing and interconnecting and its parallel solution algorithm, Int. J. Numer. Methods Engng., vol. 32 ,pp. 1205-1227, 1991.
- [14] K. FRIEDRICHS. On certain inequalities and characteristic value problems for analytic functions and for functions of two variables. Trans. Amer. Math. Soc. vol 41, pp. 321-364, 1937.
- [15] V. GIRAULT y P.A. RAVIART. Finite Element Method for Navier Stokes Equation. Theory and Algorithms. Springer Verlag, Berlin, Heidelberg, New York, 1986.
- [16] C. GREIF, G. GOLUB, y J. VARAH. Augmented lagrangian techniques for solving saddle point linear systems. Soumis a SIAM J. Matrix. Anal. and Appl., 2004.
- [17] I. HALPERIN. The product of projection operators. Acta Sci. Math. (Szeged), vol 23, pp. 96-99, 1962.

- [18] L.M. HERNANDEZ-RAMOS. Un operateur de projection pour les systemes lineaires couples, 2000.
- [19] L.M. HERNANDEZ-RAMOS. Projections alternees obliques pour les sytemes lineaires couples. Technical Report R2024, Publications du Laboratoire Jacques-Louis Lions (Paris VI). 2002.
- [20] L.M. HERNANDEZ-RAMOS. Alternating oblique projections for coupled linear systems. Numerical Algorithms. 2005.
- [21] Q. HU, Z. SHI, y D. YUf, Efficient solvers for saddle-point problems arising from domain decompositions with Lagrange multipliers, SIAM J. Numer. Anal., vol. 42 , pp. 905-933,2004.
- [22] X. JUVIGNY. Resolution des Grands Systemes Lineaires sur des Machines Massivement Paralleles. PhD thesis, Universite Pierre et Marie Curie, Paris VI, 1997.
- [23] A. KLAWONN y O. WIDLUND. Feti and Neumann-Neumann iterative substructuring methods:connections and new results. Communications on Pure and Applied Methematics, 2001.
- [24] A. QUATERNONI y A. VALLI, Numerical Approximation of Partial Diferential Equations, vol. 23 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1994.
- [25] G. STRANG, Introduction to Applied Mathematics, Wellesley-Cambridge Press, Wellesley, MA, 1986.
- [26] M. RAYDAN y R. ESCALANTE. Minimización irrestricta a gran escala. Libro en preparación.
- [27] M. RAYDAN y R. ESCALANTE. Alternating Projections Methods: Theory and Application, Caracas, 2005.
- [28] M. RAYDAN. On the Barzilai and Borwein choice of steplength for the gradient method. IMA Journal of Numerical Analysis, vol 13(3), p. 321-326, 1993.
- [29] M. ROZLOZNIK. Saddle point problems, iterative solution and preconditioning: A short overview. Resume de la memoire d'habilitation, 1994.
- [30] T. RUSTEN y R. WINTER. A Preconditioned iterative method for saddlepoint problems. SIAM J. Matrix. Anal., vol 13, pp. 887-904, 1992.
- [31] Y. SAAD. Iterative Methods for Sparse Linear Systems. PSW Publishing Company, 1996.
- [32] K.T. SMITH, D.C. SOLMON y S.L. WAGNER. Practical and mathematical aspects of the problem of reconstructing objects from radiographs. Bull. Amer. Math. Soc, vol 83, pp. 1227-1270,1977.
- [33] S. TUREK, Efcient Solvers for Incompressible Flow Problems. An Algorithmic and Computational Approach, vol. 6 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 1999.
- [34] C. F. VAN LOAN, On the method of weighting for equality constrained least squares problems, SIAM J. Numer. Anal., 23, pp. 851-864,1985.
- [35] J. VON NEUMANN. Functional operators vol. II. The geometry of orthogonal spaces. Annals of Math. Studies, vol 22, 1950.
- [36] P. WESSELING, An Introduction to Multigrid Methods, John Wiley and Sons Ltd., Chichester, UK, 1992.
- [37] M. H. WRINGHT, Interior methods for constrained optimization, Acta Numerica , pp. 341-407, 1992.