

MÉTODOS DE NEWTON Y CUASI-NEWTON
PARA EL PROBLEMA DE
COMPLEMENTARIEDAD GENERALIZADO



Hevert Vivas

Universidad del Cauca

Facultad de Ciencias Naturales, Exactas y de la Educación

Departamento de Matemáticas

Doctorado en Ciencias Matemáticas

Popayán

6 de julio de 2023

**MÉTODOS DE NEWTON Y CUASI-NEWTON
PARA EL PROBLEMA DE
COMPLEMENTARIEDAD GENERALIZADO**

Hevert Vivas

Tesis presentada como requisito parcial para optar al título de:
Doctor en Ciencias Matemáticas

Directora
Dra. Rosana Pérez Mera
Profesora de la Universidad del Cauca

Codirector
Dr. Carlos Andrés Arias Torres
Profesor de la Universidad del Cauca

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Matemáticas
Doctorado en Ciencias Matemáticas
Popayán
6 de julio de 2023

A MIS HIJOS JULIANA Y CAMILO.

Agradecimientos

Agradezco a mi familia que siempre son mi motor, a mis amigos y colegas Mauricio Macías, Wilmer Sánchez y Favián Arenas por su valiosa colaboración, a mi amigo y codirector Carlos Andrés Arias por sus excelentes aportes y críticas constructivas y principalmente, a mi mentora, directora y amiga Rosana Pérez, sin ella ni hubiese empezado ni habría terminado a tiempo.

Resumen

En este trabajo de investigación, consideramos el problema de complementariedad generalizado y, con el fin de resolverlo, lo reformulamos, no solo como un sistema de ecuaciones no lineales no diferenciable mediante una familia uniparamétrica de funciones de complementariedad, sino como un problema de minimización continuamente diferenciable.

Analizamos detalladamente el operador que define la reformulación como un sistema de ecuaciones no lineales y demostramos que los resultados obtenidos por otros autores para complementariedad no lineal se extienden naturalmente al problema de complementariedad generalizado.

Para resolver el sistema de ecuaciones no lineales, e indirectamente el problema de complementariedad generalizado, proponemos inicialmente un algoritmo global, no suave, tipo *Newton*, para el cual demostramos resultados de convergencia global y analizamos su desempeño numérico. Posteriormente, con el mismo fin, proponemos una familia de métodos secantes de cambio mínimo, desarrollamos su respectiva teoría de convergencia y analizamos su desempeño numérico.

Con el fin de globalizar la familia de métodos secantes, introducimos una *búsqueda lineal libre de derivadas*, con lo cual proponemos un algoritmo global genérico para resolver el problema de complementariedad generalizado mediante su reformulación como un problema de minimización. El carácter genérico del algoritmo se debe a que no se especifica la forma de actualizar las aproximaciones de las matrices del jacobiano generalizado. Bajo ciertas hipótesis, presentamos resultados de convergencia global para el nuevo algoritmo.

Finalmente, abordamos el problema de la actualización de las aproximaciones basándonos en la teoría secante de cambio mínimo, con lo cual obtenemos un nuevo algoritmo global tipo secante y libre de derivadas para resolver el problema de complementariedad generalizado, el cual es un caso particular del algoritmo genérico

propuesto previamente. Demostramos que cualquier sucesión generada por el nuevo algoritmo satisface las hipótesis de convergencia del algoritmo genérico heredando así, sus resultados de convergencia. Complementamos el desarrollo teórico con un análisis del desempeño numérico del algoritmo propuesto.

Palabras clave: Complementariedad generalizada, complementariedad no lineal, función de complementariedad, jacobiano generalizado, método de Newton, método cuasi-Newton, métodos secantes de cambio mínimo, convergencia q -superlineal.

Abstract

In this work, we consider the generalized complementarity problem. To solve it, we reformulate it as a nonsmooth system of equations using a one-parameter family of complementarity functions and as a minimization problem.

We analyze in detail the operator that defines the reformulation as a system of equations and we show that the results obtained by other authors for nonlinear complementarity naturally extend to the generalized complementarity problem.

In order to solve the system of nonlinear equations and indirectly the generalized complementarity problem, we initially propose a nonsmooth global *Newton*-type algorithm, for which we demonstrate global convergence results and analyze its numerical performance. After, with the same purpose, we propose a family of least-change secant methods, develop their respective theory of convergence, and analyze their numerical performance.

To globalize the family of secant methods, we introduce a *derivative-free linear search* and propose a generic global algorithm to solve the generalized complementarity problem using its reformulation as a minimization problem. The generic name of the algorithm is due to the fact that the way to update the approximations of the generalized Jacobian matrices is not specified. Under certain hypotheses, we present global convergence results for the new algorithm.

Finally, we analyze the problem of updating the approximations. We use least-change secant theory and obtain a new derivative-free secant-type global algorithm which is a particular case of the generic algorithm. We show that any sequence generated by the new algorithm satisfies the convergence hypotheses of the generic algorithm, thus it inherits its convergence results. We complement the theoretical development with an analysis of the numerical performance of the proposed algorithm.

Keywords: Generalized complementarity, nonlinear complementarity, complementa-

ity function, generalized Jacobian, Newton's method, quasi-Newton method, least-change secant methods, q -superlinear convergence.

Productos de la investigación

Artículos

- [59] *A nonsmooth Newton method for solving the generalized complementarity problem.* Numerical Algorithms, accepted. R. Pérez y C. Arias.

Ponencias

- *Método de Newton semisuave para resolver el problema de complementariedad generalizado.* Presentado en MAPI II, Medellín, Colombia, junio 8-10, 2022.
- *Método de Newton semisuave para resolver el problema de complementariedad generalizado.* Presentado en XV-International Fast Workshop on Applied and Computational Mathematics, Trujillo, Perú, enero 6-7, 2022.

Índice general

Resumen	vi
Abstract	viii
Productos de la investigación	x
1. Introducción	2
2. Existencia de soluciones	8
2.1. Funciones de valor y variable real	8
2.2. Funciones definidas sobre $\mathbb{R}^n, n > 1$	12
3. Problemas de prueba	18
4. Preliminares	22
5. Un método de Newton no suave	28
5.1. El operador Φ_λ	28
5.2. Algoritmo y resultados de convergencia	35
5.3. Experimentación numérica	41
5.3.1. Experimento 1	41
5.3.2. Experimento 2.	43

5.3.3. Experimento 3	43
5.3.4. Experimento 4.	45
6. Métodos secantes de cambio mínimo	48
6.1. Hipótesis y resultados de convergencia	50
6.2. Familia de métodos secantes de cambio mínimo para $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$	60
6.3. Resultados de convergencia adicionales	63
6.4. Pruebas Numéricas	68
6.4.1. Experimento 1	68
6.4.2. Experimento 2	69
6.4.3. Experimento 3	70
6.4.4. Experimento 4	73
7. Un método libre de derivadas	76
7.1. Búsqueda lineal libre de derivadas	77
7.2. Algoritmo	77
7.3. Familia de métodos secantes	86
7.3.1. Verificación de la Hipótesis $\bar{H}4$.	88
7.3.2. Verificación de la Hipótesis $\bar{H}5$.	92
7.4. Experimentación numérica	93
7.4.1. Experimento 1	94
7.4.2. Experimento 2	95
7.4.3. Experimento 3	97
7.4.4. Experimento 4	101
7.4.5. Experimento 5	101
8. Conclusiones	103

Introducción

Dadas $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ y $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuamente diferenciables, el problema de complementariedad generalizado PCG(F, G) consiste en encontrar un vector $\mathbf{x} \in \mathbb{R}^n$, tal que

$$F(\mathbf{x}) \geq 0, \quad G(\mathbf{x}) \geq 0, \quad F(\mathbf{x})^T G(\mathbf{x}) = 0. \quad (1.1)$$

En este contexto, un vector es no negativo si, y solo si todas sus componentes son no negativas.

La tercera condición en (1.1) es llamada condición de complementariedad y, dadas las dos primeras condiciones, implica que $F_i(\mathbf{x}) = 0$ o $G_i(\mathbf{x}) = 0$, para todo $i = 1, \dots, n$, lo que a su vez puede interpretarse, como la necesidad de un equilibrio entre las funciones que definen el problema. Esto hace que el concepto de *complementariedad* esté íntimamente relacionado con el de *sistema en equilibrio* [25].

El problema de complementariedad generalizada PCG(F, G) tiene numerosas aplicaciones en ingeniería, economía, y en una amplia clase de problemas que contienen el problema clásico de complementariedad no lineal, como caso especial. Además, juega un papel importante en el estudio de problemas de programación no lineal, de desigualdades variacionales, problemas de equilibrio y en ingeniería mecánica, etc [58].

Observemos que, si $F = G$, el PCG(F, G) se reduce a resolver un sistema de ecuaciones no lineales. Otros casos particulares de (1.1), ampliamente estudiados por diferentes autores son los siguientes:

- *Problema de complementariedad implícito* [21, 45, 46]. Cuando $G(\mathbf{x}) = \mathbf{x} - E(\mathbf{x})$, donde $E: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es continuamente diferenciable.
- *Problema de complementariedad lineal* [16, 17]. Si $F(\mathbf{x}) = M\mathbf{x} + \mathbf{q}$, $G(\mathbf{x}) = \mathbf{x}$, con $M \in \mathbb{R}^{n \times n}$ y $\mathbf{q} \in \mathbb{R}^n$.

- *Problema de complementariedad no lineal* [2, 4, 5, 64]. Cuando $G(\mathbf{x}) = \mathbf{x}$ y F es no lineal. Este problema es un referente general para tratar condiciones de optimalidad tanto de problemas matemáticos, como de desigualdades variacionales.

Es importante mencionar que en la mayoría de los métodos de solución de este tipo de problemas, se utiliza su reformulación como un sistema de ecuaciones no lineales y no diferenciable, usando una función $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$ conocida como *función de complementariedad* [34], la cual satisface la siguiente equivalencia:

$$\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0. \quad (1.2)$$

En forma análoga a los casos particulares, el PCG(F, G) se puede reformular como el siguiente sistema de ecuaciones no lineales:

$$\Phi(\mathbf{x}) = (\varphi(F_1(\mathbf{x}), G_1(\mathbf{x})), \dots, \varphi(F_n(\mathbf{x}), G_n(\mathbf{x}))) = 0, \quad (1.3)$$

donde $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$, y $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función de complementariedad. A partir de la equivalencia (1.2), se tiene que \mathbf{x}^* resuelve PCG(F, G) si, y solo si $\Phi(\mathbf{x}^*) = 0$, lo que garantiza que resolver el problema de complementariedad generalizado es equivalente a resolver su reformulación.

Alternativamente, si $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}$ denota la función de mérito natural definida por

$$\Psi(\mathbf{x}) = \frac{1}{2} \Phi(\mathbf{x})^T \Phi(\mathbf{x})$$

entonces podemos reformular el problema de complementariedad generalizado como el siguiente problema de minimización sin restricciones

$$\begin{aligned} \text{Minimizar} \quad & \Psi(\mathbf{x}). \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned} \quad (1.4)$$

Observe que cualquier solución de (1.3) es un minimizador global de Ψ en \mathbb{R}^n . Recíprocamente, cualquier solución local \mathbf{x} de (1.4), tal que $\Psi(\mathbf{x}) = 0$, es una solución de (1.3).

Los métodos tipo Newton no suave son populares para resolver problemas de complementariedad no lineal [54] [18] [34]. Estos métodos usan el concepto de *jacobiano generalizado* [13]. Algunos de ellos se han extendido a complementariedad generalizada [58] [22].

En este trabajo, consideramos la reformulación (1.3) con $\varphi = \varphi_\lambda$, la familia uniparamétrica de funciones de complementariedad introducida en [34], definida por

$$\varphi_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} - a - b, \quad \lambda \in (0, 4). \quad (1.5)$$

Para $\lambda = 2$, (1.5) se reduce a la función de *Fischer* [28] [18] y cuando λ tiende a 0, converge a un múltiplo de la función *Mínimo* [28, 47, 48].

Por lo tanto, podemos reescribir el sistema de ecuaciones no lineales (1.3) como

$$\Phi_\lambda(\mathbf{x}) = 0, \quad (1.6)$$

y el problema de minimización (1.4) como

$$\begin{aligned} \text{Minimizar} \quad & \Psi_\lambda(\mathbf{x}) = \frac{1}{2} \|\Phi_\lambda(\mathbf{x})\|_2^2, \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned} \quad (1.7)$$

cuya función objetivo es continuamente diferenciable [34].

Elegimos la familia φ_λ por dos razones. En primer lugar, por su relación con las dos funciones de complementariedad, probablemente más prominentes, como lo son la *Fischer* y la *Mínimo*.

La segunda razón es que, hasta donde sabemos, la familia φ_λ no se ha utilizado en relación con el problema de complementariedad generalizado, solo se utilizó el caso particular $\lambda = 2$ (función de *Fischer*) en [22] para proponer un método generalizado tipo Newton para resolver el $\text{PCG}(F, G)$ y la función *Mínimo* se utilizó en [58] para analizar una convergencia local de un método de tipo *Levenberg-Marquardt* para el $\text{PCG}(F, G)$.

Para resolver el $\text{PCG}(F, G)$, se han propuesto diferentes tipos de métodos dependiendo de condiciones sobre las funciones F y G . Es decir, si son suaves [38, 42, 56, 65], si F es localmente *Lipschitz* y $G(\mathbf{x}) = \mathbf{x}$ [11, 12, 26, 27, 33], o si ambas funciones son no suaves y localmente *Lipschitz* [32, 58].

En [34], los autores hacen un análisis detallado de las *propiedades del operador* que permite reformular el problema de complementariedad no lineal como un sistema de ecuaciones no lineales. Teniendo en cuenta que estos resultados fueron fundamentales para los desarrollos teóricos y prácticos posteriores presentados por dichos autores para complementariedad no lineal y dado que ese problema es un caso particular del problema de complementariedad generalizada, en este trabajo extendemos las propiedades del operador Φ_λ al contexto de complementariedad generalizada.

En [22], el autor propone un método global tipo Newton generalizado [52] para resolver el $\text{PCG}(F, G)$ basado en su reformulación como un sistema de ecuaciones no lineales mediante la función de complementariedad de *Fischer* [1].

Los buenos resultados obtenidos en [22] y el hecho de que la función de *Fischer* es un caso particular de una familia de funciones de complementariedad propuesta en [34], nos motivó a extender en esta investigación los resultados obtenidos en [22] a

todos los miembros de la familia uniparamétrica de funciones de complementariedad propuesta en [34], con lo cual presentamos un nuevo algoritmo tipo Newton para resolver el $\text{PCG}(F, G)$ indirectamente, a través de su reformulación como un sistema de ecuaciones no lineales, no diferenciable mediante la familia mencionada. Además, demostramos resultados de convergencia global y analizamos su desempeño numérico.

En [2, 4, 5], los autores proponen y desarrollan la teoría de convergencia de una familia de *métodos secantes de cambio mínimo* para resolver el problema de *complementariedad no lineal*, mediante su reformulación como un sistema de ecuaciones no lineales usando la familia uniparamétrica de funciones de complementariedad mencionada anteriormente [3, 34].

Teniendo en cuenta que los métodos secantes de cambio mínimo tienen muy buenas propiedades de convergencia, proporcionan una alternativa de trabajo en la solución de sistemas de ecuaciones no lineales y se han obtenido buenos resultados en el caso particular de problemas de complementariedad no lineal, en el desarrollo de este trabajo usamos, por primera vez, dichos métodos para resolver el problema de complementariedad generalizado. En efecto, proponemos una familia de métodos secantes de cambio mínimo local, desarrollamos su respectiva teoría de convergencia y analizamos su desempeño numérico.

Con el fin de globalizar la familia de métodos secantes, introducimos una *búsqueda lineal libre de derivadas*, con lo cual proponemos un algoritmo global genérico para resolver el problema de complementariedad generalizado, mediante su reformulación como un problema de minimización. El carácter genérico del algoritmo se debe a que no se especifica la forma de actualizar las aproximaciones de las matrices del jacobiano generalizado. Bajo ciertas hipótesis, presentamos resultados de convergencia global para el nuevo algoritmo.

Finalmente, abordamos el problema de la actualización de las aproximaciones basándonos en la teoría secante de cambio mínimo, con lo cual obtenemos un nuevo algoritmo global tipo secante y libre de derivadas para resolver el problema de complementariedad generalizado, el cual es un caso particular del algoritmo genérico. Demostramos que cualquier sucesión generada por el nuevo algoritmo satisface las hipótesis de convergencia del genérico heredando así, sus resultados de convergencia. Complementamos el desarrollo teórico con un análisis del desempeño numérico del algoritmo propuesto.

Organizamos la presentación de este documento en la siguiente forma:

En el **Capítulo 2**, analizamos el problema de complementariedad generalizado mediante casos particulares para determinar si existe solución o soluciones del problema, o por el contrario, no tiene solución.

En el **Capítulo 3**, presentamos los diez problemas de prueba que fueron usa-

dos para analizar el desempeño numérico de los algoritmos que proponemos en esta investigación.

En el **Capítulo 4**, presentamos notaciones, algunos conceptos básicos, hipótesis y propiedades necesarias en el desarrollo de este trabajo.

En el **Capítulo 5**, analizamos y extendemos a complementariedad generalizada las propiedades del operador Φ_λ investigadas detalladamente, para el caso particular de complementariedad no lineal en [34]. Adicionalmente, presentamos un nuevo algoritmo tipo Newton no suave para resolver el $\text{PCG}(F, G)$. Demostramos resultados de convergencia global para este algoritmo y analizamos su desempeño numérico.

En el **Capítulo 6**, presentamos una familia de métodos secantes de cambio mínimo para resolver el sistema de ecuaciones no lineales (1.6) y desarrollamos su respectiva teoría de convergencia.

En el **Capítulo 7**, proponemos inicialmente, un algoritmo cuasi-Newton global libre de derivadas para resolver el $\text{PCG}(F, G)$, indirectamente, a través de la solución del problema de minimización (1.7). Este algoritmo es genérico porque no se especifica la forma de actualizar las aproximaciones de las matrices del jacobiano generalizado. Bajo ciertas hipótesis, presentamos resultados de convergencia global para este algoritmo. En la segunda parte del capítulo, abordamos el problema de la actualización de las aproximaciones, con lo cual tenemos un nuevo algoritmo global secante, libre de derivadas para resolver el $\text{PCG}(F, G)$, el cual es un caso particular del algoritmo genérico propuesto. Demostramos que este nuevo algoritmo converge.

En el **Capítulo 8**, presentamos algunas conclusiones y comentarios finales.

Finalmente, presentamos la bibliografía que usamos en el desarrollo de esta investigación.

Existencia de soluciones

En este capítulo, analizamos el problema de complementariedad generalizado mediante casos particulares para determinar si existe solución o soluciones del problema, o por el contrario, el problema no tiene solución. El análisis de existencia de soluciones es algo natural y razonable cuando se desea resolver un problema en matemáticas u otras ciencias.

2.1. Funciones de valor y variable real

En primer lugar, consideramos el caso $n = 1$. Es decir, $F, G: \mathbb{R} \rightarrow \mathbb{R}$ y $x \in \mathbb{R}$. A partir de la definición (1.1), un número real x resuelve el PCG(F, G), si satisface que

$$F(x) \geq 0, \quad G(x) \geq 0 \quad \text{y} \quad F(x)^T G(x) = 0.$$

Podemos deducir que, si ocurre una de las siguientes dos posibilidades (no necesariamente son excluyentes), el PCG(F, G) tiene solución:

Si existe $x = a \in \mathbb{R}$, tal que $F(a) = 0$ y $G(a) \geq 0$, $x = a$ es solución del PCG(F, G) (Figura 2.1).

Si existe $x = a \in \mathbb{R}$, tal que $F(a) \geq 0$ y $G(a) = 0$, $x = a$ es solución del PCG(F, G).

De las dos consideraciones anteriores, se tiene que las soluciones del PCG(F, G) son *las raíces reales* de F (o de G), para las cuales el valor de G (respectivamente el de F) es no negativo.

Por lo tanto, cuando F y G están definidas sobre \mathbb{R} , el PCG(F, G) se reduce a un *problema de cálculo de raíces*. Así, el problema de complementariedad generalizado

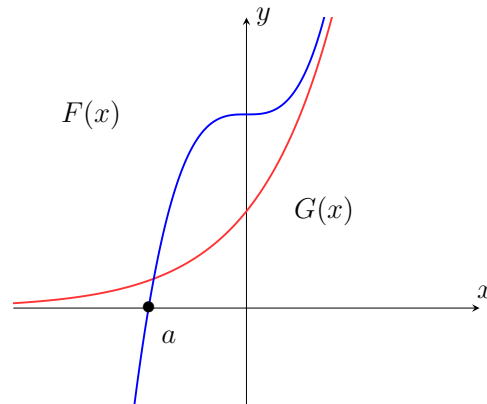


Figura 2.1: a es solución del PCG(F, G).

puede o no tener solución y en caso de tenerla, podrían ser varias, incluso infinitas como lo ilustran los siguientes ejemplos.

Ejemplo 2.1. Sean F y G definidas por:

$$F(x) = -0.125x^3 \quad y \quad G(x) = x^2 - 2.$$

En este caso, los candidatos a solución del PCG(F, G) son las dos raíces de F , $r_1 = -\sqrt{2}$ y $r_2 = \sqrt{2}$ y la raíz de G , $r_3 = 0$. Sin embargo, como se puede observar en la Figura 2.2, solamente $r_1 = -\sqrt{2}$ satisface la condición de no negatividad. Por lo tanto, esta es la única solución del problema. Este ejemplo también muestra que cero no, necesariamente, es solución del problema.

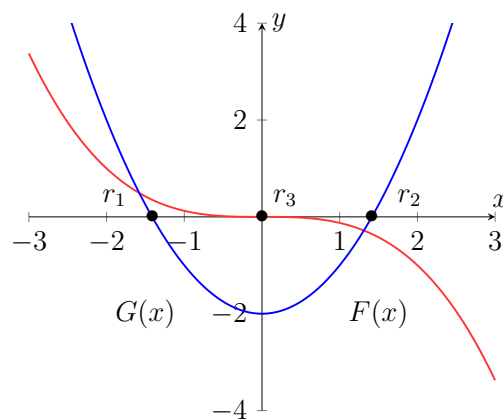


Figura 2.2: Solución única.

Ejemplo 2.2. Sean F y G definidas por: $F(x) = 2 - x^2$ y $G(x) = x^2 + 1/2$. Las únicas raíces de F son a_1 y a_2 (Figura 2.3) y puesto que $G(a_1) > 0$ y $G(a_2) > 0$, las dos son soluciones del problema.

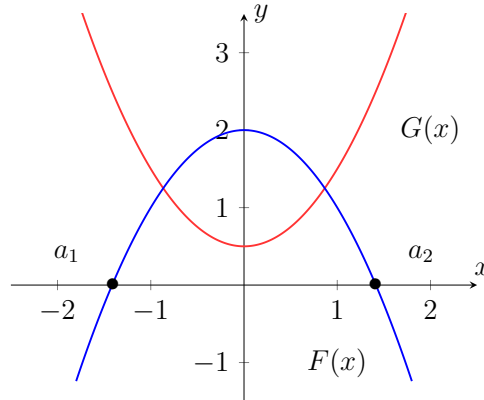


Figura 2.3: Dos soluciones.

El siguiente ejemplo ilustra el caso en el que el $\text{PCG}(F, G)$ no tiene solución.

Ejemplo 2.3. Sean F y G definidas por:

$$F(x) = 1 - x^2 \quad \text{y} \quad G(x) = x^2 - 2.$$

En este caso, hay cuatro raíces: a_1, a_2, a_3 y a_4 (Figura 2.4); sin embargo, en ninguna de estas, se cumple la condición de no negatividad. Por lo tanto, el $\text{PCG}(F, G)$ no tiene solución.

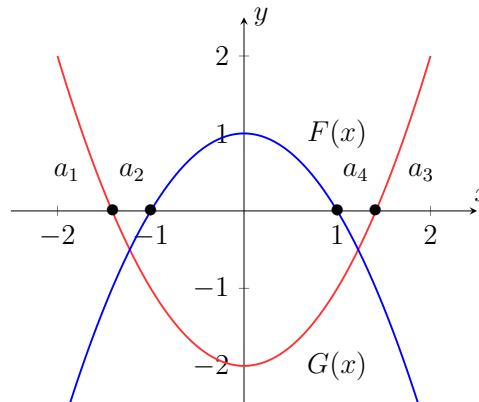


Figura 2.4: No existe solución.

Ejemplo 2.4. Sean F y G definidas por:

$$F(\mathbf{x}) = 1 - x^2 \quad \text{y} \quad G(\mathbf{x}) = x^2 - 1.$$

En la Figura 2.5 observamos que a_1 y a_2 son raíces de F y G ; es decir, $F(a_1) = G(a_2) = 0$. Por lo tanto, el $\text{PCG}(F,G)$ tiene dos soluciones (degeneradas).

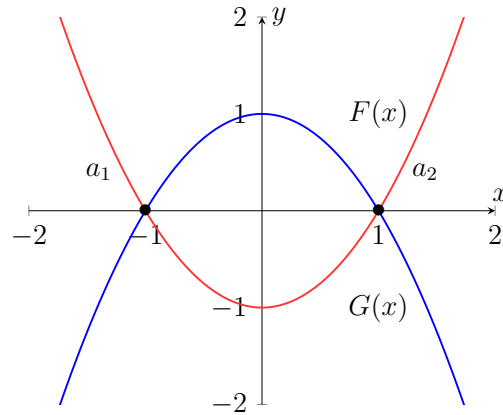


Figura 2.5: Dos soluciones degeneradas.

Ejemplo 2.5. Sean F y G definidas por $F(\mathbf{x}) = \sin 2x$ y $G(x) = e^x$. Dado que la función F tiene infinitas raíces de la forma $x_* = n\frac{\pi}{2}$, $n \in \mathbb{Z}$ y $G(n\pi) > 0$, para todo $n \in \mathbb{Z}$, el $\text{PCG}(F,G)$ tiene infinitas soluciones (Figura 2.6).

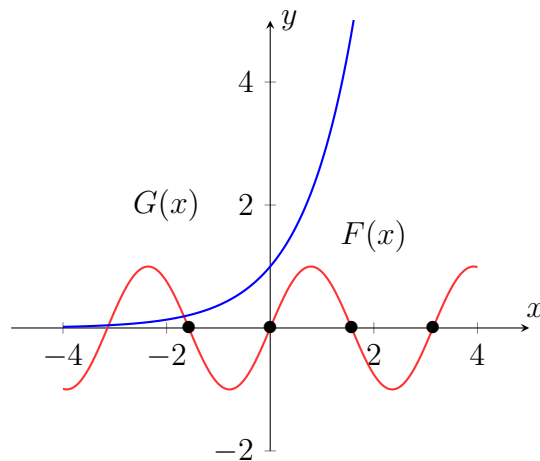


Figura 2.6: Infinitas soluciones.

2.2. Funciones definidas sobre \mathbb{R}^n , $n > 1$.

En general, si F y G son campos vectoriales sobre \mathbb{R}^n , $F(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))$ y $G(\mathbf{x}) = (G_1(\mathbf{x}), \dots, G_n(\mathbf{x}))$, se deben tener más consideraciones en el análisis de existencia de solución del PCG(F, G).

Las restricciones de no negatividad, $F(\mathbf{x}) \geq 0$ y $G(\mathbf{x}) \geq 0$, equivalentemente, $F_i(\mathbf{x}) \geq 0$ y $G_i(\mathbf{x}) \geq 0, i = 1, \dots, n$ definen igual número de regiones en \mathbb{R}^n . Si su intersección es el conjunto vacío, el PCG(F, G) no tiene solución. En otro caso, los puntos de dicha región son los candidatos a ser solución del problema, los cuales definen un conjunto llamado región factible definido por:

$$\mathcal{R} = \{\mathbf{x} \in \mathbb{R}^n : F_i(\mathbf{x}) \geq 0 \text{ y } G_i(\mathbf{x}) \geq 0, \quad i = 1, \dots, n\}.$$

La condición de complementariedad, $F(x, y)^T G(x, y) = 0$, la cual es equivalentemente a la condición $\sum_{i=1}^n F_i(\mathbf{x})G_i(\mathbf{x}) = 0$, permitirá decidir cuál (o cuáles) de los puntos de \mathcal{R} es (son) solución (soluciones) del PCG(F, G).

Teniendo en cuenta lo anterior, cuando $n = 2$, un punto $(x, y) = (a, b) \in \mathcal{R}$ es solución del PCG(F, G), si los vectores no negativos $F(a, b)$ y $G(a, b)$ satisfacen una de las dos condiciones siguientes (no mutuamente excluyentes): $F(a, b) = (0, F_2(a, b))$ y $G(a, b) = (G_1(a, b), 0)$, $F(a, b) = (F_1(a, b), 0)$ y $G(a, b) = (0, G_2(a, b))$, lo cual permite caracterizar las soluciones del PCG(F, G) como las raíces de F_i (o G_i), para las cuales $G_i(a, b) \geq 0$ (respectivamente, $F_i(a, b) \geq 0$), para $i = 1, 2$. Por lo tanto, que las raíces de F o de G no son las únicas candidatas a ser solución del PCG(F, G), como sí sucede cuando F y G están definidas sobre \mathbb{R} .

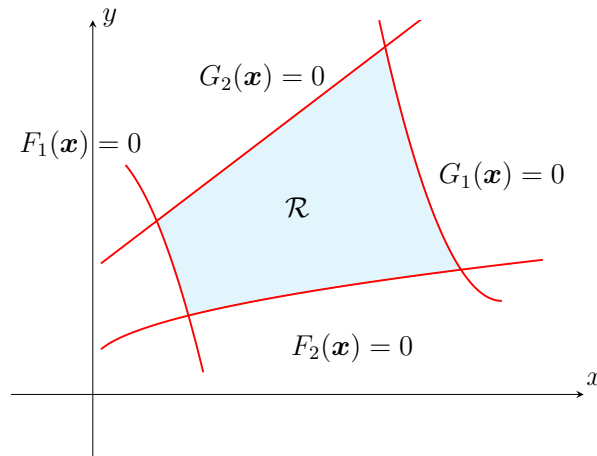


Figura 2.7: Región factible.

Gráficamente, la región factible, en el caso $n = 2$, es la intersección de cuatro regiones determinadas por las funciones componentes F y G (Figura 2.7). En este conjunto, debemos analizar los puntos que satisfacen la condición de complementariedad y así, siempre que \mathcal{R} sea no vacío, podemos intentar determinar la o las soluciones del problema.

Ejemplo 2.6. Para las funciones F y G , definidas por:

$$F(\mathbf{x}) = \begin{pmatrix} 15 - x_2 \\ -\frac{45}{2} + 2x_2 + \frac{5}{4}x_1 \end{pmatrix} \quad y \quad G(\mathbf{x}) = \begin{pmatrix} -\frac{100}{3} + 2x_1 + \frac{8}{3}x_2 \\ 20 - x_1 \end{pmatrix}.$$

Los puntos $\mathbf{P} = (20, 15)$ y $\mathbf{R} = (10, 5)$ son soluciones del $\text{PCG}(F, G)$, puesto que $F(\mathbf{P}) = (0, 32.5)$, $G(\mathbf{P}) = (46.6, 0)$, $F(\mathbf{R}) = (10, 0)$ y $G(\mathbf{R}) = (0, 10)$. Sin embargo ninguna de estas, es raíz de las funciones F o G .

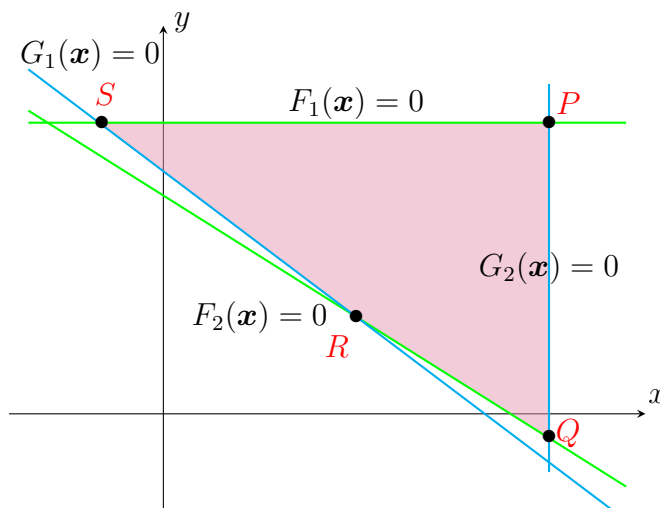


Figura 2.8: Soluciones que no son raíces.

En general, si analizamos la existencia y unicidad de soluciones del $\text{PCG}(F, G)$, nos damos cuenta que este problema puede tener solución, no tenerla, tener varias soluciones o incluso infinitas.

Existen soluciones llamadas *degeneradas* que por sus características especiales, juegan un papel importante en los métodos de solución del problema. Definimos a continuación este tipo de soluciones.

Definición 2.1. Una solución \mathbf{x}_* del $\text{PCG}(F, G)$ es llamada *degenerada*, si existe un índice $i \in \{1, 2, \dots, n\}$, tal que $F_i(\mathbf{x}_*) = G_i(\mathbf{x}_*) = 0$.

El siguiente ejemplo ilustra el caso de soluciones degeneradas para F y G definidas sobre \mathbb{R}^3 .

Ejemplo 2.7. Sean F y G definidas por:

$$F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y^2 + 1 \\ z - x^2 - y^2 + 1 \\ z \end{pmatrix} \quad y \quad G \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ z + x^2 + y^2 - 1 \\ x^2 + 1 \end{pmatrix}.$$

A partir de la condición de complementariedad aplicada componente a componente, tenemos que todos los puntos de la circunferencia intersección de los paraboloides: $z = -1 + x^2 + y^2$ y $z = 1 - x^2 - y^2$; es decir, $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, z = 0\}$, es el conjunto de soluciones (degeneradas) del PCG(F, G) como lo ilustra la Figura 2.9.

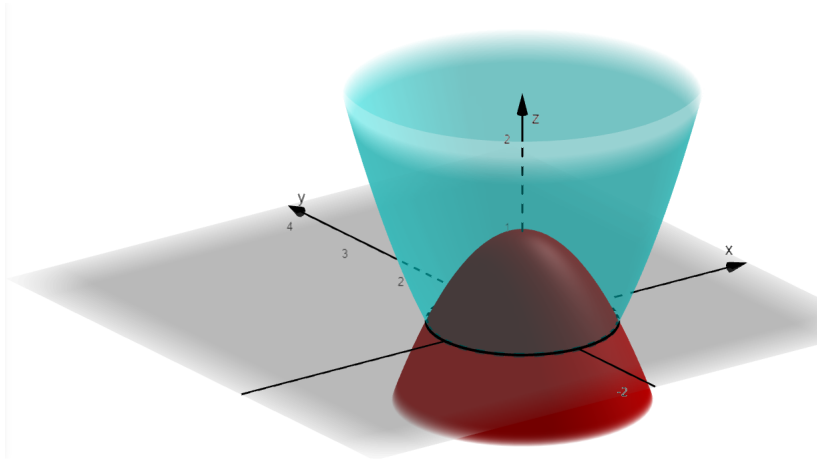


Figura 2.9: Infinitas soluciones degeneradas.

Podemos observar que, a medida que aumenta la dimensión de las funciones involucradas en el problema de complementariedad generalizado, encontrar estas soluciones analíticamente se hace más dispendioso, por lo cual se hace necesario usar métodos computacionales para este propósito.

En general, el estudio analítico del problema de existencia y unicidad de soluciones del PCG(F, G) asociado a las funciones F y G hace necesario que estas satisfagan ciertas propiedades de monotonía. Por tal motivo, definiremos el concepto de par fuertemente monótono.

Definición 2.2. Sean $X \subset \mathbb{R}^n$ un conjunto no vacío y $F, G: X \rightarrow \mathbb{R}^n$ funciones continuamente diferenciables. Decimos que el par (F, G) es fuertemente monótono

(sobre X) si existe una constante $\mu > 0$, tal que

$$(F(\mathbf{x}) - F(\mathbf{y}))^T (G(\mathbf{x}) - G(\mathbf{y})) \geq \mu \|\mathbf{x} - \mathbf{y}\|^2, \quad \forall \mathbf{x}, \mathbf{y} \in X. \quad (2.1)$$

En particular, si G es la función identidad, la **Definición 2.2** se reduce a la de monotonía fuerte para $F: X \rightarrow \mathbb{R}^n$ [3].

Para una interpretación geométrica del concepto de par fuertemente monótono, supongamos que el par (F, G) con $F, G: X \subset \mathbb{R} \rightarrow \mathbb{R}$ es fuertemente monótono y sin pérdida de generalidad que, $x > y$ entonces existe una constante $\mu > 0$, tal que

$$(F(x) - F(y))^T (G(x) - G(y)) \geq \mu |x - y|^2, \quad \forall x, y \in X.$$

lo cual es equivalente a la siguiente desigualdad

$$\frac{(F(x) - F(y)) (G(x) - G(y))}{(x - y) (x - y)} \geq \mu > 0.$$

Entonces las pendientes de la recta secante a la función F que une los puntos $(x, F(x))$ y $(y, F(y))$ tiene el mismo signo que la recta secante a la función G que une los puntos $(x, G(x))$ y $(y, G(y))$, lo que implica que ambas rectas secantes son simultáneamente crecientes o decrecientes en X . Además, dado que $\mu > 0$, geoméricamente, las rectas secantes de cualquier par de funciones fuertemente monótonas en un conjunto $X \subset \mathbb{R}$ están lejos de ser horizontales. Esto lo podemos observar en el **Ejemplo 2.8**.

En [60], los autores dan condiciones suficientes sobre las funciones F y G para garantizar la existencia y unicidad de la solución del PCG(F, G). El siguiente lema resume dichos resultados demostrados en [60].

Lema 2.1. [60] *Si el par (F, G) es fuertemente monótono sobre $X \subset \mathbb{R}^n$ entonces el PCG(F, G) tiene a lo más una solución. Además, si la función F (o G) es invertible y Lipschitz continua, el par $(G \circ F^{-1}, I)$ es fuertemente monótono y F^{-1} es continua sobre $X \subset \mathbb{R}^n$ entonces el PCG(F, G) tiene solución única.*

Ejemplo 2.8. Sean F y G definidas por:

$$F(\mathbf{x}) = x^3 \quad y \quad G(x) = \sqrt[5]{x}.$$

Un interrogante que podría presentarse en este caso es si el PCG(F, G) tiene solución y de tenerla, es única?

No es difícil mostrar que este par (F, G) es fuertemente monótono sobre el intervalo $[0, \infty) \subset \mathbb{R}$. Ahora, dado que se cumplen las hipótesis del **Lema 2.1**, concluimos

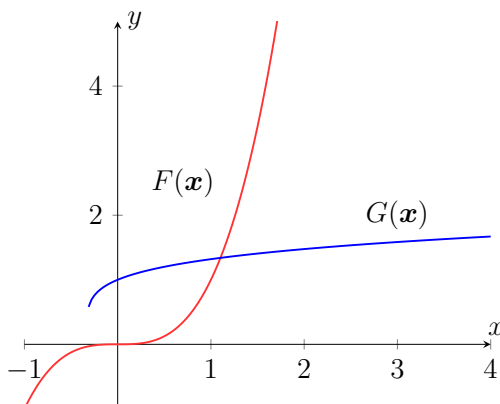


Figura 2.10: Par (F, G) fuertemente monótono.

que en este caso, el $\text{PCG}(F, G)$ tiene solución y es única. De hecho, en la Figura 2.10 se puede ver que dicha solución es $x = 0$.

Problemas de prueba

En este capítulo, presentamos diez problemas de prueba, algunos de los cuales provienen de aplicaciones, los cuales serán usados para analizar el desempeño numérico de los algoritmos que proponemos en esta investigación y describimos una estrategia que usaremos para actualizar el parámetro λ .

Para cada problema describimos los puntos iniciales y la(s) solución(es) encontrada(s). Denotaremos el vector de *unos* en \mathbb{R}^n como \mathbf{e}_n .

Problema 1 [35]. Este problema corresponde al caso particular de un PCNL(F), en el cual F es la función de *Kojima-Shindo* y G es la función identidad de orden cuatro; es decir,

$$F(\mathbf{x}) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_2^2 + x_1 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix} \quad y \quad G(\mathbf{x}) = \mathbf{x}.$$

Puntos iniciales : $\mathbf{x}_1 = (0, 0, 0, 0)$, $\mathbf{x}_2 = (1, 0, 1, 0)$, $\mathbf{x}_3 = (1, 0, 0, 0)$ y $\mathbf{x}_4 = (0, 1, 1, 0)$.
Soluciones: $\mathbf{x}_*^1 = (1, 0, 3, 0)$ y $\mathbf{x}_*^2 = (\sqrt{6}/2, 0, 0, 1/2)$.

Problema 2 [22]. Este problema fue usado para mostrar el desempeño numérico del método de Newton generalizado para el PCG(F, G) mediante su reformulación con la función de Fischer. En este caso, $F, G: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ están definidas por:

$$F(\mathbf{x}) = \begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix} \quad y \quad G(\mathbf{x}) = \begin{pmatrix} x_1^2 + 10 \\ x_2^2 + 1 \end{pmatrix}.$$

Puntos iniciales: $\mathbf{x}_5 = (10, 1)$, $\mathbf{x}_6 = 100\mathbf{e}_2$, $\mathbf{x}_7 = (1000, 10000)$ y $\mathbf{x}_8 = 10000\mathbf{e}_2$.
Solución: $\mathbf{x}_*^1 = (0, 0)$.

Los **Problemas 3, 4 y 6**, originalmente fueron planteados en [44] en el contexto de complementariedad implícita; nosotros los usamos en complementariedad generalizada.

Problema 3 [44]. Las funciones $F, G: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ están definidas por :

$$F(\mathbf{x}) = \begin{pmatrix} -\frac{100}{3} + 2x_1 + \frac{8}{3}x_2 \\ -22.5 + 2x_2 + \frac{5}{4}x_1 \end{pmatrix} \quad y \quad G(\mathbf{x}) = \begin{pmatrix} 15 - x_2 \\ 20 - x_1 \end{pmatrix}.$$

Puntos iniciales: $\mathbf{x}_9 = (0, 0)$, $\mathbf{x}_{10} = (5, 0)$ y $\mathbf{x}_{11} = (11, 0)$. Soluciones: $\mathbf{x}_*^1 = (10, 5)$ y $\mathbf{x}_*^2 = (20, 15)$.

Problema 4 [44]. Las funciones $F, G: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ están definidas por:

$$F(\mathbf{x}) = A\mathbf{x} + \mathbf{b} \quad y \quad G(\mathbf{x}) = \mathbf{x} - \varphi(\mathbf{x}),$$

donde $\varphi_i(\mathbf{x}) = -0.5 - x_i$, con A y b definidas por:

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \quad y \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Puntos iniciales: $\mathbf{x}_{12} = \mathbf{x}_1$, $\mathbf{x}_{13} = -0.5\mathbf{e}_4$ y $\mathbf{x}_{14} = -\mathbf{e}_4$.
Solución: $\mathbf{x}_*^1 = (-0.9, -1.2, -1.2, -0.9)$.

Problema 5 [49]. Sean $F, G: \mathbb{R}^5 \rightarrow \mathbb{R}^5$ definidas por:

$$G(\mathbf{x}) = F(\mathbf{x}) = c + L^{\frac{1}{b}}\mathbf{x}^{\frac{1}{b}} - \left(\frac{5000}{\left(\sum_{i=1}^5 x_i \right)} \right)^{\frac{1}{\gamma}} \begin{pmatrix} \mathbf{e}_5 - \frac{1}{\gamma \left(\sum_{i=1}^5 x_i \right)} \mathbf{x} \end{pmatrix},$$

donde $c = (10, 8, 6, 4, 2)^T$, $b = (1.2, 1.1, 1, 0.9, 0.8)^T$, $L = (5, 5, 5, 5, 5)$ y $\gamma = 1.1$.

Puntos iniciales: $\mathbf{x}_{15} = \mathbf{e}_5$, $\mathbf{x}_{16} = 10\mathbf{e}_5$ y $\mathbf{x}_{17} = 20\mathbf{e}_5$.
Solución: $\mathbf{x}_*^1 = (15.4293, 12.4986, 9.6635, 7.1651, 5.1326)^T$.

Problema 6 [44] La función F se define como en el **Problema 3**, la función $G: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ es definida por $G(\mathbf{x}) = (2.5 - x_1^2, 2.5 - x_2^2, 2.5 - x_3^2, 2.5 - x_4^2)$.

Puntos iniciales: $\mathbf{x}_{18} = \mathbf{x}_1$, $\mathbf{x}_{19} = \mathbf{x}_{13}$, $\mathbf{x}_{20} = 2\mathbf{x}_{13}$ y $\mathbf{x}_{21} = 4\mathbf{x}_{13}$.
 Solución: $\mathbf{x}_*^1 = (-0.9, -1.2, -1.2, -0.9)$.

Problema 7 [31]. Las funciones $F, G: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ están definidas por:

$$F(\mathbf{x}) = (e^{x_1} - x_2, x_1 + x_2 - 1) \quad y \quad G(\mathbf{x}) = (x_1^2 - 2x_1 + x_2 - 2, x_2 - e^{-x_1}).$$

Puntos iniciales: $\mathbf{x}_{22} = (2, 3)$, $\mathbf{x}_{23} = (3, 1)$, $\mathbf{x}_{24} = (3.5, 2)$ y $\mathbf{x}_{25} = (4, 0.5)$.
 Solución: $\mathbf{x}_*^1 = (2.7128, 0.0664)$.

Problema 8 Las funciones $F, G: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ están definidas por:

$$F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y^2 + 1 \\ z - x^2 - y^2 + 1 \\ z \end{pmatrix} \quad y \quad G \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ z + x^2 + y^2 - 1 \\ x^2 + 1 \end{pmatrix}.$$

Puntos iniciales: $\mathbf{x}_{26} = (0, 0, 0)$, $\mathbf{x}_{27} = -0.5\mathbf{e}_3$, $\mathbf{x}_{28} = -\mathbf{e}_3$ y $\mathbf{x}_{29} = -2\mathbf{e}_3$.

Conjunto de soluciones: $\{(x, y, z) \in \mathbb{R}^3: x^2 + y^2 = 1, z = 0\}$. Es decir, la circunferencia intersección entre los paraboloides $z = -1 + x^2 + y^2$ y $z = -1 - x^2 - y^2$.

Problema 9 [62]. Las funciones $F, G: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ están definidas por $F(\mathbf{x}) = A\mathbf{x} + \mathbf{q} + \Gamma(\mathbf{x})$, $G(\mathbf{x}) = \mathbf{x} - \Upsilon(\mathbf{x})$, con $A = \text{tridiag}(-I, S, -I) \in \mathbb{R}^{n \times n}$, $S = \text{tridiag}(-1, 4, -1) \in \mathbb{R}^{m \times m}$, $\mathbf{q}(\mathbf{x}) = (-1, 1, \dots, (-1)^n)$, $\Psi(\mathbf{x}) = \mathbf{x}^2$ y $\Phi(\mathbf{x}) = \mathbf{x}^3$; Aquí, $\mathbf{x}^p = (x_1^p, x_2^p, \dots, x_n^p)$, para $\mathbf{x} \in \mathbb{R}^n$, m es un entero positivo, $n = m^2$ y $\mathbf{q} \in \mathbb{R}^n$. usamos los mismos puntos iniciales de [62], con $\mathbf{x}_n = (0, 0, \dots, 0)$. Además, usamos los puntos iniciales \mathbf{x}_1 , \mathbf{x}_2 , \mathbf{x}_3 y \mathbf{x}_4 .

Problema 10 [62]. Las funciones F, G, Φ y Ψ están definidas con en el **Problema 9** con $A = \text{tridiag}(-1.5I, S, -0.5I)$, $S = \text{tridiag}(-1.5, 4, -0.5)$ y los mismos puntos iniciales usados en ese problema.

Por otra parte, mencionamos que los algoritmos propuestos se implementaron en Matlab[®] R2020a, en un computador con procesador AMD Sempron (tm) de 2.21 GHz.

Basados en trabajos previos [34] [39] [4], para métodos Newton y cuasi-Newton generalizados, en los que se muestran resultados numéricos en los que la elección $\lambda = 2$ (función de Fischer) da mejores propiedades de convergencia global y valores de λ cercanos a cero usualmente conducen a convergencia local más rápida del respectivo algoritmo, usamos en cada iteración la siguiente estrategia dinámica [34] para la escogencia del parámetro λ :

- Iniciamos con $\lambda = 2$.

- Si $\Psi_\lambda(\mathbf{x}_k) \leq \gamma_1$ entonces $\lambda = \Psi_\lambda(\mathbf{x}_k)$. Sino $\lambda = \min \{c_1 \Psi_\lambda(\mathbf{x}_k), \lambda\}$.
- Si $\Psi_\lambda(\mathbf{x}_k) \leq \gamma_2$ entonces $\lambda = \min \{c_2, \lambda\}$,

con $\gamma_1 = 10^{-2}$, $\gamma_2 = 10^{-4}$, $c_1 = 10$ y $c_2 = 10^{-8}$. Esta estrategia usa el valor $\lambda = 2$ lejos de la solución y lo reduce rápidamente, si estamos cerca de una solución del problema de complementariedad.

Capítulo 4

Preliminares

En este capítulo, presentamos notaciones, algunos conceptos básicos, hipótesis y propiedades necesarias en el desarrollo de este trabajo.

Sobre el espacio vectorial \mathbf{V} , denotamos el producto interno como $\langle \cdot, \cdot \rangle$ y la norma asociada, $\|\cdot\|$; en particular, la norma matricial infinito de una matriz $A \in \mathbb{R}^{n \times n}$, se define como en [61], por:

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \max_{1 \leq i \leq n} \{\|[A]_i\|_1\}, \quad (4.1)$$

donde $\|[A]_i\|_1 = \sum_{j=1}^n |a_{ij}|$ es la norma $\|\cdot\|_1$ aplicada a la i -ésima fila de la matriz A . En cuanto a las demás normas que aparecen en el documento, a no ser que se diga lo contrario, serán la norma euclidiana.

Para $W: \mathbb{R}^n \rightarrow \mathbb{R}$ y $U: \mathbb{R}^n \rightarrow \mathbb{R}^n$, el gradiente de W y la matriz jacobiana de U en un punto \mathbf{x} se denotarán $\nabla W(\mathbf{x})$ y $U'(\mathbf{x})$, respectivamente.

Entre los conceptos importantes en análisis semisuave, están los de B-Jacobiano y Jacobiano generalizado, que definimos a continuación.

Definición 4.1. [39]. Sea $K: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función localmente Lipschitz y sea D_K el conjunto donde K es diferenciable. Para todo $\mathbf{x} \in \mathbb{R}^n$, el conjunto dado por

$$\partial_B K(\mathbf{x}) = \left\{ \lim_{k \rightarrow \infty} K'(\mathbf{x}_k) \in \mathbb{R}^{n \times n} : \lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}, \mathbf{x}_k \in D_K \right\},$$

se conoce como B-Jacobiano de K en \mathbf{x} .

Definición 4.2. La envolvente convexa de $\partial_B K(\mathbf{x})$ se conoce como Jacobiano gene-

ralizado de K en \mathbf{x} , y se denota $\partial K(\mathbf{x})$.

Usualmente, el conjunto $\partial K(\mathbf{x})$ es difícil de calcular. Una alternativa es la siguiente estimación

$$\partial K(\mathbf{x})^T \subseteq \partial K_1(\mathbf{x}) \times \cdots \times \partial K_n(\mathbf{x}) \quad (4.2)$$

dada en [14], donde el lado derecho, se denota como $\partial_C K(\mathbf{x})$ [53], es conocida como *C-subdiferencial* de K en \mathbf{x} y es el conjunto de matrices en $\mathbb{R}^{n \times n}$, cuya i -ésima columna es el gradiente generalizado de la i -ésima componente de la función K .

Definición 4.3. [52]. Una función localmente Lipschitz y direccionalmente diferenciable $K: \mathbb{R}^n \rightarrow \mathbb{R}^n$ se dice semisuave en el punto $\mathbf{x} \in \mathbb{R}^n$, si

$$H\mathbf{d} - K'(\mathbf{x}; \mathbf{d}) = o(\|\mathbf{d}\|)$$

se satisface para todo \mathbf{d} tendiendo a cero y toda matriz $H \in \partial K(\mathbf{x} + t\mathbf{d})$ y, además

$$\lim_{H \in \partial K(\mathbf{x} + t\mathbf{d}'), \mathbf{d}' \rightarrow \mathbf{d}, t \rightarrow 0^+} \{H\mathbf{d}'\}$$

existe para cualquier $\mathbf{d} \in \mathbb{R}^n$. Es fuertemente semisuave en $\mathbf{x} \in \mathbb{R}^n$ si

$$H\mathbf{d} - K'(\mathbf{x}; \mathbf{d}) = O(\|\mathbf{d}\|^2)$$

se satisface para todo \mathbf{d} tendiendo a cero y toda matriz $H \in \partial K(\mathbf{x} + \mathbf{d})$. Aquí, $K'(\mathbf{x}; \mathbf{d})$ denota usualmente la derivada direccional de K en \mathbf{x} y en la dirección \mathbf{d} .

Definición 4.4. [52]. Suponga que $K: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es B -diferenciable en una vecindad de \mathbf{x} . Decimos que la derivada direccional $K'(\mathbf{x}; \mathbf{d})$ es semicontinua en \mathbf{x} si, para todo $\epsilon > 0$, existe un entorno N de \mathbf{x} tal que, para todo $\mathbf{x} + \mathbf{d} \in N$,

$$\|K'(\mathbf{x} + \mathbf{d}, \mathbf{d}) - K'(\mathbf{x}, \mathbf{d})\| \leq \epsilon \|\mathbf{d}\|. \quad (4.3)$$

Definición 4.5. [34]. La función $K: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es de clase LC^1 , si es diferenciable y su derivada es localmente Lipschitz continua.

Definición 4.6. [52] La función $K: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es B -diferenciable en \mathbf{x} , si es direccionalmente diferenciable en \mathbf{x} y

$$K(\mathbf{x} + \mathbf{h}) - K(\mathbf{x}) - K'(\mathbf{x}, \mathbf{h}) = o(\|\mathbf{h}\|). \quad (4.4)$$

Más aún, K es B -diferenciable de grado 2 en \mathbf{x} si y solo si

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}) + F'(\mathbf{x}; \mathbf{h}) + O(\|\mathbf{h}\|^2).$$

Además, la derivada direccional $K'(\cdot, \cdot)$ es semicontinua de grado 2 en \mathbf{x} [52] si existe una constante L y una vecindad N de \mathbf{x} , tal que para todo $\mathbf{x} + \mathbf{h} \in N$,

$$\|K'(\mathbf{x} + \mathbf{h}, \mathbf{h}) - K'(\mathbf{x}, \mathbf{h})\| \leq L \|\mathbf{h}\|^2. \quad (4.5)$$

Los siguientes dos lemas serán útiles para determinar la tasa de convergencia de nuestros algoritmos.

Lema 4.1. ([52]) *Supongamos que $K: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es direccionalmente diferenciable en una vecindad de \mathbf{x} . Las siguientes proposiciones son equivalentes*

- (1) K es semisuave en \mathbf{x} .
- (2) $K'(\cdot, \cdot)$ es semicontinua en \mathbf{x} .
- (3) Para cualquiera matriz $H \in \partial K(\mathbf{x} + \mathbf{h})$, $\mathbf{h} \rightarrow \mathbf{0}$, $H\mathbf{h} - K'(\mathbf{x}; \mathbf{h}) = o(\|\mathbf{h}\|)$.

Lema 4.2. ([52]) *Supongamos que $K: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es direccionalmente diferenciable en una vecindad de \mathbf{x} . Las siguientes proposiciones son equivalentes:*

- (1) K is semicontinua de grado 2 en \mathbf{x} .
- (2) Para cualquiera matriz $H \in \partial K(\mathbf{x} + \mathbf{h})$, $\mathbf{h} \rightarrow \mathbf{0}$, $H\mathbf{h} - K'(\mathbf{x}; \mathbf{h}) = O(\|\mathbf{h}\|^2)$.
Si (1) o (2) se cumplen, entonces K is B-diferenciable de grado 2 en \mathbf{x} .

A continuación, definimos los conceptos de BD-regularidad y R-regularidad en el contexto de complementariedad generalizada. Para ello, consideramos una solución $\mathbf{x}_* = (x_{*1}, \dots, x_{*n})$ de PCG(F, G) y el conjunto de índices,

$$\begin{aligned} \alpha &= \{i \in I: F_i(\mathbf{x}_*) > 0 = G_i(\mathbf{x}_*)\} \\ \beta &= \{i \in I: F_i(\mathbf{x}_*) = 0 = G_i(\mathbf{x}_*)\} \\ \gamma &= \{i \in I: F_i(\mathbf{x}_*) = 0 < G_i(\mathbf{x}_*)\}. \end{aligned} \quad (4.6)$$

Definición 4.7. [4]. *Sea \mathbf{x}_* una solución de PCG(F, G).*

1. Si todas las matrices $H \in \partial_B \Phi(\mathbf{x}_*)$ son no singulares, \mathbf{x}_* es llamada una solución BD-regular.
2. Sea $F'(\mathbf{x}_*)$ no singular y $K = G'(\mathbf{x}_*)F'(\mathbf{x}_*)^{-1}$. Si la submatriz ¹ $K_{\alpha\alpha}$ es no singular y el complemento de Schur de K , $K_{\beta\beta} - K_{\beta\alpha}K_{\alpha\alpha}^{-1}K_{\alpha\beta} \in \mathbb{R}^{|\beta| \times |\beta|}$ es una P-matriz ² entonces \mathbf{x}_* es llamada una solución R-regular.

¹Sea $A = (a_{ij}) \in \mathbb{R}^{m \times n}$. La matriz $A_{\alpha\beta}$ es aquella cuyos elementos a_{ij} tales que $i \in \alpha$ y $j \in \beta$.

²Una matriz $M \in \mathbb{R}^{n \times n}$ es una P-matriz, si para todo vector no nulo \mathbf{z} existe un índice $j \in \{1, \dots, n\}$, tal que $\mathbf{z}_j[M\mathbf{z}]_j > 0$.

Observe que para el caso particular $G(\mathbf{x}) = \mathbf{x}$, la **Definition 4.7** se reduce a la de BD-regularidad y R -regularidad para complementariedad no lineal.

Siguiendo [34], usaremos la notación $f_\lambda(a, b)$ para el primer término del lado derecho de (1.5) y usaremos la cota superior c_λ para el gradiente de f_λ dado en el siguiente lema.

Lema 4.3. [23]. *Sea $f_\lambda: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por*

$$f_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab}, \quad \lambda \in (0, 4). \quad (4.7)$$

Existe una constante $c_\lambda \in (0, 2)$, tal que $\|\nabla f_\lambda(a, b)\|^2 \leq c_\lambda$, para todos los vectores no nulos $(a, b) \in \mathbb{R}^2$.

Finalmente, presentamos a continuación una caracterización de la clase de P -matrices para complementariedad no lineal, que será de utilidad más adelante.

Proposición 4.1. [34]. *Una matriz de la forma $D_a + D_b N$ es no singular para todas las matrices diagonales semidefinidas positivas (negativas) $D_a, D_b \in \mathbb{R}^{n \times n}$ tales que $D_a + D_b$ es definida positiva (negativa) si, y solo si $N \in \mathbb{R}^{n \times n}$ es una P -matriz.*

El siguiente resultado da una relación entre las normas de las inversas de dos matrices cercanas y será útil en el análisis de convergencia de los algoritmos propuestos más adelante.

Lema 4.4. [20]. *Sea $\|\cdot\|$ una norma matricial inducida y $A, C \in \mathbb{R}^{n \times n}$. Si C es no singular y $\|I_n - C^{-1}A\| < 1$ entonces A es no singular y además,*

$$\|A^{-1}\| \leq \frac{\|C^{-1}\|}{1 - \|I_n - C^{-1}A\|}.$$

Exigir que en cada iteración se satisfaga la ecuación secante y que haya un cambio mínimo entre dos actualizaciones consecutivas, hace que la sucesión de matrices $\{A_k\}$ exhiba un fenómeno conocido como deteriorización acotada [20, 41], el cual garantiza que las matrices de la sucesión permanecen en una vecindad de $F'(\mathbf{x}_*)$. El siguiente es un lema técnico que muestra que si la actualización de Broyden [20], se deteriora, lo hace en forma controlada.

Lema 4.5. [20] *Sea $D \subseteq \mathbb{R}^n$ un conjunto abierto y conexo que contiene los puntos \mathbf{x}_k y \mathbf{x}_{k+1} con $\mathbf{x}_k \neq \mathbf{x}_*$. Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F'(\mathbf{x}) \in Lip_\gamma(D)$, $A_k \in \mathbb{R}^{n \times n}$, A_{k+1} definido por*

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k}.$$

Entonces ya sea para la norma de Frobenius o la norma matricial 2 se cumple que

$$\|A_{k+1} - F'(\mathbf{x}_{k+1})\| \leq \|A_k - F'(\mathbf{x}_k)\| + \frac{3\gamma}{2} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2.$$

Además, si $\mathbf{x}_* \in D$ y $F'(\mathbf{x})$ satisfacen la condición débil de Lipschitz

$$\|F'(\mathbf{x}) - F'(\mathbf{x}_*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}_*\|$$

entonces

$$\|A_{k+1} - F'(\mathbf{x}_*)\| \leq \|A_k - F'(\mathbf{x}_*)\| + \frac{\gamma}{2} (\|\mathbf{x}_{k+1} - \mathbf{x}_*\|_2 + \|\mathbf{x}_k - \mathbf{x}_*\|_2).$$

Los siguientes son dos lemas técnicos que ayudan a analizar la convergencia de sucesiones, los cuales serán útiles en la teoría de convergencia del algoritmo propuesto en el **Capítulo 7**.

Lema 4.6. [19, 37] Sean $\{a_k\}$ y $\{r_k\}$ sucesiones de términos positivos tales que $a_{k+1} \leq (1 + r_k)a_k + r_k$ y

$$\sum_{k=0}^{\infty} \|r_k\| < \infty$$

entonces $\{a_k\}$ converge.

Lema 4.7. [37] Sean $\{a_k\}$, $\{b_k\}$ y $\{\zeta_k\}$ sucesiones de términos positivos tales que

$$a_{k+1}^2 \leq (a_k + b_k)^2 - \alpha \zeta_k^2, \quad k = 0, 1, \dots,$$

donde $\alpha > 0$ es una constante. Entonces la condición

$$\sum_{k=0}^{\infty} \|b_k\|^2 < \infty$$

implica

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \zeta_i^2 = 0.$$

y la condición

$$\sum_{k=0}^{\infty} \|b_k\| < \infty$$

implica

$$\sum_{k=0}^{\infty} \zeta_k^2 < \infty.$$

Capítulo 5

Un método de Newton no suave

En este capítulo, analizamos y extendemos a complementariedad generalizada las propiedades del operador Φ_λ investigadas detalladamente para el caso particular de complementariedad no lineal en [34]. Adicionalmente, presentamos un nuevo algoritmo tipo Newton no suave para resolver el PCG(F, G) basado en su reformulación como un sistema de ecuaciones no lineales usando la familia uniparamétrica de funciones de complementariedad [2]. Demostramos resultados de convergencia global para este algoritmo y analizamos su desempeño numérico.

5.1. El operador Φ_λ

Para el análisis de Φ_λ es necesario considerar tanto la función de complementariedad φ_λ como la función auxiliar f_λ , definidas en (1.2) y (4.7), respectivamente. Iniciamos con un lema y corolario que dan cotas superiores para las derivadas parciales de la función f_λ , las cuales usaremos en resultados posteriores.

Lema 5.1. *Las derivadas parciales de la función f_λ están acotadas superiormente por 1, para todo vector distinto de cero $(a, b) \in \mathbb{R}^2$.*

Demostración. Debemos demostrar que para todo vector $(a, b) \in \mathbb{R}^2$, distinto de cero, se satisfacen las siguientes desigualdades:

$$\frac{\partial f_\lambda(a, b)}{\partial a} \leq 1 \quad \text{y} \quad \frac{\partial f_\lambda(a, b)}{\partial b} \leq 1. \quad (5.1)$$

La primera desigualdad fue probada en [2]. Análogamente, demostramos a continuación la segunda. De la desigualdad $\lambda a^2(\lambda - 4) \leq 0$, sumando $4(a - b)^2 + 4\lambda ab$ y después

de algunas manipulaciones algebraicas, obtenemos las desigualdades

$$4(a-b)^2 + 4\lambda ab + \lambda^2 a^2 - 4\lambda a^2 \leq 4f_\lambda^2(a, b)$$

$$4(a-b)^2 - 4\lambda a(a-b) + \lambda^2 a^2 \leq 4f_\lambda^2(a, b)$$

$$[-2(a-b) + \lambda a]^2 \leq 4f_\lambda^2(a, b),$$

lo que implica que $|-2(a-b) + \lambda a| \leq 2f_\lambda(a, b) > 0$ luego,

$$\left| \frac{-2(a-b) + \lambda a}{2f_\lambda(a, b)} \right| \leq 1,$$

luego

$$-1 \leq \frac{-2(a-b) + \lambda a}{2f_\lambda(a, b)} \leq 1.$$

Por lo tanto, obtenemos la segunda desigualdad en (5.1). \square

Corolario 5.1. *Las derivadas parciales de la función f_λ satisfacen la desigualdad*

$$\frac{\partial f_\lambda(a, b)}{\partial a} + \frac{\partial f_\lambda(a, b)}{\partial b} < 2,$$

para todo vector no nulo $(a, b) \in \mathbb{R}^2$.

Demostración. Por el **Lema 5.1**,

$$\frac{\partial f_\lambda(a, b)}{\partial a} + \frac{\partial f_\lambda(a, b)}{\partial b} \leq 2.$$

Con el objetivo de llegar a una contradicción, supongamos que $\frac{\partial f_\lambda(a, b)}{\partial a} + \frac{\partial f_\lambda(a, b)}{\partial b} = 2$.

Así, $\frac{\partial f_\lambda(a, b)}{\partial a} = \frac{\partial f_\lambda(a, b)}{\partial b} = 1$, lo que contradice la desigualdad $\|\nabla f_\lambda(a, b)\|^2 \leq c_\lambda < 2$ del

Lema 4.3. Por lo tanto, $\frac{\partial f_\lambda(a, b)}{\partial a} + \frac{\partial f_\lambda(a, b)}{\partial b} < 2$. \square

El siguiente lema da una expresión compacta para el gradiente de las funciones componentes de Φ_λ , cuando φ_λ es diferenciable. Esta expresión será útil para caracterizar las matrices del jacobiano generalizado de Φ_λ en \mathbf{x} .

Lema 5.2. *Sea $\Phi_{\lambda, i}(\mathbf{x}) = \varphi_\lambda(F_i(\mathbf{x}), G_i(\mathbf{x}))$, $i \in \{1, 2, \dots, n\}$, tal que $(F_i(\mathbf{x}), G_i(\mathbf{x})) \neq (0, 0)$. Entonces el gradiente de $\Phi_{\lambda, i}(\mathbf{x})$ está dado por*

$$\nabla \Phi_{\lambda, i}(\mathbf{x}) = (a_i(\mathbf{x}) - 1)\nabla F_i(\mathbf{x}) + (b_i(\mathbf{x}) - 1)\nabla G_i(\mathbf{x}),$$

donde

$$a_i(\mathbf{x}) = \frac{2(F_i(\mathbf{x}) - G_i(\mathbf{x})) + \lambda G_i(\mathbf{x})}{2\sqrt{(F_i(\mathbf{x}) - G_i(\mathbf{x}))^2 + \lambda F_i(\mathbf{x})G_i(\mathbf{x})}} \quad (5.2)$$

$$b_i(\mathbf{x}) = \frac{-2(F_i(\mathbf{x}) - G_i(\mathbf{x})) + \lambda F_i(\mathbf{x})}{2\sqrt{(F_i(\mathbf{x}) - G_i(\mathbf{x}))^2 + \lambda F_i(\mathbf{x})G_i(\mathbf{x})}}. \quad (5.3)$$

Demostración. Sea $\mathbf{x} \in \mathbb{R}^n$, tal que $(F_i(\mathbf{x}), G_i(\mathbf{x})) \neq (0, 0)$. Entonces Φ_i y φ_λ son diferenciables en \mathbf{x} . Más aún, $\nabla\Phi_{\lambda,i}(\mathbf{x}) = \nabla\varphi_\lambda(F_i(\mathbf{x}), G_i(\mathbf{x}))$. Después de algunas manipulaciones algebraicas en las que se usa la regla de la cadena, obtenemos que

$$\nabla\Phi_{\lambda,i}(\mathbf{x}) = \nabla\varphi_\lambda(F_i(\mathbf{x}), G_i(\mathbf{x})) = \begin{bmatrix} (a_i(\mathbf{x}) - 1)\frac{\partial F_i(\mathbf{x})}{\partial x_1} + (b_i(\mathbf{x}) - 1)\frac{\partial G_i(\mathbf{x})}{\partial x_1} \\ \vdots \\ (a_i(\mathbf{x}) - 1)\frac{\partial F_i(\mathbf{x})}{\partial x_n} + (b_i(\mathbf{x}) - 1)\frac{\partial G_i(\mathbf{x})}{\partial x_n} \end{bmatrix},$$

equivalentemente,

$$\nabla\Phi_{\lambda,i}(\mathbf{x}) = (a_i(\mathbf{x}) - 1)\nabla F_i(\mathbf{x}) + (b_i(\mathbf{x}) - 1)\nabla G_i(\mathbf{x}),$$

donde $a_i(\mathbf{x})$ y $b_i(\mathbf{x})$ están dadas por (5.2) y (5.3), respectivamente. \square

El siguiente resultado extiende la **Proposición 2.5** de [34] a complementariedad generalizada; es decir, da una sobreestimación del jacobiano generalizado Φ_λ en \mathbf{x} .

Proposición 5.1. *Para todo $\mathbf{x} \in \mathbb{R}^n$, tenemos que*

$$\partial\Phi_\lambda(\mathbf{x}) \subseteq D_a(\mathbf{x})F'(\mathbf{x}) + D_b(\mathbf{x})G'(\mathbf{x}),$$

donde $D_a(\mathbf{x}) = \text{diag}(a_1(\mathbf{x}) - 1, \dots, a_n(\mathbf{x}) - 1)$ y $D_b(\mathbf{x}) = \text{diag}(b_1(\mathbf{x}) - 1, \dots, b_n(\mathbf{x}) - 1)$ son matrices diagonales con $a_i(\mathbf{x})$ y $b_i(\mathbf{x})$ dadas por (5.2) y (5.3), si $(F_i(\mathbf{x}), G_i(\mathbf{x})) \neq (0, 0)$, y por $a_i(\mathbf{x}) = \xi_i, b_i(\mathbf{x}) = \chi_i$, para todo $(\xi_i, \chi_i) \in \mathbb{R}^2$, tal que $\|(\xi_i, \chi_i)\| \leq \sqrt{c_\lambda}$ si $(F_i(\mathbf{x}), G_i(\mathbf{x})) = (0, 0)$, donde c_λ es la constante dada del **Lema 4.3**.

Demostración. Por (4.2) y la continuidad Lipschitz de Φ_λ , tenemos que

$$\partial\Phi_\lambda(\mathbf{x})^T \subseteq \partial\Phi_{\lambda,1}(\mathbf{x}) \times \cdots \times \partial\Phi_{\lambda,n}(\mathbf{x}),$$

donde $\Phi_{\lambda,i}$ es la i -ésima componente de Φ_λ y $\partial\Phi_{\lambda,i}(\mathbf{x})$ denota el gradiente generalizado de $\Phi_{\lambda,i}$ en \mathbf{x} .

Si $(F_i(\mathbf{x}), G_i(\mathbf{x})) \neq (0, 0)$, la función $\Phi_{\lambda,i}$ es continuamente diferenciable. Entonces el gradiente generalizado $\partial\Phi_{\lambda,i}(\mathbf{x})$ es el conjunto cuyo único elemento es el gradiente de $\Phi_{\lambda,i}$ en \mathbf{x} , que por el **Lema 5.2** está dado por

$$\partial\Phi_{\lambda,i}(\mathbf{x}) = \{(a_i(\mathbf{x}) - 1)\nabla F_i(\mathbf{x}) + (b_i(\mathbf{x}) - 1)\nabla G_i(\mathbf{x})\},$$

donde $a_i(\mathbf{x})$ y $b_i(\mathbf{x})$ están dadas por (5.2) y (5.3), respectivamente.

Si $(F_i(\mathbf{x}), G_i(\mathbf{x})) = (0, 0)$ entonces $\Phi_{\lambda,i}$ no es diferenciable en \mathbf{x} y por lo tanto, para $H \in \partial_B\Phi_{\lambda,i}(\mathbf{x})$, tenemos

$$H = \lim_{\mathbf{x}_k \rightarrow \mathbf{x}} \nabla\Phi_{\lambda,i}(\mathbf{x}_k), \quad (5.4)$$

donde $\{\mathbf{x}_k\}$ es una sucesión de puntos en $D_{\Phi_{\lambda,i}}$ tales que $\mathbf{x}_k \rightarrow \mathbf{x}$, cuando $k \rightarrow \infty$. Así, por **Lema 5.2**, para cada $\mathbf{x}_k \in D_{\Phi_{\lambda,i}}$, la i -ésima fila del gradiente de $\Phi_{\lambda,i}$ en \mathbf{x}_k está dada por

$$\nabla\Phi_{\lambda,i}(\mathbf{x}_k) = (a_i(\mathbf{x}_k) - 1)\nabla F_i(\mathbf{x}_k)^T + (b_i(\mathbf{x}_k) - 1)\nabla G_i(\mathbf{x}_k)^T. \quad (5.5)$$

Dado que el límite en (5.5) existe, $\nabla\Phi_{\lambda,i}$, F y G son continuamente diferenciables en \mathbf{x}_k y de (5.4), tenemos que

$$\begin{aligned} H &= \lim_{\mathbf{x}_k \rightarrow \mathbf{x}} [(a_i(\mathbf{x}_k) - 1)\nabla F_i(\mathbf{x}_k) + (b_i(\mathbf{x}_k) - 1)\nabla G_i(\mathbf{x}_k)] \\ &= (\xi_i - 1)\nabla F_i(\mathbf{x}) + (\chi_i - 1)\nabla G_i(\mathbf{x}), \end{aligned}$$

donde $\xi_i = \lim_{\mathbf{x}_k \rightarrow \mathbf{x}} a_i(\mathbf{x}_k)$ y $\chi_i = \lim_{\mathbf{x}_k \rightarrow \mathbf{x}} b_i(\mathbf{x}_k)$. Por otro lado, para cada k ,

$$(a_i(\mathbf{x}_k), b_i(\mathbf{x}_k))^T = \nabla f_\lambda((F_i(\mathbf{x}_k), G_i(\mathbf{x}_k))),$$

donde f_λ es la función definida por (4.7). Por el **Lema 5.1**, para cada $i = 1, \dots, n$ y para cada $k \in N$, $\|\nabla f_\lambda((F_i(\mathbf{x}_k), G_i(\mathbf{x}_k)))\|^2 \leq c_\lambda$; pasando al límite cuando $k \rightarrow \infty$,

$$\lim_{k \rightarrow \infty} \|\nabla f_\lambda((F_i(\mathbf{x}_k), G_i(\mathbf{x}_k)))\|^2 \leq c_\lambda.$$

Entonces $\|(\xi_i, \chi_i)\| \leq \sqrt{c_\lambda}$ y la conclusión del lema se sigue fácilmente. \square

El siguiente resultado generaliza la **Proposición 2.6** en [34], al caso donde F y G son funciones continuamente diferenciables. Además, garantiza que las componentes de la diagonal $a_i(\mathbf{x}) - 1$ y $b_i(\mathbf{x}) - 1$ definidas en la **Proposición 5.1** no son positivas (es decir, para aquellos índices tales que $(F_i(\mathbf{x}), G_i(\mathbf{x})) = (0, 0)$).

Proposición 5.2. *Cualquier $H \in \partial\Phi_\lambda(\mathbf{x})$ puede escribirse en la forma*

$$H = D_a(\mathbf{x})F'(\mathbf{x}) + D_b(\mathbf{x})G'(\mathbf{x}), \quad (5.6)$$

donde $D_a(\mathbf{x})$ y $D_b(\mathbf{x})$ son matrices diagonales semidefinidas negativas tales que $D_a(\mathbf{x}) + D_b(\mathbf{x})$ es definida negativa.

Demostración. Es análoga a la demostración de la **Proposición 2.6** en [34]. Sin embargo, es importante observar que en esta prueba usamos la **Proposición 5.1** que fue probada antes de usar el **Lema 5.2**. Primero veamos que la proposición se cumple para cualquier matriz del B-jacobiano generalizado de $\Phi_\lambda(\mathbf{x})$ en \mathbf{x} , $\partial_B\Phi_\lambda(\mathbf{x})$.

Sea $H \in \partial_B\Phi_\lambda(\mathbf{x})$. Existe una sucesión \mathbf{x}_k de puntos en D_{Φ_λ} que converge a \mathbf{x} , con $(F_i(\mathbf{x}_k)G_i(\mathbf{x}_k)) \neq 0$, para todo $i = 1, 2, \dots, n$ y todo $k \in \mathbb{N}$, tales que

$$H = \lim_{\mathbf{x}_k \rightarrow \mathbf{x}} \Phi'_\lambda(\mathbf{x}_k). \quad (5.7)$$

Ahora, por la **Proposición 5.1** y el **Lema 5.1**, para cada $k \in \mathbb{N}$ existen matrices D_a^k y D_b^k en $\mathbb{R}^{n \times n}$ tales que

$$\Phi'_\lambda(\mathbf{x}_k) = D_a^k F'(\mathbf{x}_k) + D_b^k G'(\mathbf{x}_k). \quad (5.8)$$

donde $D_a^k = D_a(\mathbf{x}_k)$ y $D_b^k = D_b(\mathbf{x}_k)$. Además, dado que

$$\|(a_i(\mathbf{x}_k), b_i(\mathbf{x}_k))\| \leq \sqrt{c_\lambda}, \quad c_\lambda \in (0, 2),$$

tenemos que $(a_i(\mathbf{x}_k) - 1, b_i(\mathbf{x}_k) - 1) \neq (0, 0)$ ya que, en caso contrario, tendríamos que $(a_i(\mathbf{x}_k), b_i(\mathbf{x}_k)) = (1, 1)$, con lo cual $\|(a_i(\mathbf{x}_k), b_i(\mathbf{x}_k))\| = \sqrt{2}$, lo que contradice que $\|(a_i(\mathbf{x}_k), b_i(\mathbf{x}_k))\| < \sqrt{2}$. Por lo tanto, $a_i(\mathbf{x}_k) - 1 \leq 0$ y $b_i(\mathbf{x}_k) - 1 \leq 0$, de donde se sigue que las matrices $D_a(\mathbf{x}_k)$ y $D_b(\mathbf{x}_k)$ son semidefinidas negativas y $D_a^k + D_b^k$ es definida negativa.

Por otro lado, como $D_a(\mathbf{x}_k)$ y $D_b(\mathbf{x}_k)$ están acotados superiormente para cada $k \in \mathbb{N}$, por el Teorema de *Bolzano-Weierstrass*, existen subsucesiones $\{D_a(\mathbf{x}_k)\}_K$ y $\{D_b(\mathbf{x}_k)\}_K$ que convergen a matrices diagonales $D_a(\mathbf{x})$ y $D_b(\mathbf{x})$, respectivamente, las cuales son semidefinidas negativas con $D_a(\mathbf{x}) + D_b(\mathbf{x})$ definida negativa.

De (5.7), (5.8) y pasando al límite cuando $k \rightarrow \infty$, obtenemos

$$\begin{aligned} H &= \lim_{\mathbf{x}_k \rightarrow \mathbf{x}} \Phi'_\lambda(\mathbf{x}_k) = D_a(\mathbf{x})F'(\mathbf{x}) + D_b(\mathbf{x})G'(\mathbf{x}) \\ &= \lim_{k \rightarrow \infty} \left[D_a(\mathbf{x}_k)F'(\mathbf{x}_k) + D_b(\mathbf{x}_k)G'(\mathbf{x}_k) \right], \end{aligned}$$

donde F y G son continuamente diferenciables en \mathbb{R}^n .

Finalmente, para $H \in \partial\Phi_\lambda(\mathbf{x})$ y por definición del jacobiano generalizado y su envolvente convexa, existen números $\lambda_i \geq 0$, con $\sum_{i=1}^m \lambda_i = 1$ tales que para algún $m > 0$ y algunas matrices $H_i \in \partial_B\Phi_\lambda(\mathbf{x})$,

$$H = \sum_{i=1}^m \lambda_i H_i.$$

Ahora, puesto que cada $H_i \in \partial_B\Phi_\lambda(\mathbf{x})$ es de la forma:

$$H_i = D_a(\mathbf{x}_i)F'(\mathbf{x}) + D_b(\mathbf{x}_i)G'(\mathbf{x}), i = 1, \dots, m,$$

donde $D_a(\mathbf{x}_i)$ y $D_b(\mathbf{x}_i)$ son matrices semidefinidas negativas tales que $D_a(\mathbf{x}_i) + D_b(\mathbf{x}_i)$ es definida negativa, tenemos que

$$\begin{aligned} H &= \sum_{i=1}^m \lambda_i H_i = \sum_{i=1}^m \lambda_i D_a(\mathbf{x}_i) + \sum_{i=1}^m \lambda_i D_b(\mathbf{x}_i) \\ &= \sum_{i=1}^m \lambda_i D_a(\mathbf{x}_i)F'(\mathbf{x}) + \sum_{i=1}^m \lambda_i D_b(\mathbf{x}_i)G'(\mathbf{x}) \\ &= D_a(\mathbf{x})F'(\mathbf{x}) + D_b(\mathbf{x})G'(\mathbf{x}), \end{aligned}$$

donde $D_a(\mathbf{x}) = \sum_{i=1}^m \lambda_i D_a(\mathbf{x}_i)$ y $D_b(\mathbf{x}) = \sum_{i=1}^m \lambda_i D_b(\mathbf{x}_i)$ son matrices semidefinidas negativas tales que $D_a(\mathbf{x}) + D_b(\mathbf{x})$ es definida negativa puesto que

$$D_a(\mathbf{x}) + D_b(\mathbf{x}) = \sum_{i=1}^m \lambda_i (D_a(\mathbf{x}_i) + D_b(\mathbf{x}_i)),$$

con lo cual concluimos la demostración. \square

La siguiente proposición da una nueva caracterización de la clase de P -matrices, en el contexto de complementariedad generalizada, la cual se reduce a la presentada en [34], para complementariedad no lineal, cuando G es la función identidad.

Proposición 5.3. Sean $D_a(\mathbf{x})$ y $D_b(\mathbf{x})$ matrices diagonales semidefinidas positivas (negativas) tales que $D_a(\mathbf{x}) + D_b(\mathbf{x})$ es definida positiva (negativa); M y $N \in \mathbb{R}^{n \times n}$, con M no singular. La matriz $D_a(\mathbf{x})M + D_b(\mathbf{x})N$ es no singular si, y solo si $NM^{-1} \in \mathbb{R}^{n \times n}$ es una P -matriz.

Demostración. Claramente,

$$D_a(\mathbf{x})M + D_b(\mathbf{x})N = (D_a(\mathbf{x}) + D_b(\mathbf{x})NM^{-1})M. \quad (5.9)$$

Suponemos que $D_a(\mathbf{x})M + D_b(\mathbf{x})N$ es no singular. Como M es no singular entonces de (5.9) tenemos que $D_a(\mathbf{x}) + D_b(\mathbf{x})NM^{-1}$ también es no singular y la **Proposición 4.1** garantiza que NM^{-1} es una P -matriz.

Si NM^{-1} es una P -matriz, la **Proposición 4.1** garantiza que la matriz $D_a(\mathbf{x}) + D_b(\mathbf{x})NM^{-1}$ es no singular. Usando este hecho en (5.9) y teniendo en cuenta que M es no singular, concluimos que $D_a(\mathbf{x})M + D_b(\mathbf{x})N$ también es no singular. \square

El siguiente teorema da una condición suficiente para garantizar la no singularidad de las matrices del jacobiano generalizado de Φ_λ en una solución \mathbf{x}_* del $\text{PCG}(F, G)$.

Teorema 5.1. *Si \mathbf{x}_* es una solución R-regular del $\text{PCG}(F, G)$ entonces todas las matrices del jacobiano generalizado $\partial\Phi_\lambda(\mathbf{x}_*)$ son no singulares.*

Demostración. Sea $H \in \partial\Phi_\lambda(\mathbf{x}_*)$ arbitraria pero fija. Por la **Proposición 5.1**, existen matrices diagonales $D_a(\mathbf{x}_*)$ y $D_b(\mathbf{x}_*)$ tales que

$$H = D_a(\mathbf{x}_*)F'(\mathbf{x}_*) + D_b(\mathbf{x}_*)G'(\mathbf{x}_*). \quad (5.10)$$

Como \mathbf{x}_* es una solución R-regular de $\text{PCG}(F, G)$ tenemos que $F'(\mathbf{x}_*)$ es no singular. Así, de (5.10) $\bar{H} = HF'(\mathbf{x}_*)^{-1} = D_a(\mathbf{x}_*) + D_b(\mathbf{x}_*)G'(\mathbf{x}_*)F'(\mathbf{x}_*)^{-1} = D_a(\mathbf{x}_*) + D_b(\mathbf{x}_*)K$, donde $K = G'(\mathbf{x}_*)F'(\mathbf{x}_*)^{-1}$. A partir de aquí, la prueba es esencialmente la misma que la del **Teorema 2.8** en [34]. \square

Corolario 5.2. *Si \mathbf{x}_* es una solución R-regular de $\text{PCG}(F, G)$ entonces \mathbf{x}_* también es una solución BD regular.*

Demostración. Sea \mathbf{x}_* una solución R-regular del $\text{PCG}(F, G)$. Entonces todas las matrices en $\partial\Phi_\lambda(\mathbf{x}_*)$ son no singulares. Puesto que $\partial_B\Phi_\lambda(\mathbf{x}_*) \subseteq \partial\Phi_\lambda(\mathbf{x}_*)$ entonces si $H \in \partial_B\Phi_\lambda(\mathbf{x}_*)$, H es no singular. Por lo tanto, \mathbf{x}_* es una solución BD-regular de $\text{PCG}(F, G)$. \square

El siguiente teorema garantiza la semisuavidad de Φ_λ y da una condición suficiente para su semisuavidad fuerte.

Teorema 5.2. *La función Φ_λ es semisuave. Más aún, si F y G son de clase LC^1 entonces Φ_λ es fuertemente semisuave.*

Demostración. La demostración es análoga a la dada en [34] en el contexto de complementariedad no lineal y es una consecuencia inmediata del hecho de que φ_λ es fuertemente semisuave. \square

Finalmente, presentamos el siguiente resultado que será útil en la siguiente sección.

Teorema 5.3. *La función Ψ_λ es continuamente diferenciable con $\nabla\Psi_\lambda(\mathbf{x}) = H^T\Phi_\lambda(\mathbf{x})$ para cualquier $H \in \partial\Phi_\lambda(\mathbf{x})$.*

Demostración. Usando la **Proposición 5.2** y propiedades de la familia de funciones de complementariedad φ_λ [34], la demostración es análoga a la prueba de la **Proposición 3.4** en [24]. \square

5.2. Algoritmo y resultados de convergencia

En esta sección, presentamos un algoritmo global, no suave tipo Newton para resolver el PCG(F, G) mediante su reformulación como el problema de minimización (1.7) y presentamos resultados de convergencia global y local de este algoritmo.

Algoritmo 1

Entrada: \mathbf{x}_0 y $\lambda \in (0, 4)$, $\rho > 0$, $\sigma \in (0, 1/2)$, $p > 2$, $\epsilon \geq 0$, $k = 0$.

Salida: Aproximación a la solución del sistema (1.6).

- 1: **mientras** $\|\nabla\Psi_\lambda(\mathbf{x}_k)\| \leq \epsilon$ **hacer**
- 2: Seleccionar un elemento H_k en $\partial_B\Phi_\lambda(\mathbf{x}_k)$. Encuentre una solución \mathbf{d}_k al sistema lineal \mathbf{d}_k :

$$H_k\mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k). \quad (5.11)$$

- 3: **si** el sistema (5.11) no es soluble o si \mathbf{d}_k no cumple la condición de descenso

$$\nabla\Psi_\lambda(\mathbf{x}_k)^T\mathbf{d}_k \leq -\rho\|\mathbf{d}_k\|^p \quad (5.12)$$

entonces

- 4: **hacer** $\mathbf{d}_k = -\nabla\Psi_\lambda(\mathbf{x}_k)$.
- 5: **sino**
- 6: Encontrar el menor $i_k \in \{0, 1, \dots\}$, tal que

$$\Psi_\lambda(\mathbf{x}_k + 2^{-i_k}\mathbf{d}_k) \leq \Psi_\lambda(\mathbf{x}_k) + \sigma 2^{-i_k}\nabla\Psi_\lambda(\mathbf{x}_k)^T\mathbf{d}_k \quad (5.13)$$

- 7: **Hacer**

$$\mathbf{x}_{k+1} = \mathbf{x}_k + t_k\mathbf{d}_k,$$

donde $t_k = 2^{-i_k}$.

- 8: **fin si**
 - 9: $k = k + 1$.
 - 10: **fin mientras**
 - 11: **devolver** \mathbf{x}_*
-

El algoritmo propuesto utiliza en cada iteración la matriz H definida en (5.6) que, a su vez, usa las matrices jacobianas de F y G y puede verse como una extensión del algoritmo propuesto en [22], a todos los miembros de la familia de funciones de complementariedad φ_λ definida en (4.7).

Es importante mencionar que nuestro algoritmo difiere del propuesto en [22] en dos aspectos. En primer lugar, el algoritmo que proponemos resuelve el problema de minimización con la familia uniparamétrica de funciones de complementariedad definidas en (1.5), mientras que el propuesto en [22] usa la función de *Fischer* como función de complementariedad, por lo que en cierto sentido, nuestro algoritmo puede verse como una generalización de aquel.

En segundo lugar, los dos algoritmos difieren en la búsqueda lineal, en el sentido que la del **Algoritmo 1** garantiza que la longitud del paso satisface la condición de *Armijo* [43] mientras que el otro, que la longitud del paso satisface las condiciones de *Goldstein* [43]; así que el **Algoritmo 1** es menos restrictivo que el propuesto en [22].

Previo a los resultados de convergencia global del **Algoritmo 1**, presentamos dos lemas que serán de utilidad para las pruebas de convergencia.

Teorema 5.4. *Cada punto de acumulación de la sucesión $\{\mathbf{x}_k\}$ generada por el Algoritmo 1 es un punto estacionario de Ψ_λ .*

Demostración. Sea \mathbf{x}_* un punto de acumulación de la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 1**. Si para un conjunto infinito de índices J , la dirección dada en el Paso 2 del algoritmo es $\mathbf{d}_k = -\nabla\Psi_\lambda(\mathbf{x}_k)$, para todo $k \in J$ entonces cualquier punto límite de la subsucesión $\{\mathbf{x}_k\}_J$ es un punto estacionario de Ψ_λ [6].

Asumamos, sin perder generalidad, que \mathbf{d}_k está dado por (5.11). Probaremos que para todo k , existen constantes positivas m y M tales que

$$m \leq \|\mathbf{d}_k\| \leq M. \quad (5.14)$$

Sea k cualquier índice de la sucesión generada por el **Algoritmo 1**. Como \mathbf{d}_k está dado por (5.11), tenemos $H_k \mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k)$ entonces

$$\|\Phi_\lambda(\mathbf{x}_k)\| \leq \|H_k\| \|\mathbf{d}_k\|, \quad (5.15)$$

para una norma vectorial $\|\cdot\|$. Claramente $\|H_k\| \neq 0$, en caso contrario, por **Teorema 5.3**, tendríamos que $\nabla\Psi_\lambda(\mathbf{x}_k) = 0$; es decir, \mathbf{x}_k sería un punto estacionario de Ψ_λ y el **Algoritmo 1** se detendría. Así, de (5.15), tenemos la desigualdad

$$\frac{\|\Phi_\lambda(\mathbf{x}_k)\|}{\|H_k\|} \leq \|\mathbf{d}_k\|. \quad (5.16)$$

Supongamos que \mathbf{x}_* no es un punto estacionario de Ψ_λ ($\nabla\Psi_\lambda(\mathbf{x}_*) \neq 0$). Si para un conjunto infinito de índices J , $\{\|\mathbf{d}_k\|\}_J \rightarrow 0$, cuando $k \rightarrow \infty$ entonces, dado que el jacobiano generalizado es compacto, H_k está acotado y por (5.15), obtenemos $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}_J \rightarrow 0$.

Ahora, por la continuidad de Φ_λ y tomando el límite cuando $k \rightarrow \infty$ in (5.16), obtenemos $\Phi_\lambda(\mathbf{x}_*) = 0$, por lo cual, $\nabla\Psi_\lambda(\mathbf{x}_*) = 0$, lo que contradice que \mathbf{x}_* no es un punto estacionario. De allí que $\{\|\mathbf{d}_k\|\}$ no puede converger a 0. Por lo tanto, existe $m > 0$ tal que $m \leq \|\mathbf{d}_k\|$, para todo k .

Por otro lado, si $\|\mathbf{d}_k\|$ no es acotada superiormente entonces, ya que $p > 2$ y $\nabla\Psi_\lambda(\mathbf{x}_k)$ está acotado, tenemos que

$$\lim_{k \rightarrow \infty} \frac{\|\nabla\Psi_\lambda(\mathbf{x}_k)\| \cos \theta}{\|\mathbf{d}_k\|^{p-1}} = 0,$$

donde θ es el ángulo entre $\nabla\Psi_\lambda(\mathbf{x}_k)$ y \mathbf{d}_k . Equivalentemente,

$$\lim_{k \rightarrow \infty} \frac{\nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k}{\|\mathbf{d}_k\|^p} = 0,$$

lo que contradice (5.12). Por lo cual, existe $M > 0$, tal que $\|\mathbf{d}_k\| \leq M$.

Finalmente, dado que (5.12) se satisface en cada iteración y Ψ_λ es continuamente diferenciable, $\{\Psi_\lambda(\mathbf{x}_{k+1}) - \Psi_\lambda(\mathbf{x}_k)\} \rightarrow 0$ cuando $k \rightarrow \infty$, lo que implica por (5.12), que $\{\sigma 2^{-i_k} \|\nabla\Psi_\lambda(\mathbf{x}_k)\| \|\mathbf{d}_k\|\} \rightarrow 0$, por lo tanto

$$\{2^{-i_k} \|\nabla\Psi_\lambda(\mathbf{x}_k)\| \|\mathbf{d}_k\|\} \rightarrow 0. \quad (5.17)$$

Ahora, demostraremos que 2^{-i_k} está acotado lejos de 0. Supongamos lo contrario. Tomando una subsucesión, de ser necesario, de tal manera que en cada iteración el tamaño del paso se reduzca al menos una vez, tenemos que (usando el mismo índice) $\{2^{-i_k}\} \rightarrow 0$ y puesto que si para i_k se satisface (5.13) entonces para $i_k - 1$ no, por lo cual

$$\frac{\Psi_\lambda(\mathbf{x}_k + 2^{-(i_k-1)} \mathbf{d}_k) - \Psi_\lambda(\mathbf{x}_k)}{2^{-(i_k-1)}} > \sigma \nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k. \quad (5.18)$$

De (5.14), podemos asumir que $\{\mathbf{d}_k\} \rightarrow \bar{\mathbf{d}} \neq 0$, y pasando al límite en (5.18), obtenemos

$$\nabla\Psi_\lambda(\mathbf{x}_*)^T \bar{\mathbf{d}} \geq \sigma \nabla\Psi_\lambda(\mathbf{x}_*)^T \bar{\mathbf{d}}. \quad (5.19)$$

Por (5.12), tenemos que $\nabla\Psi_\lambda(\mathbf{x}_*)^T \bar{\mathbf{d}} \leq -\rho \|\bar{\mathbf{d}}\|^p < 0$, lo que contradice (5.19) puesto que $\sigma \in (0, 1/2)$. Por lo tanto, podemos suponer que la sucesión $\{2^{-i_k}\}$ está acotada lejos de 0. Por otra parte, (5.12) y (5.17) implican que $\{\mathbf{d}_k\} \rightarrow 0$, lo que contradice (5.14). Por lo tanto, $\nabla\Psi_\lambda(\mathbf{x}_*) = 0$. Es decir, \mathbf{x}_* es un punto estacionario de Ψ_λ . \square .

Teorema 5.5. *Si \mathbf{x}_* es un punto de acumulación aislado de la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 1**. Entonces la sucesión converge a \mathbf{x}_* .*

Demostración. Sea $\{\mathbf{x}_k\}$ la sucesión generada por **Algoritmo 1** y \mathbf{x}_* un punto de acumulación aislado de la sucesión. Por **Teorema 5.4**, \mathbf{x}_* es un punto estacionario de la función convexa Ψ_λ , con lo cual \mathbf{x}_* es un minimizador global aislado de Ψ_λ .

Sea Ω el conjunto de puntos de acumulación de $\{\mathbf{x}_k\}$. Entonces $\Omega \neq \emptyset$ ya que $\mathbf{x}_* \in \Omega$. Definimos

$$\delta = \begin{cases} \text{dist}(\mathbf{x}_*; \Omega \setminus \{\mathbf{x}_*\}) & \text{si } \Omega \setminus \{\mathbf{x}_*\} \neq \emptyset \\ 1 & \text{si } \Omega = \{\mathbf{x}_*\}. \end{cases}$$

Observemos que, $\delta > 0$ puesto que \mathbf{x}_* es un punto de acumulación aislado. Ahora, si $\Omega_1 = \{\mathbf{y} \in \mathbb{R}^n : \text{dist}(\mathbf{y}; \Omega) \leq \delta/4\}$, entonces existe \bar{k} tal que $\mathbf{x}_k \in \Omega_1$ para todo $k \geq \bar{k}$. Si

$$K = \{k \in \mathbb{N} : \text{dist}(\mathbf{x}_k; \mathbf{x}_*) \leq \delta/4\}$$

(K es distinto de vacío, pues \mathbf{x}_* es un punto límite de la sucesión) entonces $\{\mathbf{x}_k\}_K \subset \overline{B}(\mathbf{x}_*; \delta/4)$.

Como todos los puntos de la subsucesión $\{\mathbf{x}_k\}_K$ pertenecen al conjunto compacto $\overline{B}(\mathbf{x}_*; \delta/4)$ y todos sus puntos límite también son puntos límite de $\{\mathbf{x}_k\}$, concluimos que $\{\mathbf{x}_k\}_K$ converge a \mathbf{x}_* . Además, dado que $\nabla \Psi_\lambda(\mathbf{x}_*) = 0$, el **Teorema 5.4** garantiza que $\{\|\nabla \Psi_\lambda(\mathbf{x}_k)\|\}_K$ converge a cero, lo que implica, por (5.12), que la sucesión $\{\mathbf{d}_k\}$ converge al vector cero. Por lo tanto, existe $k_1 \in \mathbb{N}$, tal que si $k \in K$ y $k \geq k_1 \geq \bar{k}$ entonces $\|\mathbf{d}_k\| \leq \epsilon$, para todo ϵ . En particular, para $\epsilon = \delta/4$ tenemos $\|\mathbf{d}_k\| \leq \delta/4$.

Sea $k_2 \in K$, tal que $k_2 \geq k_1$. Por el **Algoritmo 1**, tenemos que

$$\mathbf{x}_{k_2+1} = \mathbf{x}_{k_2} + t_{k_2} \mathbf{d}_{k_2}$$

con $t_{k_2} \in (0, 1]$, donde $\|\mathbf{x}_{k_2+1} - \mathbf{x}_{k_2}\| \leq \|\mathbf{d}_{k_2}\| \leq \delta/4$. Así,

$$\begin{aligned} \text{dist}(\mathbf{x}_*; \Omega \setminus \{\mathbf{x}_*\}) &\leq \text{dist}(\mathbf{x}_{k_2+1}; \Omega \setminus \{\mathbf{x}_*\}) + \|\mathbf{x}_* - \mathbf{x}_{k_2+1}\| \\ &\leq \text{dist}(\mathbf{x}_{k_2+1}; \Omega \setminus \{\mathbf{x}_*\}) + \|\mathbf{x}_* - \mathbf{x}_{k_2}\| + \|\mathbf{x}_{k_2} - \mathbf{x}_{k_2+1}\| \\ &\leq \text{dist}(\mathbf{x}_{k_2+1}; \Omega \setminus \{\mathbf{x}_*\}) + \delta/4 + \delta/4. \end{aligned}$$

Es decir, $\text{dist}(\mathbf{x}_{k_2+1}; \Omega \setminus \{\mathbf{x}_*\}) \geq \text{dist}(\mathbf{x}_*; \Omega \setminus \{\mathbf{x}_*\}) - \delta/2 \geq \delta - \delta/2 = \delta/2$, consecuentemente, $\mathbf{x}_{k_2+1} \notin \Omega_1 \setminus \overline{B}(\mathbf{x}_*; \delta/4)$ y puesto que $\mathbf{x}_{k_2+1} \in \Omega_1$ entonces $\mathbf{x}_{k_2+1} \in \overline{B}(\mathbf{x}_*; \delta/4)$; esto es, $k_2 + 1 \in K$ y puesto que $(k_2 + 1) > k_1$, usando inducción $k_2 \in K$, para todo $k_2 > k_1$, lo que implica que $\mathbf{x}_k \in \overline{B}(\mathbf{x}_*; \delta/4)$, para todo $k_2 > k_1$. Así, $\{\mathbf{x}_k\}$ converge a \mathbf{x}_* . \square

Teorema 5.6. *Sea \mathbf{x}_* un punto de acumulación de la sucesión $\{\mathbf{x}_k\}$ generada por **Algoritmo 1**, tal que \mathbf{x}_* es una solución R-regular de PCG. Entonces la sucesión $\{\mathbf{x}_k\}$ converge a \mathbf{x}_* , la dirección de búsqueda \mathbf{d}_k finalmente viene dado por la solución del sistema lineal (5.11) y se acepta el tamaño de paso completo $t_k = 1$, para todo k suficientemente grande.*

Demostración. Por hipótesis, \mathbf{x}_* es una solución R-regular de $\text{PCG}(F, G)$ y por el **Corolario 5.2**, también es una solución BD-regular entonces la **Proposición 3** en [50] garantiza que \mathbf{x}_* es un punto de acumulación aislado. Por lo tanto, la sucesión $\{\mathbf{x}_k\}$ converge a \mathbf{x}_* por el **Teorema 5.5**.

Dado que $\{\mathbf{x}_k\}$ converge a una solución BD-regular del sistema $\Phi_\lambda(\mathbf{x}) = 0$ por el **Lema 2.6** de [52], para k suficientemente grande, H_k es no singular, por lo que el sistema (5.11) siempre tiene una solución para k suficientemente grande. Veamos que satisface la desigualdad

$$\nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{s}_k \leq -\rho_1 \|\mathbf{d}_k\|^2, \quad (5.20)$$

para algún $\rho_1 > 0$.

De (5.11),

$$\|\mathbf{s}_k\| \leq \|H_k^{-1}\| \|\Phi_\lambda(\mathbf{x}_k)\|. \quad (5.21)$$

Además, dado que $\nabla\Psi_\lambda(\mathbf{x}_k) = H_k^T \Phi_\lambda(\mathbf{x}_k)$ y $H_k \mathbf{s}_k = -\Phi_\lambda(\mathbf{x}_k)$, obtenemos

$$\nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{s}_k = -\|\Phi_\lambda(\mathbf{x}_k)\|^2. \quad (5.22)$$

Combinando (5.21) y (5.22), tenemos que

$$\nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{s}_k \leq -\frac{\|\mathbf{d}_k\|^2}{N^2}, \quad (5.23)$$

donde N es una cota superior de $\|H_k^{-1}\|$ (que existe por **Lema 2.6** de [52]). Tomando $\rho_1 = 1/N^2$ en (5.23) obtenemos (5.20). Dado que la sucesión $\{\|\mathbf{s}_k\|\}$ converge a 0. De (5.20) se sigue que (5.12) se cumple para cualquier $p > 2$ y cualquier $\rho > 0$.

Finalmente, veamos que se alcanza el tamaño de paso completo; es decir, $i_k = 0$. Primero, Ψ_λ es de clase LC^1 según el **Teorema 2.3** en [34]. Con esta hipótesis y el **Teorema 3.2** en [23], se garantiza que a partir de un cierto k , $t_k = 1$ \square

A continuación, presentamos un resultado que garantiza la tasa de convergencia del algoritmo.

Teorema 5.7. *Si \mathbf{x}_* es un punto de acumulación de la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 1**, tal que \mathbf{x}_* es una solución R-regular del $\text{PCG}(F, G)$ entonces la sucesión*

$\{\mathbf{x}_k\}$ generada por el **Algoritmo 1** converge q -superlinealmente a \mathbf{x}_* . Además, si F y G son de clase LC^1 entonces la tasa de convergencia es q -cuadrática.

Demostración. Dado que \mathbf{x}_* es R -regular y por **Lema 2.6** de [52], H es no singular y $\|H_k^{-1}\| \leq N$, para alguna constante positiva N . Entonces $t_k = 1$ se acepta a partir de un cierto valor de k , la sucesión $\{\mathbf{x}_k\}$ está bien definida y

$$\mathbf{x}_{k+1} = \mathbf{x}_k - H_k^{-1}\Phi_\lambda(\mathbf{x}_k).$$

Restando \mathbf{x}_* y usando una norma $\|\cdot\|$, tenemos

$$\|\mathbf{x}_{k+1} - \mathbf{x}_*\| = \|\mathbf{x}_k - \mathbf{x}_* - H_k^{-1}\Phi_\lambda(\mathbf{x}_k)\|.$$

Después de algunas manipulaciones algebraicas, obtenemos

$$\begin{aligned} \|\mathbf{x}_{k+1} - \mathbf{x}_*\| &= \|\mathbf{x}_k - \mathbf{x}_* - H_k^{-1}\Phi_\lambda(\mathbf{x}_k)\| \\ &\leq \|H_k^{-1}[\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}_*) - \Phi'_\lambda(\mathbf{x}_*; \mathbf{x}_k - \mathbf{x}_*)]\| \\ &\quad + \|H_k^{-1}[H_k(\mathbf{x}_k - \mathbf{x}_*) - \Phi'_\lambda(\mathbf{x}_*; \mathbf{x}_k - \mathbf{x}_*)]\|. \end{aligned} \quad (5.24)$$

Para acotar los dos términos del lado derecho en (5.24), debemos tener en cuenta que Φ_λ es semisuave (**Teorema 5.2**) y, por lo tanto es direccionalmente diferenciable. Entonces, por (4.4), $\|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}_*) - \Phi'_\lambda(\mathbf{x}_*; \mathbf{x}_k - \mathbf{x}_*)\| = o\|\mathbf{x}_k - \mathbf{x}_*\|$, y por el **Lema 4.1**,

$$\|H_k(\mathbf{x}_k - \mathbf{x}_*) - \Phi'_\lambda(\mathbf{x}_*; \mathbf{x}_k - \mathbf{x}_*)\| = o\|\mathbf{x}_k - \mathbf{x}_*\|,$$

usando las dos últimas igualdades y puesto que $\|H_k^{-1}\| \leq N$ en (5.24), concluimos que

$$\|\mathbf{x}_{k+1} - \mathbf{x}_*\| = o\|\mathbf{x}_k - \mathbf{x}_*\|, \quad (5.25)$$

lo que muestra que la sucesión $\{\mathbf{x}_k\}$ converge q -superlinealmente a \mathbf{x}_* .

Ahora, supongamos que F y G son de clase LC^1 . Entonces, por el **Teorema 5.2**, también es fuertemente semisuave en \mathbf{x}_* . Además, por la **Definición 4.6**, la derivada direccional $\Phi'_\lambda(\cdot, \cdot)$ es *semicontinua de grado 2*. Entonces, de (4.5) y por el **Lema 4.2**, tenemos que

$$\|H_k(\mathbf{x}_k - \mathbf{x}_*) - \Phi'_\lambda(\mathbf{x}_*; \mathbf{x}_k - \mathbf{x}_*)\| = O\|\mathbf{x}_k - \mathbf{x}_*\|^2$$

y $\|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}_*) - \Phi'_\lambda(\mathbf{x}_*; \mathbf{x}_k - \mathbf{x}_*)\| = O\|\mathbf{x}_k - \mathbf{x}_*\|^2$. Así, por (5.24),

$$\|\mathbf{x}_{k+1} - \mathbf{x}_*\| = O\|\mathbf{x}_k - \mathbf{x}_*\|^2,$$

lo que muestra que la sucesión $\{\mathbf{x}_k\}$ converge q -cuadráticamente a \mathbf{x}_* . \square

5.3. Experimentación numérica

En esta sección, analizamos el desempeño numérico del **Algoritmo 1** propuesto en **Sección 5.2**, al que llamamos **Algoritmo NG**. Para ello, lo comparamos con el algoritmo propuesto en [22] que llamamos **Algoritmo NGS**. Además, incorporamos al **Algoritmo NG** el procedimiento dinámico para actualizar el parámetro λ propuesto en [34], obteniendo un nuevo algoritmo que llamamos **Algoritmo NGD**, el cual también usamos en la comparación.

Los parámetros usados en los algoritmos son: $\epsilon = 10^{-4}$, $N = 100$, $\rho = 10^{-8}$, $p = 2.1$ y $\sigma = 10^{-4}$. Hicimos cuatro tipos de experimentos. Primero, comparamos el desempeño de los **Algoritmos NGD** y **NGS** en términos del número de iteraciones. Segundo, analizamos el desempeño global de estos dos algoritmos. Tercero, analizamos el comportamiento del **Algoritmo NG** variando el parámetro λ . Finalmente, en el cuarto experimento, comparamos los **Algoritmos NG** y **NGD**.

5.3.1. Experimento 1

En este experimento, comparamos el número de iteraciones de los algoritmos **NGD** y **NGS**. En la Tabla 5.1, reportamos los resultados obtenidos al ejecutar los algoritmos para resolver cada uno de los problemas descritos en el **Capítulo 3**.

P	\mathbf{x}_0	NGD		NGS	
		k	Ψ	k	Ψ
1	\mathbf{x}_1	12	1.2×10^{-10}	12	3.8×10^{-8}
	\mathbf{x}_2	5	1.6×10^{-11}	5	5.8×10^{-10}
	\mathbf{x}_3	6	1.4×10^{-8}	6	1.8×10^{-7}
	\mathbf{x}_4	12	1.9×10^{-12}	–	–
2	\mathbf{x}_1	14	1.4×10^{-6}	13	2.8×10^{-6}
	\mathbf{x}_2	17	3.1×10^{-6}	17	1.5×10^{-6}
	\mathbf{x}_3	24	1.3×10^{-6}	23	2.6×10^{-6}
	\mathbf{x}_4	24	1.8×10^{-6}	23	3.6×10^{-6}
3	\mathbf{x}_1	5	1.1×10^{-6}	5	1.6×10^{-6}
	\mathbf{x}_2	5	1.5×10^{-6}	5	3.9×10^{-6}
	\mathbf{x}_3	5	2.1×10^{-10}	5	2.3×10^{-8}
4	\mathbf{x}_1	6	2.4×10^{-12}	6	7.5×10^{-7}
	\mathbf{x}_2	5	2.7×10^{-10}	5	1.1×10^{-6}
	\mathbf{x}_3	5	3.8×10^{-10}	5	1.7×10^{-6}

5	\mathbf{x}_1	49	4.2×10^{-6}	51	2.8×10^{-6}
	\mathbf{x}_2	29	3.4×10^{-6}	26	4.1×10^{-6}
	\mathbf{x}_3	39	2.8×10^{-6}	36	3.3×10^{-6}
6	\mathbf{x}_1	16	0.8×10^{-11}	7	0.5×10^{-9}
	\mathbf{x}_{13}	13	0.4×10^{-11}	6	0.1×10^{-9}
	$10\mathbf{x}_{13}$	14	0.6×10^{-12}	5	0.2×10^{-9}
	$20\mathbf{x}_{13}$	12	0.2×10^{-11}	6	0.2×10^{-9}
7	\mathbf{x}_{22}	13	0.1×10^{-13}	-	0.4×10^{-11}
	\mathbf{x}_{23}	11	0.3×10^{-11}	4	0.3×10^{-5}
	\mathbf{x}_{24}	14	0.7×10^{-12}	5	0.2×10^{-8}
	\mathbf{x}_{25}	15	0.1×10^{-12}	5	0.6×10^{-8}
	\mathbf{x}_{26}	9	0	-	0.7
8	\mathbf{x}_{27}	13	0.1×10^{-11}	34	0.2×10^{-5}
	\mathbf{x}_{28}	-	0.6×10^{-6}	32	0.3×10^{-5}
9	\mathbf{x}_1	9	1.4×10^{-6}	-	1.1×10^{-12}
	\mathbf{x}_2	13	0	-	0
	\mathbf{x}_3	22	4.8×10^{-6}	-	5×10^{-8}
	\mathbf{x}_4	18	9.4×10^{-7}	-	1.6×10^{-11}
10	\mathbf{x}_1	12	2.8×10^{-7}	-	1.1×10^{-12}
	\mathbf{x}_2	16	4.9×10^{-6}	-	3.7×10^{-12}
	\mathbf{x}_3	16	2.2×10^{-5}	-	3.6×10^{-6}
	\mathbf{x}_4	25	3.2×10^{-6}	-	1.6×10^{-6}

Tabla 5.1: Resultados del Experimento 1.

De la Tabla 5.1, observamos que los algoritmos tienen un comportamiento similar en cuanto al número de iteraciones. Sin embargo, cuando comparamos el valor de la función de mérito, $\Psi(\mathbf{x})$, en la iteración final, el **Algoritmo NGD** alcanzó, en 90 % de los experimentos, un valor inferior al alcanzado por el **Algoritmo NGS**, lo que sugiere que probablemente el **Algoritmo NGD** da una mejor aproximación a la solución de $\text{PCG}(F, G)$ en aproximadamente el mismo número de iteraciones que el **Algoritmo NGS**.

Por otro lado, este experimento nos muestra que la estrategia de variar λ en forma dinámica, implementada en **Algoritmo NGD**, da mejores resultados de convergencia cuando las iteraciones están cerca de la solución \mathbf{x}_* . Prueba de ello es la rápida disminución del valor de $\Psi(\mathbf{x})$ en las últimas iteraciones.

5.3.2. Experimento 2.

En este experimento, estudiamos el desempeño global de los **Algoritmos NGD** y **NGS**. Para ello, generamos cien puntos iniciales aleatorios para cada uno de los diez problemas. Para los **Problemas 1** a **4**, tomamos cada componente de los puntos iniciales en el intervalo $[-30, 30]$ y para el **Problema 5** en el intervalo $[1, 50]$. Los resultados de este experimento se resumen en la Tabla 5.2, en la que \bar{k} es el promedio de las iteraciones de los cien experimentos para cada problema (P).

P	NGD		NGS	
	\bar{k}	$S(\%)$	\bar{k}	$S(\%)$
1	10	96	9	87
2	14	98	14	98
3	6	100	6	97
4	6	100	6	100
5	34	100	35	100
6	65	44		
7	25	98		
8	38	74		
9	22	99		
10	38	74		

Tabla 5.2: Tasa de éxito para los Algoritmos NGD y NGS.

El **Algoritmo NGD** parece ser un algoritmo robusto. Queremos resaltar el desempeño del algoritmo que proponemos para el primer problema ya que este se considera un problema difícil [34]. En este problema, el **Algoritmo NGD** tuvo una tasa de éxito de 96 % mientras que su contraparte, **Algoritmo NGS**, tuvo una tasa de éxito de 87 %, por lo que pensamos que la estrategia de escoger λ en forma dinámica combinada con una búsqueda lineal menos restrictiva, como una búsqueda lineal tipo Armijo, funcionaría muy bien y podría conducir a un algoritmo más robusto y eficiente.

5.3.3. Experimento 3

Los experimentos anteriores muestran que la estrategia de escoger λ en forma dinámica es una buena opción para complementar el **Algoritmo NG**, no obstante una pregunta razonable es la siguiente, ¿cuál es el rendimiento de este algoritmo sin la estrategia de escoger λ en forma dinámica ? es decir, ¿cuál es el comportamiento

del algoritmo con el mismo valor de λ durante toda la ejecución? y ¿cuál es la mejor opción de λ para ejecutar el algoritmo? En este experimento, buscamos la respuesta a estos interrogantes. Para ello, ejecutamos el algoritmo con diferentes valores de λ .

En la Figura 5.1, mostramos el número de iteraciones requeridas por el **Algoritmo NG** para resolver cada uno de los diez problemas cuando se ejecutó con 38 valores de λ igualmente espaciados en el intervalo $(0, 4)$.

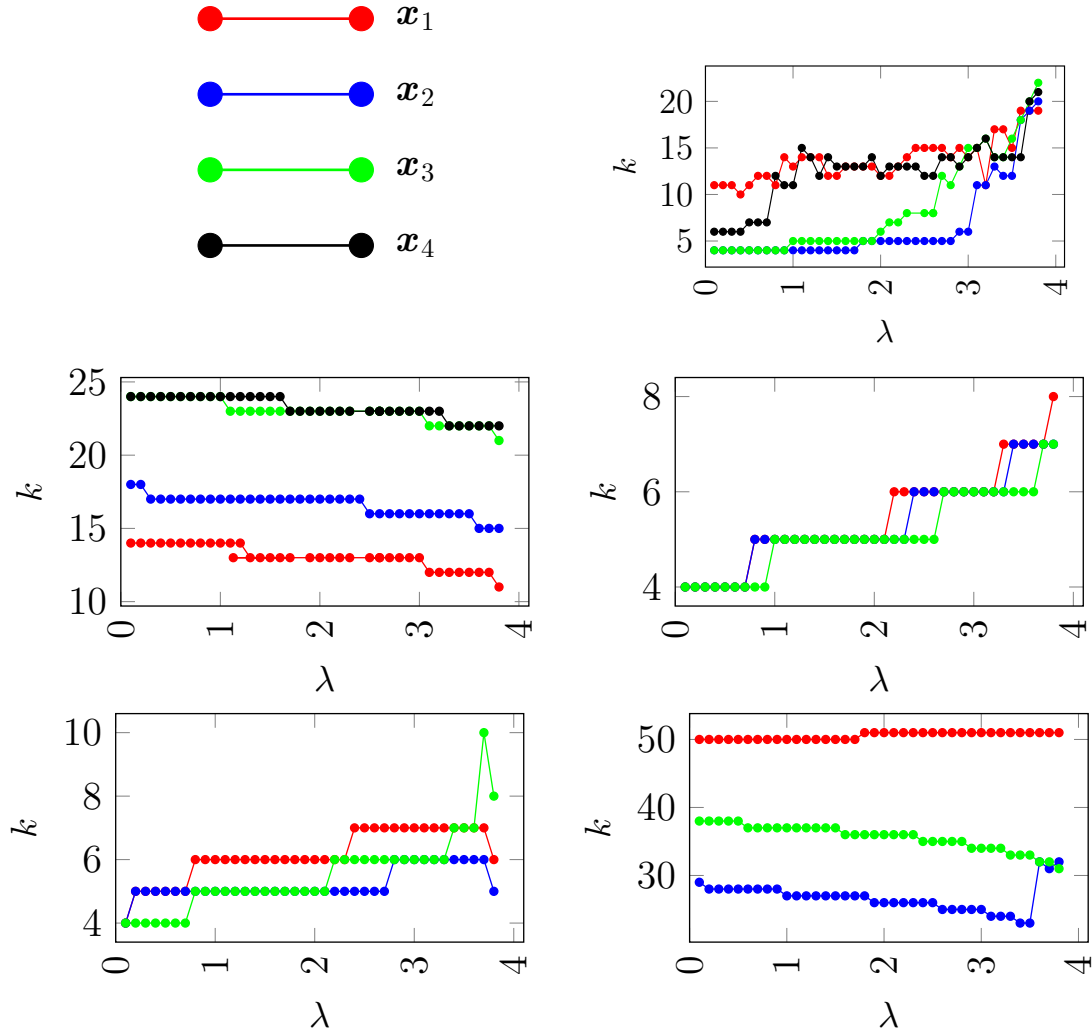


Figura 5.1: Desempeño del Algoritmo NG para diferentes valores de λ .

Los resultados de la Figura 5.1, muestran que el **Algoritmo NG** convergió en todos los experimentos y para todos los valores de λ , esto refleja en cierto sentido la robustez del algoritmo. Por otro lado, podemos ver que el **Algoritmo NG** tuvo su mejor desempeño para valores de λ al inicio o al final del intervalo $(0, 4)$.

Comparando los resultados de la Figura 5.1 con los de la Tabla 5.1 podemos concluir que en todos los experimentos es posible encontrar un valor de λ para el cual el **Algoritmo NG** es más rápido que **Algoritmo NGS**, sin embargo, no existe un valor único de λ que funcione bien para todos los problemas; por lo tanto, la estrategia de escoger λ en forma dinámica parece ser la mejor opción para complementar el **Algoritmo NG**.

5.3.4. Experimento 4.

En este experimento, resumimos los resultados obtenidos en los **Experimentos 1 y 3** para los **Algoritmos NGD y NG**, respectivamente. Esto con el objetivo de comparar el mejor y peor desempeño del **Algoritmo NG** con los respectivos del **Algoritmo NGD**.

P	\mathbf{x}_0	NG					NGD
		λ_B	λ_W	k_B	k_W	\bar{k}	k
1	\mathbf{x}_1	(0, 0.4)	(3.5, 3.8)	11	19	15	12
	\mathbf{x}_2	(0, 1.7)	(3.7, 3.8)	4	20	12	5
	\mathbf{x}_3	(0, 0.9)	(3.7, 3.8)	4	22	13	6
	\mathbf{x}_4	(0, 0.4)	(3.7, 3.8)	6	21	13.5	12
2	\mathbf{x}_1	(3.7, 3.8)	(0, 1.2)	11	14	12.5	9
	\mathbf{x}_2	(3.53, 3.8)	(0, 0.2)	15	17	16	12
	\mathbf{x}_3	(3.7, 3.8)	(0, 1.0)	21	24	22.5	19
	\mathbf{x}_4	(3.2, 3.8)	(0, 1.6)	22	24	23	19
3	\mathbf{x}_1	(0, 0.7)	(3.7, 3.8)	4	8	6	5
	\mathbf{x}_2	(0, 0.7)	(3.3, 3.8)	4	7	5.5	5
	\mathbf{x}_3	(0, 0.9)	(3.6, 3.7)	4	7	5.5	5
4	\mathbf{x}_1	(0, 0.1)	(2.3, 3.7)	4	7	5.5	6
	\mathbf{x}_2	(0, 0.1)	(2.7, 3.7)	4	6	5	5
	\mathbf{x}_3	(0, 0.7)	(3.6, 3.7)	4	10	7	5
5	\mathbf{x}_1	(0, 1.7)	(1.7, 3.8)	50	51	50.5	49
	\mathbf{x}_2	(3.3, 3.5)	(0, 0.9)	23	28	25.5	29
	\mathbf{x}_3	(3.6, 3.8)	(0, 0.5)	32	38	35	39

Tabla 5.3: Desempeño de los Algoritmos **NG** y **NGD**.

Los resultados se muestran en la Tabla 5.3, donde λ_B indica el subintervalo de valores de λ para los cuales, el **Algoritmo NG** muestra su mejor desempeño y λ_W indica el subintervalo de valores de λ para los cuales obtuvo su peor desempeño. Análogamente, k_B y k_W son las iteraciones para el **Algoritmo NG** en estos subintervalos, respectivamente.

En esta tabla, podemos observar que para todos los problemas, excepto para el **Problema 2** es posible encontrar un valor de λ para el cual el **Algoritmo NG** resolvió el correspondiente $\text{PCG}(F, G)$ en menos iteraciones que el **Algoritmo NGD**. Sin embargo, no hay un único valor de λ para el cual, el **Algoritmo NG** funcione mejor que **Algoritmo NGD**, en todos los experimentos.

Por otro lado, observamos que, en el peor de los casos, el **Algoritmo NG** necesitó un número considerablemente mayor de iteraciones para resolver el $\text{PCG}(F, G)$ que el **Algoritmo NGD**, lo que aumenta significativamente su promedio de iteraciones .

Con las observaciones anteriores, podemos concluir que cuando no tenemos disponible el mejor valor de λ para el **Algoritmo NG**, la estrategia de escoger λ en forma dinámica es una buena opción. Sin embargo, hicimos algunas pruebas preliminares variando el λ inicial con la estrategia propuesta en [34] y los resultados mostraron que al hacer algunos cambios en el valor inicial de λ , el **Algoritmo NGD** puede tener un mejor rendimiento, por lo que creemos que este es un problema abierto en esta área.

Capítulo 6

Métodos secantes de cambio mínimo

En este capítulo, presentamos una familia de *métodos secantes de cambio mínimo* para resolver el sistema de ecuaciones no lineales (1.6) y desarrollamos su respectiva teoría de convergencia. Iniciamos proponiendo un algoritmo cuasi-Newton genérico, en el que la matriz B_k es una aproximación a una matriz del jacobiano generalizado de $\Phi_\lambda(\mathbf{x}_k)$. En este algoritmo, no se especifica cómo se actualizan las matrices B_k , por ello, decimos que es genérico. En la segunda parte del capítulo, presentaremos una propuesta de actualización.

Siguiendo la filosofía de los métodos cuasi-Newton, aproximamos las matrices H_k del jacobiano generalizado por matrices B_k . Proponemos aproximaciones de la siguiente forma:

$$B_k = \begin{pmatrix} [B_k]_1 \\ \vdots \\ [B_k]_n \end{pmatrix} \quad (6.1)$$

con

$$[B_k]_i = \begin{cases} (a_i(\mathbf{x}) - 1) [A_k]_i + (b_i(\mathbf{x}) - 1) [C_k]_i, & i \notin \beta \\ (\chi_i - 1) [A_k]_i + (\xi_i - 1) [C_k]_i, & i \in \beta, \end{cases} \quad (6.2)$$

donde las matrices

$$A_k = \begin{pmatrix} [A_k]_1 \\ \vdots \\ [A_k]_n \end{pmatrix} \quad \text{y} \quad C_k = \begin{pmatrix} [C_k]_1 \\ \vdots \\ [C_k]_n \end{pmatrix}$$

son aproximaciones de las matrices jacobianas $F'(\mathbf{x}_k)$ y $G'(\mathbf{x}_k)$, respectivamente. Los términos $a_i(\mathbf{x})$, $b_i(\mathbf{x})$, χ_i y ξ_i están definidos en (5.1) y β está definido en (4.6).

Es importante mencionar que las matrices B_k conservan la estructura e información de las H_k dada en la **Proposición 5.2**. En forma compacta, las matrices B_k pueden verse de la siguiente forma:

$$B_k = D_a(\mathbf{x}_k)A_k + D_b(\mathbf{x}_k)C_k. \quad (6.3)$$

Desde este punto de vista, esta aproximación es estructurada ya que mantiene la estructura de las matrices H_k , utilizando parte de la información previamente calculada.

Los algoritmos y resultados de convergencia contenidos en este capítulo pueden verse como una generalización, al problema de complementariedad generalizado, de los presentados en [39] para complementariedad no lineal. En este sentido, veremos que las demostraciones son análogas a aquellas realizadas en ese documento.

El algoritmo cuasi-Newton genérico que proponemos es el siguiente.

Algoritmo 2 Algoritmo cuasi-Newton genérico.

Entrada: $\mathbf{x}_0 \in \mathbb{R}^n$ y $\lambda \in (0, 4)$, $A_0, C_0 \in \mathbb{R}^{n \times n}$, $\epsilon > 0$ y $k = 0$.

Salida: Aproximación a la solución del sistema (1.6).

- 1: **mientras** $\|\Phi_\lambda(\mathbf{x}_k)\| > \epsilon$ **hacer**
- 2: Calcule B_k como en (6.3), una aproximación a H_k . Encuentre una solución \mathbf{d}_k al sistema lineal

$$B_k \mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k).$$

- 3: Hacer

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k,$$

- 4: **fin mientras**

- 5: Actualizar B_k como en (6.1) y tal que B_{k+1} sea no singular. Volver al paso 1.

- 6: **devolver** \mathbf{x}_*
-

Algunos comentarios sobre el algoritmo son los siguientes.

Si $\lambda = 2$ o λ tiende a 0, y G es la función identidad, el **Algoritmo 2**, se reduce a los presentados en [39] para el problema de complementariedad no lineal.

Si G es la función identidad, el **Algoritmo 2** se reduce a los presentados en [2] para el problema de complementariedad no lineal.

Cuando λ tiende a 0, el **Algoritmo 2** se reduce al presentado en [31], para el problema de complementariedad generalizada.

6.1. Hipótesis y resultados de convergencia

Bajo las siguientes hipótesis demostramos que la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 2** está bien definido y converge linealmente a una solución del sistema $\Phi_\lambda(\mathbf{x}) = 0$.

H1. Existe $\mathbf{x}_* \in \mathbb{R}^n$, tal que $\Phi_\lambda(\mathbf{x}_*) = \mathbf{0}$.

H2. Las matrices jacobianas de F y G son *Lipschitz continuas* (sin pérdida de generalidad suponemos que la constante es la misma) en una vecindad de $\mathbf{x}_* \in \mathbb{R}^n$. Es decir, existen constantes $\gamma > 0$ y $\delta > 0$ tales que

$$\|F'(\mathbf{x}) - F'(\mathbf{x}_*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}_*\| \quad \text{y} \quad \|G'(\mathbf{x}) - G'(\mathbf{x}_*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}_*\|,$$

para todo $\mathbf{x} \in \mathcal{B}(\mathbf{x}_*; \delta)$, donde $\|\cdot\|$ denota cualquier norma en \mathbb{R}^n y también su respectiva norma matricial inducida.

H3. \mathbf{x}_* es una solución R-regular del PCG(F, G).

A partir del Teorema 5.1, el **Lema 2.6** de [52] y la hipótesis **H3** podemos deducir que las matrices del jacobiano generalizado son no singulares y sus inversas son acotadas; es decir, para cada $H_* \in \partial\Phi_\lambda(\mathbf{x}_*)$, existe una constante positiva μ , tal que

$$\|H_*^{-1}\| \leq \mu. \quad (6.4)$$

Sean $H \in \partial\Phi_\lambda(\mathbf{x})$ y $H_* \in \partial\Phi_\lambda(\mathbf{x}_*)$, con H y H_* definidas en (5.6) y B en (6.1).

Los dos lemas siguientes garantizan que para cada \mathbf{x} suficientemente cerca de la solución del PCG(F, G), $\|H - B\|_\infty$ y $\|H - H_*\|_\infty$ están acotadas superiormente.

Lema 6.1. Sean $F, G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ cuyas matrices jacobianas satisfacen **H2**, $\epsilon, \delta_F, \delta_G$, constantes positivas; H y B definidas en (5.6) y (6.1), respectivamente. Entonces, para cada $\mathbf{x} \in \mathcal{B}(\mathbf{x}_*; \epsilon)$, $A \in \mathcal{B}(F'(\mathbf{x}_*); \delta_F)$ y $C \in \mathcal{B}(G'(\mathbf{x}_*); \delta_G)$, existe una constante $\theta > 0$ tal que

$$\|H - B\|_\infty \leq \theta. \quad (6.5)$$

Demostración. Por la **Definición 4.1**, tenemos que

$$\|H - B\|_\infty = \max_{1 \leq i \leq n} \{ \| [H]_i - [B]_i \|_1 \}.$$

De (5.6) y (6.1),

$$[H - B]_i = \begin{cases} (a_i - 1)(\nabla F_i(\mathbf{x})^T - [A]_i) + (b_i - 1)(\nabla G_i(\mathbf{x})^T - [C]_i), & i \notin \beta \\ (\chi_i - 1)(\nabla F_i(\mathbf{x})^T - [A]_i) + (\xi_i - 1)(\nabla G_i(\mathbf{x})^T - [C]_i), & i \in \beta, \end{cases}$$

Supongamos que el máximo se alcanza en la fila j . Consideremos dos casos.

1. Si $j \notin \beta$, donde β esta definido en (4.6), entonces

$$\begin{aligned}
\|H - B\|_\infty &= \left\| [H]_j - [B]_j \right\|_1 \\
&= \left\| (a_j - 1)(\nabla F_j(\mathbf{x})^T - [A]_j) + (b_j - 1)(\nabla G_j(\mathbf{x})^T - [C]_j) \right\|_1 \\
&\leq |a_j - 1| \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 + |b_j - 1| \left\| \nabla G_j(\mathbf{x})^T - [C]_j \right\|_1 \\
&\leq (|a_j| + 1) \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 + (|b_j| + 1) \left\| \nabla G_j(\mathbf{x})^T - [C]_j \right\|_1.
\end{aligned}$$

Del **Lema 5.1** y la **Proposición 5.1**, $|a_j| < \sqrt{2}$ y $|b_j| < \sqrt{2}$, por lo cual

$$\begin{aligned}
\|H - B\|_\infty &\leq (1 + \sqrt{2}) \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 + \\
&\quad (1 + \sqrt{2}) \left\| \nabla G_j(\mathbf{x})^T - [C]_j \right\|_1.
\end{aligned} \tag{6.6}$$

Por otra parte, usando equivalencia entre las normas $\|\cdot\|_\infty$ y $\|\cdot\|_1$, la hipótesis **H2** y el hecho de que $A \in \mathcal{B}(F'(\mathbf{x}_*); \delta_F)$, obtenemos que

$$\begin{aligned}
\left\| \nabla F_j(\mathbf{x}) - [A]_j \right\|_1 &\leq n \left\| \nabla F_j(\mathbf{x}) - [A]_j \right\|_\infty \\
&\leq n \left(\left\| \nabla F_j(\mathbf{x}) - \nabla F_j(\mathbf{x}_*) \right\|_\infty + \left\| \nabla F_j(\mathbf{x}_*) - [A]_j \right\|_\infty \right) \\
&\leq n \left(\gamma \|\mathbf{x} - \mathbf{x}_*\|_\infty + \left\| \nabla F_j(\mathbf{x}_*) - [A]_j \right\|_\infty \right) \\
&\leq n (\gamma \|\mathbf{x} - \mathbf{x}_*\|_\infty + \delta_F) \\
&\leq n (\epsilon\gamma + \delta_F),
\end{aligned}$$

de donde

$$\left\| \nabla F_j(\mathbf{x}) - [A]_j \right\|_1 \leq n(\epsilon\gamma + \delta_F).$$

Análogamente, puesto que $C \in \mathcal{B}(F'(\mathbf{x}_*); \delta_G)$, obtenemos

$$\left\| \nabla G_j(\mathbf{x}) - [C]_j \right\|_1 \leq n(\epsilon\gamma + \delta_G).$$

Por lo tanto, de (6.6) se tiene que

$$\|H - B\|_\infty \leq 2(1 + \sqrt{2})n(\epsilon\gamma + \delta) =: \theta_1, \tag{6.7}$$

donde $\delta = \max\{\delta_F, \delta_G\}$.

2. Si $j \in \beta$,

$$\begin{aligned}
\|H - B\|_\infty &= \left\| [H]_j - [B]_j \right\|_1 \\
&= \left\| (\chi_j - 1)(\nabla F_j(\mathbf{x})^T - [A]_j) + (\xi_j - 1)(\nabla G_j(\mathbf{x})^T - [C]_j) \right\|_1 \\
&\leq |\chi_j - 1| \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 + |\xi_j - 1| \left\| \nabla G_j(\mathbf{x})^T - [C]_j \right\|_1 \\
&\leq (|\chi_j| + 1) \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 + (|\xi_j| + 1) \left\| \nabla G_j(\mathbf{x})^T - [C]_j \right\|_1.
\end{aligned}$$

Ahora, por la **Proposición 5.1** y por equivalencia entre las normas $\|\cdot\|_2$ y $\|\cdot\|_1$, se tiene que $\|(\chi, \xi)\|_1 \leq \sqrt{2} \|(\chi, \xi)\|_2 \leq \sqrt{2}\sqrt{c\lambda} < 2$ de donde $|\chi_j| < 2$ y $|\xi_j| < 2$. Procediendo como en el caso anterior, podemos ver que

$$\|H - B\|_\infty \leq 6n(\epsilon\gamma + \delta) : = \theta_2, \quad (6.8)$$

donde $\delta = \max\{\delta_F, \delta_G\}$. Finalmente, de (6.7) y (6.8), puesto que $\theta_1 < \theta_2$ y definiendo $\theta = \theta_2 = 6n(\epsilon\gamma + \delta)$, concluimos que

$$\|H - B\|_\infty \leq \theta. \quad (6.9)$$

□

Lema 6.2. Sea $\mathbf{x} \in \mathbb{R}^n$. Si $\mathbf{x}_* \in \mathbb{R}^n$ satisface la hipótesis **H1**, F' y G' satisfacen **H2** entonces para cada $H \in \partial\Phi_\lambda(\mathbf{x})$ y $H_* \in \partial\Phi_\lambda(\mathbf{x}_*)$ existe una constante positiva $\alpha > 0$, tal que si $\|\mathbf{x} - \mathbf{x}_*\|_\infty < \epsilon$ entonces $\|H - H_*\|_\infty < \alpha$.

Demostración. Supongamos que $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}_* \in \mathbb{R}^n$ satisface **H1** y consideremos la matriz $H - H_*$. Sumando y restando $D_a F'(\mathbf{x}_*) + D_b G'(\mathbf{x}_*)$ obtenemos

$$\begin{aligned}
H - H_* &= (D_a - I)F'(\mathbf{x}) + (D_b - I)G'(\mathbf{x}) - (D_{\hat{a}} - I)F'(\mathbf{x}_*) - (D_{\hat{b}} - I)G'(\mathbf{x}_*) \\
&= -[F'(\mathbf{x}) - F'(\mathbf{x}_*) + G'(\mathbf{x}) - G'(\mathbf{x}_*)] + D_a[F'(\mathbf{x}) - F'(\mathbf{x}_*)] + D_a F'(\mathbf{x}_*) + \\
&\quad + D_b[G'(\mathbf{x}) - G'(\mathbf{x}_*)] + D_b G'(\mathbf{x}_*) - D_{\hat{a}} F'(\mathbf{x}_*) - D_{\hat{b}} G'(\mathbf{x}_*).
\end{aligned}$$

Dado que F' y G' satisfacen **H2**, por el **Lema 5.1** y la **Proposición 5.1**, tenemos que

$$\begin{aligned}
\|H - H_*\|_\infty &\leq 2\gamma \|\mathbf{x} - \mathbf{x}_*\|_\infty + 2\sqrt{2}\gamma \|\mathbf{x} - \mathbf{x}_*\|_\infty + 4\sqrt{2}M \\
&\leq (2 + 2\sqrt{2})\gamma\epsilon + 4\sqrt{2}M : = \hat{\theta},
\end{aligned}$$

donde $M = \max\{\|F'(\mathbf{x}_*)\|_\infty, \|G'(\mathbf{x}_*)\|_\infty\}$. Luego, existe $\hat{\theta} : = (2 + 2\sqrt{2})\gamma\epsilon + 4\sqrt{2}M$,

tal que

$$\|H - H_*\|_\infty < \hat{\theta}, \quad (6.10)$$

con lo cual queda demostrado el lema. \square

Observemos que, en realidad, θ es una función que depende de ϵ y δ y $\hat{\theta}$ de ϵ ; es decir, $\theta = \theta(\epsilon, \delta)$ y $\hat{\theta} = \hat{\theta}(\epsilon)$.

El siguiente lema será útil en la demostración del **Teorema 6.1**, el cual establece condiciones suficientes para la convergencia lineal del **Algoritmo 2**.

Lema 6.3. Sean $r \in (0, 1)$, $\mathbf{x} \in \mathbb{R}^n$ y para cada A y C en $\mathbb{R}^{n \times n}$, definamos

$$\mathcal{Q}(\mathbf{x}, A, C) = \mathbf{x} - B^{-1}\Phi_\lambda(\mathbf{x}), \quad (6.11)$$

donde la matriz B está definida por (6.1). Si \mathbf{x}_* satisface la hipótesis **H3** entonces existen constantes $\epsilon, \delta > 0$ tales que,

$$\text{si } \|\mathbf{x} - \mathbf{x}_*\|_\infty \leq \epsilon, \quad \|C - G'(\mathbf{x}_*)\|_\infty \leq \delta \quad \text{y} \quad \|A - F'(\mathbf{x}_*)\|_\infty \leq \delta,$$

la función \mathcal{Q} está bien definida y

$$\|\mathcal{Q}(\mathbf{x}, A, C) - \mathbf{x}_*\|_\infty \leq r \|\mathbf{x} - \mathbf{x}_*\|_\infty. \quad (6.12)$$

Demostración. Sean $r \in (0, 1)$, $\epsilon_1 < \frac{r}{96n\mu\gamma}$ y $\delta < \frac{r}{96n\mu}$, donde γ es la constante de Lipschitz y μ es la constante en (6.4).

Para probar que \mathcal{Q} está bien definida se debe garantizar que B^{-1} existe. Para ello, consideremos la desigualdad

$$\|B - H_*\|_\infty \leq \|B - H\|_\infty + \|H - H_*\|_\infty. \quad (6.13)$$

Si $\|\mathbf{x} - \mathbf{x}_*\|_\infty < \epsilon_1$, $\|C - G'(\mathbf{x}_*)\|_\infty \leq \delta$ y $\|A - F'(\mathbf{x}_*)\|_\infty \leq \delta$, dado que el primer sumando en el lado derecho de (6.13) está acotado por $\theta(\epsilon_1, \delta)$ en (6.9), tenemos que

$$\|B - H\|_\infty \leq \theta(\epsilon_1, \delta) = 6n(\epsilon_1\gamma + \delta) < 6n \left(\frac{r\gamma}{96n\mu\gamma} + \frac{r}{96n\mu} \right) = \frac{r}{8\mu}.$$

Ahora, puesto que el segundo sumando de (6.13) está acotado por $\hat{\theta}(\epsilon_2)$ en (6.10), tomando

$$\epsilon_2 < \frac{\frac{r}{8(2+2\sqrt{2})\mu} - 4\sqrt{2M}}{\gamma},$$

donde γ es la constante de Lipschitz y μ es la constante en (6.4), obtenemos que

$$\begin{aligned} \|H - H_*\|_\infty &\leq \hat{\theta}(\epsilon_2) = (2 + 2\sqrt{2})\gamma\epsilon_2 + 4\sqrt{2}M \\ &< (2 + 2\sqrt{2}) \left[\gamma \left(\frac{r}{8(2 + 2\sqrt{2})\mu} - 4\sqrt{2}M \right) + 4\sqrt{2}M \right] \\ &= \frac{r}{8\mu}. \end{aligned}$$

En consecuencia, si tomamos $\epsilon := \min\{\epsilon_1, \epsilon_2\}$, podemos ver que

$$\|B - H\|_\infty < \frac{r}{8\mu} \quad \text{y} \quad \|H - H_*\|_\infty < \frac{r}{8\mu}, \quad (6.14)$$

de donde, sustituyendo (6.14) en (6.13), obtenemos que

$$\begin{aligned} \|B - H_*\|_\infty &\leq \|B - H\|_\infty + \|H - H_*\|_\infty \\ &< \frac{r}{8\mu} + \frac{r}{8\mu} = \frac{r}{4\mu}. \end{aligned}$$

Por lo tanto, de (6.4) se tiene que

$$\|B - H_*\|_\infty \leq \frac{r}{4\mu} < \frac{1}{4\|H_*^{-1}\|_\infty}.$$

Luego

$$\|H_*^{-1}B - I_n\|_\infty = \|H_*^{-1}(B - H_*)\|_\infty \leq \|H_*^{-1}\|_\infty \|B - H_*\|_\infty < \frac{1}{4},$$

por lo cual, $\|H_*^{-1}B - I_n\|_\infty < 1$ y por el **Lema 4.4**, la función \mathcal{Q} está bien definida. Más aún,

$$\|B^{-1}\| \leq \frac{\|H^{-1}\|}{1 - \|I_n - H^{-1}B\|} < \frac{\mu}{1 - \frac{1}{4}} = \frac{4}{3}\mu. \quad (6.15)$$

Ahora, para probar (6.12), sustraemos \mathbf{x}_* en (6.11), aplicamos $\|\cdot\|_\infty$ y realizamos algunas manipulaciones algebraicas. En efecto,

$$\begin{aligned}
\|\mathcal{Q}(\mathbf{x}, A, C) - \mathbf{x}_*\|_\infty &= \|\mathbf{x} - \mathbf{x}_* - B^{-1}\Phi_\lambda(\mathbf{x})\|_\infty \\
&= \|(\mathbf{x} - \mathbf{x}_*) - B^{-1}\Phi_\lambda(\mathbf{x}) + B^{-1}H_*(\mathbf{x} - \mathbf{x}_*) - B^{-1}H_*(\mathbf{x} - \mathbf{x}_*)\|_\infty \\
&= \|B^{-1}(B - H_*)(\mathbf{x} - \mathbf{x}_*) - B^{-1}[\Phi_\lambda(\mathbf{x}) + H_*(\mathbf{x} - \mathbf{x}_*)]\|_\infty \\
&\leq \|B^{-1}\|_\infty \|(B - H_*)(\mathbf{x} - \mathbf{x}_*) - [\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}_*) + H_*(\mathbf{x} - \mathbf{x}_*)]\|_\infty \\
&\leq \frac{4}{3}\mu \left[\|B - H_*\|_\infty \|\mathbf{x} - \mathbf{x}_*\|_\infty + \|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}_*) + H_*(\mathbf{x} - \mathbf{x}_*)\|_\infty \right].
\end{aligned}$$

De donde obtenemos la desigualdad,

$$\|\mathcal{Q}(\mathbf{x}, A, C) - \mathbf{x}_*\|_\infty \leq \frac{4}{3}\mu \left[\frac{r}{4\mu} + \frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}_*) + H_*(\mathbf{x} - \mathbf{x}_*)\|_\infty}{\|\mathbf{x} - \mathbf{x}_*\|_\infty} \right] \|\mathbf{x} - \mathbf{x}_*\|_\infty. \quad (6.16)$$

En [34], los autores demuestran que $\Phi_\lambda(\mathbf{x})$ es semisuave, por lo cual [51],

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_*} \frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}_*) + H(\mathbf{x} - \mathbf{x}_*)\|_\infty}{\|\mathbf{x} - \mathbf{x}_*\|_\infty} = 0,$$

por lo cual para todo ρ , existe $\epsilon_3 > 0$, tal que si $\|\mathbf{x} - \mathbf{x}_*\|_\infty < \epsilon_3$ entonces

$$\frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}_*) + H(\mathbf{x} - \mathbf{x}_*)\|_\infty}{\|\mathbf{x} - \mathbf{x}_*\|_\infty} < \rho.$$

En particular, para $\rho = \frac{r}{8\mu}$ tenemos que

$$\frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}_*) + H(\mathbf{x} - \mathbf{x}_*)\|_\infty}{\|\mathbf{x} - \mathbf{x}_*\|_\infty} < \frac{r}{8\mu}. \quad (6.17)$$

De (6.17), (6.14) y (6.16) obtenemos que

$$\|\mathcal{Q}(\mathbf{x}, A) - \mathbf{x}_*\|_\infty < \frac{4}{3}\mu \left[\frac{r}{4\mu} + \frac{r}{8\mu} \right] \|\mathbf{x} - \mathbf{x}_*\|_\infty = \frac{r}{2} \|\mathbf{x} - \mathbf{x}_*\|_\infty < r \|\mathbf{x} - \mathbf{x}_*\|_\infty.$$

Completando así la demostración del **Lema 6.3**. \square

El siguiente teorema es análogo al *Teorema de las dos vecindades* del caso diferenciable [20]. En este caso, tenemos tres vecindades: una de la solución en la que está el punto inicial, otra de la matriz jacobiana de F en la solución y la tercera de la matriz

jacobiana de G en la solución, en la que están sus aproximaciones respectivas.

Teorema 6.1. *Dado $r \in (0, 1)$, existen $\epsilon > 0$ y $\delta > 0$ tales que si*

$$\|\mathbf{x}_0 - \mathbf{x}_*\|_\infty \leq \epsilon, \quad \|A_k - F'(\mathbf{x}_*)\|_\infty \leq \delta \quad y \quad \|C_k - G'(\mathbf{x}_*)\|_\infty \leq \delta$$

entonces la sucesión $\{\mathbf{x}_k\}$ generada por

$$\mathbf{x}_{k+1} = \mathcal{Q}(\mathbf{x}_k, A_k, C_k) = \mathbf{x}_k - B_k^{-1} \Phi_\lambda(\mathbf{x}_k),$$

donde B_k es la matriz cuyas filas están definidas por (6.1), está bien definida, converge a \mathbf{x}_* y satisface

$$\|\mathbf{x}_{k+1} - \mathbf{x}_*\|_\infty \leq r \|\mathbf{x}_k - \mathbf{x}_*\|_\infty,$$

para todo $k = 0, 1, \dots$

Demostración. Consideremos la función \mathcal{Q} definida en (6.11). Así,

$$\mathbf{x}_{k+1} = \mathcal{Q}(\mathbf{x}_k, A_k, C_k) = \mathbf{x}_k - B_k^{-1} \Phi_\lambda(\mathbf{x}_k),$$

con B_k definida por (6.1). Sean $r \in (0, 1)$, $\epsilon_1 \in (0, \epsilon)$ y $\delta_1 \in (0, \delta)$, donde ϵ y δ son las constantes del **Lema 6.3**.

Usaremos *inducción matemática* sobre k .

- Para $k = 0$. Si $\|\mathbf{x}_0 - \mathbf{x}_*\|_\infty \leq \epsilon_1 \leq \epsilon$, dado que

$$\|A_0 - F'(\mathbf{x}_*)\|_\infty \leq \delta_1 \leq \delta \quad y \quad \|C_0 - G'(\mathbf{x}_*)\|_\infty \leq \delta_1 \leq \delta, \quad (6.18)$$

por el **Lema 6.3**,

$$\mathcal{Q}(\mathbf{x}_0, A_0, C_0) = \mathbf{x}_1 = \mathbf{x}_0 - B_0^{-1} \Phi_\lambda(\mathbf{x}_0)$$

está bien definido y

$$\|\mathbf{x}_1 - \mathbf{x}_*\|_\infty \leq r \|\mathbf{x}_0 - \mathbf{x}_*\|_\infty.$$

- *Hipótesis inductiva.* Supongamos que para $k = 0, \dots, m-1$, el **Teorema 6.1** es verdadero; esto es, dado $r \in (0, 1)$ existen $\epsilon > 0$ y $\delta > 0$ tales que, si

$$\|\mathbf{x}_0 - \mathbf{x}_*\|_\infty \leq \epsilon, \quad \|A_0 - F'(\mathbf{x}_*)\|_\infty \leq \delta \quad y \quad \|C_0 - G'(\mathbf{x}_*)\|_\infty \leq \delta$$

entonces

$$\mathbf{x}_{k+1} = \mathcal{Q}(\mathbf{x}_k, A_k, C_k) = \mathbf{x}_k - B_k^{-1} \Phi_\lambda(\mathbf{x}_k),$$

está bien definida y satisface

$$\|\mathbf{x}_{k+1} - \mathbf{x}_*\|_\infty \leq r \|\mathbf{x}_k - \mathbf{x}_*\|_\infty.$$

Luego,

$$\|\mathbf{x}_m - \mathbf{x}_*\|_\infty \leq r \|\mathbf{x}_{m-1} - \mathbf{x}_*\|_\infty. \quad (6.19)$$

De (6.19), la hipótesis inductiva y puesto que $r < 1$, se tiene

$$\begin{aligned} \|\mathbf{x}_m - \mathbf{x}_*\|_\infty &\leq r \|\mathbf{x}_{m-1} - \mathbf{x}_*\|_\infty \\ &\leq r^m \|\mathbf{x}_0 - \mathbf{x}_*\|_\infty \\ &\leq r^m \epsilon \leq \epsilon. \end{aligned}$$

Así, por el **Lema 6.3**,

$$\mathbf{x}_{m+1} = \mathcal{Q}(\mathbf{x}_m, A_m, C_m) = \mathbf{x}_m - B_m^{-1} \Phi_\lambda(\mathbf{x}_m)$$

está bien definido y satisface

$$\|\mathbf{x}_{m+1} - \mathbf{x}_*\|_\infty \leq r \|\mathbf{x}_m - \mathbf{x}_*\|_\infty$$

con lo cual se concluye la demostración del **Teorema 6.1**. \square

El teorema siguiente da una condición suficiente para convergencia q -superlineal del **Algoritmo 2**. Su demostración, es una extensión a complementariedad generalizada, de la hecha en [2] para el caso particular de complementariedad no lineal. En la demostración, usamos $\|\cdot\| = \|\cdot\|_\infty$, aunque sabemos que los resultados de convergencia q -superlineal son independientes de la norma.

Teorema 6.2. *Sea B_k la matriz cuyas filas están definidas por (6.1). Supongamos que son válidas las hipótesis **H1** a **H3** y que la sucesión $\{\mathbf{x}_k\}$ generada por*

$$\mathbf{x}_{k+1} = \mathbf{x}_k - B_k^{-1} \Phi_\lambda(\mathbf{x}_k) \quad (6.20)$$

satisface que $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}_$. Si*

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} = 0, \quad (6.21)$$

donde $\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ entonces $\{\mathbf{x}_k\}$ converge q -superlinealmente a \mathbf{x}_ .*

Demostración. Puesto que Φ_λ es semisuave y $\Phi_\lambda(\mathbf{x}^*) = 0$, tenemos que

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_k) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0. \quad (6.22)$$

Por otra parte,

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq \|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|, \quad (6.23)$$

además,

$$\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right| \leq \|\Phi_\lambda(\mathbf{x}_k) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|. \quad (6.24)$$

Al reemplazar (6.23) en (6.22) y utilizar (6.24), se deduce que

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right|}{\|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} = 0,$$

luego

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right|}{\|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} = 0.$$

Usando la definición de límite, tenemos que, en particular, para $\rho = \frac{1}{2}$, existe $\epsilon > 0$, tal que si $\|\mathbf{x}_k - \mathbf{x}^*\| < \epsilon$ entonces,

$$-\frac{1}{2} < \frac{\|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|}{\|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} < \frac{1}{2},$$

de donde,

$$\frac{1}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| < \|\Phi_\lambda(\mathbf{x}_k)\| < \frac{3}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|.$$

A partir de la desigualdad de la izquierda y por (6.23), obtenemos

$$\|\Phi_\lambda(\mathbf{x}_k)\| > \frac{1}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \geq \frac{1}{2 \|H_*^{-1}\|} \|\mathbf{x}_k - \mathbf{x}^*\|. \quad (6.25)$$

Por otra parte, de (6.20), tenemos la ecuación

$$0 = B_k \mathbf{d}_k + \Phi_\lambda(\mathbf{x}_k), \quad (6.26)$$

donde $\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$. Sumando y restando la expresión $H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})$ en (6.26) y reorganizando los términos, obtenemos

$$-\Phi_\lambda(\mathbf{x}_{k+1}) = B_k \mathbf{d}_k - H_* \mathbf{d}_k + \Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1}).$$

Aplicando norma a ambos lados y luego la desigualdad triangular, deducimos que

$$\begin{aligned} \|\Phi_\lambda(\mathbf{x}_{k+1})\| &= \|B_k \mathbf{d}_k - H_* \mathbf{d}_k + \Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})\| \\ &\leq \|(B_k - H_*) \mathbf{d}_k\| + \|\Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})\| \end{aligned}$$

de donde

$$\frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|} \leq \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} + \frac{\|\Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|}.$$

El primer sumando del término derecho converge a cero por la hipótesis (6.21). El segundo sumando converge a cero por la *semisuavidad* de Φ_λ . Por lo tanto,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|} = 0. \quad (6.27)$$

Ahora, de la expresión (6.27), (6.25) y (6.4), se tiene que

$$\begin{aligned} 0 &= \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|} \geq \frac{1}{2 \|H_*^{-1}\|} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|s_k\|} \\ &= \frac{1}{2 \|H_*^{-1}\|} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_{k+1} - \mathbf{x}^* + \mathbf{x}^* - \mathbf{x}_k\|} \\ &\geq \frac{1}{2\mu} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_{k+1} - \mathbf{x}^*\| + \|\mathbf{x}_k - \mathbf{x}^*\|} \\ &= \frac{1}{2\mu} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|}}{\frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} + 1}. \end{aligned}$$

Por lo tanto,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0,$$

de donde la sucesión $\{\mathbf{x}_k\}$ converge q -superlinealmente a \mathbf{x}^* . □

6.2. Familia de métodos secantes de cambio mínimo para $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$.

El **Algoritmo 2** es genérico, ya que él no dice cómo actualizar las matrices B_k . En esta Sección, abordamos este aspecto, para lo cual, usaremos la teoría de métodos secantes de cambio mínimo desarrollada en [41], la cual también fue usada para el problema de complementariedad no lineal en [39]. Extendemos los resultados de este último, al problema de complementariedad generalizada; en este sentido, veremos que las pruebas de convergencia son análogas a aquellas presentadas en [39].

Recordemos que los métodos cuasi-Newton difieren entre sí, en la forma de actualizar la matriz B_k en cada iteración. Para el caso particular del **Algoritmo 2**, dicha actualización conduce a las actualizaciones de las matrices A_k y C_k , respectivamente. Entre los algoritmos cuasi-Newton “prácticos”, se encuentran los secantes de cambio mínimo, en los cuales la actualización debe satisfacer la llamada ecuación secante [20, 41]. Así, para actualizar la matriz B_k en el **Algoritmo 2**, se debe garantizar que las actualizaciones A_{k+1} y C_{k+1} , satisfagan las ecuaciones secantes:

$$A_{k+1}(\mathbf{x}_{k+1} - \mathbf{x}_k) = F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k) \quad (6.28)$$

$$C_{k+1}(\mathbf{x}_{k+1} - \mathbf{x}_k) = G(\mathbf{x}_{k+1}) - G(\mathbf{x}_k). \quad (6.29)$$

Además, el cambio entre A_k y A_{k+1} , C_k y C_{k+1} , debe ser *mínimo*. Es decir,

$$\|A_{k+1} - A_k\| = \min_{A \in V_F} \|A - A_k\|$$

$$\|C_{k+1} - C_k\| = \min_{C \in V_G} \|C - C_k\|,$$

donde V_F y V_G son conjuntos de matrices que satisfacen las *ecuaciones secantes* (6.28) y (6.29), respectivamente.

Exigir en cada iteración que se satisfaga la *ecuación secante* y que haya un *cambio mínimo* entre dos actualizaciones consecutivas, hace que las sucesiones $\{A_k\}$ y $\{C_k\}$ exhiban un fenómeno conocido como *deteriorización acotada* [2, 20], el cual garantiza que las matrices de las sucesiones permanecen en una vecindad de $F'(\mathbf{x}^*)$ y $G'(\mathbf{x}^*)$, respectivamente, lo cual es esencial para demostrar convergencia *lineal*.

En cada iteración de un algoritmo secante de cambio mínimo, los conjuntos $V_F = V_F(\mathbf{x}, \mathbf{y})$ y $V_G = V_G(\mathbf{x}, \mathbf{y})$ están determinados por los vectores \mathbf{x} y \mathbf{y} , donde

$$V_F = V_F(\mathbf{x}, \mathbf{y}) = \{A \in S \subset \mathbb{R}^{n \times n} : A(\mathbf{y} - \mathbf{x}) = F(\mathbf{y}) - F(\mathbf{x})\} \quad (6.30)$$

$$V_G = V_G(\mathbf{x}, \mathbf{y}) = \{C \in S \subset \mathbb{R}^{n \times n} : C(\mathbf{y} - \mathbf{x}) = G(\mathbf{y}) - G(\mathbf{x})\} \quad (6.31)$$

Dado que necesitamos la matriz en V_F “más cercana” a A y la matriz en V_G “más cercana” a C , es natural pensar en la *proyección ortogonal* de estas matrices sobre V_F y V_G , denotadas por $P_{V_F}(A)$ y $P_{V_G}(C)$, respectivamente. Así, teniendo en cuenta que

$$\|P_V(X) - X\| = \inf_{Y \in V} \|Y - X\|, \quad (6.32)$$

y que V_F y V_G son *conjuntos cerrados*, podemos garantizar que $P_{V_F}(A) \in V_F$ y $P_{V_G}(C) \in V_G$.

Cada una de estas proyecciones son únicas porque V_F y V_G son *conjuntos convexos* [36] y su cambio, medido en alguna norma, será

$$\|P_{V_F}(A) - A\| = \min_{A_k \in V_F} \|A - A_k\| \quad y \quad \|P_{V_G}(C) - C\| = \min_{C_k \in V_G} \|C - C_k\|. \quad (6.33)$$

Por lo cual, de (6.32) y (6.33), concluimos que

$$A_{k+1} = P_{V_F}(A_k) \quad y \quad C_{k+1} = P_{V_G}(C_k).$$

Diferentes *actualizaciones secante de cambio mínimo* se obtienen al variar la norma matricial en $\mathbb{R}^{n \times n}$ o al variar el subespacio $S \subset \mathbb{R}^{n \times n}$. Así, se genera una *familia de métodos secantes de cambio mínimo*. Por ejemplo, si denotamos $\mathbf{y}_k = F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k)$, $\hat{\mathbf{y}}_k = G(\mathbf{x}_{k+1}) - G(\mathbf{x}_k)$ y $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$, tenemos

- Actualización “buena” de Broyden [9].

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k} \quad C_{k+1} = C_k + \frac{(\hat{\mathbf{y}}_k - C_k \mathbf{s}_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k}, \quad (6.34)$$

resultan al elegir la *norma de Frobenius* y $S = \mathbb{R}^{n \times n}$.

- Actualización “mala” de Broyden [8].

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{y}_k^T A_k}{\mathbf{y}_k^T A_k \mathbf{s}_k} \quad C_{k+1} = C_k + \frac{(\hat{\mathbf{y}}_k - C_k \mathbf{s}_k) \hat{\mathbf{y}}_k^T C_k}{\hat{\mathbf{y}}_k^T C_k \mathbf{s}_k}, \quad (6.35)$$

resulta al elegir la *norma de Frobenius* y

$$V_F = V_F(\mathbf{x}, \mathbf{y}) = \{A \in \mathbb{R}^{n \times n} : A^{-1}(\mathbf{y} - \mathbf{x}) = F(\mathbf{y}) - F(\mathbf{x})\},$$

$$V_G = V_G(\mathbf{x}, \mathbf{y}) = \{C \in \mathbb{R}^{n \times n} : C^{-1}(\mathbf{y} - \mathbf{x}) = G(\mathbf{y}) - G(\mathbf{x})\}.$$

- Actualización de Schubert [57].

$$A_{k+1} = A_k + \frac{F(\mathbf{x}_{k+1})\mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k} \quad C_{k+1} = C_k + \frac{G(\mathbf{x}_{k+1})\mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k}.$$

resulta al elegir la *norma de Frobenius*.

A continuación, presentamos un algoritmo *secante de cambio mínimo* para resolver el PCG(F, G).

Algoritmo 3 secante de cambio mínimo.

Entrada: $\mathbf{x}_0 \in \mathbb{R}^n$ y $\lambda \in (0, 4)$, A_0 y C_0 arbitrarios y $\epsilon > 0$.

Salida: Aproximación a la solución del sistema (1.6).

- 1: **mientras** $\|\Phi_\lambda(\mathbf{x}_k)\| > \epsilon$ **hacer**
- 2: Calcule $D_a(\mathbf{x}_k)$ y $D_b(\mathbf{x}_k)$ como en la **Proposición 5.1**.

$$B_k = D_a(\mathbf{x}_k)A_k + D_b(\mathbf{x}_k)C_k, \quad (6.36)$$

Encuentre una solución \mathbf{d}_k al sistema lineal

$$B_k \mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k).$$

- 3: Hacer

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k,$$

- 4: **fin mientras**
- 5: Actualizar A_k y C_k como en (6.1) y tal que

$$\begin{aligned} A_{k+1} &= P_{V_F(\mathbf{x}_k, \mathbf{x}_{k+1})}(A_k) \\ C_{k+1} &= P_{V_G(\mathbf{x}_k, \mathbf{x}_{k+1})}(C_k). \end{aligned}$$

Volver al paso 1.

- 6: **devolver** \mathbf{x}_*
-

Con el fin de desarrollar la teoría de convergencia del **Algoritmo 3**, suponemos una hipótesis adicional a las tres anteriores.

H4. Para cada $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n$ suficientemente cercanos al punto \mathbf{x}_* , existen matrices $A \in V_F(\mathbf{x}, \mathbf{z})$, $C \in V_G(\mathbf{x}, \mathbf{z})$ y constantes α y $\bar{\alpha}$, tales que

$$\|A - F'(\mathbf{x}_*)\| \leq \alpha\sigma(\mathbf{x}, \mathbf{z}) \quad y \quad \|C - G'(\mathbf{x}_*)\| \leq \bar{\alpha}\sigma(\mathbf{x}, \mathbf{z}),$$

donde $\sigma(\mathbf{x}, \mathbf{z}) = \max\{\|\mathbf{x} - \mathbf{x}_*\|, \|\mathbf{z} - \mathbf{x}_*\|\}$.

6.3. Resultados de convergencia adicionales

Lema 6.4. *Supongamos que las hipótesis **H1** a **H4** se verifican. Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ suficientemente cercanos a \mathbf{x}_* , $A, C \in \mathbb{R}^{n \times n}$. Entonces existe una constante $\gamma > 0$, tal que*

$$\|P_{V_F(\mathbf{x}, \mathbf{y})}(A) - F'(\mathbf{x}_*)\| \leq \|A - F'(\mathbf{x}_*)\| + \gamma \|\mathbf{x} - \mathbf{x}_*\| \quad (6.37)$$

y

$$\|P_{V_G(\mathbf{x}, \mathbf{y})}(C) - G'(\mathbf{x}_*)\| \leq \|C - G'(\mathbf{x}_*)\| + \gamma \|\mathbf{x} - \mathbf{x}_*\|,$$

siempre y cuando $\mathbf{x}, \mathbf{y} \in \mathcal{B}(\mathbf{x}_*; \epsilon)$, $A \in \mathcal{B}(F'(\mathbf{x}_*); \delta)$, $C \in \mathcal{B}(G'(\mathbf{x}_*); \delta)$ y

$$\|\mathbf{y} - \mathbf{x}_*\| \leq \|\mathbf{x} - \mathbf{x}_*\|.$$

Demostración. Es análoga a la demostración del **Lema 3.1** en [40]. \square

En el siguiente lema, demostramos que las actualizaciones de las matrices A_k y C_k , pueden deteriorarse pero en forma controlada.

Lema 6.5. *Supongamos que las hipótesis **H1** a **H4** se verifican. Sean A_+ la proyección ortogonal de A sobre el conjunto $V_F(\mathbf{x}, \mathbf{z})$, C_+ la proyección ortogonal de C sobre $V_G(\mathbf{x}, \mathbf{z})$, \widehat{A} la proyección ortogonal de $F'(\mathbf{x})$ sobre $V_F(\mathbf{x}, \mathbf{z})$ y \widehat{C} la proyección ortogonal de $G'(\mathbf{x})$ sobre $V_G(\mathbf{x}, \mathbf{z})$ entonces*

$$\|A_+ - F'(\mathbf{x})\| \leq \|A - F'(\mathbf{x})\| + \beta_1 \sigma(\mathbf{x}, \mathbf{z})$$

y

$$\|C_+ - G'(\mathbf{x})\| \leq \|C - G'(\mathbf{x})\| + \beta_2 \sigma(\mathbf{x}, \mathbf{z}), \quad (6.38)$$

donde $\alpha_1 > 0$, $\alpha_2 > 0$ y $\sigma(\mathbf{x}, \mathbf{z}) = \max\{\|\mathbf{x} - \mathbf{x}_*\|, \|\mathbf{z} - \mathbf{x}_*\|\}$.

Demostración. Por la desigualdad triangular, tenemos

$$\|A_+ - F'(\mathbf{x})\| \leq \|A_+ - \widehat{A}\| + \|\widehat{A} - F'(\mathbf{x})\|. \quad (6.39)$$

Dado que la matriz A_+ satisface que

$$\|A_+ - \widehat{A}\| = \min_{A \in V_F(\mathbf{x}, \mathbf{z})} \|A - \widehat{A}\|,$$

tenemos la desigualdad:

$$\|A_+ - \widehat{A}\| \leq \|A - \widehat{A}\| \leq \|A - F'(\mathbf{x})\| + \|\widehat{A} - F'(\mathbf{x})\|. \quad (6.40)$$

Reemplazando (6.40) en (6.39),

$$\|A_+ - F'(\mathbf{x})\| \leq \|A - F'(\mathbf{x})\| + 2 \left\| \widehat{A} - F'(\mathbf{x}) \right\|.$$

Por la hipótesis **H4**, $\left\| \widehat{A} - F'(\mathbf{x}) \right\| \leq \alpha_1 \sigma(\mathbf{x}, \mathbf{z})$ entonces

$$\|A_+ - F'(\mathbf{x})\| \leq \|A - F'(\mathbf{x})\| + \beta_1 \sigma(\mathbf{x}, \mathbf{z}),$$

donde $\beta_1 = 2\alpha_1$. Análogamente, tomando $\beta_2 = 2\alpha_2$, se demuestra (6.38). \square

El siguiente resultado garantiza que el **Algoritmo 3** está bien definido y converge. Su prueba es análoga a la de un resultado similar para complementareidad no lineal demostrado en [39].

Teorema 6.3. *Supongamos que las hipótesis **H1** a **H4** se verifican. Sean $\{A_k\}$ y $\{C_k\}$ las sucesiones definidas por el **Algoritmo 3**. Dado $r \in (0, 1)$, existen constantes positivas $\bar{\epsilon}$ y $\bar{\delta}$ tales que, si*

$$\|\mathbf{x}_0 - \mathbf{x}_*\| \leq \bar{\epsilon}, \quad \|A_0 - F'(\mathbf{x}_*)\| \leq \bar{\delta} \quad y \quad \|C_0 - G'(\mathbf{x}_*)\| \leq \bar{\delta},$$

la sucesión $\{\mathbf{x}_k\}$ generada por

$$\mathbf{x}_{k+1} = \mathbf{x}_k - B_k^{-1} \Phi(\mathbf{x}_k),$$

está bien definida, converge a \mathbf{x}_* y además,

$$\|\mathbf{x}_{k+1} - \mathbf{x}_*\| \leq r \|\mathbf{x}_k - \mathbf{x}_*\|.$$

Demostración. Sean ϵ y δ las constantes del **Lema 6.3**, $\bar{\delta} \in (0, \delta)$ y

$$\bar{\epsilon} < \min \{\epsilon, \delta\} \quad y \quad \bar{\epsilon} < \min \{\epsilon, \delta_r\}, \quad (6.41)$$

donde

$$\delta = \frac{\delta - \bar{\delta}}{\gamma} \quad y \quad \delta_r = (\delta - \bar{\delta}) \frac{1 - r}{\gamma},$$

con, γ la constante de (6.37).

A partir de las desigualdades (6.41), podemos deducir que

$$\bar{\delta} + \gamma \bar{\epsilon}_0 < \bar{\delta} + \gamma \delta_0 = \bar{\delta} + \frac{\gamma(\delta - \bar{\delta})}{\gamma} = \delta$$

y

$$\bar{\delta} + \frac{\bar{\epsilon}\gamma}{1-r} < \bar{\delta} + \frac{\delta_r\gamma}{1-r} < \bar{\delta} + \frac{\gamma(\delta - \bar{\delta})(1-r)}{\gamma(1-r)} = \delta. \quad (6.42)$$

Usaremos inducción sobre k .

1. Para $k = 0$. Si

$$\|\mathbf{x}_0 - \mathbf{x}_*\| \leq \bar{\epsilon}, \quad \|A_0 - F'(\mathbf{x}_*)\| \leq \bar{\delta} \quad \text{y} \quad \|C_0 - G'(\mathbf{x}_*)\| \leq \bar{\delta},$$

entonces por el **Lema 6.3**, \mathbf{x}_1 está bien definido y satisface que

$$\|\mathbf{x}_1 - \mathbf{x}_*\| \leq r \|\mathbf{x}_0 - \mathbf{x}_*\|,$$

con lo cual, el **Teorema 6.3** es verdadero para $k = 0$.

2. *Hipótesis inductiva*. Supongamos que el **Teorema 6.3** es verdadero para $k = 0, 1, \dots, m-1$ entonces, si

$$\|\mathbf{x}_{m-1} - \mathbf{x}_*\| \leq \bar{\epsilon},$$

$$\mathbf{x}_m = \mathbf{x}_{m-1} - B_{m-1}^{-1}\Phi(\mathbf{x}_{m-1}),$$

está bien definida y

$$\|\mathbf{x}_m - \mathbf{x}_*\| \leq r \|\mathbf{x}_{m-1} - \mathbf{x}_*\|.$$

3. Por hipótesis inductiva y dado que $0 < r < 1$, tenemos que

$$\|\mathbf{x}_m - \mathbf{x}_*\| \leq r \|\mathbf{x}_{m-1} - \mathbf{x}_*\| \leq r^m \|\mathbf{x}_0 - \mathbf{x}_*\| < r^m \bar{\epsilon} < \bar{\epsilon}.$$

De (6.37) y $0 < r < 1$,

$$\begin{aligned} \|A_m - F'(\mathbf{x}_*)\| &\leq \|A_{m-1} - F'(\mathbf{x}_*)\| + \gamma \|\mathbf{x}_{m-1} - \mathbf{x}_*\| \\ &\leq \bar{\delta} + \gamma \bar{\epsilon} \sum_{j=0}^{m-1} r^j \\ &\leq \bar{\delta} + \gamma \bar{\epsilon} \sum_{j=0}^{\infty} r^j \\ &< \bar{\delta} + \gamma \frac{\bar{\epsilon}}{1-r} \leq \delta, \end{aligned}$$

donde la última desigualdad se obtiene de (6.42); así, $\|A_m - F'(\mathbf{x}_*)\| < \delta$. En

forma análoga, se demuestra que $\|C_m - G'(\mathbf{x}_*)\| < \delta$. Así, por el **Lema 6.3**, \mathbf{x}_m está bien definido y satisface

$$\|\mathbf{x}_{m+1} - \mathbf{x}_*\| \leq r \|\mathbf{x}_m - \mathbf{x}_*\|.$$

Por lo tanto, el **Teorema 6.3** es verdadero para $k = m$. \square

Lema 6.6. *Bajo las hipótesis del Teorema 6.3,*

$$\lim_{k \rightarrow \infty} \|A_{k+1} - A_k\| = 0 \quad y \quad \lim_{k \rightarrow \infty} \|C_{k+1} - C_k\| = 0.$$

Demostración. Es análoga a la del **Teorema 3.3** en [41]. \square

Lema 6.7. *Bajo las hipótesis del Teorema 6.3,*

$$\lim_{k \rightarrow \infty} \|B_{k+1} - B_k\| = 0. \quad (6.43)$$

Demostración. De (6.1) y (6.2), tenemos que

$$[\Delta B_k]_i = \begin{cases} [(\bar{a}_i - 1) [A_{k+1}]_i + (\bar{b}_i - 1) [C_{k+1}]_i \\ - [(a_i - 1) [A_k]_i + (b_i - 1) [C_k]_i], & i \notin \beta \\ [(\bar{\chi}_i - 1) [A_{k+1}]_i + (\bar{\xi}_i - 1) [C_{k+1}]_i \\ - [(\chi_i - 1) [A_k]_i + (\xi_i - 1) [C_k]_i], & i \in \beta. \end{cases}$$

Procediendo en forma análoga a la demostración del **Lema 6.1** y usando la *norma infinito*, tenemos que

$$\|B_{k+1} - B_k\|_\infty = \max_{1 \leq i \leq n} \{ |[B_{k+1}]_i - [B_k]_i| \},$$

por lo cual

$$\|B_{k+1} - B_k\|_\infty \leq (1 + \sqrt{2}) \|A_{k+1} - A_k\|_\infty + (1 + \sqrt{2}) \|C_{k+1} - C_k\|_\infty,$$

de donde se deduce, usando los dos primeros límites de (6.43) y la equivalencia entre normas, que

$$\lim_{k \rightarrow \infty} \|B_{k+1} - B_k\| = 0.$$

\square

El siguiente resultado da una condición suficiente para convergencia q -superlineal.

Teorema 6.4. *Supongamos que las hipótesis **H1** a **H4** se verifican. Sean $\{A_k\}$ y $\{C_k\}$ las sucesiones generadas por el **Algoritmo 3** y $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}_*$. Si*

$$\lim_{k \rightarrow \infty} \frac{\|(B_{k+1} - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} = 0 \quad (6.44)$$

entonces la sucesión $\{\mathbf{x}_k\}$ converge q -superlinealmente.

Demostración. Es una aplicación directa de los **Teoremas 6.2** y **6.3**. En efecto,

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} &\leq \lim_{k \rightarrow \infty} \frac{\|(B_k - B_{k+1}) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} + \lim_{k \rightarrow \infty} \frac{\|(B_{k+1} - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} \\ &\leq \lim_{k \rightarrow \infty} \|B_k - B_{k+1}\| + \lim_{k \rightarrow \infty} \frac{\|(B_{k+1} - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|}. \end{aligned}$$

Por el **Lema 6.7** y la hipótesis (6.44),

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} = 0.$$

Por lo tanto, concluimos que la sucesión $\{\mathbf{x}_k\}$ converge q -superlinealmente. \square

Por los resultados demostrados en esta sección, tenemos que, bajo las hipótesis **H1** a **H4**, la familia de métodos secantes de cambio mínimo generada por (6.28) y (6.29) proporcionan algoritmos locales y superlinealmente convergentes, lo cual se resume en el siguiente teorema.

Teorema 6.5. *Supongamos que las hipótesis **H1** a **H4** se verifican. Sean $\{A_k\}$ y $\{C_k\}$ las sucesiones definidas por el **Algoritmo 3**. Dado $r \in (0, 1)$ existen constantes positivas $\bar{\epsilon}$ y $\bar{\delta}$, tales que si*

$$\|\mathbf{x}_0 - \mathbf{x}_*\| \leq \bar{\epsilon}, \quad \|A_k - F'(\mathbf{x}_*)\| \leq \bar{\delta} \quad y \quad \|C_k - G'(\mathbf{x}_*)\| \leq \bar{\delta},$$

entonces la sucesión $\{\mathbf{x}_k\}$ converge local y linealmente a \mathbf{x}_* . Si

$$\lim_{k \rightarrow \infty} \frac{\|(B_{k+1} - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} = 0 \quad (6.45)$$

entonces la sucesión $\{\mathbf{x}_k\}$ converge q -superlinealmente a \mathbf{x}_* .

Demostración. Es una aplicación directa de los **Teoremas 6.2**, **6.3** y **6.4**. \square

6.4. Pruebas Numéricas

En esta sección, analizamos numéricamente el comportamiento local del **Algoritmo 3**. Actualizamos las matrices A_k y C_k usando las fórmulas buena de Broyden (**Bb**), mala de Broyden (**Bm**) y Schubert (**Sch**), con lo cual tenemos tres versiones del algoritmo. Para cada una de ellas, analizamos radio, velocidad y región de convergencia y usamos dos formas de elegir el parámetro λ : $\lambda = 10^{-5}$ y la estrategia de λ dinámico dada en [34].

Implementamos los algoritmos y funciones de prueba en Matlab[®] R2020a, en un computador con procesador AMD Sempron (tm) de 2.21 GHz. Los parámetros usados son $\epsilon = 10^{-6}$ y $N = 100$, este último indica el número máximo de iteraciones permitido en los algoritmos. Para los experimentos consideramos los diez problemas descritos en el **Capítulo 3**.

6.4.1. Experimento 1

En el primer experimento, estimamos el radio de convergencia de las variantes del **Algoritmo 3**, que llamaremos Broyden bueno (**Bb**), Broyden malo (**Bm**) y Schubert (**Sch**), para $\lambda = 10^{-5}$, y Broyden bueno Dinámico (**BbD**), Broyden malo dinámico (**BmD**) y Schubert Dinámico (**SchD**), cuando se elige λ usando la estrategia dinámica.

El radio de convergencia es el máximo radio (r) de una vecindad centrada en una solución \mathbf{x}_* ($B_r(\mathbf{x}_*)$), tal que si tomamos un punto inicial en esta vecindad, el algoritmo converge. Para la norma infinito, esta región es el conjunto $B_r(\mathbf{x}_*) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_*\|_\infty \leq r\}$. Vale la pena mencionar que para este experimento, la solución considerada en cada caso fue \mathbf{x}_*^1 (ver **Capítulo 3**).

Consideramos un radio inicial r_0 , tal que para 1000 puntos aleatorios en $B_{r_0}(\mathbf{x}_*)$, el algoritmo respectivo converge. Luego, incrementamos el radio en 0.1 unidades y obtenemos un radio r_1 . Nuevamente, generamos 1000 puntos aleatorios en $B_{r_1}(\mathbf{x}_*)$.

Si no hay convergencia para alguno de esos puntos entonces el radio de convergencia será r_0 ; en otro caso, continuamos el proceso hasta que el radio es estimado.

Los resultados de este experimento se muestran en la **Tabla 6.1** para los dos valores del parámetro λ escogidos. La tabla contiene la siguiente información: problema (**P**), el valor del parámetro λ , con el cual se corrieron los algoritmos; radio de convergencia (r) y promedio de iteraciones (\bar{k}) en la cual cada variante del algoritmo convergió.

En general, en la **Tabla 6.1** observamos radios de convergencia más amplios para

P		$\lambda = 10^{-5}$.			$\lambda = 2$ (dinámico).		
		Bb	Bm	Sch	BbD	BmD	SchD
P1	r	0.8	0.2	0.6	0.6	0.5	0.8
	\bar{k}	10.2	18.1	14.2	13.5	9.6	8.2
P2	r	10	10	10	10	10	0.6
	\bar{k}	3	2.9	5.3	13.2	7	2
P3	r	1.1	0.9	0.9	1.2	1.3	0.6
	\bar{k}	9.1	7.9	10.3	14.2	11.5	7.1
P4	r	10	10	0.4	0.9	0.9	10
	\bar{k}	17.3	17.2	64.6	12.3	11.9	16.6
P5	r	1.7	1.4	1.6	1.9	1.5	1.7
	\bar{k}	6.5	5.8	13.1	9.1	7.5	6.1
P6	r	10	5.4	0.8	13	5.4	10
	\bar{k}	26.7	11.6	12.3	13.3	12.8	9.9
P7	r	10	10	10	10	10	1.9
	\bar{k}	2.9	2.8	6.7	9.4	6.6	2
P8	r	10	10	0.2	10	10	0.2
	\bar{k}	3.4	3.4	19.7	40.7	5.4	7.6
P9	r	0.4	0.4	0.2	0.4	0.4	0.2
	\bar{k}	10.3	8.9	13.1	12.7	7.6	8.9
P10	r	0.4	0.4	0.2	0.5	0.4	0.2
	\bar{k}	10.7	9	13.9	10.4	7.6	6.8

Tabla 6.1: Resultados del Experimento 1. Radio de convergencia.

las versiones del **Algoritmo 3**, con las actualizaciones buena y mala de Broyden, para las dos formas de elegir el parámetro λ ; esto sugiere que la escogencia de la actualización influye en el radio de convergencia de cada problema. Respecto a la elección del valor del parámetro λ , no hay diferencias significativas, excepto para los **Problemas 4 y 6**.

6.4.2. Experimento 2

En el segundo experimento, analizamos la velocidad de convergencia de las tres variantes del **Algoritmo 3**. Para ello, calculamos el cociente R_k definido por:

$$R_k = \frac{\|\mathbf{x}_{k+1} - \mathbf{x}_*\|}{\|\mathbf{x}_k - \mathbf{x}_*\|}.$$

Realizamos el experimento para los **Problemas 1 y 2**, con el punto inicial \mathbf{x}_1 , las actualizaciones mencionadas anteriormente.

Los resultados aparecen en la **Tabla 6.2**, que contiene en la primera fila el valor del parámetro λ utilizado, en la primera columna el valor de la iteración k y en las otras columnas el valor de R_k para cada método utilizado.

Tabla 6.2: Resultados del Experimento 2.

	$\lambda = 10^{-5}$.			$\lambda = 2$ (dinámico).		
k	Bb	Bm	Sch	BbD	BmD	SchD
1	0.471796	0.471796	0.443930	0.471796	0.471796	—
2	0.162536	0.162536	0.341794	0.162536	0.162536	—
3	0.149192	0.149192	0.328350	0.149192	0.149192	—
4	0.028053	0.028053	0.119448	0.257447	0.257447	—
5	0.000828	0.000828	0.167303	0.546383	0.546383	—
6	0.000001	0.000001	0.132742	0.000795	0.000795	—
7				0.000099	0.000099	—
8				0.000042	0.000042	—
9						
10			⋮			
11						
12						
13			0.170199			
14			0.132456			

En estas tablas se puede apreciar que la tasa de convergencia de los métodos propuestos sigue el patrón de convergencia q -superlineal.

6.4.3. Experimento 3

En el tercer experimento, ilustramos la región de convergencia de las tres versiones del **Algoritmo 3** para los **Problemas 4, 5 y 8**. Para ello, a cada una de las componentes de la solución \mathbf{x}_* le asociamos el siguiente conjunto:

$$X_i^* = \{y = x_{i,*} + 0.1t : t = -10, -9, \dots, 9, 10\}. \quad (6.46)$$

Posteriormente, usando (6.46), ejecutamos los algoritmos con todos los puntos del producto cartesiano $\prod_{i=1}^n X_i^*$, con lo cual, cada algoritmo se ejecuta con 201^n puntos iniciales. Destacamos el hecho de que para el **Problema 8** (dimensión $n = 3$), el número de puntos iniciales es $8'120.601$.

Figura 6.1: Región y porcentaje de convergencia del **Algoritmo 3**, para el **Problema 4**, con $\lambda = 10^{-5}$.

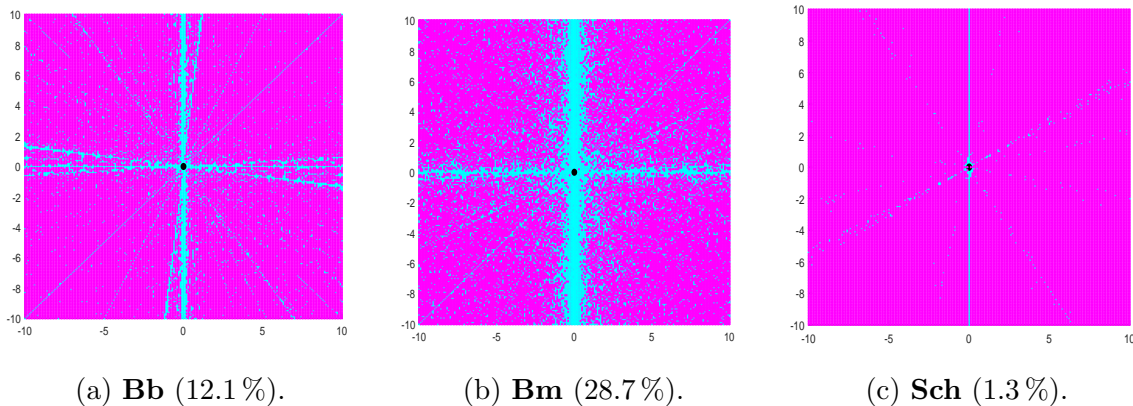


Figura 6.2: Región y porcentaje de convergencia del **Algoritmo 3**, para el **Problema 4**, con $\lambda = 10^{-5}$.

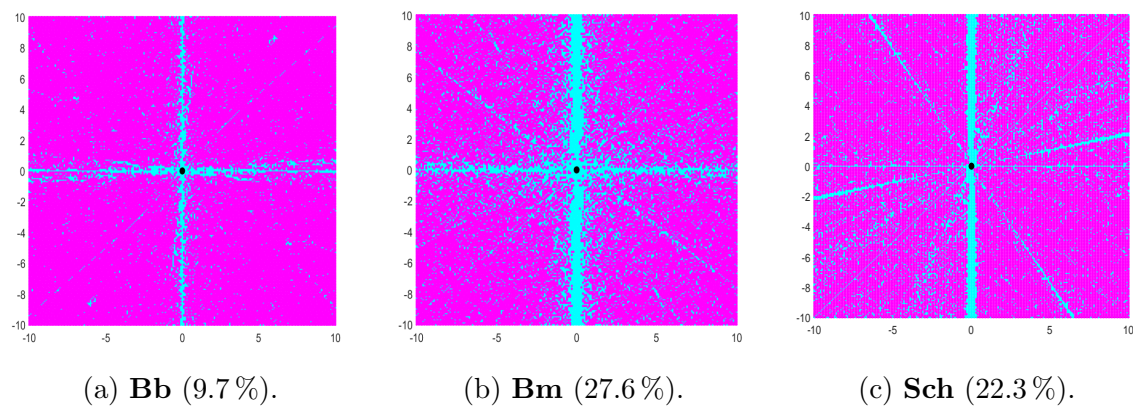


Figura 6.3: Región y porcentaje de convergencia del **Algoritmo 3**, para el **Problema 5**, con λ dinámico.

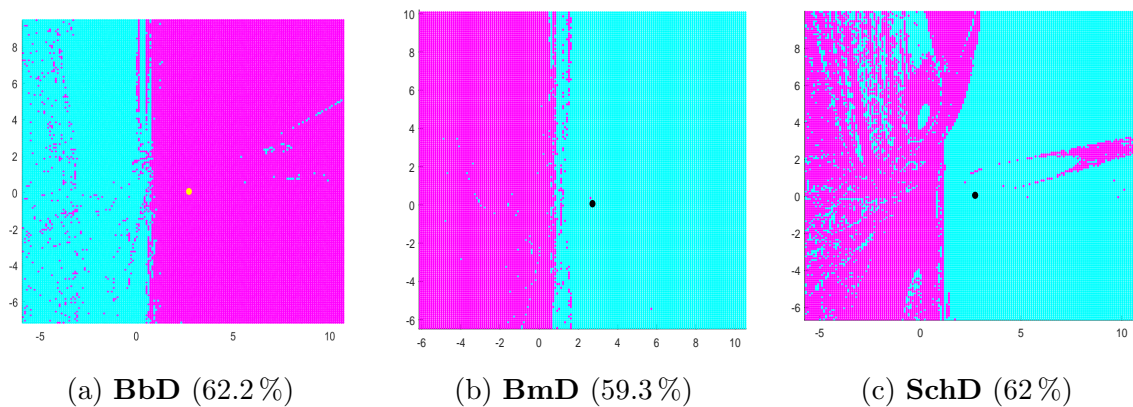


Figura 6.4: Región y porcentaje de convergencia del **Algoritmo 3**, para el **Problema 8**, con $\lambda = 10^{-5}$.

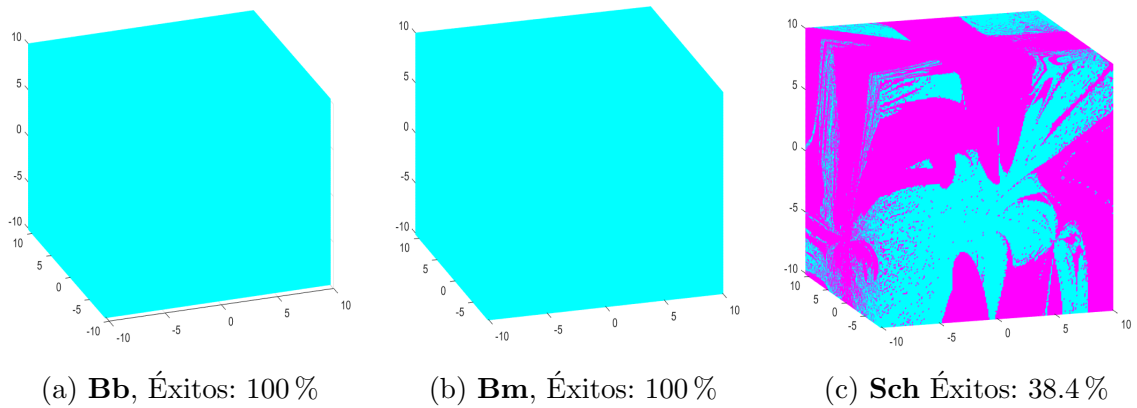
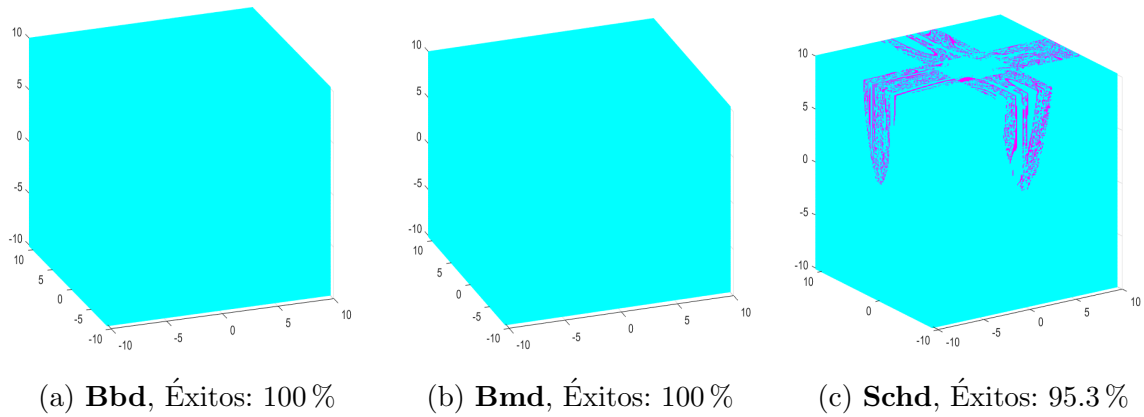


Figura 6.5: Región y porcentaje de convergencia del **Algoritmo 3**, para el **Problema 8**, con λ dinámico.



Para la región de convergencia, coloreamos en azul aquellos puntos para los cuales los algoritmos convergieron y en magenta, los que divergieron. En las **Figuras 6.1 a 6.5** presentamos los resultados obtenidos.

En estas figuras, podemos observar que para el **Problema 4**, la elección de λ dinámico mejora la región de convergencia mientras que para el **Problema 5**, cualquiera de las dos formas de λ da resultados similares. Para el **Problema 8**, la selección de λ dinámico resulta ser la mejor opción ya que para $\lambda = 10^{-5}$ la actualización de Schubert converge solamente para el 38.4% de los casos, comparada con las otras actualizaciones.

En general, vemos la necesidad de implementar una estrategia de globalización en cada uno de estos métodos.

6.4.4. Experimento 4

En el cuarto experimento, analizamos la eficiencia y robustez de los algoritmos; es decir, la capacidad que tienen para resolver un problema desde diferentes puntos iniciales; para ello, usamos fórmulas que permiten comparar la eficiencia y robustez de los algoritmos entre sí y las ejecutamos con todos los puntos iniciales señalados en el Capítulo 3. Estas fórmulas son: índice de robustez (R), de eficiencia (E) y eficiencia combinada (C), definidas de la siguiente forma [7, 10, 55]:

$$\begin{aligned} R_j &= t_j/n_j \\ E_j &= \sum_{i=1, r_{ij} \neq 0}^m (r_{ib}/r_{ij})/t_j \\ C_j &= \sum_{i=1, r_{ij} \neq 0}^m (r_{ib}/r_{ij})/n_j, \end{aligned}$$

donde r_{ij} es el número de iteraciones requeridas para resolver el problema i por el método j , $r_{ib} = \min_j r_{ij}$, t_j es el número de éxitos obtenidos por el método j y n_j (38 para todo j) es el número de problemas ejecutados por el método j .

En otras palabras, un algoritmo es más robusto que otro si resuelve más problemas; es más eficiente que otro, si al resolver un problema, en cierta medida, requiere de un número menor de iteraciones y finalmente, el índice combinado establece una relación entre los dos índices anteriores y nos da una medida de qué tan equilibrado es el algoritmo. Un algoritmo es más robusto, eficiente o equilibrado, en la medida en que estos índices estén más cercano a 1.

	Bb	Bm	Sch	BbD	BmD	SchD
R	0.87	0.92	0.74	0.79	0.92	0.89
E	0.78	0.78	0.62	0.69	0.52	0.87
C	0.68	0.72	0.46	0.55	0.48	0.78

Tabla 6.3: Índices de robustez(R), eficiencia (E) y combinado (C).

En la **Tabla 6.3** observamos que el más robusto es el **Algoritmo 3**, con la actualización mala de Broyden; el más eficiente y a su vez, más equilibrado es el **Algoritmo 3**, con la actualización de Schubert y la estrategia de λ dinámico (**SchD**), ya que su índice de eficiencia (0.87) y el combinado (0.78) son los mayores.

Un método libre de derivadas

Entre los métodos numéricos para resolver sistemas de ecuaciones no lineales, los cuasi-Newton se consideran entre los más eficaces; sin embargo, no existe mucha literatura sobre convergencia global de ellos. La principal dificultad parece estar en que las búsquedas lineales que requieren el cálculo de derivadas son inapropiadas para este tipo de métodos. Una alternativa que ha resultado efectiva es una búsqueda lineal libre de derivadas [37].

Motivados por lo mencionado en el párrafo anterior y los buenos resultados obtenidos en [37] para un algoritmo libre de derivadas que resuelve sistemas de ecuaciones diferenciables, en este capítulo proponemos inicialmente, un algoritmo cuasi-Newton global libre de derivadas para resolver el $\text{PCG}(F, G)$, indirectamente, a través de la solución del problema de minimización (1.7). Este algoritmo es genérico porque no se especifica la forma de actualizar las aproximaciones de las matrices del jacobiano generalizado. Bajo ciertas hipótesis, presentamos resultados de convergencia global para este algoritmo.

En la segunda parte del capítulo, abordamos el problema de la actualización de las aproximaciones, basándonos en la teoría secante de cambio mínimo desarrollada en el **Capítulo 6**, con lo cual tenemos un nuevo algoritmo global secante, libre de derivadas para resolver el $\text{PCG}(F, G)$, el cual es un caso particular del algoritmo genérico propuesto en la primera parte. Demostramos que cualquier sucesión generada por la nueva propuesta algorítmica satisface las hipótesis de convergencia del algoritmo genérico heredando así, sus resultados de convergencia.

7.1. Búsqueda lineal libre de derivadas

Una estrategia para resolver problemas de minimización sin restricciones utilizando un método de búsqueda direccional consiste en encontrar una dirección en la que la función objetivo decrezca y dar un paso en esa dirección. El proceso de elección del “tamaño del paso”, se conoce como búsqueda lineal [20].

Por otra parte, la función objetivo del problema de minimización (1.7) es una suma de cuadrados y el método que pretendemos usar para resolverlo es cuasi-Newton. En este proceso, se presentan varios inconvenientes: en primer lugar, este tipo de métodos usa aproximaciones al jacobiano de la función objetivo y el gradiente de esta depende a su vez, de las matrices del jacobiano generalizado, lo cual hace difícil la minimización. Además, el cálculo de derivadas puede ser muy costoso o el paso cuasi-Newton puede no ser una dirección de descenso.

Una alternativa es usar una estrategia de globalización que no requiera del cálculo de derivadas, ni imponga una condición de descenso suficiente sobre el valor de la función objetivo. Tal estrategia se conoce como búsqueda lineal libre de derivadas.

Uno de los primeros trabajos al respecto fue [29], que sirvió de base para el trabajo en convergencia global de métodos tipo Broyden para sistemas no lineales irrestrictos dado por [37]. La idea básica de este tipo de estrategia para resolver el sistema de ecuaciones no lineales $K(\mathbf{x}) = 0$, donde K es un campo vectorial continuamente diferenciable, consiste en lo siguiente: dada la iteración \mathbf{x}_k y una dirección de búsqueda \mathbf{d}_k , la iteración siguiente es de la forma $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{d}_k$, $\lambda > 0$, la cual satisface

$$\|K(\mathbf{x}_k + \lambda \mathbf{d}_k)\| \leq (1 + \eta_k) \|K(\mathbf{x}_k)\| - \alpha \lambda^2 \|\mathbf{d}_k\|^2 \quad (7.1)$$

para algunas constantes $\alpha \in (0, 1)$ y $\eta_k > 0$. La sucesión $\{\eta_k\}$ es tal que

$$\sum_{k=0}^{\infty} \eta_k \leq \eta < \infty. \quad (7.2)$$

Por la continuidad de K , la condición (7.1) se satisface para λ suficientemente pequeño, y es llamada una búsqueda lineal de descenso aproximado en norma, puesto que

$$\|K(\mathbf{x}_k + \lambda \mathbf{d}_k)\| \leq (1 + \eta_k) \|K(\mathbf{x}_k)\|.$$

7.2. Algoritmo y resultados de convergencia

El algoritmo libre de derivadas genérico que proponemos es el siguiente.

Algoritmo 4

Entrada: $\mathbf{x}_0 \in \mathbb{R}^n$, B_0 no singular, $\lambda \in (0, 4)$, $\sigma_1, \sigma_2 > 0$, $\beta, \rho \in (0, 1)$, $p > 2$, $\epsilon \geq 0$, $\{\eta_k\}$ una sucesión de términos positivos que satisfice (7.2) y $k = 0$.

Salida: Aproximación a la solución del sistema (1.6).

1: **mientras** $\|\Phi_\lambda(\mathbf{x}_k)\| > \epsilon$ **hacer**

2: Sea B_k una aproximación a H_k . Encuentre una solución \mathbf{d}_k al sistema de ecuaciones lineales:

$$B_k \mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k). \quad (7.3)$$

3: **si**

$$\|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_2 \|\mathbf{d}_k\|^2 \quad (7.4)$$

entonces

4: hacer $t_k = 1$ y pasar al Paso 7

5: **sino**

6: Haga $t_k = \max\{\beta^l : l = 0, 1, 2, \dots\}$, tal que

$$\|\Phi_\lambda(\mathbf{x}_k + t_k \mathbf{d}_k)\| \leq \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_1 \|t_k \mathbf{d}_k\|^2 + \eta_k \|\Phi_\lambda(\mathbf{x}_k)\|. \quad (7.5)$$

7: Actualizar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_k,$$

8: **fin si**

9: Actualizar B_k , tal que B_{k+1} sea no singular, $k = k + 1$. Volver al Paso 1.

10: **fin mientras**

11: **devolver** \mathbf{x}_*

Algunos comentarios respecto al algoritmo son los siguientes. La búsqueda lineal del **Algoritmo 4** no solo es libre de derivadas [29,37] sino, no monótona [30]. Es decir, usaremos un criterio para el tamaño del paso que puede verse como una generalización de la regla de *Armijo* [20], en el sentido de que la condición que implica un descenso monótono de la función objetivo, se puede relajar y, aún así, establecer convergencia global.

El sistema (7.3) siempre tiene solución porque las matrices B_k son no singulares para todo k . En este sentido, el algoritmo está bien definido.

La búsqueda lineal (7.5) es de *descenso aproximado de la norma* puesto que

$$\|\Phi_\lambda(\mathbf{x}_{k+1})\| \leq (1 + \eta_k) \|\Phi_\lambda(\mathbf{x}_k)\|,$$

para cada k , con η_k que satisface (7.2). De (7.5), es fácil ver que

$$\sigma_1 \|t_k \mathbf{d}_k\|^2 \leq \|\Phi_\lambda(\mathbf{x}_k)\| - \|\Phi_\lambda(\mathbf{x}_{k+1})\| + \eta_k \|\Phi_\lambda(\mathbf{x}_k)\|. \quad (7.6)$$

Ahora analizaremos la convergencia del algoritmo. A continuación, presentamos las hipótesis bajo las cuales desarrollamos la teoría de convergencia del **Algoritmo 4**.

H0 El conjunto de nivel Ω , definido por

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n : \|\Phi_\lambda(\mathbf{x})\| \leq e^\eta \|\Phi_\lambda(\mathbf{x}_0)\|\}, \quad (7.7)$$

es acotado, para alguna constante positiva η .

H1 Existe $\mathbf{x}_* \in \mathbb{R}^n$, tal que $\Phi_\lambda(\mathbf{x}_*) = \mathbf{0}$.

H2 Las matrices jacobianas de F y G son *Lipschitz continuas* (con la misma constante) en una vecindad de $\mathbf{x}_* \in \mathbb{R}^n$. Es decir, existen constantes $\gamma > 0$ y $\delta > 0$ tales que

$$\|F'(\mathbf{x}) - F'(\mathbf{x}_*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}_*\| \quad y \quad \|G'(\mathbf{x}) - G'(\mathbf{x}_*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}_*\|,$$

para todo $\mathbf{x} \in \mathcal{B}(\mathbf{x}_*; \delta)$, donde $\|\cdot\|$ denota cualquier norma en \mathbb{R}^n y también su respectiva norma matricial inducida.

H3 \mathbf{x}_* es una solución R-regular del PCG (F, G) .

H4 $\{B_k\}$ es una sucesión de matrices no singulares y acotadas tal que $\|B_k^{-1}\| \leq T$, para todo k , con T una constante positiva.

H5 Las matrices B_k satisfacen que

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = 0,$$

donde $H_* \in \partial\Phi_\lambda(\mathbf{x}_*)$ y $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k = t_k \mathbf{d}_k$.

Los siguientes son tres lemas técnicos que serán útiles en el desarrollo de la teoría de convergencia del método propuesto.

Lema 7.1. [19, 37] *La sucesión $\{\mathbf{x}_k\}$ generada por el Algoritmo 4 está contenida en el conjunto Ω definido en (7.7).*

Demostración. La demostración es análoga a la hecha en [37] para un sistema de ecuaciones no lineales diferenciable. \square

Lema 7.2. [37] *Si la hipótesis $\bar{\mathbf{H0}}$ se verifica y $\{\mathbf{x}_k\}$ es la sucesión generada por el Algoritmo 4 entonces*

$$\sum_{k=0}^{\infty} \|\mathbf{s}_k\|^2 < \infty, \quad (7.8)$$

donde $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$.

Demostración. Por hipótesis y el **Lema 7.1**, la sucesión $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ es acotada. Puesto que η_k satisface (7.2) entonces, haciendo la sumatoria desde cero a infinito en (7.6), obtenemos (7.8). \square

Lema 7.3. *Si $\{\mathbf{x}_k\}$ es una sucesión generada por el Algoritmo 4 entonces $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ converge.*

Demostración. Si existiera \bar{k} , tal que $t_k = 1$, para todo $k > \bar{k}$ entonces por (7.4), $\|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_2 \|\mathbf{d}_k\|^2 \leq \|\Phi_\lambda(\mathbf{x}_k)\|$, por lo cual $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ sería decreciente y por el **Lema 7.1**, acotada, de donde $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ también sería convergente.

Ahora, si existiera \bar{k} , tal que t_k satisface (7.5), para todo $k > \bar{k}$ entonces

$$\|\Phi_\lambda(\mathbf{x}_k + t_k \mathbf{d}_k)\| \leq (1 + \eta_k) \|\Phi_\lambda(\mathbf{x}_k)\| + \eta_k,$$

y por el **Lema 4.6**, con $a_k = \|\Phi_\lambda(\mathbf{x}_k)\|$ y $r_k = \eta_k$, la sucesión $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ converge. \square

El siguiente lema garantiza que las direcciones generadas por el **Algoritmo 4** son acotadas.

Lema 7.4. *Supongamos que las hipótesis $\bar{\mathbf{H0}}$ a $\bar{\mathbf{H4}}$ se verifican. Entonces*

$$\|\mathbf{d}_k\| \leq T \|\Phi_\lambda(\mathbf{x}_k)\|. \quad (7.9)$$

Demostración. Por (7.3) y la hipótesis $\bar{\mathbf{H4}}$, tenemos que

$$\|\mathbf{d}_k\| = \|-B_k^{-1} \Phi_\lambda(\mathbf{x}_k)\| \leq \|B_k^{-1}\| \|\Phi_\lambda(\mathbf{x}_k)\| \leq T \|\Phi_\lambda(\mathbf{x}_k)\|,$$

lo cual implica que $\|\mathbf{d}_k\| \leq T \|\Phi_\lambda(\mathbf{x}_k)\|$. \square

En el siguiente teorema, demostramos la convergencia de la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 4**.

Teorema 7.1. *Sea $\{\mathbf{x}_k\}$ una sucesión generada por el Algoritmo 4. Supongamos que las hipótesis $\bar{\mathbf{H0}}$ a $\bar{\mathbf{H4}}$ se verifican. Entonces la sucesión $\{\mathbf{x}_k\}$ converge.*

Demostración. Por el **Lema 7.4**

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| = \|t_k \mathbf{d}_k\| \leq t_k T \|\Phi_\lambda(\mathbf{x}_k)\|.$$

Usando (7.5), el **Lema 7.2** y la desigualdad

$$\sigma_1 (t_k \|\mathbf{d}_k\|)^2 \leq (1 + \eta_k) \|\Phi_\lambda(\mathbf{x}_k)\| - \|\Phi_\lambda(\mathbf{x}_{k+1})\|,$$

obtenemos que

$$\begin{aligned} \sum_{k=0}^{\infty} (t_k \|\mathbf{d}_k\|)^2 &\leq \sum_{k=0}^{\infty} \left(\frac{1 + \eta_k}{\sigma_1} \|\Phi_\lambda(\mathbf{x}_k)\| - \frac{1}{\sigma_1} \|\Phi_\lambda(\mathbf{x}_{k+1})\| \right) \\ &= \sum_{k=0}^{\infty} \frac{1}{\sigma_1} (\|\Phi_\lambda(\mathbf{x}_k)\| - \|\Phi_\lambda(\mathbf{x}_{k+1})\|) + \sum_{k=0}^{\infty} \frac{\eta_k}{\sigma_1} \|\Phi_\lambda(\mathbf{x}_k)\| \\ &\leq \frac{1}{\sigma_1} \|\Phi_\lambda(\mathbf{x}_0)\| + \sum_{k=0}^{\infty} \frac{\eta_k}{\sigma_1} e^\eta \|\Phi_\lambda(\mathbf{x}_0)\| \\ &\leq \left(\frac{1}{\sigma_1} + \sum_{k=0}^{\infty} \frac{\eta_k}{\sigma_1} e^\eta \right) \|\Phi_\lambda(\mathbf{x}_0)\|. \end{aligned}$$

Por otra parte, dado que los términos $(t_k \|\mathbf{d}_k\|)^2$ son no negativos, la serie

$$\sum_{k=0}^{\infty} (t_k \|\mathbf{d}_k\|)^2$$

converge.

Para demostrar que $\{\mathbf{x}_k\}$ es convergente, definamos $m \geq l$ y consideremos

$$\|\mathbf{x}_m - \mathbf{x}_l\|^2 \leq \sum_{k=l}^{m-1} (t_k \|\mathbf{d}_k\|)^2 \leq \sum_{k=l}^{\infty} (t_k \|\mathbf{d}_k\|)^2.$$

Teniendo en cuenta que

$$\sum_{k=l}^{\infty} (t_k \|\mathbf{d}_k\|)^2 \leq \sum_{k=0}^{\infty} (t_k \|\mathbf{d}_k\|)^2 - \sum_{k=0}^{l-1} (t_k \|\mathbf{d}_k\|)^2$$

tiende a cero cuando l tiende a infinito entonces, para todo $\epsilon > 0$, existe l suficientemente grande tal que $\|\mathbf{x}_m - \mathbf{x}_l\|^2 \leq \epsilon$ siempre que $m \geq l$, por lo cual $\|\mathbf{x}_m - \mathbf{x}_l\|$

tiende a cero cuando l tiende a infinito. Esto significa que $\{\mathbf{x}_k\}$ es una sucesión de *Cauchy* y por lo tanto, converge. \square

El siguiente lema garantiza que los puntos de acumulación de la sucesión generada por el algoritmo son soluciones del PCG(F, G).

Lema 7.5. *Supongamos que las hipótesis $\bar{\mathbf{H}}0$ a $\bar{\mathbf{H}}4$ se verifican. Entonces todo punto de acumulación de la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 4** es solución de (1.6).*

Demostración. Del **Lema 7.3** tenemos que la sucesión $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ converge. Luego, es suficiente demostrar que existe un punto de acumulación de la sucesión $\{\mathbf{x}_k\}$, el cual es solución de (1.6).

Si existen infinitos k , para los cuales t_k es determinado por el Paso 3 del **Algoritmo 4** entonces, para esos valores de k , se cumple que

$$\|\Phi_\lambda(\mathbf{x}_{k+1})\| \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_2 \|t_k \mathbf{d}_k\|^2 \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| \leq \rho^{k+1} \|\Phi_\lambda(\mathbf{x}_0)\|,$$

por lo cual $\|\Phi_\lambda(\mathbf{x}_k)\| \leq \rho^k \|\Phi_\lambda(\mathbf{x}_0)\|$. Pasando al límite cuando k tiende a infinito, se tiene que

$$\lim_{k \rightarrow \infty} \|\Phi_\lambda(\mathbf{x}_k)\| = 0, \quad (7.10)$$

dato que $0 < \rho < 1$, con lo cual existe un punto \mathbf{x}_* tal que la sucesión $\{\mathbf{x}_k\}_{k \in K} \subset \Omega$ converge a \mathbf{x}_* . Por continuidad de Φ_λ y de la norma, nuevamente, pasando al límite cuando k tiende a infinito $\Phi_\lambda(\mathbf{x}_*) = 0$.

Por otro lado, si t_k es determinado por (7.24) para todo k suficientemente grande, por la hipótesis $\bar{\mathbf{H}}4$ y el **Lema 7.4**, sin pérdida de generalidad, podemos asumir que las sucesiones $\{\mathbf{d}_k\}$ y $\{\mathbf{B}_k\}$ convergen a algún vector $\bar{\mathbf{d}}$ y alguna matriz $\tilde{H} \in \partial\Phi_\lambda(\mathbf{x}_*)$, respectivamente. Tomando el límite en (7.3), cuando k tiende a infinito, tenemos que

$$\tilde{H}\bar{\mathbf{d}} + \Phi_\lambda(\mathbf{x}_*) = 0. \quad (7.11)$$

Sea $\bar{t} = \limsup_{k \in K} t_k$. Entonces $\bar{t} \geq 0$ y $\bar{t}\bar{\mathbf{d}} = 0$. Si $\bar{\mathbf{d}} = 0$ entonces de (7.11), $\Phi_\lambda(\mathbf{x}_*) = 0$. Si $\bar{t} = 0$ ($\limsup_{k \in K} t_k = 0$) entonces para $k \in K$ suficientemente grande, $t'_k = \frac{t_k}{\beta}$ no satisface (7.5), de ahí que $\|\Phi_\lambda(\mathbf{x}_k + t'_k \mathbf{d}_k)\| > \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_1 \|t'_k \mathbf{d}_k\|^2 + \eta_k \|\Phi_\lambda(\mathbf{x}_k)\|$, por lo cual

$$\left\| \Phi_\lambda(\mathbf{x}_k + t'_k \mathbf{d}_k) \right\| - \|\Phi_\lambda(\mathbf{x}_k)\| \geq -\sigma_1 \|t'_k \mathbf{d}_k\|^2. \quad (7.12)$$

Multiplicando (7.12) por $\frac{\|\Phi_\lambda(\mathbf{x}_k + t'_k \mathbf{d}_k)\| + \|\Phi_\lambda(\mathbf{x}_k)\|}{t'_k}$ y tomando límite, cuando

$k \in K$ tiende a infinito, obtenemos que $\tilde{H}^T \Phi_\lambda(\mathbf{x}_*) \bar{\mathbf{d}} \geq 0$. De esto último y por (7.11), se sigue que $\Phi_\lambda(\mathbf{x}_*) = 0$, con lo cual completamos la demostración. \square

El siguiente teorema permite garantizar que para k suficientemente grande la dirección calculada en el Paso 2 del **Algoritmo 4** es de descenso; más aún, para k suficientemente grande $t_k = 1$ será aceptado en el Paso 3 del algoritmo.

Teorema 7.2. *Sea $\{\mathbf{x}_k\}$ una sucesión generada por el **Algoritmo 4**. Supongamos que las hipótesis $\bar{\mathbf{H0}}$ a $\bar{\mathbf{H4}}$ se satisfacen. Entonces existe $\bar{k} \geq 0$ tal que*

$$\|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_2 \|\mathbf{d}_k\|^2 < \rho \|\Phi_\lambda(\mathbf{x}_k)\|, \quad (7.13)$$

siempre que $k \geq \bar{k}$.

Demostración. Por el lema anterior, existe un índice $\bar{k} \geq 0$ tal que (7.10) se satisface, para todo $k \geq \bar{k}$. Haciendo algunas manipulaciones algebraicas obtenemos que

$$\begin{aligned} \|\mathbf{x}_k + \mathbf{d}_k - \mathbf{x}_*\| &= \|\mathbf{x}_k + \mathbf{d}_k - \mathbf{x}_* - H_*^{-1} \Phi_\lambda(\mathbf{x}_k) + H_*^{-1} \Phi_\lambda(\mathbf{x}_k)\| \\ &\leq \|\mathbf{x}_k - \mathbf{x}_* - H_*^{-1} \Phi_\lambda(\mathbf{x}_k)\| + \|\mathbf{d}_k + H_*^{-1} \Phi_\lambda(\mathbf{x}_k)\| \\ &\leq \|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}_*) + \Phi_\lambda(\mathbf{x}_*) - \Phi_\lambda(\mathbf{x}_k)\| \\ &\quad + \|H_*^{-1}\| \|H_* \mathbf{d}_k + \Phi_\lambda(\mathbf{x}_k)\|. \end{aligned}$$

Ahora, usando el hecho de que existe una constante μ tal que $\|H_*^{-1}\| \leq \mu$, obtenemos que

$$\begin{aligned} \|\mathbf{x}_k + \mathbf{d}_k - \mathbf{x}_*\| &\leq \mu \left[\frac{\|H_*(\mathbf{x}_k - \mathbf{x}_*) + \Phi_\lambda(\mathbf{x}_*) - \Phi_\lambda(\mathbf{x}_k)\| \|\mathbf{x}_k - \mathbf{x}_*\|}{\|\mathbf{x}_k - \mathbf{x}_*\|} \right. \\ &\quad \left. + \frac{\|(B_k - H_*) \mathbf{d}_k\| \|\mathbf{d}_k\|}{\|\mathbf{d}_k\|} \right]. \end{aligned} \quad (7.14)$$

Por la continuidad Lipschitz de Φ_λ y por (7.14), tenemos que

$$\begin{aligned} \|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| &= \|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k) - \Phi_\lambda(\mathbf{x}_*)\| \\ &\leq M_2 \|\mathbf{x}_k + \mathbf{d}_k - \mathbf{x}_*\| \end{aligned} \quad (7.15)$$

$$\begin{aligned} &\leq M_2 \mu \left[\frac{\|H_*(\mathbf{x}_k - \mathbf{x}_*) + \Phi_\lambda(\mathbf{x}_*) - \Phi_\lambda(\mathbf{x}_k)\| \|\mathbf{x}_k - \mathbf{x}_*\|}{\|\mathbf{x}_k - \mathbf{x}_*\|} \right. \\ &\quad \left. + \frac{\|(B_k - H_*) \mathbf{d}_k\| \|\mathbf{d}_k\|}{\|\mathbf{d}_k\|} \right]. \end{aligned} \quad (7.16)$$

Puesto que la función Φ_λ es *semisuave* en \mathbf{x}^* [51, 59] entonces

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}^*) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0, \quad (7.17)$$

donde $H_* \in \partial\Phi_\lambda(\mathbf{x}^*)$. Ahora, como $\Phi_\lambda(\mathbf{x}^*) = 0$ entonces

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_k) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0. \quad (7.18)$$

Por otra parte, dado que

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq \|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|,$$

y

$$\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right| \leq \|\Phi_\lambda(\mathbf{x}_k) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|.$$

tenemos que,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right|}{\|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} = 0,$$

por lo cual,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right|}{\|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} = 0.$$

De (7.18), en particular, tenemos que para $\bar{\epsilon} = \frac{1}{2}$, existe $\bar{\delta} > 0$, tal que si $\|\mathbf{x}_k - \mathbf{x}^*\| < \bar{\delta}$ entonces

$$-\frac{1}{2} < \frac{\|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|}{\|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} < \frac{1}{2},$$

de donde

$$\frac{1}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| < \|\Phi_\lambda(\mathbf{x}_k)\| < \frac{3}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|,$$

por lo cual

$$\|\Phi_\lambda(\mathbf{x}_k)\| > \frac{1}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \geq \frac{1}{2 \|H_*^{-1}\|} \|\mathbf{x}_k - \mathbf{x}^*\|; \quad (7.19)$$

es decir, existe $m > 0$, tal que

$$\|\Phi_\lambda(\mathbf{x}_k)\| \geq m \|\mathbf{x}_k - \mathbf{x}^*\|. \quad (7.20)$$

Por último, denotando

$$\Delta_1 = \frac{\|H_*(\mathbf{x}_k - \mathbf{x}_*) + \Phi_\lambda(\mathbf{x}_*) - \Phi_\lambda(\mathbf{x}_k)\|}{\|\mathbf{x}_k - \mathbf{x}_*\|} \quad y \quad \Delta_2 = \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} \quad (7.21)$$

y de (7.9), (7.15) y (7.20), se sigue que

$$\Delta \leq C_1\mu[\Delta_1 \|\mathbf{x}_k - \mathbf{x}_*\| + \Delta_2 \|\mathbf{d}_k\|] - \rho m \|\mathbf{x}_k - \mathbf{x}^*\| + \sigma_2 T_1 \|\Phi_\lambda(\mathbf{x}_k)\|^2,$$

donde $\Delta = \|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| - \rho \|\Phi_\lambda(\mathbf{x}_k)\| + \sigma_2 \|\mathbf{d}_k\|^2$.

Ahora, por la hipótesis $\bar{\mathbf{H}}1$, (7.17) y pasando al límite cuando $k \rightarrow \infty$, obtenemos

$$\Delta = \|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| - \rho \|\Phi_\lambda(\mathbf{x}_k)\| + \sigma_2 \|\mathbf{d}_k\|^2 \leq 0;$$

es decir, se satisface la primera desigualdad en (7.13). Para esos mismos valores de k , el Paso 2 del **Algoritmo 4** garantiza que se alcanza el paso completo en la actualización. La segunda desigualdad se concluye de manera inmediata. \square

El siguiente teorema garantiza que la tasa de convergencia local del **Algoritmo 4** es q -superlineal.

Teorema 7.3. *Sea $\{\mathbf{x}_k\}$ una sucesión generada por el **Algoritmo 4**. Supongamos que las hipótesis $\bar{\mathbf{H}}0$ a $\bar{\mathbf{H}}5$ se satisfacen. Entonces la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 4** converge q -superlinealmente a una solución del sistema de ecuaciones no lineales $\Phi_\lambda(\mathbf{x}) = 0$.*

Demostración. Por el **Lema 7.5**, la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 4** converge a una solución de $\Phi_\lambda(\mathbf{x}) = 0$, y el Teorema 7.2 garantiza que a partir de un cierto \bar{k} , el **Algoritmo 4** alcanza el paso completo en la actualización; es decir, $t_k = 1$, para todo $k > \bar{k}$, por lo cual $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$. Para demostrar que la tasa de convergencia es superlineal usaremos un argumento similar al usado en [3].

A partir de (7.3) y después de algunas manipulaciones algebraicas, obtenemos

$$\begin{aligned} \|\Phi_\lambda(\mathbf{x}_{k+1})\| &= \|B_k \mathbf{d}_k - H_* \mathbf{d}_k + \Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})\| \\ &\leq \|(B_k - H_*) \mathbf{d}_k\| + \|\Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})\| \end{aligned}$$

de donde

$$\frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|} \leq \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} + \frac{\|\Phi_\lambda(\mathbf{x}_k) + H_* \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|}.$$

Usando la hipótesis $\bar{\mathbf{H}}3$, la semisuavidad de la función Φ_λ (que implica semicontinuidad [52]) y pasando al límite cuando $\mathbf{x}_k \rightarrow \mathbf{x}^*$, obtenemos que

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|} = 0. \quad (7.22)$$

De (7.22), (7.19) y la cota para $\|H_*^{-1}\|$, deducimos que

$$\begin{aligned}
0 &= \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{d}_k\|} \geq \frac{1}{2\|H_*^{-1}\|} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{d}_k\|} \\
&= \frac{1}{2\|H_*^{-1}\|} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_{k+1} - \mathbf{x}^* + \mathbf{x}^* - \mathbf{x}_k\|} \\
&\geq \frac{1}{2\mu} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_{k+1} - \mathbf{x}^*\| + \|\mathbf{x}_k - \mathbf{x}^*\|} \\
&= \frac{1}{2\mu} \lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|}}{\frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} + 1}.
\end{aligned}$$

Por lo tanto,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0;$$

es decir, la sucesión $\{\mathbf{x}_k\}$ converge q -superlinealmente a \mathbf{x}_* . \square

7.3. Familia de métodos secantes

En el **Algoritmo 4**, las matrices B_k son solo aproximaciones a las H_k , lo que hace de él un algoritmo genérico. En esta sección, proponemos un algoritmo, en el cual actualizamos dichas matrices en forma estructurada, teniendo en cuenta la teoría secante desarrollada en el capítulo anterior.

Por otra parte, las actualizaciones buena y mala de Broyden (6.34) y (6.35), tienen una estructura particular, que se puede generalizar. En efecto, ellas son casos particulares de las actualizaciones:

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{v}_k^T}{\mathbf{v}_k^T \mathbf{s}_k} \quad C_{k+1} = C_k + \frac{(\hat{\mathbf{y}}_k - C_k \mathbf{s}_k) \hat{\mathbf{v}}_k^T}{\hat{\mathbf{v}}_k^T \mathbf{s}_k},$$

con $\mathbf{v}_k = \hat{\mathbf{v}}_k = \mathbf{s}_k$, para Broyden buena y $\mathbf{v}_k = A_k \mathbf{s}_k$, $\hat{\mathbf{v}}_k = C_k \mathbf{s}_k$, para Broyden mala. Teniendo en cuenta esto, el algoritmo que proponemos a continuación permite actualizaciones tipo Broyden, con otras elecciones de \mathbf{v}_k y $\hat{\mathbf{v}}_k$.

El algoritmo es el siguiente.

Algoritmo 5

Entrada: $\mathbf{x}_0 \in \mathbb{R}^n$, $A_0, C_0 \in \mathbb{R}^{n \times n}$ no singulares, $\lambda \in (0, 4)$, $\sigma_1, \sigma_2 > 0$, $\beta, \rho, r \in (0, 1)$, $p > 2$, $\epsilon \geq 0$ $\{\eta_k\}$ una sucesión de términos positivos que satisface (7.2) y $k = 0$.

Salida: Aproximación a la solución del sistema (1.6).

- 1: **mientras** $\|\Phi_\lambda(\mathbf{x}_k)\| > \epsilon$ **hacer**
- 2: Sea $B_k = D_a(\mathbf{x}_k)A_k + D_b(\mathbf{x}_k)C_k + r_k I$, con $r_k \in [0, 1]$, tal que B_k es no singular.
- 3: Encuentre una solución \mathbf{d}_k al sistema de ecuaciones lineales:

$$B_k \mathbf{d}_k = -\Phi_\lambda(\mathbf{x}_k).$$

- 4: **si**

$$\|\Phi_\lambda(\mathbf{x}_k + \mathbf{d}_k)\| \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_2 \|\mathbf{d}_k\|^2 \quad (7.23)$$

entonces

- 5: hacer $t_k = 1$ y pasar al Paso 7

- 6: **sino**

- 7: Haga $t_k = \max\{\beta^l : l = 0, 1, 2, \dots\}$, tal que

$$\|\Phi_\lambda(\mathbf{x}_k + t_k \mathbf{d}_k)\| \leq \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_1 \|t_k \mathbf{d}_k\|^2 + \eta_k \|\Phi_\lambda(\mathbf{x}_k)\|. \quad (7.24)$$

- 8: Actualizar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_k,$$

- 9: **fin si**

- 10: Actualización de B_k : actualice A_k y C_k usando (6.34) o (6.35). Hacer $k = k + 1$ y volver al Paso 1.

- 11: **fin mientras**

- 12: **devolver** \mathbf{x}_*

Respecto a la convergencia del **Algoritmo 5**, teniendo en cuenta que es un caso particular del **Algoritmo 4** y que la teoría de este último fue desarrollada en la sección anterior bajo las hipótesis $\bar{\mathbf{H}}0$ a $\bar{\mathbf{H}}5$, para garantizar su convergencia basta verificar estos supuestos. Esto lo abordamos a continuación.

En primer lugar, observamos que las hipótesis $\bar{\mathbf{H}}0$ - $\bar{\mathbf{H}}3$ son hipótesis del nuevo problema. Por lo tanto, las asumimos como verdaderas sin verificación y solo verificaremos las hipótesis $\bar{\mathbf{H}}4$ y $\bar{\mathbf{H}}5$.

7.3.1. Verificación de la Hipótesis $\bar{\mathbf{H}}4$.

Recordemos la Hipótesis $\bar{\mathbf{H}}4$: $\{B_k\}$ es una sucesión de matrices no singulares y acotadas, tal que $\|B_k^{-1}\| \leq T$, para todo k , con T una constante positiva.

Por la forma en que se construyen las matrices B_k en el algoritmo, estas son no singulares; más aún, por la **Proposición 5.1** y la propiedad de deteriorización acotada de las matrices A_k y C_k dada en el **Lema 4.5**, las matrices B_k son acotadas. Falta verificar que la inversa de B_k es uniformemente acotada, para todo k . Para ello, demostramos algunos resultados previos.

Lema 7.6. Sean \mathbf{s}_k , $\{A_k\}$ y $\{C_k\}$ sucesiones generadas por el **Algoritmo 5**. Defina

$$M_{k+1} = \int_0^1 F'(\mathbf{x}_k + t\mathbf{d}_k)dt, \quad N_{k+1} = \int_0^1 G'(\mathbf{x}_k + t\mathbf{d}_k)dt, \quad (7.25)$$

$$\zeta_k = \frac{\|A_k\mathbf{s}_k - \mathbf{y}_k\|}{\|\mathbf{s}_k\|} \quad y \quad \hat{\zeta}_k = \frac{\|C_k\mathbf{s}_k - \hat{\mathbf{y}}_k\|}{\|\mathbf{s}_k\|}.$$

Si

$$\sum_{k=0}^{\infty} \|\mathbf{s}_k\|^2 < \infty \quad (7.26)$$

entonces ζ_k y $\hat{\zeta}_k$ convergen a cero.

Demostración. De (7.25), se deduce que $\mathbf{y}_k = M_{k+1}\mathbf{s}_k$ y $\hat{\mathbf{y}}_k = N_{k+1}\mathbf{s}_k$.

Además, por la hipótesis **H2**, deducimos que

$$\begin{aligned} \|M_{k+1} - M_k\| &\leq \int_0^1 \left\| F'(\mathbf{x}_k + t\mathbf{d}_k) - F'(\mathbf{x}_{k-1} + t\mathbf{d}_{k-1}) \right\|_F dt \\ &\leq \sqrt{2}\gamma \int_0^1 \|\mathbf{x}_k + t\mathbf{d}_k - \mathbf{x}_{k-1} + t\mathbf{d}_{k-1}\| dt \\ &= \sqrt{2}\gamma \int_0^1 \|(\mathbf{x}_k - \mathbf{x}_{k-1}) + t(\mathbf{d}_k - \mathbf{d}_{k-1})\| dt \\ &= \sqrt{2}\gamma \int_0^1 \|t_{k-1}\mathbf{d}_{k-1} + t(\mathbf{d}_k - \mathbf{d}_{k-1})\| dt \\ &\leq \sqrt{2}\gamma \int_0^1 (t_{k-1}\|\mathbf{d}_{k-1}\| + t\|\mathbf{d}_k - \mathbf{d}_{k-1}\|) dt \end{aligned}$$

$$\begin{aligned}
&= \sqrt{2}\gamma \left[t_{k-1} \|\mathbf{d}_{k-1}\| + \frac{1}{2} \|\mathbf{d}_k - \mathbf{d}_{k-1}\| \right] \\
&\leq \sqrt{2}\gamma \left[\|\mathbf{d}_{k-1}\| + \frac{1}{2} \|\mathbf{d}_k\| + \frac{1}{2} \|\mathbf{d}_{k-1}\| \right] \\
&= \frac{1}{2} \sqrt{2}\gamma [3 \|\mathbf{d}_{k-1}\| + \|\mathbf{d}_k\|], \tag{7.27}
\end{aligned}$$

análogamente,

$$\|N_{k+1} - N_k\| \leq \frac{1}{2} \sqrt{2}\gamma [3 \|\mathbf{d}_{k-1}\| + \|\mathbf{d}_k\|]. \tag{7.28}$$

Luego,

$$\zeta_k = \frac{\|A_k \mathbf{s}_k - \mathbf{y}_k\|}{\|\mathbf{s}_k\|} = \frac{\|M_{k+1} \mathbf{s}_k - A_k \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = \frac{\|(M_{k+1} - A_k) \mathbf{d}_k\|}{\|\mathbf{d}_k\|}. \tag{7.29}$$

Análogamente,

$$\hat{\zeta}_k = \frac{\|A_k \mathbf{s}_k - \hat{\mathbf{y}}_k\|}{\|\mathbf{s}_k\|} = \frac{\|N_{k+1} \mathbf{s}_k - C_k \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = \frac{\|(N_{k+1} - C_k) \mathbf{d}_k\|}{\|\mathbf{d}_k\|}. \tag{7.30}$$

Definamos:

$$a_k = \|A_k - M_k\|_F \quad y \quad b_k = \|M_{k+1} - M_k\|_F.$$

Por propiedades de la norma de *Frobenius* y de la norma dos, tenemos que

$$\begin{aligned}
a_{k+1}^2 &= \|A_{k+1} - M_{k+1}\|_F^2 \\
&= \left\| A_k + (\mathbf{y}_k - A_k \mathbf{s}_k) \frac{\mathbf{v}_k^T}{\mathbf{v}_k^T \mathbf{s}_k} - M_{k+1} \right\|_F^2 \\
&= \left\| (A_k - M_{k+1}) \left(I - \frac{\mathbf{v}_k^T}{\mathbf{v}_k^T \mathbf{s}_k} \right) \right\|_F^2 \\
&= \left\| (A_k - M_{k+1}) - (A_k - M_{k+1}) \left(I - \frac{\mathbf{v}_k^T}{\mathbf{v}_k^T \mathbf{s}_k} \right) \right\|_F^2 \\
&= \|A_k - M_{k+1}\|_F^2 - \frac{\|(A_k - M_{k+1}) \mathbf{s}_k \mathbf{v}_k^T\|_F^2}{\|\mathbf{v}_k\|_F^2 \|\mathbf{s}_k\|_F^2}
\end{aligned}$$

$$\begin{aligned}
&= \|A_k - M_{k+1}\|^2 - \frac{\|(A_k \mathbf{s}_k - \mathbf{y}_k) \mathbf{s}_k^T\|^2}{\|\mathbf{s}_k\|^2} \\
&= \|A_k - M_{k+1}\|^2 - \zeta_k^2 \\
&\leq (\|A_k - M_k\| + \|M_k - M_{k+1}\|)^2 - \zeta_k^2 \\
&= (a_k + b_k)^2 - \zeta_k^2.
\end{aligned}$$

Así, por (7.26) y (7.27), deducimos que $\{b_k\}$ satisface las hipótesis del **Lema 4.7**; es decir,

$$\sum_{k=0}^{\infty} \|b_k\| < \infty \quad y \quad \sum_{k=0}^{\infty} \|b_k\|^2 < \infty.$$

En forma análoga, para

$$\hat{a}_k = \|C_k - N_k\|_F \quad y \quad \hat{b}_k = \|N_{k+1} - N_k\|_F,$$

se tiene que

$$\hat{a}_{k+1}^2 \leq (\hat{a}_k + \hat{b}_k)^2 - \hat{\zeta}_k^2,$$

por (7.26) y (7.28), $\{\hat{b}_k\}$ también satisface las hipótesis del **Lema 4.7**; es decir,

$$\sum_{k=0}^{\infty} \|\hat{b}_k\| < \infty \quad y \quad \sum_{k=0}^{\infty} \|\hat{b}_k\|^2 < \infty;$$

por lo tanto,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \hat{\zeta}_i^2 = 0 = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \zeta_i^2, \quad \sum_{k=0}^{\infty} \hat{\zeta}_k^2 < \infty \quad y \quad \sum_{k=0}^{\infty} \zeta_k^2 < \infty,$$

y en particular, las sucesiones $\{\zeta_k\}$ y $\{\hat{\zeta}_k\}$ tienden a cero. \square

Lema 7.7. Sea $B_k = D_a(\mathbf{x}_k)A_k + D_b(\mathbf{x}_k)C_k + r_k I$, con $r_k \in [0, 1]$, $k = 0, 1, \dots$. Defina

$$\hat{B}_k = D_a(\mathbf{x}_{k-1})M_k + D_b(\mathbf{x}_{k-1})N_k + r_{k-1}I, \quad k \geq 1. \quad (7.31)$$

Entonces

$$\bar{\zeta}_k = \frac{\|(\hat{B}_{k+1} - B_k) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} = \frac{\|(\hat{B}_{k+1} - B_k) \mathbf{s}_k\|}{\|\mathbf{s}_k\|}$$

tiende a cero cuando k tiende a infinito.

Demostración. Por hipótesis y (7.31) obtenemos que

$$\begin{aligned}
\frac{\left\| \left(\hat{B}_{k+1} - B_k \right) \mathbf{s}_k \right\|}{\|\mathbf{s}_k\|} &= \frac{\|D_a(\mathbf{x}_k)M_{k+1}\mathbf{s}_k + D_b(\mathbf{x}_k)N_{k+1}\mathbf{s}_k - (D_a(\mathbf{x}_k)A_k + D_a(\mathbf{x}_k)C_k)\mathbf{s}_k\|}{\|\mathbf{s}_k\|} \\
&= \frac{\|D_a(\mathbf{x}_k)(M_{k+1} - A_k)\mathbf{s}_k + D_b(\mathbf{x}_k)(N_{k+1} - C_k)\mathbf{s}_k\|}{\|\mathbf{s}_k\|} \\
&\leq \|D_a(\mathbf{x}_k)\| \frac{\|(M_{k+1} - A_k)\mathbf{s}_k\|}{\|\mathbf{s}_k\|} + \|D_b(\mathbf{x}_k)\| \frac{\|(N_{k+1} - C_k)\mathbf{s}_k\|}{\|\mathbf{s}_k\|} \\
&= \|D_a(\mathbf{x}_k)\| \zeta_k + \|D_b(\mathbf{x}_k)\| \hat{\zeta}_k,
\end{aligned} \tag{7.32}$$

donde la última igualdad se obtiene por (7.29) y (7.30). Por lo tanto, el **Lema 7.6** garantiza que ζ_k tiende a cero, cuando k tiende a infinito. \square

El siguiente lema permite garantizar que el tamaño de las direcciones no crece en forma descontrolada.

Lema 7.8. *Supongamos que las hipótesis $\bar{\mathbf{H}}1$ a $\bar{\mathbf{H}}3$ se verifican. Entonces*

$$\|\mathbf{d}_k\| \leq T \|\Phi_\lambda(\mathbf{x}_k)\|.$$

Demostración. Del **Lema 7.3**, tenemos que la sucesión $\{\|\Phi_\lambda(\mathbf{x}_k)\|\}$ converge. Así, es suficiente demostrar que existe un punto de acumulación de la sucesión $\{\mathbf{x}_k\}$, el cual es solución de (1.6). Si existen infinitos k para los cuales t_k es determinado como en el paso 3 del **Algoritmo 5** entonces, para infinitos valores de k , se cumple que

$$\begin{aligned}
\|\Phi_\lambda(\mathbf{x}_{k+1})\| &\leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| - \sigma_2 \|t_k \mathbf{d}_k\|^2 \leq \rho \|\Phi_\lambda(\mathbf{x}_k)\| \\
&\leq \rho^{k+1} \|\Phi_\lambda(\mathbf{x}_0)\|,
\end{aligned}$$

por lo cual $\|\Phi_\lambda(\mathbf{x}_k)\| \leq \rho^k \|\Phi_\lambda(\mathbf{x}_0)\|$ y puesto que $0 < \rho < 1$ entonces, pasando al límite cuando k tiende a infinito, se tiene que

$$\lim_{k \rightarrow \infty} \|\Phi_\lambda(\mathbf{x}_k)\| = 0,$$

por lo cual existe un punto $\bar{\mathbf{x}}$ tal que la sucesión $\{\mathbf{x}_k\}_{k \in K} \subset \Omega$ converge a $\bar{\mathbf{x}}$ y además, $\Phi_\lambda(\bar{\mathbf{x}}) = 0$. Por lo tanto, para todo k suficientemente grande, el **Lema 6.3** garantiza que las matrices B_k , con $r_k = 0$, son no singulares y, por (6.15), existe una

constante positiva μ tal que $\|B_k^{-1}\| \leq \frac{4}{3}\mu$. Así, por (7.3) y (7.32), tenemos que

$$\begin{aligned} \|\mathbf{d}_k\| &= \left\| \hat{B}_{k+1}^{-1} \left(\hat{B}_{k+1} - B_k \right) \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_k) \right\| \\ &\leq \left\| \hat{B}_{k+1}^{-1} \right\| \left\| \left(\hat{B}_{k+1} - B_k \right) \mathbf{d}_k \right\| + \|\Phi_\lambda(\mathbf{x}_k)\| \\ &\leq \frac{4}{3}\mu \left(\bar{\zeta}_k \|\mathbf{d}_k\| + \|\Phi_\lambda(\mathbf{x}_k)\| \right), \end{aligned}$$

y puesto que $\bar{\zeta}_k$ tiende a cero, cuando k tiende a infinito, usando el concepto de límite de una sucesión, existe una constante positiva T tal que $\|\mathbf{d}_k\| \leq T \|\Phi_\lambda(\mathbf{x}_k)\|$.

Por otra parte, supongamos que t_k es determinado por (7.24), para todo k suficientemente grande. Por el **Lema 7.7**, existe una subsucesión $\{\bar{\zeta}\}_{k \in K}$ de $\{\bar{\zeta}_k\}$ que converge a cero y puesto que la sucesión $\{\mathbf{x}_k\}_{k \in K}$ es acotada, sin perder generalidad, podemos asumir que converge a algún $\bar{\mathbf{x}}$. Además, la sucesión $\{r_k\}_{k \in K}$ converge a cero, dado que el **Lema 6.3** garantiza que las matrices B_k , con $r_k = 0$, son no singulares para todo k suficientemente grande. Además, el **Lema 7.2** implica que $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ tiende a cero cuando k tiende a infinito, de ahí que \hat{B}_{k+1} tiende a $D_a(\bar{\mathbf{x}})F'(\bar{\mathbf{x}}) + D_b(\bar{\mathbf{x}})G'(\bar{\mathbf{x}}) = \bar{H} \in \partial_B \Phi_\lambda(\bar{\mathbf{x}})$ y, por (6.15), existe una constante positiva μ tal que $\left\| \hat{B}_{k+1}^{-1} \right\| \leq \frac{4}{3}\mu$. De esta manera, por (7.3) y (7.32), obtenemos que

$$\begin{aligned} \|\mathbf{d}_k\| &= \left\| \hat{B}_{k+1}^{-1} \left(\hat{B}_{k+1} - B_k \right) \mathbf{d}_k - \Phi_\lambda(\mathbf{x}_k) \right\| \\ &\leq \left\| \hat{B}_{k+1}^{-1} \right\| \left\| \left(\hat{B}_{k+1} - B_k \right) \mathbf{d}_k \right\| + \|\Phi_\lambda(\mathbf{x}_k)\| \\ &\leq \frac{4}{3}\mu \left(\bar{\zeta}_k \|\mathbf{d}_k\| + \|\Phi_\lambda(\mathbf{x}_k)\| \right), \end{aligned}$$

y puesto que $\bar{\zeta}_k$ tiende a cero cuando k tiende a infinito, usando el concepto de límite de una sucesión, existe una constante positiva T tal que $\|\mathbf{d}_k\| \leq T \|\Phi_\lambda(\mathbf{x}_k)\|$. \square

7.3.2. Verificación de la Hipótesis $\bar{\mathbf{H}}5$.

Hasta aquí, hemos verificado que la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 5** satisface las hipótesis $\bar{\mathbf{H}}0$ a $\bar{\mathbf{H}}4$. Falta verificar la tasa de convergencia del algoritmo. Para ello, debemos garantizar que la hipótesis $\bar{\mathbf{H}}5$ se cumple. Es decir, que las matrices

B_k satisfacen el siguiente límite:

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = 0,$$

donde $H_* \in \partial\Phi_\lambda(\mathbf{x}_*)$ y $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k = t_k \mathbf{d}_k$.

Sumando y restando la matriz \hat{B}_{k+1} , definida en (7.32), y usando desigualdad triangular, tenemos lo siguiente:

$$\begin{aligned} \frac{\|(B_k - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} &\leq \frac{\|(\hat{B}_{k+1} - H_*) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} + \frac{\|(B_k - \hat{B}_{k+1}) \mathbf{d}_k\|}{\|\mathbf{d}_k\|} \\ &\leq \|\hat{B}_{k+1} - H_*\| + \frac{\|(B_k - \hat{B}_{k+1}) \mathbf{d}_k\|}{\|\mathbf{d}_k\|}. \end{aligned} \quad (7.33)$$

Dado que la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 5** converge a una solución del problema. Es decir, \mathbf{x}_k tiende a \mathbf{x}_* , cuando k tiende a infinito, con lo cual r_k tiende a cero y, en consecuencia, \hat{B}_{k+1} tiende a H_* . Por otro lado, el **Lema 7.7** garantiza que, el segundo sumando en (7.33) tiende a cero, lo que permite concluir que

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = 0,$$

con lo cual se verifica la hipótesis $\bar{\mathbf{H}}5$.

Dado que las cinco hipótesis se satisfacen, concluimos que la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 5** satisface los resultados obtenidos en la **Sección 7.2**. Por lo tanto, la sucesión generada por este algoritmo está bien definida y converge a la solución \mathbf{x}_* de (1.6). Además, para todo k suficientemente grande, el tamaño de paso $t_k = 1$ es aceptado, y la sucesión converge q -superlinealmente a \mathbf{x}_* .

7.4. Experimentación numérica

En esta sección, analizamos el desempeño numérico del **Algoritmo 5**, propuesto en la **Sección 7.3**, al que llamamos **Algoritmo SCM**. Para ello, lo comparamos con dos algoritmos propuestos en los **Capítulos 5** y **6**, respectivamente. El primero, es el **Algoritmo 1** actualizando el parámetro λ mediante la escogencia dinámica propuesta en [34], el cual llamaremos **Algoritmo (NGD)**. El segundo, es el **Algoritmo 3** que llamaremos **Algoritmo SCML**.

Hicimos cinco tipos de experimentos. Primero, comparamos el desempeño de los

Algoritmos SCM y NGD, en términos del tiempo (CPU) y número de iteraciones. En el segundo, tercero y cuarto, analizamos el desempeño global, la influencia del parámetro λ y la tasa de convergencia numérica del **Algoritmo SCM**, respectivamente. Finalmente, en el quinto experimento, comparamos el Algoritmo secante local (**SCML**) con el global (**SCM**), para evidenciar el impacto de introducir una búsqueda lineal en el algoritmo local. Vale la pena mencionar que para los experimentos numéricos usamos los diez problemas descritos en el **Capítulo 3**.

7.4.1. Experimento 1

En este experimento, comparamos el desempeño de los algoritmos **SCM** y **NGD**, en cuanto al tiempo (CPU) y número de iteraciones. Los resultados obtenidos están contenidos en la **Tabla 7.1** que contiene la siguiente información: Problema, puntos iniciales, actualizaciones secantes (Bb y Bm), tiempo CPU y número de iteraciones.

		Bb		Bm		NGD	
P	\mathbf{x}_0	t	k	t	k	t	k
P1	\mathbf{x}_1	10^{-2}	13	10^{-2}	56	—	—
	\mathbf{x}_2	—	—	10^{-2}	54	2.3×10^{-1}	4
	\mathbf{x}_3	0×10^{-2}	6	0×10^{-2}	6	0.42	6
	\mathbf{x}_4	0×10^{-2}	74	10^{-2}	51	1.1	11
P2	\mathbf{x}_1	3×10^{-2}	12	10^{-2}	12	0.39	10
	\mathbf{x}_2	—	—	—	—	6×10^{-1}	13
	\mathbf{x}_3	—	—	0×10^{-2}	67	0.9	19
	\mathbf{x}_4	0×10^{-2}	26	0×10^{-2}	26	0.7	20
P3	\mathbf{x}_1	10^{-2}	7	10^{-2}	7	2×10^{-1}	5
	\mathbf{x}_2	0×10^{-2}	6	0×10^{-2}	6	2×10^{-1}	5
	\mathbf{x}_3	0×10^{-2}	6	0×10^{-2}	6	2×10^{-1}	5
P4	\mathbf{x}_1	10^{-2}	7	0×10^{-2}	7	0.17	5
	\mathbf{x}_2	10^{-2}	6	0×10^{-2}	6	2×10^{-1}	4
	\mathbf{x}_3	0×10^{-2}	6	0×10^{-2}	6	0.28	4
P5	\mathbf{x}_1	—	—	10^{-2}	86	9×10^{-1}	8
	\mathbf{x}_2	0×10^{-2}	7	0×10^{-2}	7	3×10^{-1}	4
	\mathbf{x}_3	0×10^{-2}	10	0×10^{-2}	9	0.43	5
P6	\mathbf{x}_1	10^{-2}	9	0×10^{-2}	3	5×10^{-1}	6
	\mathbf{x}_2	0×10^{-2}	7	0×10^{-2}	7	0.26	5
	\mathbf{x}_3	0×10^{-2}	6	0×10^{-2}	6	0.17	4
	\mathbf{x}_4	0×10^{-2}	7	0×10^{-2}	7	2×10^{-1}	5

P7	\mathbf{x}_1	10^{-2}	8	10^{-2}	8	2×10^{-1}	5
	\mathbf{x}_2	0×10^{-2}	5	10^{-2}	5	0.18	4
	\mathbf{x}_3	0×10^{-2}	6	0×10^{-2}	6	0.15	4
P8	\mathbf{x}_1	4×10^{-2}	20	4×10^{-2}	20	—	—
	\mathbf{x}_2	10^{-2}	6	0×10^{-2}	6	1.2	22
	\mathbf{x}_3	0×10^{-2}	6	0×10^{-2}	6	6×10^{-1}	13
	\mathbf{x}_4	0×10^{-2}	6	0×10^{-2}	7	1×10^{-1}	18
P9	\mathbf{x}_1	0×10^{-2}	2	10^{-2}	2	—	—
	\mathbf{x}_2	0×10^{-2}	1	0×10^{-2}	1	0×10^{-2}	1
	\mathbf{x}_3	3×10^{-2}	2	0×10^{-2}	2	—	—
	\mathbf{x}_4	10^{-2}	7	0×10^{-2}	3	10^{-2}	5
P10	\mathbf{x}_1	3×10^{-2}	2	10^{-2}	2	—	—
	\mathbf{x}_2	0×10^{-2}	5	0×10^{-2}	3	10^{-2}	4
	\mathbf{x}_3	10^{-2}	2	0×10^{-2}	2	—	—
	\mathbf{x}_4	0×10^{-2}	6	0×10^{-2}	3	0.17	5

Tabla 7.1: Resultados del Experimento 1.

A partir de la **Tabla 7.1**, podemos observar lo siguiente: en términos del tiempo (CPU), el **Algoritmo SCM** presentó en el 100 % de los casos un tiempo similar con las dos actualizaciones y mucho menor que el del **Algoritmo NGD**. En cuanto al número de iteraciones, el **Algoritmo SCM**, con la actualización **Bb**, convergió en el 88.9 % de los casos y con la actualización **Bm**, lo hizo en el 97.2 % mientras que el **Algoritmo NGD** convergió en el 72.2 % de los casos, pero en menos iteraciones para cada punto inicial.

7.4.2. Experimento 2

En el segundo experimento, analizamos el desempeño global del **Algoritmo SCM**. Para ello, generamos 100 puntos iniciales aleatorios con componentes en el intervalo $[-30, 30]$, para los **Problemas 1, 2, 3, 4, 6, 7**; en el intervalo $[0, 50]$ para el **Problema 5** y en el intervalo $[-2, 2]$, para los problemas restantes.

Los resultados obtenidos se presentan en la **Tabla 7.2**, en la cual se observa que el **Algoritmo SCM**, con la actualización **Bm**, presentó un porcentaje de éxito mayor o igual que los otros dos. Respecto al número de iteraciones, el **Algoritmo NGD** convergió en menos iteraciones, sin embargo su tiempo CPU fue mucho mayor que el usado por los otros dos algoritmos. Por otra parte, al comparar los resultados

obtenidos del **Algoritmo SCM**, con las dos actualizaciones, observamos que tuvo un mejor desempeño con la actualización **Bm**, ya que convergió en menos iteraciones para 7 de los 10 problemas considerados.

Tabla 7.2: Algoritmo SCM para 100 puntos iniciales aleatorios. $\lambda_0 = 2$.

P		Bb	Bm	NGD
P1	<i>E</i>	0.78	0.92	0.91
	<i>t</i>	0.72	1.1	60.2
	<i>k</i>	52.7	56.4	10.5
P2	<i>E</i>	0.33	0.94	1.00
	<i>t</i>	0.01	0.29	47.9
	<i>k</i>	14.8	36.7	41.3
P3	<i>E</i>	1.00	1.00	1.00
	<i>t</i>	0.04	0.03	21.76
	<i>k</i>	5.8	5.8	4.9
P4	<i>E</i>	1.00	1.00	1.00
	<i>t</i>	0.07	0.01	20.0
	<i>k</i>	6.8	6.8	5.0
P5	<i>E</i>	1.00	1.00	1.00
	<i>t</i>	0.03	0.04	50.8
	<i>k</i>	13.7	14.0	5.7
P6	<i>E</i>	0.96	0.98	1.00
	<i>t</i>	2.9	0.2	36.9
	<i>k</i>	31.1	30.3	9.3
P7	<i>E</i>	0.99	1.00	0.93
	<i>t</i>	0.07	0.06	18.5
	<i>k</i>	43	32	43
P8	<i>E</i>	1.00	1.00	1.00
	<i>t</i>	0.03	0.03	121.0
	<i>k</i>	5.1	4.9	23.8
P9	<i>E</i>	0.77	0.94	—
	<i>t</i>	2.68	2.92	—
	<i>k</i>	63.0	61.5	—
P10	<i>E</i>	0.84	0.93	—
	<i>t</i>	2.5	3.1	—
	<i>k</i>	63.0	60.2	—

7.4.3. Experimento 3

En el tercer experimento, con el fin de identificar valores o intervalos de λ fijo, para los cuales hay buen desempeño del **Algoritmo SCM**, lo ejecutamos con distintos valores del parámetro. Exactamente, tomando 38 valores de λ igualmente espaciados en el intervalo $(0, 4)$, desde $\lambda = 0.1$ hasta $\lambda = 3.8$. Es importante mencionar que para este experimento no usamos la escogencia dinámica de λ .

Los resultados de este experimento se ilustran en las Figuras 7.1 a 7.9, donde el eje de abscisas indica los valores del parámetro λ , el de las ordenadas, el número de iteraciones (k) requeridas por el **Algoritmo SCM**, en sus dos actualizaciones, para resolver cada uno de los diez problemas. Usamos los colores rojo, azul, verde y negro para indicar el número de iteraciones del algoritmo, a partir de los puntos \mathbf{x}_1 , \mathbf{x}_2 , \mathbf{x}_3 o \mathbf{x}_4 , respectivamente.

Las figuras permiten observar que el intervalo donde se presenta el menor número de iteraciones es $[0.1, 2.0]$ y el peor, $[3.0, 3.8]$.

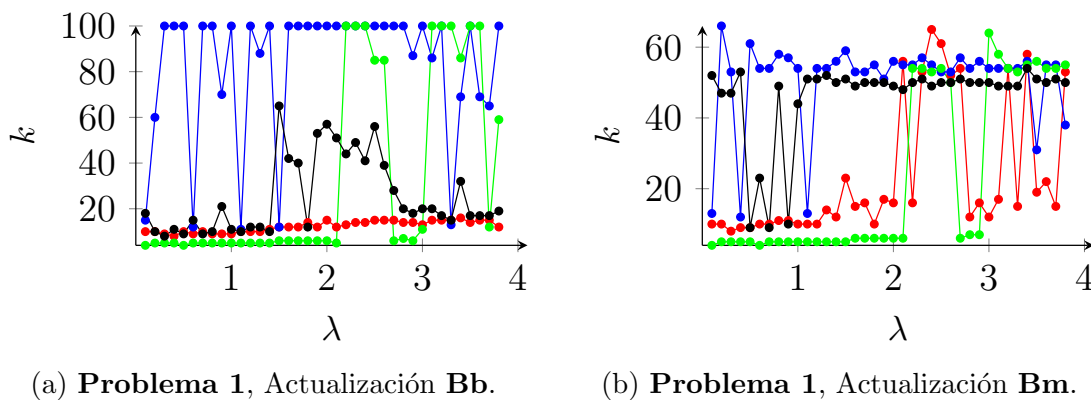
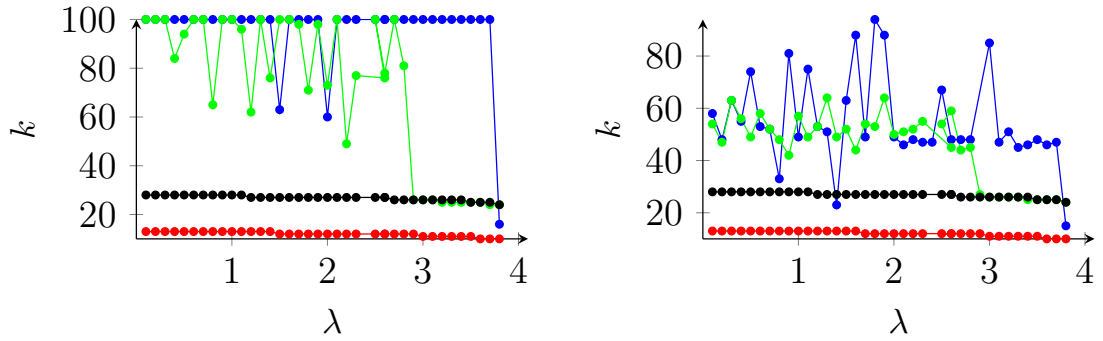
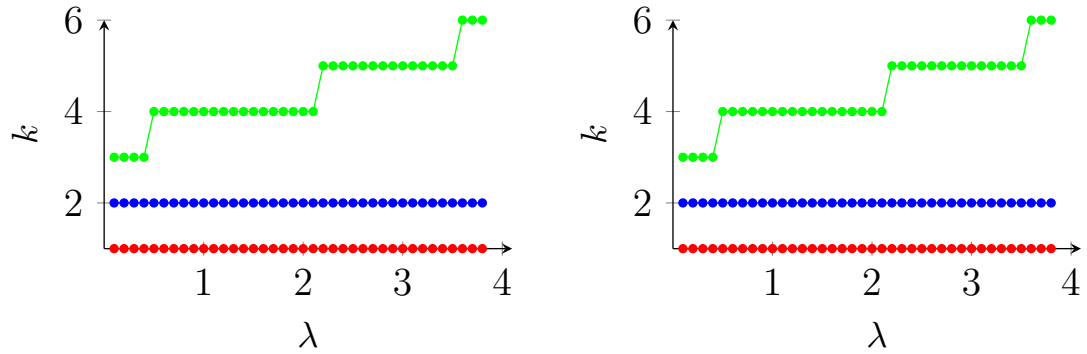


Figura 7.1: Experimento 3.



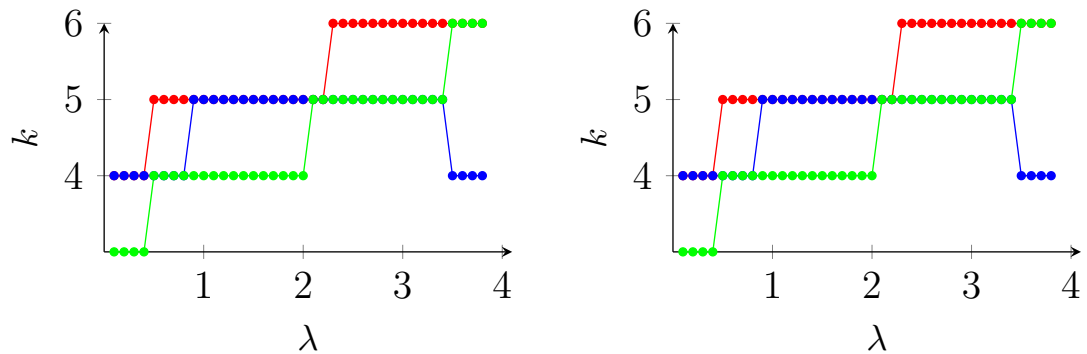
(a) Problema 2, Actualización Bb. (b) Problema 2, Actualización Bm.

Figura 7.2: Experimento 3.



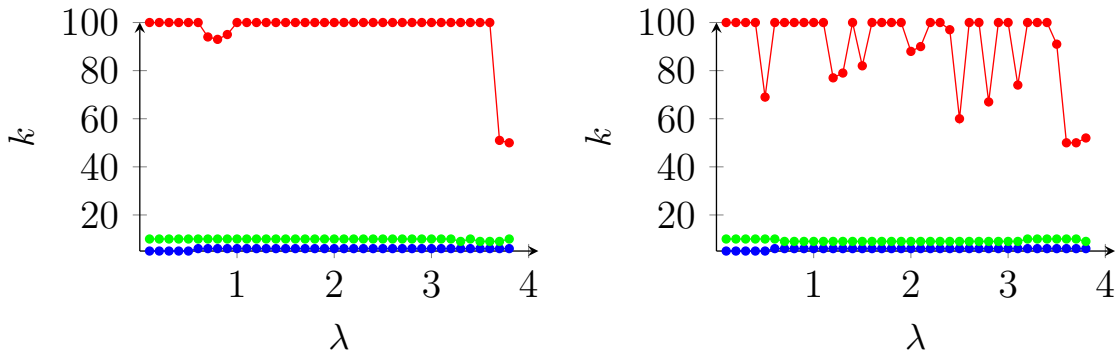
(a) Problema 3, Actualización Bb. (b) Problema 3, Actualización Bm.

Figura 7.3: Experimento 3.



(a) Problema 4, Actualización Bb. (b) Problema 4, Actualización Bm.

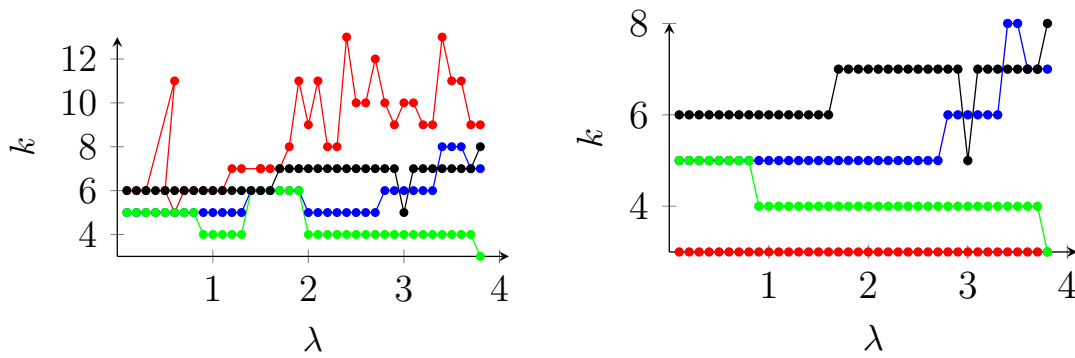
Figura 7.4: Experimento 3.



(a) Problema 5, Actualización Bb.

(b) Problema 5, Actualización Bm.

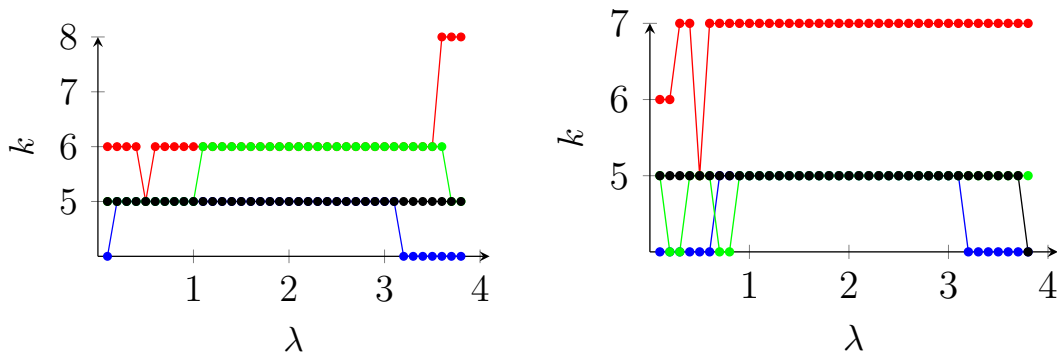
Figura 7.5: Experimento 3.



(a) Problema 6, Actualización Bb.

(b) Problema 6, Actualización Bm.

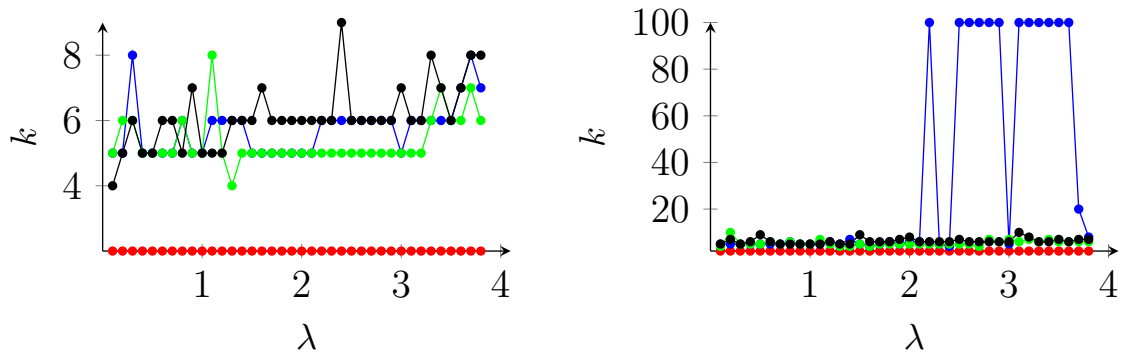
Figura 7.6: Experimento 3.



(a) Problema 7, Actualización Bb.

(b) Problema 7, Actualización Bm.

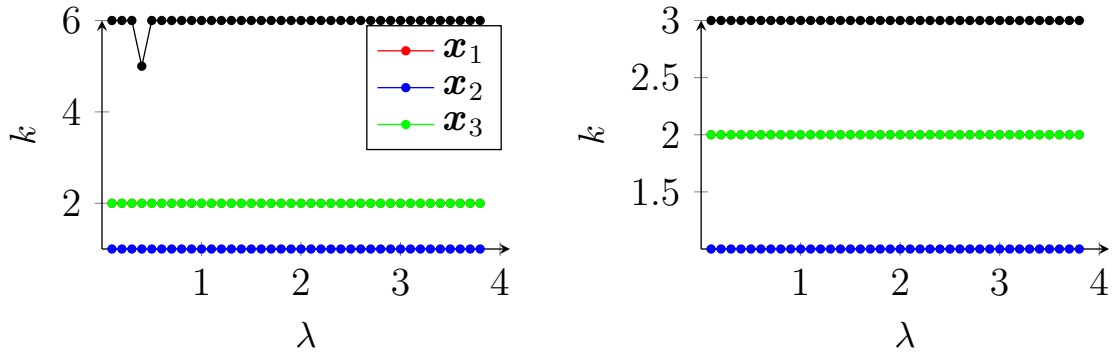
Figura 7.7: Experimento 3.



(a) Problema 8, Actualización Bb.

(b) Problema 8, Actualización Bm.

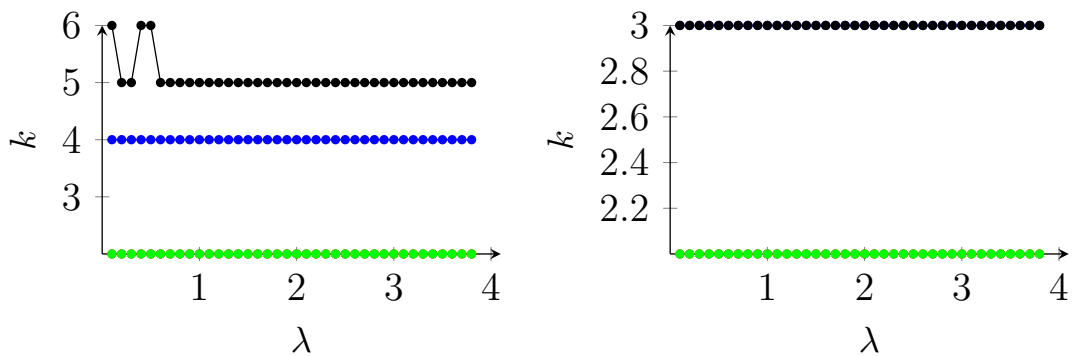
Figura 7.8: Experimento 3.



(a) Problema 9, Actualización Bb.

(b) Problema 9, Actualización Bm.

Figura 7.9: Experimento 3.



(a) Problema 10, Actualización Bb.

(b) Problema 10, Actualización Bm.

Figura 7.10: Experimento 3.

7.4.4. Experimento 4

En el cuarto experimento, para corroborar el orden de convergencia teórico del algoritmo propuesto, con las actualizaciones **Bb** y **Bm**, consideramos el llamado *orden computacional de convergencia*, [15] [63]. Para ello, analizamos, para cada p , los cocientes:

$$\frac{\|\mathbf{x}_{k+1} - \mathbf{x}_k\|}{\|\mathbf{x}_k - \mathbf{x}_{k-1}\|^p}, \quad k = 1, 2, 3, \dots,$$

donde $p = 1, 2, 3, \dots$, es llamado orden computacional de convergencia, el cual se aproxima mediante la fórmula:

$$p \approx p_c = \frac{\log(\|\mathbf{x}_{k+1} - \mathbf{x}_k\| / \|\mathbf{x}_k - \mathbf{x}_{k-1}\|)}{\log(\|\mathbf{x}_k - \mathbf{x}_{k-1}\| / \|\mathbf{x}_{k-1} - \mathbf{x}_{k-2}\|)},$$

usando las cuatro últimas aproximaciones durante el proceso iterativo.

Para el experimento, consideramos los **Problemas 1, 2, 3 y 7** y el punto inicial \mathbf{x}_1 . Los resultados lo presentamos en la **Tabla 7.3**, donde se observa, en general, que el orden computacional de convergencia es q -superlineal, coincidiendo con el resultado teórico obtenido en este capítulo.

Es importante mencionar que el orden computacional de convergencia depende del punto inicial y del problema [63]. En nuestro caso, experimentamos con más puntos iniciales y el comportamiento fue análogo al presentado en la **Tabla 7.3**, para el punto \mathbf{x}_1 , por ello, no incluimos más tablas.

Tabla 7.3: Orden de convergencia computacional del **Algoritmo SCM**.

P	actualización	p_c
P1	Bb	1.526
	Bm	2.926
P2	Bb	1.102
	Bm	1.054
P3	Bb	16.901
	Bm	17.179
P7	Bb	0.498
	Bm	1.066

7.4.5. Experimento 5

En este experimento, comparamos el Algoritmo secante local (**Algoritmo SCML**) con el global (**Algoritmo SCM**).

Tabla 7.4: Comparación entre los Algoritmos SCML y SCM.

P	Algoritmo	Bb	Bm
P1	SCML	36	53
	SCM	71	95
P2	SCML	4	4
	SCM	9	86
P3	SCML	100	100
	SCM	100	100
P4	SCML	100	100
	SCM	100	100
P5	SCML	100	100
	SCM	100	100
P6	SCML	73	66
	SCM	85	93
P7	SCML	36	69
	SCM	95	100
P8	SCML	100	100
	SCM	100	100
P9	SCML	33	32
	SCM	86	91
P10	SCML	36	24
	SCM	79	94

Los resultados obtenidos los presentamos en la **Tabla 7.4**, la cual contiene cuatro columnas, con la siguiente información: problema (P), **Algoritmo** (SCML y SCM), porcentaje de éxitos de cada algoritmo, con la actualización buena de Broyden (**Bb**) y, con la actualización mala de Broyden (**Bm**). Los 100 puntos iniciales usados fueron generado en el intervalo $[-60, 60]$.

Se observa claramente un porcentaje de éxito mayor o igual cuando se usa el algoritmo globalizado, como era de esperarse, ya que el objetivo de introducir una búsqueda lineal en un algoritmo local, es potenciar sus buenas propiedades de convergencia.

Conclusiones

Con el fin de resolver el problema de complementariedad generalizada vía su reformulación como un sistema de ecuaciones no lineales, no diferenciables y como un problema de optimización sin restricciones, en este trabajo de investigación propusimos un algoritmo cuasi-Newton local y dos algoritmos globales, uno tipo Newton y otro cuasi-Newton. En particular, de este último, propusimos una versión genérica y analizamos en detalle una versión particular del mismo.

Para todas las propuestas algorítmicas, desarrollamos la respectiva teoría de convergencia demostrando que bajo hipótesis razonables, y clásicas para esta clase de métodos, nuestros algoritmos resuelven el problema en cuestión y poseen las características de convergencia deseadas; es decir, q -cuadrática para el método tipo Newton y q -superlineal para los cuasi-Newton. De igual manera, realizamos pruebas numéricas con todos los algoritmos, comparándolos entre sí e identificando sus fortalezas y debilidades.

Vale la pena mencionar que para estructurar los algoritmos de nuestra propuesta investigativa analizamos, en primer lugar, el operador Ψ_λ que define la reformulación del PCG(F, G) como un sistema de ecuaciones no lineales, demostrando que los resultados obtenidos por otros autores, para el caso particular de complementariedad no lineal, se extienden naturalmente al problema de complementariedad generalizado.

Respecto al algoritmo cuasi-Newton global, destacamos que es no monótono y libre de derivadas, características que disminuyeron considerablemente el costo computacional en comparación con el método tipo Newton propuesto. De igual manera, resaltamos que para este algoritmo establecimos condiciones, lo más generales posibles, sobre las matrices B_k , permitiéndonos desarrollar una teoría de convergencia general para esta clase de métodos.

Con el ánimo de mostrar que las hipótesis requeridas sobre las matrices B_k son en realidad alcanzables, hicimos una propuesta de actualización explícita de estas matrices. Para ello recurrimos a las actualizaciones secantes y a la teoría desarrollada en el **Capítulo 5**, aspectos que nos permitieron construir actualizaciones secantes estructuradas y con ello una versión particular del algoritmo cuasi-Newton genérico.

Es importante decir que en esta investigación se abordaron todos los problemas propuestos en el proyecto de tesis doctoral que dio lugar a la misma, dando cumplimiento así a los objetivos planteados. Más aún, pudimos desarrollar una teoría más general, para resolver el $\text{PCG}(F,G)$, que la estipulada inicialmente.

Para finalizar, planteamos algunas preguntas que surgieron de esta investigación y que podrían dar lugar a trabajos futuros:

1. ¿Es posible extender la teoría desarrollada en esta investigación a otras familias de funciones de complementariedad?
2. Basados en los experimentos numéricos ¿Es posible proponer una modificación a la estrategia de λ dinámico que mejore la eficiencia de los algoritmos?
3. ¿Es posible realizar un análisis de complejidad de los algoritmos propuestos?
4. ¿Existen estrategias diferenciables que permitan resolver el $\text{PCG}(F,G)$?
5. ¿Es posible extender la teoría desarrollada a problemas de complementariedad definidos sobre funciones o conjuntos más generales?
6. ¿Es posible incluir estrategias de búsqueda direccional inexacta a los algoritmos propuestos con el fin de abordar problemas de gran tamaño?

Bibliografía

- [1] R. Andreani, A. Friedlander, and S. A. Santos, *On the resolution of the generalized nonlinear complementarity problem*, SIAM Journal on Optimization **12** (2002), no. 2, 303–321.
- [2] F. Arenas, H. J. Martínez, and R. Pérez, *Least change secant update methods for nonlinear complementarity problem*, Ingeniería y Ciencia **11** (2015), no. 21, 11–36.
- [3] ———, *El problema de complementariedad No lineal: Teoría, aplicaciones y nuevos algoritmos para su solución*, Programa Editorial Univalle, 2019.
- [4] C. A. Arias, H. J. Martínez, and R. Pérez, *A global quasi-Newton algorithms for nonlinear complementarity problems*, Pacific Journal of Optimization **13** (2017), no. 1, 1–15.
- [5] ———, *A nonsmooth global quasi-newton method for nonlinear complementarity problems*, Pacific Journal Of Optimization **13** (2017), no. 1, 1–15.
- [6] D. P. Bertsekas, *Constrained optimization and lagrange multiplier methods*, Academic press, 2014.
- [7] I. D. L. Bogle and J. D. Perkins, *A new sparsity preserving quasi-newton update for solving nonlinear equations*, SIAM Journal on Scientific and Statistical Computing **11** (1990), no. 4, 621–630.
- [8] P. N. Brown and Y. Saad, *Convergence theory of nonlinear newton–krylov algorithms*, SIAM Journal on Optimization **4** (1994), no. 2, 297–330.
- [9] C. G. Broyden, *A class of methods for solving nonlinear simultaneous equations*, Mathematics of computation **19** (1965), no. 92, 577–593.

-
- [10] S. Buhmiller and N. Krejic, *A new smoothing quasi-newton method for nonlinear complementarity problems*, Journal of Computational and Applied Mathematics **211** (2008), no. 2, 141–155.
- [11] Y.-Y. Chen and Y. Gao, *Two new levenberg-marquardt methods for nonsmooth nonlinear complementarity problems*, Sci. Asia **40** (2014), 89.
- [12] A. Chu, S. Du, and Y. Su, *A new smoothing conjugate gradient method for solving nonlinear nonsmooth complementarity problems*, Algorithms **8** (2015), no. 4, 1195–1209.
- [13] F. H. Clarke, *Necessary conditions for nonsmooth problems in optimal control and the calculus of variations*, Ph.D. thesis, University of Washington, 1973.
- [14] Frank H Clarke, *Optimization and nonsmooth analysis*, Montreal, 1990.
- [15] A. Cordero, J. L. Hueso, E. Martínez, and J. R. Torregrosa, *Increasing the convergence order of an iterative method for nonlinear systems*, Applied Mathematics Letters **25** (2012), no. 12, 2369–2374.
- [16] R. W. Cottle and G. B. Dantzig, *A generalization of the linear complementarity problem*, Journal of Combinatorial Theory **8** (1970), no. 1, 79–90.
- [17] R. W. Cottle, J-S. Pang, and R. E. Stone, *The linear complementarity problem*, SIAM, 2009.
- [18] T. De Luca, F. Facchinei, and C. Kanzow, *A semi-smooth equation approach to the solution of nonlinear complementarity problems*, Mathematical Programming **75** (1996), 407–439.
- [19] J. E. Dennis and J. J. Moré, *A characterization of superlinear convergence and its application to quasi-newton methods*, Mathematics of computation **28** (1974), no. 126, 549–560.
- [20] J. E. Dennis Jr and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, SIAM, 1996.
- [21] I. C. Dolcetta and U. Mosco, *Implicit complementarity problems and quasi-variational inequalities*, in *Variational Inequalities and Complementarity Problems: Theory and Applications*, eds., Wiley (1980), 75–87.
- [22] S.-q. Du, *Generalized Newton method for a kind of complementarity problem*, Abstract and Applied Analysis, vol. 2014, Hindawi, 2014.

-
- [23] F. Facchinei, *Minimization of SC1 functions and the Maratos effect*, Operations Research Letters **17** (1995), no. 3, 131–137.
- [24] F. Facchinei and J. Soares, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM Journal on Optimization **7** (1997), no. 1, 225–247.
- [25] M. C. Ferris and C. Kanzow, *Complementarity and related problems: A survey*, Tech. report, Department of Computer Sciences, UW-Madison, Math Prog Technical Reports, 1998.
- [26] A. Fischer, *Solution of monotone complementarity problems with locally lipschitzian functions*, Mathematical Programming **76** (1997), no. 3, 513–532.
- [27] A. Fischer, V. Jeyakumar, and T. D. Luc, *Solution point characterizations and convergence analysis of a descent algorithm for nonsmooth continuous complementarity problems*, Journal of Optimization Theory and Applications **110** (2001), no. 3, 493–513.
- [28] A. Fischer and C. Kanzow, *On finite termination of an iterative method for linear complementarity problems*, Mathematical Programming **74** (1996), no. 3, 279–292.
- [29] A. Griewank, *The global convergence of partitioned bfgs on problems with convex decompositions and lipschitzian gradients*, Mathematical Programming **50** (1991), no. 1-3, 141–175.
- [30] L. Grippo, F. Lampariello, and S. Lucidi, *A nonmonotone line search technique for newton's method*, SIAM Journal on Numerical Analysis **23** (1986), no. 4, 707–716.
- [31] G. Z. T. Grueso, *Un algoritmo local secante de cambio mínimo para complementariedad generalizada*, Master's thesis, Universidad del Cauca, 2021.
- [32] W-Z. Gu and M. A. Tawhid, *Further application of-differentiability to generalized complementarity problems based on generalized fisher-burmeister functions*, Abstract and Applied Analysis, vol. 2014, Hindawi, 2014.
- [33] H. Jiang, *Unconstrained minimization approaches to nonlinear complementarity problems*, Journal of Global Optimization **9** (1996), no. 2, 169–181.
- [34] C Kanzow and H. Kleinmichel, *A new class of semi-smooth Newton-type methods for nonlinear complementarity problems*, Computational Optimization and Applications **11** (1998), no. 3, 227–251.

- [35] M. Kojima and S. Shindo, *Extension of Newton and quasi-Newton methods to systems of PC1 equations*, Journal of the Operations Research Society of Japan **29** (1986), no. 4, 352–375.
- [36] E. Kreyszig, *Introductory functional analysis with applications*, vol. 17, John Wiley & Sons, 1991.
- [37] D-H. Li and M. Fukushima, *A derivative-free line search and global convergence of broyden-like method for nonlinear equations*, Optimization methods and software **13** (2000), no. 3, 181–201.
- [38] J. Long and S. Zeng, *A new filter-levenberg–marquardt method with disturbance for solving nonlinear complementarity problems*, Applied mathematics and computation **216** (2010), no. 2, 677–688.
- [39] V. L. R. Lopes, J. M. Martínez, and R. Pérez, *On the local convergence of quasi-Newton methods for nonlinear complementarity problems*, Applied Numerical Mathematics **30** (1999), no. 1, 3–22.
- [40] J. M. Martínez, *Local convergence theory of inexact newton methods based on structured least change updates*, mathematics of computation **55** (1990), no. 191, 143–167.
- [41] ———, *On the relation between two local convergence theories of least-change secant update methods*, Mathematics of Computation **59** (1992), no. 200, 457–481.
- [42] T. Ni and P. Wang, *A smoothing-type algorithm for solving nonlinear complementarity problems with a non-monotone line search*, Applied Mathematics and Computation **216** (2010), no. 7, 2207–2214.
- [43] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer, 1999.
- [44] J. V. Outrata and J. Zowe, *A newton method for a class of quasi-variational inequalities*, Computational Optimization and Applications **4** (1995), 5–21.
- [45] J-S. Pang, *The implicit complementarity problem*, Nonlinear Programming 4, Elsevier, 1981, pp. 487–518.
- [46] ———, *On the convergence of a basic iterative method for the implicit complementarity problem.*, Journal of Optimization Theory y Applications **37** (1982), no. 2, 149–162.
- [47] ———, *Newton’s method for B-differentiable equations*, Mathematics of operations research **15** (1990), no. 2, 311–341.

- [48] ———, *A b -differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity, and variational inequality problems*, *Mathematical Programming* **51** (1991), no. 1, 101–131.
- [49] J-S. Pang and L. Qi, *Nonsmooth equations: motivation and algorithms*, *SIAM Journal on optimization* **3** (1993), no. 3, 443–465.
- [50] ———, *Nonsmooth equations: motivation and algorithms*, *SIAM Journal on optimization* **3** (1993), no. 3, 443–465.
- [51] ———, *Nonsmooth equations: motivation and algorithms*, *SIAM Journal on optimization* **3** (1993), no. 3, 443–465.
- [52] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, *Mathematics of operations research* **18** (1993), no. 1, 227–244.
- [53] ———, *C-differentiability, C-differential operators and generalized Newton methods*, *AMR 96/5*, Tech. report, Applied Mathematics Report, University of New South Wales, Sydney, 1996.
- [54] L. Qi and J. Sun, *A nonsmooth version of newton's method*, *Mathematical programming* **58** (1993), no. 1, 353–367.
- [55] N. M. Ralevic and D. Cebic, *The properties of modifications of newton's method for solving nonlinear equations*, 2013 IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY), 2013, pp. 341–345.
- [56] S.-P. Rui and C.-X. Xu, *A smoothing inexact newton method for nonlinear complementarity problems*, *Journal of Computational and Applied Mathematics* **233** (2010), no. 9, 2332–2338.
- [57] L. K. Schubert, *Modification of a quasi-newton method for nonlinear equations with a sparse jacobian*, *Mathematics of Computation* **24** (1970), no. 109, 27–30.
- [58] L. Song and Y. Gao, *On the local convergence of a levenberg-marquardt method for nonsmooth nonlinear complementarity problems*, *ScienceAsia* **43** (2017), 377.
- [59] H. Vivas, R. Pérez, and C. A. Arias, *A nonsmooth newton method for solving the generalized complementarity problem*, [Manuscript submitted for publication].
- [60] Z. Wan, M. Yuan, and C. Wang, *A partially smoothing jacobian method for nonlinear complementarity problems with $p0$ function*, *Journal of Computational and Applied Mathematics* **286** (2015), 158–171.

-
- [61] David S Watkins, *Fundamentals of matrix computations*, John Wiley & Sons, 2004.
- [62] S-L. Wu and P. Guo, *Modulus-based matrix splitting algorithms for the quasi-complementarity problems*, Applied Numerical Mathematics **132** (2018), 127–137.
- [63] X.-Y. Xiao, *New techniques to develop higher order iterative methods for systems of nonlinear equations*, Computational and Applied Mathematics **41** (2022), no. 6, 243.
- [64] H. Xu, *Point-based set-valued approximations, c -differential operators and applications*, Optimization **52** (2003), no. 2, 127–143.
- [65] H. Yu and D. Pu, *Smoothing levenberg–marquardt method for general nonlinear complementarity problems under local error bound*, Applied Mathematical Modelling **35** (2011), no. 3, 1337–1348.