



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA

Aceleración de Métodos de Punto Fijo.

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela por el **Br. Eliézer Peñaloza** para optar al título de Licenciado en Matemática.

Tutor: Luis Manuel Hernández Ramos.

Caracas, Venezuela

Accleración de Métodos de Punto Fijo.

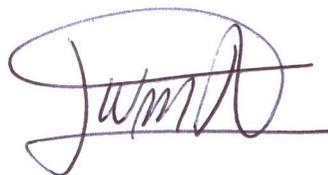
Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado “**Accleración de Métodos de Punto Fijo**”, presentado por el **Br. Eliézer Peñaloza**, titular de la Cédula de Identidad **18.780.687**, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de **Licenciado en Matemática**.



Luis Manuel Hernández Ramos
Tutor



Carmen Da Silva
Jurado



William La Cruz
Jurado

Resumen

Los métodos de punto fijo son usados para solucionar algunos problemas en cálculo numérico. Hay problemas y métodos que pueden verse como esquemas de punto fijo, tales como: el problema de factibilidad dividida (LSFP), la proyección de Landweber, y algunos de los métodos estacionarios para la solución de sistemas de ecuaciones lineales. Recientemente se desarrolló el método tipo residual, el cual ha demostrado ser más eficiente en comparación a otros métodos, este último fue usado para replantear los problemas anteriormente mencionados, y se mostró que los mismos también pueden ser vistos como problemas de tipo residual. Además, se realizaron experimentos numéricos, aplicando los métodos clásicos y tipo residual, a través de análisis y comparación de los resultados obtenidos se determinó que los métodos clásicos también pueden ser resueltos como tipo residual; además de ver que este esquema resultó ser una aceleración más eficiente (en términos de recursos computacionales) que la estrategia de punto fijo.

Índice general

Introducción	1
Capítulo 1. Operadores de punto fijo.	3
1.1. Preliminares	3
1.2. Teoremas de existencia de punto fijo	5
1.3. Iteración de Punto Fijo.	6
1.4. Aplicaciones de los Métodos de Punto Fijo.	7
Capítulo 2. Métodos del Tipo Residual para Problemas de Punto Fijo.	16
2.1. Algunos Métodos del Tipo Residual.	16
2.2. DF-SANE	19
2.3. Métodos tipo residual para problemas de punto fijo	20
Capítulo 3. Experimentos numéricos.	23
3.1. Proyección sobre algunos conjuntos.	24
3.2. Problema de factibilidad lineal dividida (LSFP).	26
3.3. Proyección de Landweber	30
3.4. Métodos Estacionarios.	34
Conclusión	37
Bibliografía	38

Introducción

Muchos problemas de las matemáticas pueden ser resueltos hallando los puntos fijos de un cierto operador T no expansivo. Generalmente, métodos iterativos son utilizados para hallar los puntos fijos de T con ayuda del computador, entre ellos, el más básico y el más utilizado es el método basado en la iteración de Banach-Picard [4] y [6]. Sin embargo, en algunas ocasiones, esta iteración puede tener una velocidad de convergencia lenta lo cual pudiera tener consecuencias a la hora de su resolución computacional, ya que cada iteración tiene consigo un costo en términos de tiempo de ejecución y memoria.

Por otra parte, los métodos tipo residual, SANE, DF-SANE, etc., son métodos recientes desarrollados por La Cruz, Martínez y Raydan [12],[13] para la resolución de problemas lineales o no lineales del tipo $F(x) = 0$. La característica principal de este tipo de métodos es que en la dirección de búsqueda no utilizan ninguna información del gradiente de la función, sino que simplemente utilizan el valor de la función $\pm F(x)$. Esto combinado con estrategias de globalización que garantizan su convergencia.

Anteriormente, Hernández-Ramos [9] y Hernández-Ramos, Escalante, y Raydan [10], enfocando los métodos de proyecciones alternantes y simultáneas como una iteración de punto fijo, aceleraron dicha iteración, con métodos de tipo residual, con resultados muy satisfactorios. Quedando abierta la generalización de este tipo de métodos para una clase más amplia de problemas de punto fijo.

En este trabajo, pretendemos resolver distintos problemas de punto fijo que se encuentran en la bibliografía, aplicando una aceleración del esquema de iteración de Banach-Picard, por medio de uno de los métodos de tipo residual como lo es el DF-SANE; y así realizar una verificación numérica de la factibilidad de este esquema y comprobaciones en cuanto a eficiencia computacional. Entre estos problemas resueltos se encuentran: el problema de factibilidad lineal dividida, el problema de factibilidad dividida, etc. Cabe destacar que estos son problemas con mucha aplicabilidad dentro de las matemáticas y que cualquier avance en la resolución numérica, supondría un beneficio en las áreas en las que se aplican.

El documento a continuación se divide de la siguiente manera: Primeramente, un marco teórico acerca de puntos fijos: definición, existencia y unicidad, iteración de Banach-Picard, algunas aplicaciones importantes, etc. Seguidamente, un capítulo en donde se describen brevemente los métodos de tipo residual, específicamente el DF-SANE, y la propuesta de aceleración de este trabajo. Luego, se describen experimentos numéricos en varios de los problemas que están asociados a operadores de punto fijo, haciendo la comparación entre la iteración clásica de Banach-Picard con la aceleración propuesta en este trabajo. Y finalmente nuestras conclusiones y conjeturas.

Operadores de punto fijo.

1.1. Preliminares

DEFINICIÓN 1.1. Un espacio métrico es un conjunto M , no vacío, de objetos (que llamaremos puntos) dotados de una función $\rho : X \times X \rightarrow \mathbb{R}$ que satisface las siguientes condiciones:

- $\rho(x, y) \geq 0$ para todo $x, y \in X$.
- $\rho(x, y) = 0$ si $x = y$.
- $\rho(x, y) = \rho(y, x)$ para todo $x, y \in X$.
- $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ para todo $x, y, z \in X$.

DEFINICIÓN 1.2. Un espacio normado es un par $(X, \|\cdot\|)$, formado por un espacio vectorial y una aplicación $\|\cdot\| : X \rightarrow \mathbb{R}$, llamada norma, con las siguientes propiedades:

- $\|x\| \geq 0$.
- $\|x\| = 0$ si y solo si $x = 0$.
- $\|\alpha x\| = |\alpha| \|x\|$.
- $\|x + y\| \leq \|x\| + \|y\|$.

Para todo $x, y \in X$.

DEFINICIÓN 1.3. Un espacio normado $(X, \|\cdot\|)$ se denomina **espacio de Banach** si es completo respecto a la norma inducida.

DEFINICIÓN 1.4. Sea V un espacio vectorial, y sea $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ una función. Se llama producto interno en V si cumple con las siguientes propiedades para todo $x, y, w \in V$, con $\alpha \in \mathbb{R}$:

- $\langle x, x \rangle \geq 0$
- $\langle x, x \rangle = 0$ si y solo si $x = 0$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x + y, w \rangle = \langle x, w \rangle + \langle y, w \rangle$

DEFINICIÓN 1.5. Si X es un espacio de Banach con la norma de producto interno se dice que $(X, \|\cdot\|)$ es un **espacio de Hilbert**.

DEFINICIÓN 1.6 (Punto Interior). Sea X un subconjunto de \mathbb{R}^n , y supongamos que existe $a \in X$. Entonces se denomina a un punto interior si existe una bola abierta con centro en a y contenida en X .

DEFINICIÓN 1.7. Un conjunto X de \mathbb{R}^n es un abierto si todos sus puntos son interiores.

DEFINICIÓN 1.8. Un conjunto X de \mathbb{R}^n es un conjunto cerrado si su complemento es un conjunto abierto de \mathbb{R}^n .

DEFINICIÓN 1.9. Un conjunto C se dice que es convexo si y solo si cumple:

$$\forall x, y \implies \lambda x + (1 - \lambda)y \in C, \lambda \in (0, 1).$$

DEFINICIÓN 1.10. La proyección ortogonal de u sobre el subespacio $X(v) \subset \mathbb{R}^n$ es el vector

$$\hat{u} = \frac{\langle u, v \rangle}{\langle v, v \rangle} v.$$

DEFINICIÓN 1.11. Un operador $T : H \rightarrow D$ es monótono si para todo $x, y \in D$

$$\langle Tx - Ty, x - y \rangle \geq 0.$$

DEFINICIÓN 1.12. Sea $\lambda \in [0, 2]$. El operador $T_\lambda = I + \lambda(I - T)$ es llamado relajación del operador $T : H \rightarrow D$. λ es llamado el parametro relajador. Si $\lambda \in (0, 2)$ entonces T_λ es llamado relajación estricta de T .

DEFINICIÓN 1.13. Un operador $T : H \rightarrow D$ es firmemente no expansivo si para todo $x, y \in D$

$$\langle Tx - Ty, x - y \rangle \geq \|Tx - Ty\|^2.$$

DEFINICIÓN 1.14. Un operador $T : H \rightarrow D$ es relajado firmemente no expansivo si T es una relajación de un operador firmemente no expansivo.

Algunos problemas de las matemáticas aplicadas se reducen a encontrar los puntos fijos de un cierto operador T . Los puntos fijos de T se definen de la siguiente manera:

DEFINICIÓN 1.15. Sea H un espacio de Hilbert, sea $X \subseteq H$ un subespacio de H y sea $T : X \rightarrow H$ un operador. Un punto fijo del operador T , es un punto z tal que $T(z) = z$. Al conjunto $Fix T = \{z \in X : T(z) = z\}$ se le conoce con el nombre de conjunto de los puntos fijos de T .

A continuación se enunciarán una serie de definiciones que caracterizan a los operadores y que son necesarias al considerar los teoremas de existencia y unicidad de los puntos fijos que se darán en la próxima sección.

DEFINICIÓN 1.16. Decimos que un operador $T : X \rightarrow H$ es:

- No expansivo, si

$$\|Tx - Ty\| \leq \|x - y\|,$$

para todo $x, y \in X$.

- Estrictamente no expansivo, si

$$\|Tx - Ty\| < \|x - y\|, \text{ o si } x - y = Tx - Ty,$$

para todo $x, y \in X$.

- Una α -contracción, donde $\alpha \in (0, 1)$ o, pequeña contracción, si

$$\|Tx - Ty\| \leq \alpha \|x - y\|,$$

para todo $x, y \in X$.

Para más detalles, ver por ejemplo: [4] y [6].

1.2. Teoremas de existencia de punto fijo

En esta sección se enunciarán varios teoremas que garantizan la existencia de puntos fijos. Primeramente, el conocido como el teorema de punto fijo de Banach, el cual es ampliamente usado en varias áreas de la matemática.

TEOREMA 1.17 (Banach,1922). *Sea X un espacio con métrica completa y $T : X \rightarrow X$ una contracción. Entonces T tiene exactamente un punto fijo $x^* \in X$. Además, para cualquier x , la órbita $\{T^k(x)\}_{k=0}^{\infty}$ converge a x^* con una tasa de progresión geométrica.*

Como puede apreciarse, el teorema de Banach es válido en los espacios con una métrica completa, por lo tanto, es válido en espacios de Hilbert, que es nuestro caso de interés. Además de garantizar la existencia y unicidad del punto fijo bajo las condiciones descritas, el Teorema de Banach sugiere un método iterativo para hallar el punto fijo, esto es, mediante la generación de la órbita $\{T^k(x)\}_{k=0}^{\infty}$. Esta órbita se genera con la iteración:

$$(1.1) \quad x^{(k+1)} = Tx^{(k)},$$

para un $x_0 \in X$ dado, donde H es un espacio de Hilbert, con $X \subset H$, y $T : X \rightarrow X$.

Este teorema es bastante restrictivo ya que demanda que el operador T sea una contracción. Hay problemas prácticos en los que se requiere encontrar puntos fijos de operadores que no son contracciones, sino simplemente se trata de operadores no expansivos. Esto sucede, por ejemplo, cuando hay más de un punto fijo.

Los teoremas a continuación garantizan la existencia (más no la unicidad) de puntos fijos de T en condiciones menos restrictivas. Sin embargo, bajo esas condiciones el esquema iterativo que da el teorema de Banach no necesariamente converge.

TEOREMA 1.18 (Brouwer,1912). *Sea $X \subseteq \mathbb{R}^n$ un conjunto no vacío, compacto y convexo, y sea $T : X \rightarrow X$ continuo. Entonces T tiene un punto fijo.*

TEOREMA 1.19 (Schauder,1930). *Sea X un subconjunto de un espacio de Banach, no vacío, compacto y convexo, y $T : X \rightarrow X$ continuo. Entonces T tiene un punto fijo.*

Hay otro teorema el cual es aplicable cuando en un espacio no hay compacidad, pero tal espacio es acotado.

TEOREMA 1.20 (Browder-Göhde-Kirk,1965). *Sea X no vacío, cerrado, acotado, convexo, y subconjunto de un conjunto de Banach uniformemente convexo (es decir, de un espacio de Hilbert H), y $U : X \rightarrow X$ no expansivo. Entonces U tiene un punto fijo.*

Para mayor información, consultar: [5], y [6].

1.3. Iteración de Punto Fijo.

El matemático francés Picard, fué de los primeros en estudiar los esquemas iterativos que conocemos hoy en día como las iteraciones de Picard. Muchos de los métodos iterativos

pueden escribirse a través de la siguiente relación:

$$(1.2) \quad x^{(k+1)} = T_k x^{(k)},$$

para un $x_0 \in X$ dado, donde H es un espacio de Hilbert, con $X \subset H$, y $T_k : X \rightarrow X$ es una sucesión de operadores. La sucesión generada por la relación (1.2), es conocida con el nombre de sucesión de aproximación. Nuestro trabajo estará enfocado en el caso particular $T_k = T$, para $k \geq 0$, que es el caso de la iteración sugerida por el teorema de Banach. La sucesión generada por (1.2) es conocida como la órbita de T .

En diversas aplicaciones, mediante la relación (1.2), se genera una sucesión que deberá converger a la solución del problema de interés. Si el operador T es continuo, y la sucesión $\{T_k x^0\}$ converge a un cierto x^* , entonces $Tx^* = x^*$, es decir, convergerá a un punto fijo de T . Posteriormente se estudiarán bajo que condiciones la sucesión que define el método iterativo (1.2) converge.

Para una mayor comprensión, ver, por ejemplo: [3] y [4].

1.4. Aplicaciones de los Métodos de Punto Fijo.

En esta parte veremos como resolver algunos problemas a través de métodos iterativos de punto fijo. Para ello consideraremos la siguiente notación:

- H un espacio de Hilbert.
- $\langle \cdot, \cdot \rangle$ un producto escalar en H .
- $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ la norma inducida por $\langle \cdot, \cdot \rangle$.
- $D \subset H$ un subconjunto cerrado y convexo.
- $P_D(x)$ es la proyección del vector x sobre el conjunto D .

1.4.1. Problema de Minimización Convexa. Sea $f : H \rightarrow \mathbb{R}$ una función convexa. El problema de minimización con restricciones viene dado de la siguiente manera:

$$(1.3) \quad \min_{x \in D} f(x).$$

Es decir, hay que hallar x^* , tal que $f(x^*) \leq f(x)$ para toda $x \in D$. El punto x^* es llamado el minimizador de la función f en D , o solución óptima del problema (1.3). El valor $f^* = f(x^*)$ es llamado el mínimo de la función f en D .

1.4.1.1. *Operador asociado.* Si f es diferenciable, entonces el problema (1.3) es equivalente a hallar el punto fijo del operador $T : D \rightarrow D$.

$$T(x) = P_D(x - \gamma \nabla f(x)),$$

donde $\gamma > 0$. Además si ∇f es Lipschitz continua, entonces T es un operador relajado firmemente no expansivo para $\gamma \in [0, \frac{2}{L}]$.

1.4.1.2. *Iteración de punto fijo.* La iteración de punto fijo dado para este problema, es:

$$x_{k+1} = T(x_{k+1}) = P_D(x_k - \gamma \nabla f(x_k)),$$

con x_0 dado.

1.4.1.3. *Convergencia.* Si $\gamma \in (0, \frac{2}{L})$ donde L es la constante de Lipschitz del gradiente de f , la iteración anterior converge a un punto fijo, y por lo tanto a la solución del problema. Además, si f es fuertemente convexo, entonces el problema (1.3) tiene solución única.

1.4.2. Desigualdad Variacional. Sea $F : D \rightarrow H$ un operador monótono. El problema de desigualdad variacional consiste en hallar un $x \in D$ tal que:

$$(1.4) \quad \langle F(x), y - x \rangle \geq 0,$$

para todo $y \in D$.

Se puede observar que el problema de desigualdad variacional generaliza al problema de minimización convexa (1.3) y a otros problemas que trataremos aquí.

1.4.2.1. *Operador asociado.* El problema (1.4) es equivalente a hallar un punto fijo del operador $T : D \rightarrow D$.

$$T(x) = P_D(x - \gamma F(x)),$$

dónde $\gamma > 0$.

Si F es estrictamente monótona y Lipschitz continua, el operador T tendrá un único punto fijo $x^* \in D$, que coincidirá con la única solución del problema de desigualdad variacional en este caso.

1.4.2.2. *Iteración de punto fijo.* La iteración del punto fijo en el problema de desigualdad variacional es:

$$x_{k+1} = T(x_{k+1}) = P_D(x_k - \gamma F(x_k)).$$

1.4.2.3. *Convergencia.* Si $\gamma \in (0, \frac{2}{L})$, donde L es la constante de Lipschitz de F , el esquema iterativo anterior converge al punto fijo del operador, en consecuencia converge a la solución del problema planteado.

1.4.3. Problema de Factibilidad Convexa. Sea $C_i \subset H$, $i \in I = \{1, \dots, m\}$ subconjuntos no vacíos, cerrados, y convexos. El problema de factibilidad convexa (**CFP**), tiene la siguiente forma:

Encontrar $x \in C = \bigcap_{i \in I} C_i$, si dicho punto existe.

El **CFP** se puede formular como la minimización de la función de proximidad convexa $f : H \rightarrow \mathbb{R}$.

$$f(x) = \frac{1}{2} \sum_{i \in I} \omega_i d^2(x, C_i),$$

donde $\omega = (\omega_1, \dots, \omega_m)^T \in \mathbb{R}^m$ es un vector de pesos, es decir, $\omega_i > 0$, $i \in I$, $\sum_{i \in I} \omega_i = 1$. Y $d(x, C_i)$ es la distancia del punto x al conjunto C_i .

1.4.3.1. *Operador asociado.* Por las condiciones necesarias y suficientes de optimización para el problema de minimización convexa sin restricciones, la minimización de f equivale a la resolución del siguiente problema:

Hallar el punto fijo (si existe) del operador $T : H \rightarrow H$, definido de la siguiente manera:

$$T(x) = \sum_{i \in I} \omega_i P_{C_i}(x).$$

Este operador es el usado en el método de Cimmino, llamado método de proyecciones simultaneas.

Otra manera para resolver el problema de factibilidad convexa es mediante el siguiente operador:

$$T(x) = P_{C_m} \dots P_{C_1}(x),$$

donde P_{C_i} , $i \in I$, es la proyección sobre el conjunto convexo C_i . Dicho operador es empleado en el Método de Proyecciones Alternantes.

1.4.3.2. *Iteración de punto fijo.* En el caso del método de Cimmino, la iteración de punto fijo para el **CFP** viene dada por:

$$x_{k+1} = T(x_{k+1}) = \sum_{i \in I} \omega_i P_{C_i}(x_k),$$

con x_0 dado, y en el caso del método de Proyecciones Alternantes, la iteración de punto fijo viene dada por:

$$x_{k+1} = T(x_k) = P_{C_m} \dots P_{C_1}(x_k),$$

con x_0 dado.

1.4.3.3. *Convergencia.* Si la intersección C de los conjuntos C_1, \dots, C_m es no vacía, entonces $C = \text{Fix}(T)$, y la órbita de la sucesión $\{T^k x\}$ converge a un elemento de C . Si C es vacío, la sucesión converge al mínimo de la función f , donde

$$f(x) = \sum_{i \in I} \omega_i d^2(x, C_i).$$

Si además cada C_i es un subespacio o una variedad lineal (subespacio afín), tendremos:

$$\lim_{k \rightarrow \infty} T^k(x_0) = P_{\cap_{i \in I} C_i}(x_0).$$

Es decir, en el caso en que los C_i sean subespacios no va a converger a cualquier punto factible, sino a aquel más cercano del x_0 en la intersección de los C_i .

1.4.4. Problema de Factibilidad Lineal. Sea $A \in \mathbb{R}^{m \times n}$, y $b \in \mathbb{R}^m$. El problema de factibilidad lineal (**LFP**) es de la forma:

$$\text{Encontrar } x \in \mathbb{R}^n \text{ con } Ax \leq b, \text{ si tal } x \text{ existe.}$$

Cabe destacar que el **LFP**, es un caso especial del **CFP**.

1.4.5. Problema de Factibilidad Dividida. Sean $C \in H_1$, y $Q \in H_2$ subconjuntos cerrados y convexos de espacios de Hilbert, H_1 , y H_2 respectivamente. El problema de factibilidad dividida (**SFP**), tiene la forma:

$$\text{Encontrar un } x \in C, \text{ tal que } Ax \in Q, \text{ si tal } x \text{ existe,}$$

donde $A : H_1 \rightarrow H_2$, es un operador lineal acotado.

El **SFP** también puede verse como una minimización de la función de proximidad $f : C \rightarrow \mathbb{R}$

$$f(x) = \frac{1}{2} \|P_Q(Ax) - Ax\|^2.$$

1.4.5.1. *Operador.* Por las condiciones de optimalidad necesarias y suficientes para el problema de minimización convexa restringido, la minimización de f es equivalente al problema:

$$\text{Hallar el punto fijo (si existe) del operador } T : C \rightarrow C$$

definido así:

$$T(x) = P_C(x + \gamma A(P_Q(Ax) - Ax)),$$

para $\gamma > 0$.

1.4.5.2. *Iteración de punto fijo.* La iteración para el problema Factibilidad Dividida es:

$$x_{k+1} = T(x_{k+1}) = P_C(x_k + \gamma A(P_Q(Ax_k) - Ax_k)).$$

Esta iteración también es conocida como el algoritmo CQ .

1.4.5.3. *Convergencia.* Si $\gamma \in \left(0, \frac{2}{\rho(AA^T)}\right)$, entonces el esquema iterativo anterior converge a algún punto fijo, y de esta forma a las solución del problema.

1.4.6. Problema de factibilidad dividida lineal. Sea $C \subset \mathbb{R}^n$ cerrado y convexo. Sea $A \in \mathbb{R}^{n \times m}$, y sea $b \in \mathbb{R}^m$. El problema de factibilidad dividida lineal (**LSFP**) tiene la forma:

$$\text{Encontrar } x \in C \text{ con } Ax \leq b,$$

si tal x existe, el **LSFP** es una caso particular del (**SFP**).

Sea $r(x) = (\rho_1(x), \rho_2(x), \dots, \rho_m(x)) = Ax - b$ el vector residual, y sea $r_+(x) = \max\{0, r_x\}$ la parte positiva de $r(x)$. El **LSFP** puede ser formulado como una minimización de una función de proximidad convexa $f : C \rightarrow \mathbb{R}$, que está definida de la siguiente manera:

$$f(x) = \frac{1}{2} \sum_{i=1}^m \nu_i (\alpha_i^T x - \beta_i)^2 = \frac{1}{2} \sum_{i=1}^m \nu_i (\rho_i)_+,$$

donde $\nu = (\nu_1, \nu_2, \dots, \nu_m)^T \in \mathbb{R}_{++}^m$, $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})^T$ es la i -ésima fila de A , con $i = 1, 2, \dots, m$, y $b = (\beta_1, \beta_2, \dots, \beta_m)^T$. Si \mathbb{R}^m posee producto escalar $\langle \cdot, \cdot \rangle_\nu$, definido por $\langle x, y \rangle_\nu = x^T V y$, donde $V = \text{diag}(\nu)$, entonces, se puede probar $P_Q(Ax) - Ax = -V r_+(x)$.

1.4.6.1. *Operador.* En consecuencia la minimización de f puede verse como este problema:

Encontrar un punto fijo (si existe del operador $T : C \rightarrow C$), definido por:

$$T(x) = P_C(x - \gamma A^T V r_+(x)),$$

con $\gamma > 0$ para todo x , y $V = \text{diag}(\nu)$. El algoritmo que genera la resolución de este problema es conocido como el algoritmo **CQ**.

1.4.6.2. *Iteración de punto fijo.* La iteración para el **LSFP** es:

$$x_{k+1} = T(x_{k+1}) = P_C(x_k - \gamma A^T V r_+(x_k)),$$

con x_0 dado.

1.4.6.3. *Convergencia.* Si $\gamma \in \left(0, \frac{2}{\rho(AA^T)}\right)$, donde $\rho(AA^T)$ es el autovalor más grande para toda $x \in C$, la iteración vista anteriormente converge a un punto fijo, en consecuencia también a la solución del problema.

1.4.7. Proyección de Landweber. Si $C \subset \mathbb{R}^n$, sea $Q = b$. Entonces resolver el **SFP**, es equivalente a resolver el sistema de ecuaciones lineales $Ax = b$. Si C es un conjunto propio de \mathbb{R}^n , tal como un cono no negativo, las soluciones que se buscan de $Ax = b$, son aquellas que están dentro de C , si las hay. Es un caso particular del algoritmo **CQ**, conocido como la proyección de Landweber, a su vez también es un caso particular del **SFP**.

1.4.7.1. *Operador.* Así como en el **LSFP** la solución del problema viene dada en hallar los puntos fijos del operador asociado, el cual es:

$$T(x) = P_C(x + \gamma A^T(b - Ax)),$$

$\gamma \in \left(0, \frac{2}{\rho(AA^T)}\right)$, donde $\rho(AA^T)$ es el autovalor más grande para toda $x \in C$, y P_C es la proyección sobre el conjunto C .

1.4.7.2. *Iteración de punto fijo.* El Método de Landweber proyectado, es un caso particular del algoritmo **CQ**. La iteración viene dada por:

$$x_{n+1} = T(x_{n+1}) = P_C(x_n + \gamma A^T(b - Ax_n)).$$

1.4.7.3. *Convergencia.* Como el método de Landweber al ser un caso particular del **LSFP**, converge bajo la misma condición de convergencia del caso anterior, es decir, si $\gamma \in \left(0, \frac{2}{\rho(AA^T)}\right)$.

Para mayores detalles sobre todos estos problemas e iteraciones asociadas, se puede consultar, por ejemplo: [3], [4], [5], y [6].

1.4.8. Métodos Estacionarios para Sistemas de Ecuaciones Lineales.

1.4.8.1. *Sistemas de ecuaciones lineales (S.E.L).* Los sistemas de ecuaciones lineales surgen en muchos temas de la matemática aplicada y computación científica. Con frecuencia se dan de forma natural, pero hay casos en que los mismos pueden tener su origen en la aproximación de sistemas de ecuaciones no lineales, o mediante ecuaciones diferenciales, etc.

Una ecuación lineal, es un polinomio con variables de primer grado. Cuando hay 2 ó más ecuaciones lineales es conocido como un sistema de ecuaciones lineales; el cual puede ser visto de la siguiente manera:

DEFINICIÓN 1. Un sistema de ecuaciones lineales (o S.E.L), es un problema en el cual debe hallarse un vector x , que cumpla con la siguiente igualdad

$$Ax = b,$$

donde $A \in \mathbb{R}^{m \times n}$, y $b \in \mathbb{R}^m$.

1.4.8.2. *Solución de S.E.L. Mediante métodos iterativos estacionarios.* En muchos casos resulta de gran dificultad hallar la solución exacta de un S.E.L, a veces basta con tener una buena aproximación a dicha solución. Para ello suelen utilizarse métodos iterativos que trabajan realizando aproximaciones sucesivas a la solución, a partir de un iterado inicial. Algunos de estos métodos pueden interpretarse como métodos de punto fijo.

Analizando la ecuación $Ax = b$, podremos reescribir la misma de la siguiente manera, $0 = -Ax + b$, ahora consideremos una matriz Q , invertible, y sumemos Qx de ambos lados de la igualdad, nos queda $Qx = Qx - Ax + b$, además $Qx = (Q - A)x + b$, luego multiplicando por Q^{-1} , tenemos $x = (I - Q^{-1}A)x + Q^{-1}b$, de esta manera, surge naturalmente una expresión iterativa de la siguiente forma:

$$(1.5) \quad x_{k+1} = (I - Q^{-1}A)x_k + Q^{-1}b, \quad k = 1, 2, \dots$$

Si observamos la expresión (1.5), es una iteración de punto fijo. Además es estacionario porque $(I - Q^{-1}A)$, y $Q^{-1}b$ son constantes en cada iteración.

1.4.8.3. *Teoremas de convergencia.* Con la expresión (1.5) puede realizarse un análisis sobre la convergencia del método. Supongamos que x es la solución a la ecuación (1.5), y A es no singular, entonces tenemos que:

$$\begin{aligned} x_k - x &= (I - Q^{-1}A)x_{k-1} + Q^{-1}b - x \\ &= (I - Q^{-1}A)x_{k-1} - (I - Q^{-1}A)x \\ &= (I - Q^{-1}A)(x_{k-1} - x). \end{aligned}$$

De esta forma tenemos un vector de error e_k .

$$(1.6) \quad e_k = (I - Q^{-1}A)(x_{k-1} - x).$$

Queremos que e_k sea más que pequeño a medida que k incremente su valor. La ecuación (1.6) muestra que e_k es menor que e_{k-1} , siempre que $(I - Q^{-1}A)$ también sea pequeño. A su vez, I será cercano a $Q^{-1}A$, en consecuencia A será más cercano a Q . En consecuencia, se desprende el siguiente teorema:

TEOREMA 1.21 (Radio Espectral). *la sucesión generada por $x_{k+1} = (I - Q^{-1}A)x_k + Q^{-1}b$, sea x_0 cualquiera. La condición necesaria y suficiente para la convergencia de la sucesión, es que todos los autovalores de $(I - Q^{-1}A)$, deben estar dentro de la bola unitaria $|z| < 1$, del plano complejo.*

En pocas palabras, el método converge si $\rho(I - Q^{-1}A) < 1$. Donde $\rho(A)$, es el radio espectral de la matriz A .

1.4.9. Método de Richardson. De acuerdo con la expresión (1.5), para este método sustituimos la matriz I en Q , de esta manera queda la siguiente expresión:

$$x_{k+1} = \underbrace{(I - A)}_{T_R} x_k + b.$$

El cual para que haya convergencia es suficiente que ocurra que:

$$\rho(T_R) = \rho(I - A) < 1.$$

1.4.10. Método de Jacobi. Para este método separemos la matriz A , en tres partes, es decir: $A = L + D + U$, donde L es la parte inferior de A , D es la matriz diagonal de A , y U es la parte superior de la matriz A , sin embargo, la matriz A queda de la siguiente forma:

$$\text{Sea } A = M + N. \text{ Donde } M = D, N = L + U.$$

Sustituyendo A en la expresión (1.5), nos queda

$$x_{k+1} = \underbrace{M^{-1}N}_{T_J} x_k + M^{-1}b.$$

Luego sustituyendo M , y N , tenemos

$$x_{k+1} = \underbrace{-D^{-1}(L + U)}_{T_J} x_k + D^{-1}b.$$

El cual para que haya convergencia es suficiente que ocurra que:

$$\rho(T_J) = \rho(-D^{-1}(L + U)) < 1.$$

1.4.11. Método de Gauss-Seidel. En este método también haremos algo similar que en la parte anterior, podemos ver la matriz $A = L + D + U$. También puede escribirse la matriz A así:

Sea $A = M - N$. Donde $M = D + L$, $N = -U$.

Al aplicar el cambio de A , tenemos la siguiente expresión

$$x_{k+1} = \underbrace{M^{-1}N}_{T_{G-S}} x_k + M^{-1}b.$$

Ahora si sustituimos los valores de M , y N . Finalmente resulta lo siguiente

$$x_{k+1} = - \underbrace{(L + D)^{-1}U}_{T_{G-S}} x_k + (L + D)^{-1}b.$$

El cual para que haya convergencia es suficiente con satisfacer que:

$$\rho(T_{G-S}) = \rho(-(L + D)^{-1}U) < 1.$$

Para más informacion, consultar: [7], [8], y [11].

Capítulo 2

Métodos del Tipo Residual para Problemas de Punto Fijo.

Consideremos la siguiente ecuación:

$$(2.1) \quad F(x) = b,$$

donde $F : H \rightarrow H$, y H es un espacio de Hilbert. El residual de (2.1) en un punto x viene dado por la siguiente expresión:

$$r(x) = b - F(x).$$

Un método iterativo tipo residual, es un método que utiliza el residual como dirección de búsqueda. En general, los métodos de tipo residual presentan la forma:

$$(2.2) \quad x_{k+1} = x_k \pm \alpha_k r(x_k),$$

con $\alpha_k \in \mathbb{R}, \forall k$.

2.1. Algunos Métodos del Tipo Residual.

2.1.1. Método de Richardson. En la resolución del problema $Ax = b$, el residual en un punto x dado vendría siendo

$$r(x) = b - Ax.$$

El Método de Richardson utiliza el residual como dirección de búsqueda, esto puede apreciarse en su iteración:

$$x_{k+1} = x_k + \alpha r(x_k),$$

con $r(x_k) = b - Ax_k$. Como vimos en (1.4.9), la iteración del método de Richardson viene dada por la siguiente ecuación:

$$x_{k+1} = \underbrace{(I - A)}_{T_R} x_k + b.$$

donde $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$, y T_R es la matriz de iteración asociada al método.

2.1.2. Método de Mínimo Descenso. Sea $A \in \mathbb{R}^{n \times n}$ una matriz cuadrada y sea $b \in \mathbb{R}^n$ un vector. El método de Mínimo Descenso o Método de Cauchy es un método que se utiliza para minimizar la siguiente función cuadrática:

$$(2.3) \quad q(x) = \langle x, Ax \rangle - 2 \langle x, b \rangle.$$

La iteración del método de mínimo descenso tiene la siguiente forma:

$$x_{k+1} = x_k + t_k v_k,$$

donde $t_k = \langle v_k, b - Ax_k \rangle / \langle v_k, Av_k \rangle$, y v_k es el gradiente negativo de q evaluado en x_k , es decir, $v_k = -\nabla q(x_k)$. En el caso del problema de sistemas de ecuaciones lineales $Ax = b$, con A simétrica y definida positiva resulta que este gradiente negativo es el residual $r_k = -\nabla q(x_k) = b - Ax_k$ y el mínimo de q coincide con la solución del sistema de ecuaciones lineales $Ax = b$.

2.1.2.1. *Convergencia del Método.* En esta parte veremos bajo que condiciones converge el método del mínimo descenso. Tomaremos en cuenta que q debe ser una función cuadrática, A una matriz simétrica y positivo definida (SPD), y q es de la forma

$$q(x) = x^T A x + b x.$$

Supongamos que q tiene un mínimo en x_* y definamos una función $E(x)$ de la forma

$$E(x) = \frac{1}{2}(x - x_*)^T A(x - x_*).$$

Vea que $E(x)$ posee un mínimo en x_* .

LEMA 2.1. *Sea $\{x_k\}$ una sucesión generada por el método de descenso para aproximar el mínimo de una función q cuadrática con matriz A SPD y sea $g_k = \nabla q(x_k)$. Entonces*

$$E(x_{k+1}) = \left\{ 1 - \frac{(g_k^T g_k)^2}{g_k^T A g_k g_k^T A^{-1} g_k} \right\} E(x_k).$$

LEMA 2.2 (Desigualdad de Kantorovich.). *Sea A una matriz tal que $A \in \mathbb{R}^{n \times n}$, y SPD, con λ_n y λ_1 , como los autovalores más grande y más pequeño de A respectivamente, entonces*

$$\frac{(x^T x)^2}{(x^T A x)(x^T A^{-1} x)} \geq \frac{4(\lambda_1 \lambda_n)}{(\lambda_1 + \lambda_n)^2}.$$

TEOREMA 2.3 (Convergencia del método de mínimo descenso para el caso cuadrático.).
 Para cualquier $x_0 \in \mathbb{R}^n$, el método de mínimo descenso converge al mínimo de x_* de q , y

$$E(x_{k+1}) \leq \left(\frac{(\lambda_n - \lambda_1)}{(\lambda_n + \lambda_1)} \right)^2 E(x_k),$$

donde λ_1 , y λ_n , son los autovalores más pequeño y más grande respectivamente.

DEMOSTRACIÓN 1. Usando los lemas 2.1 y 2.2, tenemos que

$$E(x_{k+1}) = \left\{ 1 - \frac{(g_k^T g_k)^2}{g_k^T A g_k g_k^T A^{-1} g_k} \right\} E(x_k).$$

$$E(x_{k+1}) \leq \left\{ 1 - \frac{4(\lambda_1 \lambda_n)}{(\lambda_1 + \lambda_n)^2} \right\}.$$

Haciendo algunas operaciones, se tiene el resultado.

2.1.3. Método del Gradiente Espectral. El método del gradiente espectral fue propuesto por Barzilai y Borwein [1], también fue analizado por Raydan [14] para funciones cuadráticas estrictamente convexas. Este método es un método tipo gradiente, por lo que su iteración tiene la forma:

$$(2.4) \quad x_{k+1} = x_k - \alpha_k g_k,$$

con $g_k = \nabla q(x_k)$, sin embargo, difiere del método de Cauchy en la escogencia del paso α_k .

El método del gradiente espectral es no monótono, es decir, propone una nueva longitud de paso $\alpha_k = \frac{1}{\delta_k}$ que no garantiza descenso de la función objetivo. Esta longitud de paso es obtenida al resolver el sistema lineal sobredimensionado:

$$(2.5) \quad y_{k+1} = \delta_k s_{k-1},$$

donde $s_{k-1} = x_k - x_{k-1}$, y $y_{k+1} = y_{k-1} = g_k - g_{k-1}$, con $g_k = \nabla q(x_k)$. Para los métodos espectrales, el escalar δ_k viene de resolver el sistema (2.5) en el sentido de los mínimos cuadrados, como:

$$(2.6) \quad \delta_k = \frac{s_{k-1}^t y_{k-1}}{s_{k-1}^t s_{k-1}},$$

si $s_{k-1} \neq 0$. Esta forma de escoger δ_k está inspirada en el cálculo de un paso Cuasi-Newton $SCN = -A_k^{-1} \nabla q(x_k)$, donde $A_k \in \mathbb{R}^{n \times n}$ es una aproximación del Hessiano de q en x_k . Esta aproximación, usualmente se escoge de forma tal que satisfaga la ecuación de la secante,

$$(2.7) \quad A_k s_{k-1} = y_{k-1},$$

restringiendo A_{k+1} a la familia de múltiplos escalares de la identidad y exigiendo que la ecuación de la secante se satisfaga en el sentido de los mínimos cuadrados.

Si consideramos el caso cuando q es cuadrática, entonces (2.6) puede escribirse como:

$$(2.8) \quad \delta_{k+1} = \frac{s_k^t A_k s_k}{s_k^t s_k}.$$

En este caso, δ_{k+1} resulta ser el cociente de Rayleigh de A evaluado en s_k . Igualmente, se cumple:

$$(2.9) \quad \delta_{k+1} = \frac{g_k^t A_k g_k}{g_k^t g_k},$$

que es el cociente de Rayleigh evaluado en g_k .

Para funciones cuadráticas estrictamente convexas, se presenta el siguiente resultado de convergencia:

TEOREMA 2.4 ([9]). *Sea $f(x)$ una función cuadrática estrictamente convexa. Sea $\{x_k\}$ la sucesión generada por el método del gradiente espectral y x_* el único minimizador de q . Entonces, o bien $x_j = x_*$ para algún j finito, o la sucesión $\{x_k\}$ converge a x_* .*

2.2. DF-SANE

DF-SANE es un método tipo residual para hallar los ceros de una función $F(x)$. Las iteraciones tipo residual (SANE ó DF-SANE) planteadas y estudiadas en [12] y [13], para

problemas finito dimensionales, tienen la siguiente forma:

$$(2.10) \quad x_{k+1} = x_k \pm \alpha_k F(x_k),$$

donde $\alpha_k > 0$, es el tamaño de paso y la dirección de búsqueda es $F(x_k)$ ó $-F(x_k)$, dependiendo de cuál sea la dirección de descenso para la función de mérito,

$$(2.11) \quad f(x) = \|F(x)\|^2 = \langle F(x), F(x) \rangle.$$

Estas ideas son muy efectivas y son esquemas competitivos para la resolución de sistemas no lineales en gran escala, cuando los tamaños de paso son elegidos convenientemente. La convergencia de (2.10) es alcanzada para sistemas no lineales generales, cuando se asocia con las búsquedas lineales (line search) no monótonas y libre de derivadas, como se ha descrito en [13]. Cuando el Jacobiano de $F(x)$ es simétrico y positivo semidefinido, habrá convergencia en el método puro, es decir, sin estrategias de globalización, y moviéndose a lo largo de la dirección $-F(x)$, la cuál, en este caso será siempre una dirección de descenso.

Para la elección de un tamaño de paso $\alpha_k > 0$, hay muchas opciones que nos garantizan la convergencia. Una de las más conocidas es la elección no monótona del paso espectral, la posee propiedades interesantes, y que está definida como:

$$(2.12) \quad \alpha_k = \frac{\langle s_{k-1}, s_{k-1} \rangle}{\langle s_{k-1}, y_{k-1} \rangle},$$

donde $s_{k-1} = x_k - x_{k-1}$, y $y_{k-1} = F(x_k) - F(x_{k-1})$. Mediante la aplicación del tamaño de paso (2.12), se reduce el trabajo computacional, la cual acelera la convergencia del método, y se utilizan las dos últimas iteraciones, de tal forma que, incorpora la información de primer orden de la dirección de búsqueda.

El algoritmo 1 produce una sucesión $\{x_k\}$ que converge globalmente, i.e, para cualquier iterado inicial x_0 , a un vector x_* tal que $F(x_*) = 0$.

Para mas detalles ver: [2], [9], [12], y [13].

2.3. Métodos tipo residual para problemas de punto fijo

La propuesta principal de este trabajo es la utilización de los métodos de tipo residual para acelerar la iteración de punto fijo.

Algoritmo 1 DFSANE para $F(x)$, en su forma pura.

1: Dado $x_0 \in H, \alpha_0 \in \mathbb{R}, \alpha_0 \neq 0$
 2:
 3: **for** $k = 1, \dots$ **do**
 4: $x_{k+1} = x_k - \alpha_k F(x_k)$
 5: $s_k = x_{k+1} - x_k$
 6: $y_k = F(x_{k+1}) - F(x_k)$
 7: $\alpha_{k+1} = \frac{\langle s_k, s_k \rangle}{\langle s_k, y_k \rangle}$
 8:
 9: **end for**

La iteración de punto fijo de Banach-Picard es un método de tipo residual para hallar los ceros de la ecuación $F(x) = x - T(x)$. En efecto, esta iteración de punto fijo puede verse como:

$$\begin{aligned} x_{k+1} &= T(x_k) \\ &= x_k - (x_k - T(x_k)) \\ &= x_k - F(x_k), \end{aligned}$$

donde $F(x) = x - T(x)$. Es decir, es un método de tipo residual con paso constante igual a uno. En nuestra opinión, pensamos que el paso constante igual a uno no es una elección inteligente y que pudiera ser sustituida por una elección de paso variable que utilizara otro tipo de información que acelerara la convergencia, como por ejemplo la elección del paso espectral de los métodos SANE, DF-SANE, etc.

2.3.1. Aceleración espectral de los métodos de punto fijo. La propuesta general de este trabajo es la de acelerar los métodos de punto fijo con la forma:

$$x_{k+1} = Tx_k.$$

mediante los métodos de tipo residual, en particular DF-SANE. Por la definición de punto fijo, tenemos que un vector x es un punto fijo de un operador T si verifica:

$$x = Tx,$$

esto equivale evidentemente, que los puntos fijos de un operador T son los ceros de la función F dada por:

$$F(x) = (I - T)x,$$

en donde I es el operador identidad ($Ix = x$).

Respecto a otros métodos para hallar ceros de funciones, los métodos de tipo residual presentan ventaja al momento de su aplicación, que no requieren del uso de la matriz Jacobiana o derivada de F , ni de aproximaciones del mismo. Esto hace que para su aplicación sea solo necesaria la evaluación de F en cada iterado x_k . En el caso de los métodos de punto fijo esto es simplemente evaluar $F(x_k) = x_k - Tx_k$.

En efecto, un método de tipo residual para hallar los puntos fijos de un operador T viene dado por la siguiente expresión:

$$x_{k+1} = x_k \pm \alpha_k F(x_k) = x_k \pm \alpha_k (x_k - Tx_k),$$

con x_0 el iterado inicial dado.

Este esquema de aceleración ya fue probado con éxito por Hernandez-Ramos [9], para el caso particular de operadores de punto fijo asociados a los métodos de proyecciones alternantes. En este trabajo se propone la extensión de este esquema a otros tipos de problemas de punto fijo que resultan de otros que son prácticos.

Experimentos numéricos.

En este capítulo, a través de experimentos numéricos compararemos los resultados obtenidos en la solución de algunos problemas de punto fijo, aplicando los métodos clásicos, y los métodos acelerados mediante la estrategia propuesta que utilizan los métodos tipo residual, específicamente el DF-SANE. Los problemas que fueron tomados para tales experimentos numéricos son:

- (1) El problema de factibilidad lineal dividida (LSFP).
- (2) La proyección de Landweber.
- (3) Problemas asociados con métodos estacionarios.

Los problemas: 1,2,3 son un caso particular de los problemas: Desigualdad Varicional, Minimización Convexa, como también del Problema de Factibilidad Dividida o **CFP**, etc. Es por ello que no son usados en los experimentos numéricos. Además cabe destacar que el **CFP** fue resuelto por Hernandez-Ramos en su tesis doctoral [9].

Para la implementación de estos algoritmos se utilizó el MATLAB versión 7.8.0 corriendo en una máquina con procesador AMD FX(tm)-6100Six-Core Processor. Como criterio de parada de los algoritmos iterativos se tomó que la norma Euclídea del vector residual fuera menor o igual a 10^{-8} . En los problemas en los que se requiere proyectar sobre conjuntos convexos se usaron como prueba: semiespacios, hiperplanos, esferas, y cajas.

Con el uso de tablas y gráficas comparativas podremos visualizar dichos resultados. Para entender la tablas, y gráficas se dispuso de la siguiente nomenclatura:

- M. Clásico: se refiere a que se está empleando el método clásico o el algoritmo clásico para resolver el problema.
- M.T.R: se refiere al uso del método tipo residual para resolver el problema.
- PPCMP: problema de proyección sobre una caja usando la matriz de Poisson.
- PPCVA: problema de proyección sobre una caja usando valores aleatorios.

- PPEMP: problema de proyección sobre una esfera usando matiz de Poisson.
- PPEVA: problema de proyección sobre una esfera usando valores aleatorios.
- PPHMP: problema de proyección sobre un hiperplano usando la matriz de Poisson.
- PPHVA: problema de proyección sobre un hiperplano usando valores aleatorios.
- PPSEMP: problema de proyección sobre un semiespacio usando la matriz de Poisson.
- PPSEVA: problema de proyección sobre un semiespacio usando valores aleatorios.
- MP: Matriz de Poisson.
- iter: número de iteraciones realizadas por el algoritmo.
- tiempo: tiempo que requiere el algoritmo para resolver el problema, medido en segundos.
- m y n : son las dimensiones de la matriz.
- Error: se refiere al error absoluto de cada iteración.
- EF: es el número de iteraciones realizadas por el algoritmo del método.

3.1. Proyección sobre algunos conjuntos.

Antes de continuar con los resultados obtenidos en los experimentos numéricos, se mostrará como proyectar sobre algunos conjuntos como: cajas, esferas, hiperplanos y semiespacios. Debido a que estos conjuntos fueron los usados como conjuntos convexos para realizar las proyecciones que se necesitan en tales experimentos.

DEFINICIÓN 3.1 (Caja Cerrada). Sean dos vectores $l, u \in \mathbb{R}^n$ con la propiedad de que $l_i \leq u_i$ para $i = 1, \dots, n$. La caja cerrada de \mathbb{R}^n definida por los vectores l, u es el conjunto:

$$B_{l,u} = \{x \in \mathbb{R}^n : l_i \leq x_i \leq u_i, i = 1, \dots, n\}.$$

DEFINICIÓN 3.2 (Esfera). Dado un vector $c \in \mathbb{R}^n$ y dado un valor r . Se define la esfera con centro en c y de radio r como el conjunto:

$$E(c, r) = \{x \in \mathbb{R}^n : \|x - c\| \leq r\}.$$

DEFINICIÓN 3.3 (Hiperplano). Un Hiperplano de \mathbb{R}^n es el conjunto de soluciones a una ecuación lineal o equivalentemente el conjunto de la forma $H(a, c) = \{x \in \mathbb{R}^n : \langle a, x \rangle = c\}$, y $c \in \mathbb{R}$ una constante.

DEFINICIÓN 3.4 (Semiespacio). Un Hiperplano de \mathbb{R}^n es el conjunto de soluciones a un conjunto de la forma $H(a, c) = \{x \in \mathbb{R}^n : \langle a, x \rangle \leq c\}$, y $c \in \mathbb{R}$ una constante.

3.1.1. Proyección de un punto sobre una caja. Para proyectar un punto cualquiera $x \in \mathbb{R}^n$ en la caja cerrada $B_{l,u}$ se procede como sigue:

Para todo $i = 1, \dots, n$:

- Si $x_i \leq l_i$ entonces $p_i = l_i$;
- Si $x_i \geq u_i$ entonces $p_i = u_i$;
- Si $l_i \leq x_i$ y $x_i \leq u_i$ entonces $p_i = x_i$;

3.1.2. Proyección de un punto en una esfera. Para proyectar un punto cualquiera $x \in \mathbb{R}^n$ en la esfera $E(c, r)$, primero se calcula la distancia de x al centro c de la esfera. Es decir, $d = \|x - c\|$.

- Si la distancia d es menor al radio de la esfera, es decir, si $d \leq r$, significa que el punto x está dentro de la esfera y por lo tanto la proyección es el mismo punto x , i.e. $p = x$.
- En caso contrario, si el punto x está fuera de la esfera se calcula la proyección p de la siguiente manera:

$$p = c + \frac{r}{d}(x - c).$$

3.1.3. Proyección de un punto en un hiperplano. Para proyectar un punto cualquiera $x \in \mathbb{R}^n$ en el hiperplano $H(a, c)$ se procede como sigue:

- Si $\langle a, x \rangle = c$ entonces $p = x$.
- Si $\langle a, x \rangle \neq c$ entonces

$$p = x + \frac{c - \langle a, x \rangle}{\langle a, a \rangle} a.$$

3.1.4. Proyección de un punto en un Semiespacio. Para proyectar un punto cualquiera $x \in \mathbb{R}^n$ en el Semiespacio $H(a, c)$ se procede como sigue:

- Si $\langle a, x \rangle \leq c$ entonces $p = x$.
- Si $\langle a, x \rangle > c$ entonces

$$p = x + \frac{c - \langle a, x \rangle}{\langle a, a \rangle} a.$$

3.2. Problema de factibilidad lineal dividida (LSFP).

De acuerdo a lo expuesto en (1.4.6), el LSFP consiste en hallar $x \in C$ (si existe), tal que $Ax \leq b$. Tenemos que la iteración para el LSFP, llamado algoritmo CQ , viene dada por la siguiente expresión:

$$x_{n+1} = P_C (x_n - \gamma A^T V r_+(x_n)),$$

donde $\gamma > 0$, y $\gamma \in \left(0, \frac{2}{\rho(AA^T)}\right)$. Y P_C es la proyección sobre el conjunto convexo C . Como mencionamos previamente dicho conjunto C puede ser: una caja, hiperplano, esfera, etc. En particular para los experimentos numéricos se tomaron para las restricciones $Ax \leq b$, valores de A y b siguientes:

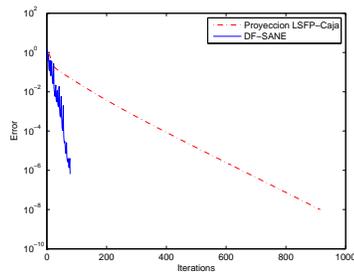
- A es una matriz de Poisson o aleatoria.
- $b = (4, \dots, 4)^t$ o aleatorio.
- $x_0 = (3, \dots, 3)^t$ o aleatorio.
- $k = 30000$ es el máximo número de iteraciones.
- $s = 1/\rho(AA^T)$.

Los convexos usados para C fueron:

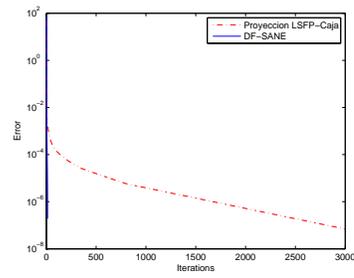
- Caja: el conjunto convexo C , tal que $C = \{x \in H : x_i \leq 4\}$.
- Esfera: $c = (-4, \dots, -4)^t$ o aleatorio, como centro de la esfera $r = 0,01$, el valor del radio de la esfera. Es decir, C , tal que $C = \{x \in H : \|x - c\| \leq 0,01\}$.
- Hiperplano. $c = (-4, \dots, -4)^t$ o aleatorio, como el vector director del hiperplano, y $r = 0,01$. Es decir, C , tal que $C = \{x \in H : \langle x, y \rangle = 0,01\}$.
- Semiespacio: $c = (-4, \dots, -4)^t$ o aleatorio, como el vector director del hiperplano, y $r = 0,01$. Es decir, C , tal que $C = \{x \in H : \langle x, y \rangle \leq 0,01\}$.

CUADRO 1. Tabla de resultados asociada al LSFP.

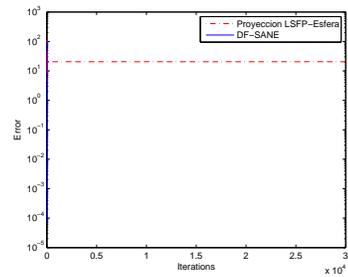
Problema			M. Clásico		M. T .R.	
Nombre	n	m	EF	Tiempo	EF	Tiempo
PPCMP	25	25	2858	1.633	373	0.336
PPCMP	100	100	905	0.336	49	0.016
PPCMP	2500	2500	914	0.735	72	0.015
PPCMP	10000	10000	915	122.3	77	10.42
PPCVA	5	7	140	0.017	3	1e-3
PPCVA	20	30	635	0.073	6	0.028
PPCVA	50	70	731	0.159	8	0.001
PPCVA	100	150	1026	0.541	10	0.005
PPCVA	500	700	3000	89.13	12	0.327
PPEMP	25	25	*	*	11	0.05
PPEMP	100	100	*	*	13	0.002
PPEMP	2500	2500	*	*	10	0.016
PPEMP	10000	10000	*	*	14	0.094
PPEVA	5	7	*	*	6	0.001
PPEVA	20	30	*	*	4	1e-3
PPEVA	50	70	*	*	4	0.002
PPEVA	100	150	*	*	4	0.003
PPEVA	500	700	*	*	3	0.073
PPHMP	25	25	1379	0.28	3	0.027
PPHMP	100	100	2127	0.367	3	0.011
PPHMP	2500	2500	7375	10.87	3	0.005
PPHMP	10000	10000	12819	106.7	3	0.024
PPHVA	5	7	47	0.007	3	1e-3
PPHVA	20	30	58	0.03	3	1e-3
PPHVA	50	70	140	0.05	3	0.012
PPHVA	100	150	82	0.041	3	0.002
PPHVA	500	700	87	2.666	3	0.092
PPSEMP	25	25	836	0.128	2	0.009
PPSEMP	100	100	2095	0.342	2	0.014
PPSEMP	2500	2500	11898	0.016	2	0.005
PPSEMP	10000	10000	23743	195.95	2	0.014
PPSEVA	5	7	17	0.042	2	1e-4
PPSEVA	20	30	2	0.001	1	1e-4
PPSEVA	50	70	15	0.004	3	1e-4
PPSEVA	100	150	16	0.012	3	0.003
PPSEVA	500	700	16	0.1	3	0.094



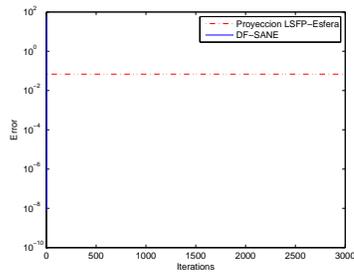
(A)



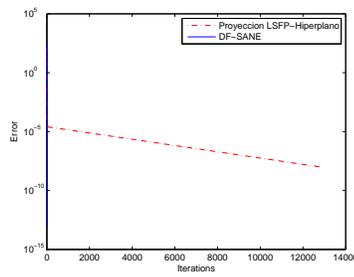
(B)



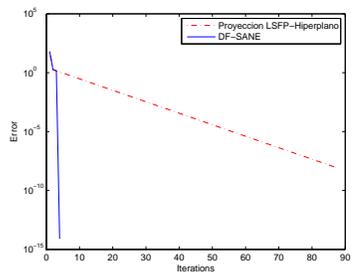
(C)



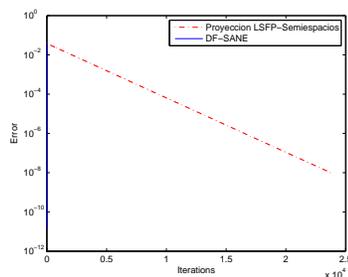
(D)



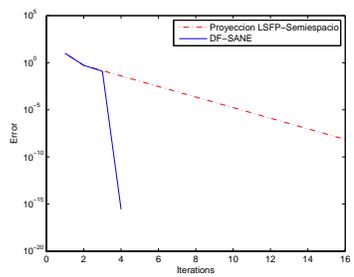
(E)



(F)



(G)



(H)

FIGURA 3.1. Gráficas resultantes de aplicar los esquemas de punto fijo y los tipo residual. En la derecha usando matriz de Poisson e izquierda usando matriz con valores aleatorios.

3.2.1. Resultados. De acuerdo con el cuadro 1 , junto con las gráficas 3.1 obtenidas, puede apreciarse claramente que la estrategia asociada al DF-SANE es superior al esquema clásico para la solución del LSFP, requiriendo de menos tiempo de ejecución e iteraciones para la la resolución del problema en comparación con la estrategia dada por la de punto fijo, tanto para matrices de Poisson como para matrices con valores aleatorio. Por un lado, hay gráficas en la figura 3.1 , tales como: C, D, y E en las que el número de iteraciones que requirió el DF-SANE es tan pequeño, que la gráfica asociada al mismo está encima del eje Y. Además, de acuerdo a la tabla 1, en la proyección sobre una esfera puede apreciarse *, el cual se debe a problemas de mal condicionamiento, afectando de tal manera que para resolver el problema usando la estrategia de punto fijo no bastaba con el número de máximo de iteraciones programadas, sin llegar a la solución del mismo, sin embargo el DF-SANE sí logró resolverlo.

3.3. Proyección de Landweber

Por lo visto en (1.4.7), la proyección de Landweber consiste en hallar $x \in C$ (si existe), tal que $Ax = b$, tenemos que la iteración asociada al método es de la forma:

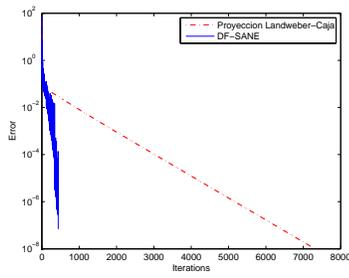
$$x_{n+1} = P_C(x_n + \gamma A^T(b - Ax_n)),$$

con $\gamma > 0$ y $\gamma \in \left(0, \frac{2}{\rho(AA^T)}\right)$, donde P_C es la proyección sobre el conjunto C . Al igual que en el LSFP, el conjunto C es un convexo.

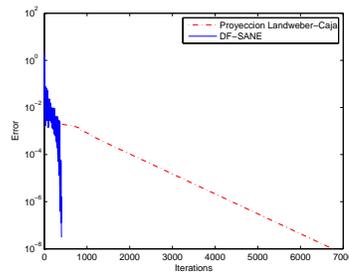
Para entender la tabla de resultados obtenidos que pertenecen a esta parte, respecto a los problemas a los cuales aplicamos ambos algoritmos, son los mismos que usados en el LSFP, por lo tanto los términos que se emplean también coinciden con los de la tabla anterior.

CUADRO 2. Tabla de resultados asociada a la Proyección de Landweber.

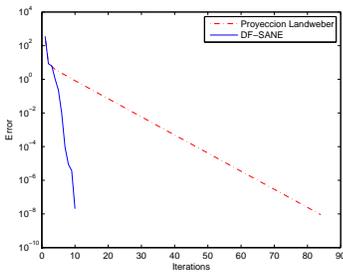
Problema			M. Clásico		M. T .R.	
Nombre	n	m	EF	Tiempo	EF	Tiempo
PPCMP	22	25	468	0.065	47	0.014
PPCMP	97	100	7311	1.03	439	0.1
PPCMP	2497	2500	*	*	*	*
PPCMP	9997	10000	*	*	*	*
PPCVA	5	7	914	0.059	37	0.004
PPCVA	20	30	6079	0.436	174	0.017
PPCVA	50	70	25323	2.781	11261	1.332
PPCVA	100	150	30000	5.710	25352	4.846
PPCVA	500	700	*	*	*	*
PPEMP	22	25	24	0.0027	8	1e-3
PPEMP	97	100	49	0.0059	10	1e-3
PPEMP	2497	2500	92	0.029	9	0.004
PPEMP	9997	10000	90	0.138	9	0.014
PPEVA	5	7	10	0.003	6	0.001
PPEVA	20	30	7	0.001	5	1e-4
PPEVA	50	70	7	0.002	4	1e-4
PPEVA	100	150	7	0.001	4	1e-4
PPEVA	500	700	6	0.04	3	0.02
PPHMP	22	25	391	0.05	50	0.01
PPHMP	97	100	4152	0.467	177	0.043
PPHMP	2497	2500	*	*	*	*
PPHMP	9997	10000	*	*	*	*
PPHVA	5	7	639	0.051	41	0.004
PPHVA	20	30	13943	1.227	173	0.018
PPHVA	50	70	29003	3.186	286	0.036
PPHVA	100	150	*	*	504	0.079
PPHVA	500	700	*	*	992	6.385
PPSEMP	22	25	1756	0.171	133	0.039
PPSEMP	97	100	25325	2.63	2325	0.50
PPSEMP	2497	2500	*	*	*	*
PPSEMP	9997	10000	*	*	*	*
PPSEVA	5	7	796	0.058	59	0.005
PPSEVA	20	30	4634	0.371	167	0.017
PPSEVA	50	70	30000	3.076	215	0.023
PPSEVA	100	150	*	*	1533	0.232
PPSEVA	500	700	*	*	*	*



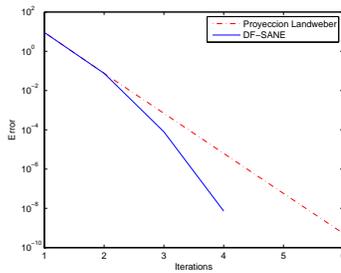
(A)



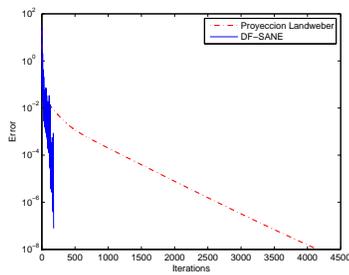
(B)



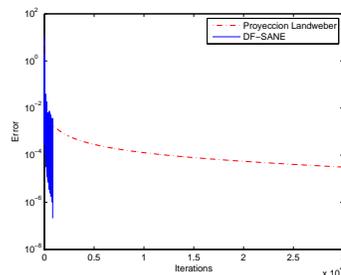
(C)



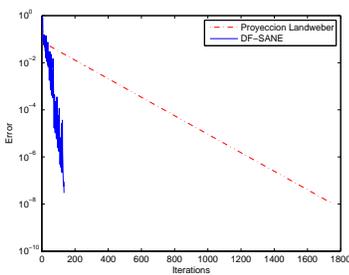
(D)



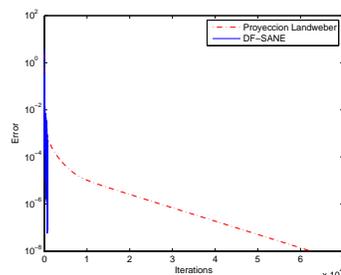
(E)



(F)



(G)



(H)

FIGURA 3.2. Gráficas obtenidas de la proyección de Landweber. Para la matriz de Poisson derecha, y para matriz de valores aleatorios izquierda.

3.3.1. Resultados. De acuerdo con los resultados obtenidos en los experimentos y mostrados en el cuadro 2, además de gráficas 3.2 , la estrategia usada por el DF-SANE para la resolución de estos problemas fue óptima en tiempo e iteraciones en comparación con la estrategia dada por la de punto fijo, la cual es usualmente es aplicada para resolver estos problemas, ésta última en algunos casos para resolver el problema requirió del uso del número máximo de iteraciones, inversión de mucho tiempo, y sin lograr resolver el problema. Además hay casos en los cuales ambos métodos no pudieron resolver el problema, esto puede deberse por el mal condicionamiento de las matrices usadas.

3.4. Métodos Estacionarios.

De acuerdo con lo parte 1.4.8, se tiene que los métodos estacionarios buscan resolver sistemas de ecuaciones lineales, es decir:

$$\text{Hallar } x \text{ (si existe), tal que } Ax = b.$$

En general la iteración para resolver este problema viene dado por:

$$x_{k+1} = Mx_k + d,$$

donde M es la matriz de iteración del método. Para los experimentos empleamos 3 tipos de métodos, lo cuales son: Richardson, Jacobi y Gauss-Seidel.

Tenemos:

- A es una matriz de Poisson.
- $x_0 = (1, \dots, n)^t$.
- $b = (1, \dots, 1)^t$.
- $k = 35000$ es el máximo número de iteraciones.

CUADRO 3. Tabla de resultados obtenidos al aplicar el método de Richardson y DF-SANE.

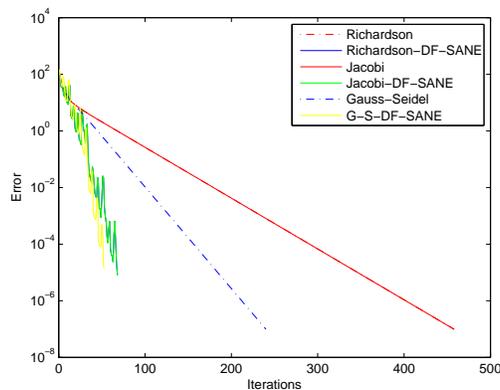
Problema			M. Clásico		M. T .R.	
Nombre	n	m	EF	Tiempo	EF	Tiempo
MP	25	25	127	0.013	37	0.016
MP	100	100	458	0.063	67	0.016
MP	2500	2500	100000	$8.17e^{-6}$	5434	2.25
MP	10000	10000	100000	0.4	39246	$3.6e^{-5}$

CUADRO 4. Tabla de resultados obtenidos al aplicar el método de Jacobi y DF-SANE.

Problema			M. Clásico		M. T .R.	
Nombre	n	m	EF	Tiempo	EF	Tiempo
MP	25	25	127	0.02	37	0.016
MP	100	100	458	0.097	67	0.032
MP	2500	2500	12061	6.22	6091	3.17
MP	10000	10000	48778	131.27	41289	92.2

CUADRO 5. asociado método de Gauss-Seidel, y la aceleración a través del DF-SANE.

Problema			M. Clásico		M. T .R.	
Nombre	n	m	EF	Tiempo	EF	Tiempo
MP	25	25	68	0.04	22	1e-3
MP	100	100	240	0.019	52	0.016
MP	2500	2500	6223	0.65	3735	0.52
MP	10000	10000	25123	8.16	19649	7.89



(A)

FIGURA 3.3. Gráfica resultante al ejecutar los algoritmos de cada método clásico(Richardson, Jacobi y Gauss-Seidel.) con su respectiva aplicación DF-SANE.

3.4.1. Resultados Empleando el Método de Richardson. En este método se empleó $\gamma = 1/(\lambda_{max} + \lambda_{min})$. De acuerdo con la tabla 3 y la gráfica 3.3, los resultados obtenidos fueron que el esquema DF-SANE es superior a la estrategia clásica del método de Richardson, haciendo menos iteraciones y disminuyendo el tiempo de resolución. En la misma tabla 33, puede apreciarse que aparece * esto se debe a que ninguno de los métodos pudo resolver el problema, como consecuencia de mal condicionamiento de las matrices.

3.4.2. Resultados Empleando el Método de Jacobi. La tabla 4 y la gráfica 3.3 muestran como claramente la estrategia que aplica el DF-SANE para resolver problemas, es superior a la usada por método clásico, reduciendo el tiempo e iteraciones. En la misma tabla 4, puede apreciarse que aparece * esto se debe a que ninguno de los métodos pudo resolver el problema, como consecuencia de mal condicionamiento de las matrices.

3.4.3. Resultados Empleando el Método de Gauss- Seidel. De acuerdo con lo obtenido en el cuadro 5, y la gráfica 3.3 puede observarse que el DF-SANE es más óptimo que el algoritmo clásico del método, puesto que reduce el tiempo y número de iteraciones. En la misma tabla 5, puede apreciarse que aparece * esto se debe a que ninguno de los métodos pudo resolver el problema, como consecuencia de mal condicionamiento de las matrices.

Conclusión

En este trabajo, vimos como las técnicas de tipo residual para la resolución de problemas $F(x) = 0$ tales como DF-SANE, pueden ser utilizadas para acelerar la convergencia de las iteraciones de punto fijo en un conjunto importante de problemas. De hecho, vimos como la misma iteración de Banach-Picard, utilizada frecuentemente para resolver los problemas de punto fijo, es en si misma, un método de tipo residual a paso constante, lo cual se supuso que no era una alternativa inteligente como tamaño de paso.

En una muestra bastante significativa de problemas, se realizaron experimentos numéricos académicos para comparar ambos métodos, tanto la iteración de Banach-Picard, como el DF-SANE; y se mostró que: siempre que el método clásico de Banach-Picard mostraba convergencia hacia la solución, su aceleración también lo hacía. Además, hubo ocasiones que, para una cantidad máxima tolerada de iteraciones, el método clásico no cumplía con el criterio de parada por tolerancia y la aceleración si lo hacía. Adicionalmente, en términos de número de iteraciones, la versión acelerada alcanzó siempre los criterios de tolerancia de manera más rápida, y como en la mayoría de los problemas el costo computacional está asociado a las evaluaciones del operador T , en términos de tiempo casi siempre era también más eficiente en términos de tiempo computacional. En resumen, la versión acelerada es mucho más eficiente en términos computacionales de tiempo y número de iteraciones que la versión clásica en la casi totalidad de los problemas.

Como problema abierto queda demostrar bajo que condiciones la versión acelerada por DF-SANE converge. En este sentido, basados en las numerosas experiencias numéricas realizadas, nos atrevemos a conjeturar que: si el método clásico converge, su aceleración por métodos de tipo residual también convergerá, siempre y cuando el tamaño de paso permanezca acotado. Se sospecha además, que la convergencia de la versión acelerada será más rápida en un conjunto grande de operadores T o bajo condiciones no muy fuertes.

Bibliografía

- [1] Barzilai, Jonathan y Borwein, Jonathan M, *Two-point step size gradient methods*, IMA Journal of Numerical Analysis, Vol. 8, No. 1, pp. 141-148, 1988.
- [2] Bertsekas, Dimitri P, *Nonlinear programming*, 1999, Athena scientific.
- [3] Byrne, Charles, *A unified treatment of some iterative algorithms in signal processing and image reconstruction*, Inverse Problems, Vol. 20, No. 1, pp. 103, 2004, IOP Publishing.
- [4] Byrne, Charles L, *Applied iterative methods*, AK Peters Wellesley, 2008.
- [5] Cegielski, Andrzej, *Methods for finding fixed points of nonexpansive operators in a Hilbert space*, Leiter des Instituts für Mathematik, Ilmenau, 2008.
- [6] Cegielski, Andrzej, *Iterative methods for fixed point problems in Hilbert spaces*, Springer, 2013.
- [7] Cheney, E y Kincaid, David, *Numerical mathematics and computing*, Cengage Learning, 2012.
- [8] Heath, Michael T, *Scientific computing*, McGraw-Hill, 2001.
- [9] Hernández-Ramos, Luis Manuel, *Métodos de proyecciones alternantes vía optimización numérica.*, Tesis de Doctorado, Universidad Central de Venezuela, 2011.
- [10] Hernández-Ramos, Luis M y Escalante, René y Raydan, Marcos, *Unconstrained optimization techniques for the acceleration of alternating projection methods*, Numerical functional analysis and optimization, Vol. 32, No. 10, pp. 1041-1066, Taylor & Francis, 2011.
- [11] Kincaid, David Ronald y Cheney, Elliott Ward, *Numerical analysis: mathematics of scientific computing*, vol. 2, American Mathematical Soc., 2002.
- [12] La Cruz, William y Martínez, José y Raydan, Marcos, *Spectral residual method without gradient information for solving large-scale nonlinear systems of equations*, Mathematics of Computation, Vol. 75, No. 255, pp. 1429-1448, 2006.
- [13] La Cruz, William y Raydan, Marcos, *Nonmonotone spectral methods for large-scale nonlinear systems*, Optimization Methods and Software, Vol. 18, No. 5, pp. 583-599, Taylor & Francis, 2003.
- [14] Raydan, Marcos, *On the Barzilai and Borwein choice of steplength for the gradient method*, IMA Journal of Numerical Analysis, Vol.13, No. 3, pp. 321-326, 1993.