

**Métodos secantes de cambio mínimo
para el problema de complementariedad no lineal**

Favián Enrique Arenas Aparicio

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Matemáticas
Maestría en Ciencias Matemáticas
Popayán
2013

Métodos secantes de cambio mínimo para el problema de complementariedad no lineal

Favián Enrique Arenas Aparicio

Trabajo de investigación presentado como requisito parcial para optar al
título de Magíster en Ciencias Matemáticas

Directora

Dra. Rosana Pérez Mera
Profesora de la Universidad del Cauca

Codirector

Dr. Héctor Jairo Martínez Romero
Profesor de la Universidad del Valle

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Matemáticas
Maestría en Ciencias Matemáticas
Popayán
2013

Nota de aceptación

Director: **Dra. Rosana Pérez Mera**

Jurado: **Dr. Luis Eduardo Montoya D.**

Jurado: **Mg. Ana Milena Plaza B.**

Popayán, 29 de abril de 2013

Agradecimientos

*¿Qué pagaré a Jehová por todos sus beneficios para conmigo?
(Sal. 116:12)*

A mi Dios por bendecirme y porque hizo realidad este logro.

A mi familia que fue para mí una gran fuente de ánimo. A mi madre, a mi comprensiva esposa y a mi querida hija, por su apoyo.

A mi directora de tesis, *Dra. Rosana Pérez Mera*, por su esfuerzo y dedicación, porque con sus conocimientos, su experiencia, su paciencia y su motivación logró que yo pudiera terminar esta empresa con éxito. Ella me dió ánimo cuando lo necesité y me enseñó humildad cuando me faltó.

A mi codirector, *Dr. Héctor Jairo Martínez Romero* porque con su visión sobre muchos aspectos de la investigación y con su ayuda y aportes me orientó y me mostró cómo pensar como investigador.

Al *Dr. Luis Eduardo Montoya* y a la *Mg. Ana Milena Plaza* quienes con sus sugerencias acertadas, producto de su experiencia académica e investigativa, dieron más realce a este trabajo.

Al *Mg. José Ignacio Téllez R.* por sus sugerencias en algunos aspectos gráficos y de redacción en este trabajo.

A todos mis profesores porque han aportado un grano de arena a mi formación.

A mis colegas del Departamento de Matemáticas de la Universidad del Cauca por su apoyo, especialmente a mis compañeros de clases *Hevert Vivas* y *Mauricio Macias*.

Favián Enrique Arenas Aparicio

Tabla de Contenidos

Tabla de Contenidos	II
Índice de figuras	IV
Índice de tablas	VI
1. Introducción	1
2. La Función de Kanzow	5
3. El PCNL y su reformulación mediante la <i>función de Kanzow</i>	18
4. Algoritmo y teoría de convergencia	25
4.1. Hipótesis	26
4.2. Resultados de convergencia	27

4.3. Familia de métodos secante de cambio mínimo para $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$	40
4.4. Resultados de convergencia adicionales	43
5. Pruebas numéricas	50
6. Comentarios finales	60
Bibliografía	61

Índice de figuras

2.1. <i>Función de Kanzow para algunos valores de λ.</i>	6
2.2. <i>Función α_{\min}.</i>	11
2.3. <i>Función $\ M\ _2$.</i>	12
2.4. <i>Desigualdad triangular.</i>	13
3.1. <i>Gráficas de $F_1(x, y) = 0$ y $F_2(x, y) = 0$.</i>	19
5.1. <i>Problema Kojima-Shindo.</i>	57
5.2. <i>Problema Kojima-Josephy.</i>	57
5.3. <i>Problema Sistema Trigonométrico.</i>	58
5.4. <i>Problema Nash-Cornout.</i>	58
5.5. <i>Problema Mathiesen Modificado.</i>	58
5.6. <i>Problema Sistema Tridiagonal.</i>	59

5.7. <i>Problema Sistema Trigonométrico Exponencial.</i>	59
5.8. <i>Problema Rosenbrock.</i>	59

Índice de tablas

5.1. <i>Comportamiento local de los Algoritmos 5.1 y 5.2.</i>	56
--	----

Capítulo 1

Introducción

Dada $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))$ continuamente diferenciable, el Problema de Complementariedad No Lineal (PCNL) consiste en encontrar un vector $\mathbf{x} \in \mathbb{R}^n$ que satisfaga las tres condiciones siguientes,

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

En este contexto, la expresión $\mathbf{y} \geq 0$ para $\mathbf{y} \in \mathbb{R}^n$ significa que $y_i \geq 0$ para todo $i = 1, \dots, n$.

Relacionados con este problema, están los de inecuaciones variacionales, complementariedad lineal, complementariedad mixta [34], complementariedad horizontal, entre otros [41].

El problema de complementariedad no lineal surge en diversas aplicaciones como por ejemplo, problemas de contacto mecánico y fricción [2], problemas de mecánica estructural y diseño estructural, problemas de lubricación elasto-hidrodinámicos [25], problemas de equilibrio de tráfico [8], así como en problemas relacionados con modelos de equilibrio económico [18]. La importancia del PCNL en las áreas de física, ingeniería y economía se debe al hecho de que el concepto de complementariedad es sinónimo de la noción de sistema en equilibrio.

En los últimos años, se han estudiado diferentes técnicas para resolver el problema de complementariedad no lineal; una de éstas consiste en reformularlo como un sistema de ecuaciones no lineales, no diferenciable mediante el uso de ciertas funciones especiales

llamadas *funciones de complementariedad* [47].

Una función $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

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

se denomina *función de complementariedad*.

Geoméricamente, a partir de la equivalencia (1.2), se infiere que la traza de la función φ obtenida por la intersección con el plano \mathbf{xy} , es la curva formada por los semiejes positivos \mathbf{x} y \mathbf{y} , la cual no es diferenciable en $(0, 0)$. Esta falta de suavidad en la curva implica la no suavidad en la función φ .

Para reformular el PCNL como un sistema de ecuaciones no lineales, se considera una *función de complementariedad* φ y se define $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ por,

$$\Phi(\mathbf{x}) = \begin{pmatrix} \varphi(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi(x_n, F_n(\mathbf{x})) \end{pmatrix}. \quad (1.3)$$

Como consecuencia de la no suavidad de φ , el sistema de ecuaciones no lineales

$$\Phi(\mathbf{x}) = 0 \quad (1.4)$$

no es diferenciable.

De la definición de *función de complementariedad* dada en (1.2), se deduce fácilmente que un vector \mathbf{x}_* resuelve el sistema (1.4), si y sólo si, \mathbf{x}_* resuelve el PCNL.

En la literatura sobre problemas de complementariedad no lineal, existen numerosas *funciones de complementariedad*, pero las más utilizadas han sido la *función mínimo* [37] y la *función de Fischer - Burmeister* [19], definidas respectivamente por

$$\varphi(a, b) = \min\{a, b\}, \quad \varphi(a, b) = \sqrt{a^2 + b^2} - a - b. \quad (1.5)$$

La *función mínimo* se puede escribir en forma equivalente como

$$\varphi(a, b) = \frac{1}{2}((a + b) - |a - b|),$$

que permite observar su no diferenciability en los puntos de la forma (a, a) . En cuanto a la *función de Fischer - Burmeister*, la no diferenciability se presenta en el punto

$(0, 0)$.

En 1998, *Kanzow* y *Kleinmichel* [23] presentaron la *función de complementariedad* φ_λ definida por,

$$\varphi_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} - a - b, \quad (1.6)$$

donde el parámetro $\lambda \in (0, 4)$. Analizaremos esta función detalladamente en el siguiente capítulo.

Observemos que la *función de Fischer - Burmeister* es un caso particular de φ_λ , correspondiente a $\lambda = 2$. En lo que sigue, denotaremos por Φ_λ la función definida en (1.4) y obtenida mediante la *función de complementariedad* φ_λ .

Una vez reformulado el PCNL por el sistema de ecuaciones no lineales (1.4), se han propuesto diferentes algoritmos para su solución, entre los que se incluyen métodos tipo Newton no suave [40] [43], tipo cuasi Newton no suave [28] [6] [27], entre otros [1] [42] [14] [38] [30] [7].

En este trabajo de investigación, analizamos detalladamente la *función de complementariedad* φ_λ y proponemos un método *cuasi Newton no diferenciable* para resolver el PCNL usando el sistema $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$, para dicho método desarrollamos la teoría de convergencia local respectiva. Además, generamos una *familia de métodos secante de cambio mínimo* para resolver el PCNL vía su reformulación $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$, bajo ciertas hipótesis demostramos que esta familia proporciona *algoritmos local y superlinealmente convergentes*.

Organizamos la presentación de este documento en la siguiente forma.

En el Capítulo **2**, analizamos detalladamente las propiedades de la función φ_λ , mediante manipulaciones algebraicas y matriciales, obtenemos cotas de gran utilidad para el desarrollo teórico del Capítulo **4**.

En el Capítulo **3**, reformulamos el PCNL como un sistema de ecuaciones no lineales usando la función Φ_λ y caracterizamos un subconjunto del jacobiano generalizado de Φ_λ en un punto \mathbf{x} .

En el Capítulo **4**, en la primera parte proponemos un nuevo algoritmo *cuasi Newton* genérico para resolver el sistema de ecuaciones no lineales, no diferenciable $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$ y desarrollamos la teoría de convergencia local para el mismo. En la segunda parte, generamos una familia de *métodos secante de cambio mínimo* siguiendo las reglas de [31] para métodos de este tipo. Demostramos que dicha familia proporciona *algoritmos local*

y *superlinealmente* convergentes.

En el Capítulo 5, analizamos numéricamente el comportamiento local de los algoritmos propuestos en el Capítulo 4, para lo cual utilizamos 8 *problemas de prueba* propuestos en [37] [29]. Cuatro de los cuales son problemas de aplicación al *equilibrio económico* y a la *teoría de juegos*.

En el Capítulo 6, hacemos algunos comentarios finales y propuestas para trabajos futuros sobre el tema.

Capítulo 2

La Función de Kanzow

Las llamadas *funciones de complementariedad*, definidas en el capítulo anterior, permiten reformular el PCNL como un sistema de ecuaciones no lineales. Debido a que nuestro interés es la solución del PCNL mediante su reformulación como un sistema no lineal vía la *función de complementariedad* φ_λ , dedicamos este capítulo al estudio detallado de dicha función y sus propiedades. Los resultados obtenidos serán de gran ayuda en los desarrollos teóricos presentados en el capítulo siguiente.

Definición 2.1. La función $\varphi_\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por,

$$\varphi_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} - a - b, \quad (2.1)$$

donde, el parámetro $\lambda \in (0, 4)$ se denomina *función de Kanzow*.

En la definición de φ_λ no es conveniente incluir los extremos del intervalo en el que varía el parámetro λ ya que, la inclusión de $\lambda = 0$, aumentaría los puntos de no diferenciabilidad de φ_λ porque φ_0 , definida por,

$$\varphi_0(a, b) = |a - b| - (a + b) = -2 \min \{a, b\},$$

no es diferenciable en todos los puntos de la forma (a, a) . Para el extremo $\lambda = 4$,

$$\varphi_4(a, b) = |a + b| - (a + b),$$

luego, para todo (a, b) , con $a \geq 0$ y $b \geq 0$, tenemos que φ_4 es la *función nula*, la cual no contribuye a la solución del PCNL.

En la Figura 2.1, ilustramos la *función de Kanzow* para algunos valores de λ .

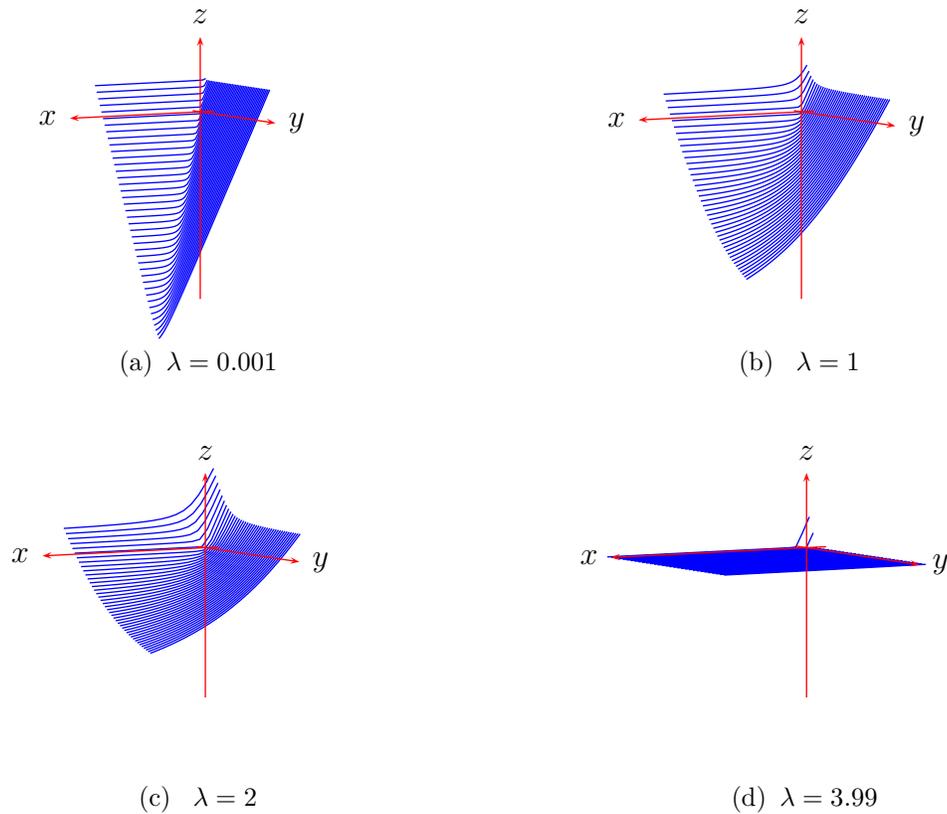


Figura 2.1: *Función de Kanzow para algunos valores de λ .*

Una pregunta natural que surge de la **Definición 2.1** es la siguiente,

¿La función φ_λ está bien definida, para todo λ en $(0, 4)$?

La respuesta a este interrogante es afirmativa. Para verificarlo, basta probar que

$$(a - b)^2 + \lambda ab \geq 0,$$

para todo vector $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ y para todo $\lambda \in (0, 4)$. Consideramos dos casos,

1) Si $a = b = 0$ entonces,

$$(a - b)^2 + \lambda ab = 0, \quad (2.2)$$

para todo $\lambda \in (0, 4)$.

2) En cualquier otro caso, $\begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y $(a - b)^2 + \lambda ab$ se puede escribir como la forma cuadrática

$$(a - b)^2 + \lambda ab = \begin{pmatrix} a & b \end{pmatrix} K \begin{pmatrix} a \\ b \end{pmatrix}, \quad (2.3)$$

donde

$$K = \begin{pmatrix} 1 & \frac{\lambda}{2} - 1 \\ \frac{\lambda}{2} - 1 & 1 \end{pmatrix}, \quad (2.4)$$

la cual es una *matriz simétrica* con valores propios que son funciones de λ ,

$$\alpha_1 = 2 - \frac{\lambda}{2} \quad \text{y} \quad \alpha_2 = \frac{\lambda}{2}.$$

Dado que $\lambda \in (0, 4)$, dichos valores propios son positivos y, por lo tanto, la matriz K es *definida positiva*¹, es decir,

$$\begin{pmatrix} a & b \end{pmatrix} K \begin{pmatrix} a \\ b \end{pmatrix} = (a - b)^2 + \lambda ab > 0. \quad (2.5)$$

De 1) y 2) concluimos que φ_λ está bien definida.

Presentamos en los **Lemas 2.1 y 2.2** dos propiedades interesantes del primer sumando en la *función de Kanzow*.

Lema 2.1. Sea $\lambda \in (0, 4)$. La función $G_\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$G_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} \quad (2.6)$$

es una norma en \mathbb{R}^2 .

¹Una matriz $A \in \mathbb{R}^{n \times n}$ es *definida positiva* [44], si y sólo si, $\mathbf{x}^T A \mathbf{x} > 0$, para todo vector no nulo $\mathbf{x} \in \mathbb{R}^n$. En particular, una *matriz simétrica* A , es *definida positiva*, si y sólo si, sus valores propios son positivos [36].

Demostración. Por (2.3),

$$G_\lambda(a, b) = \sqrt{(a-b)^2 + \lambda ab} = \sqrt{(a \ b) K \begin{pmatrix} a \\ b \end{pmatrix}} = \left\| \begin{pmatrix} a \\ b \end{pmatrix} \right\|_K.$$

La última igualdad se tiene debido a que K es una matriz *simétrica y definida positiva*, por lo tanto, define una norma sobre \mathbb{R}^2 [45].

Sin embargo, con el objetivo de familiarizar al lector con la función G_λ , realizamos la prueba utilizando la definición de *norma vectorial* [17]. Para ello, verifiquemos que G_λ satisface las propiedades de una norma sobre \mathbb{R}^2 .

- $G_\lambda(a, b) \geq 0$. Esta propiedad se deduce de (2.2) y (2.5).
- Supongamos que $G_\lambda(a, b) = 0$. Dado que la matriz K es *definida positiva*, tiene *descomposición de Cholesky* [46], es decir, existe una matriz no singular L tal que $K = L^T L$.

Luego,

$$0 = (a \ b) K \begin{pmatrix} a \\ b \end{pmatrix} = (a \ b) L^T L \begin{pmatrix} a \\ b \end{pmatrix} = \left\| L \begin{pmatrix} a \\ b \end{pmatrix} \right\|_2^2$$

ya que L es no singular, tenemos que,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

El recíproco de ésta propiedad lo deducimos de (2.2).

- Sea $\beta \in \mathbb{R}$,

$$G_\lambda(\beta a, \beta b) = \sqrt{(\beta a \ \beta b) K \begin{pmatrix} \beta a \\ \beta b \end{pmatrix}} = \sqrt{\beta^2 (a \ b) K \begin{pmatrix} a \\ b \end{pmatrix}} = |\beta| G_\lambda(a, b).$$

- Sean $\begin{pmatrix} a \\ b \end{pmatrix}$ y $\begin{pmatrix} c \\ d \end{pmatrix}$ en \mathbb{R}^2 .

$$\begin{aligned}
(G_\lambda(a+c, b+d))^2 &= (a+c, b+d) K \begin{pmatrix} a+c \\ b+d \end{pmatrix} \\
&= (a+c, b+d) \left[K \begin{pmatrix} a \\ b \end{pmatrix} + K \begin{pmatrix} c \\ d \end{pmatrix} \right] \\
&= (G_\lambda(a, b))^2 + (G_\lambda(c, d))^2 + 2(a, b) K \begin{pmatrix} c \\ d \end{pmatrix}. \quad (2.7)
\end{aligned}$$

Usando siguiente la desigualdad, conocida como *desigualdad Wielandt* [22] en (2.7),

$$(a \ b) K \begin{pmatrix} c \\ d \end{pmatrix} \leq \sqrt{\left[(a \ b) K \begin{pmatrix} a \\ b \end{pmatrix} \right] \left[(c \ d) K \begin{pmatrix} c \\ d \end{pmatrix} \right]} = G_\lambda(a, b) G_\lambda(c, d)$$

tenemos

$$(G_\lambda(a+c, b+d))^2 \leq (G_\lambda(a, b) + G_\lambda(c, d))^2,$$

la cual conduce a la desigualdad triangular.

Por lo anterior, concluimos que G_λ define una norma vectorial sobre \mathbb{R}^2 . ★

El siguiente resultado es válido para toda forma cuadrática, en particular para G_λ .

Lema 2.2. *La función G_λ definida por (2.6) satisface las siguientes desigualdades, para todo $(a, b) \in \mathbb{R}^2$,*

$$\sqrt{\alpha_{\min}} \left\| \begin{pmatrix} a \\ b \end{pmatrix} \right\|_2 \leq G_\lambda(a, b) \leq \sqrt{\alpha_{\max}} \left\| \begin{pmatrix} a \\ b \end{pmatrix} \right\|_2, \quad (2.8)$$

donde α_{\min} y α_{\max} son los valores propios mínimo y máximo respectivamente, de la matriz K definida en (2.4).

Demostración. Teniendo en cuenta que $G_\lambda(a, b) = \sqrt{(a \ b) K \begin{pmatrix} a \\ b \end{pmatrix}}$ con K dada por (2.4) es una matriz *simétrica* y $\alpha_{\min} > 0$, y $\alpha_{\max} > 0$ son sus valores propios mínimo y

máximo, respectivamente, concluimos²

$$\alpha_{\min} \left\| \begin{pmatrix} a \\ b \end{pmatrix} \right\|_2^2 \leq [G_\lambda(a, b)]^2 \leq \alpha_{\max} \left\| \begin{pmatrix} a \\ b \end{pmatrix} \right\|_2^2,$$

de donde obtenemos las desigualdades (2.8). ★

Utilizando la primera desigualdad en (2.8), obtenemos la siguiente cota para el recíproco de $G_\lambda(a, b)$, para todo $(a, b) \neq (0, 0)$

$$\frac{1}{G_\lambda(a, b)} = \frac{1}{\sqrt{(a-b)^2 + \lambda ab}} \leq \frac{1}{\sqrt{\alpha_{\min}} \sqrt{a^2 + b^2}}, \quad (2.9)$$

la cual usaremos en la demostración del **Lema 2.5**.

Observemos que el valor α_{\min} depende de λ , en efecto,

$$\alpha_{\min}(\lambda) = \begin{cases} \frac{\lambda}{2}, & \text{si } \lambda \in (0, 2), \\ 2 - \frac{\lambda}{2}, & \text{si } \lambda \in [2, 4), \end{cases}$$

La Figura 2.2 ilustra la función $\alpha_{\min}(\lambda)$, para todo $\lambda \in (0, 4)$.

Es importante observar que la función φ_λ no es diferenciable en $(0, 0)$. Para cualquier otro vector de \mathbb{R}^2 , el vector gradiente de φ_λ está definido por

$$\nabla \varphi_\lambda(a, b) = \begin{pmatrix} \frac{2(a-b) + \lambda b}{2G_\lambda(a, b)} - 1 \\ \frac{-2(a-b) + \lambda a}{2G_\lambda(a, b)} - 1 \end{pmatrix} = \nabla G_\lambda(a, b) - \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (2.10)$$

²Si $A \in \mathbb{R}^{n \times n}$ es una matriz *simétrica* con valores propios $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, entonces, para todo $\mathbf{x} \in \mathbb{R}^n$,

$$\lambda_1 \mathbf{x}^T \mathbf{x} \leq \mathbf{x}^T A \mathbf{x} \leq \lambda_n \mathbf{x}^T \mathbf{x}.$$

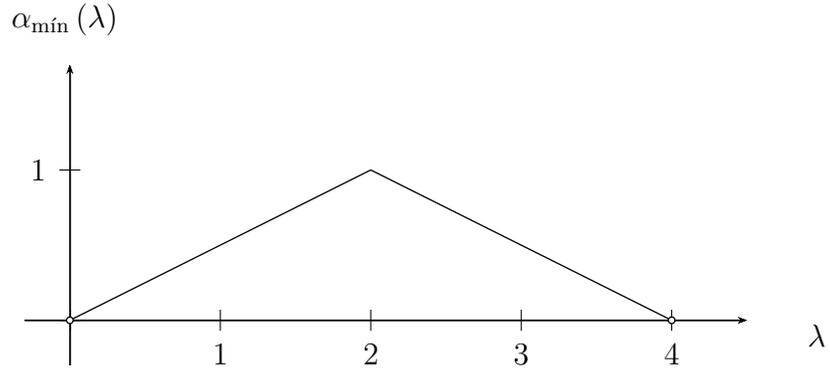


Figura 2.2: Función α_{\min} .

Para su uso posterior, denotaremos las derivadas parciales de G_λ de la siguiente manera:

$$\chi(a, b) = \frac{2(a - b) + \lambda b}{2G_\lambda(a, b)} \quad \text{y} \quad \psi(a, b) = \frac{-2(a - b) + \lambda a}{2G_\lambda(a, b)}. \quad (2.11)$$

En [23], los autores demuestran que

$$\|\nabla G_\lambda(a, b)\|_2 \leq \sqrt{2}. \quad (2.12)$$

A partir de (2.10), encontramos una expresión matricial para $\nabla G_\lambda(a, b)$.

$$\begin{aligned} \nabla G_\lambda(a, b) &= \frac{1}{2G_\lambda(a, b)} \begin{pmatrix} 2(a - b) + \lambda b \\ -2(a - b) + \lambda a \end{pmatrix} \\ &= \frac{1}{2G_\lambda(a, b)} \begin{pmatrix} 2a + (\lambda - 2)b \\ (\lambda - 2)a + 2b \end{pmatrix} \\ &= \frac{1}{2G_\lambda(a, b)} \begin{pmatrix} 2 & \lambda - 2 \\ \lambda - 2 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \end{aligned}$$

por lo tanto,

$$\nabla G_\lambda(a, b) = \frac{1}{2G_\lambda(a, b)} M \begin{pmatrix} a \\ b \end{pmatrix}, \quad (2.13)$$

donde,

$$M = \begin{pmatrix} 2 & \lambda - 2 \\ \lambda - 2 & 2 \end{pmatrix}. \quad (2.14)$$

Analizando esta matriz, observamos que $M^T M = M^2$ tiene valores propios

$$\rho_1 = \lambda^2 \quad \text{y} \quad \rho_2 = (4 - \lambda)^2.$$

Luego,

$$\|M\|_2 = \sqrt{\max\{|\rho_1|, |\rho_2|\}} = \sqrt{\max\{\lambda^2, (4 - \lambda)^2\}} = \max\{\lambda, 4 - \lambda\}.$$

Dado que $\|M\|_2$ depende de λ y que $\lambda \in (0, 4)$, se tiene que

$$\|M\|_2 < 4, \quad (2.15)$$

lo cual ilustramos en la Figura 2.3.

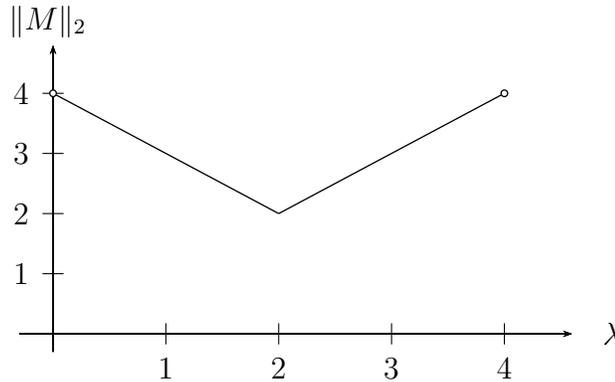


Figura 2.3: Función $\|M\|_2$.

Otra de las propiedades importantes de la *función de Kanzow* es que es *uniformemente continua*; es más, es *Lipschitz continua*, como lo confirma el siguiente lema.

Lema 2.3. *La función de Kanzow es Lipschitz continua con constante $2\sqrt{2}$. Es decir, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$*

$$|\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{y})| \leq 2\sqrt{2} \|\mathbf{x} - \mathbf{y}\|_2.$$

Demostración. Supongamos que el vector $(0, 0)$ no está en el segmento $[\mathbf{x}, \mathbf{y}]$. El *teorema del valor medio* [3] garantiza que existe un vector \mathbf{z} en el segmento abierto (\mathbf{x}, \mathbf{y}) tal

que,

$$\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{y}) = \nabla\varphi_\lambda(\mathbf{z})^T (\mathbf{x} - \mathbf{y}).$$

Así, usando la desigualdad de *Cauchy-Schwartz*, tenemos que

$$|\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{y})| \leq \|\nabla\varphi_\lambda(\mathbf{z})\|_2 \|\mathbf{x} - \mathbf{y}\|_2.$$

Por (2.12)

$$|\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{y})| \leq \sqrt{2} \|\mathbf{x} - \mathbf{y}\|_2. \quad (2.16)$$

Si el vector $(0,0)$ está en el segmento $[\mathbf{x}, \mathbf{y}]$, elegimos un vector $\mathbf{w} \notin [\mathbf{x}, \mathbf{y}]$, tal que $\|\mathbf{x} - \mathbf{w}\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_2$ y $\|\mathbf{y} - \mathbf{w}\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_2$, como se indica en la Figura 2.4. Aplicamos, el resultado (2.16) a los segmentos $[\mathbf{x}, \mathbf{w}]$ y $[\mathbf{w}, \mathbf{y}]$ junto con la desigualdad triangular y obtenemos

$$\begin{aligned} |\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{y})| &\leq |\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{w}) + \varphi_\lambda(\mathbf{w}) - \varphi_\lambda(\mathbf{y})| \\ &\leq |\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{w})| + |\varphi_\lambda(\mathbf{w}) - \varphi_\lambda(\mathbf{y})| \\ &\leq \sqrt{2} (\|\mathbf{x} - \mathbf{w}\|_2 + \|\mathbf{y} - \mathbf{w}\|_2) \\ &\leq 2\sqrt{2} \|\mathbf{x} - \mathbf{y}\|_2. \end{aligned} \quad (2.17)$$

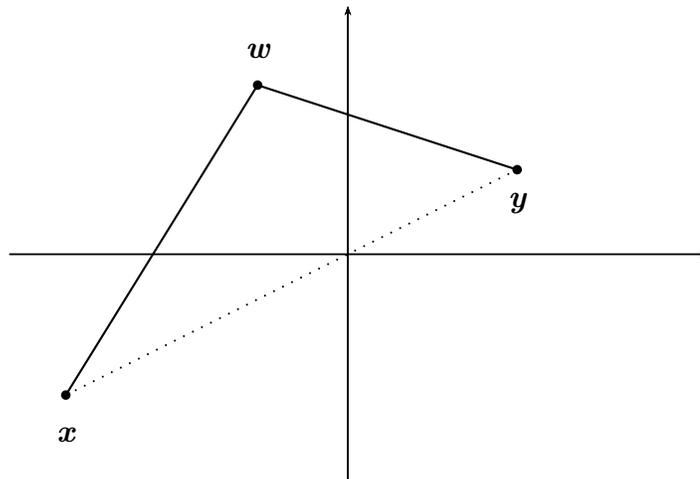


Figura 2.4: *Desigualdad triangular.*

De (2.16) y (2.17) concluimos que φ_λ es *Lipschitz continua* con constante $2\sqrt{2}$. ★

Corolario 2.1. La función G_λ es Lipschitz continua, es decir, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, se tiene que

$$|G_\lambda(\mathbf{x}) - G_\lambda(\mathbf{y})| \leq 2\sqrt{2} \|\mathbf{x} - \mathbf{y}\|_2. \quad (2.18)$$

Demostración. Es similar a la demostración del lema anterior. ★

El **Lema 2.4** proporciona cotas para cada una de las derivadas parciales de G_λ y el **Lema 2.5** garantiza que el gradiente de φ_λ también es una función localmente Lipschitz continua.

Lema 2.4. Sea $\lambda \in (0, 4)$. Las derivadas parciales de G_λ definidas en (2.11) satisfacen las siguientes desigualdades

$$|\chi(a, b)| \leq 1 \quad y \quad |\psi(a, b)| \leq \sqrt{2}, \quad (2.19)$$

para todo $(a, b) \neq (0, 0)$

Demostración. Iniciamos con el análisis de la función χ . Dado que $\lambda \in (0, 4)$, se tiene que

$$b^2\lambda(\lambda - 4) \leq 0,$$

equivalentemente

$$b^2\lambda^2 - 4b^2\lambda \leq 0.$$

Adicionamos los términos $4a^2 + 4ab\lambda - 8ab + 4b^2$ y después de algunas manipulaciones algebraicas obtenemos las siguientes desigualdades

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

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

por (2.5), se tiene $(a - b)^2 + ab\lambda \geq 0$

$$\begin{aligned} |2(a - b) + b\lambda| &\leq 2\sqrt{(a - b)^2 + ab\lambda} \\ \frac{|2(a - b) + b\lambda|}{2\sqrt{(a - b)^2 + ab\lambda}} &\leq 1 \\ |\chi(a, b)| &\leq 1. \end{aligned}$$

Para acotar la función ψ , aplicamos (2.12) y obtenemos

$$|\psi(a, b)| \leq \|G_\lambda(a, b)\| \leq \sqrt{2},$$

lo cual completa la prueba. ★

Lema 2.5. Sean φ_λ la función de Kanzow, \mathbf{w} un vector no nulo de \mathbb{R}^2 , $\mathcal{B}(\mathbf{w}; \epsilon)$ la bola que no contiene a $(0, 0)$, con $0 < \epsilon < \frac{1}{2} \|\mathbf{w}\|_2$. Entonces existe $\eta > 0$ tal que para todo $\mathbf{u}, \mathbf{v} \in \mathcal{B}(\mathbf{w}; \epsilon)$,

$$\|\nabla\varphi_\lambda(\mathbf{u}) - \nabla\varphi_\lambda(\mathbf{v})\|_2 \leq \eta \|\mathbf{u} - \mathbf{v}\|_2. \quad (2.20)$$

Demostración. Sean $\mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} c \\ d \end{pmatrix}$. De (2.10) y (2.13),

$$\begin{aligned} \|\nabla\varphi_\lambda(\mathbf{u}) - \nabla\varphi_\lambda(\mathbf{v})\|_2 &= \|\nabla G_\lambda(a, b) - \nabla G_\lambda(c, d)\|_2 \\ &\leq \frac{1}{2} \left\| \frac{1}{G_\lambda(a, b)} M \mathbf{u} - \frac{1}{G_\lambda(c, d)} M \mathbf{v} \right\|_2 \\ &\leq \frac{1}{2} \|M\|_2 \left\| \frac{1}{G_\lambda(a, b)} \mathbf{u} - \frac{1}{G_\lambda(c, d)} \mathbf{v} \right\|_2. \end{aligned} \quad (2.21)$$

Además, de (2.15) tenemos que

$$\|M\|_2 < 4,$$

para todo $\lambda \in (0, 4)$. Luego, de (2.21) y (2.15) tenemos la desigualdad

$$\|\nabla\varphi_\lambda(\mathbf{u}) - \nabla\varphi_\lambda(\mathbf{v})\|_2 \leq 2 \left\| \frac{1}{G_\lambda(a, b)} \mathbf{u} - \frac{1}{G_\lambda(c, d)} \mathbf{v} \right\|_2.$$

Sumamos y restamos la expresión $\frac{1}{G_\lambda(a, b)} \mathbf{v}$ para obtener

$$\|\nabla\varphi_\lambda(\mathbf{u}) - \nabla\varphi_\lambda(\mathbf{v})\|_2 \leq 2 \frac{1}{G_\lambda(a, b)} \|\mathbf{u} - \mathbf{v}\|_2 + 2 \left| \frac{1}{G_\lambda(a, b)} - \frac{1}{G_\lambda(c, d)} \right| \|\mathbf{v}\|_2. \quad (2.22)$$

Del segundo sumando, tenemos

$$\left| \frac{1}{G_\lambda(a, b)} - \frac{1}{G_\lambda(c, d)} \right| = \left| \frac{G_\lambda(c, d) - G_\lambda(a, b)}{G_\lambda(a, b)G_\lambda(c, d)} \right|,$$

como la función G_λ es *Lipschitz continua* (**Corolario 2.18**),

$$\left| \frac{1}{G_\lambda(a, b)} - \frac{1}{G_\lambda(c, d)} \right| \leq \frac{2\sqrt{2}}{G_\lambda(a, b)G_\lambda(c, d)} \|\mathbf{u} - \mathbf{v}\|_2. \quad (2.23)$$

Luego, de las desigualdades(2.22) y (2.23), tenemos

$$\|\nabla\varphi(\mathbf{u}) - \nabla\varphi(\mathbf{v})\|_2 \leq 2 \left(\frac{1}{G_\lambda(a, b)} + \frac{2\sqrt{2}\|\mathbf{v}\|_2}{G_\lambda(a, b)G_\lambda(c, d)} \right) \|\mathbf{u} - \mathbf{v}\|_2$$

y por (2.9),

$$\|\nabla\varphi(\mathbf{u}) - \nabla\varphi(\mathbf{v})\|_2 \leq \frac{2}{\sqrt{\alpha_{\min}}} \left(\frac{1}{\sqrt{a^2 + b^2}} + \frac{2\sqrt{2}\sqrt{c^2 + d^2}}{\sqrt{\alpha_{\min}}\sqrt{a^2 + b^2}\sqrt{c^2 + d^2}} \right) \|\mathbf{u} - \mathbf{v}\|_2$$

es decir,

$$\|\nabla\varphi(\mathbf{u}) - \nabla\varphi(\mathbf{v})\|_2 \leq \frac{2}{\sqrt{\alpha_{\min}}\sqrt{a^2 + b^2}} \left(1 + \frac{2\sqrt{2}}{\sqrt{\alpha_{\min}}} \right) \|\mathbf{u} - \mathbf{v}\|_2.$$

Por otro lado, $\mathbf{u} \in \mathcal{B}(\mathbf{w}; \epsilon)$, por lo tanto

$$\|\mathbf{u} - \mathbf{w}\|_2 < \epsilon,$$

de donde

$$-\epsilon + \|\mathbf{w}\|_2 < \|\mathbf{u}\|_2,$$

luego

$$\|\mathbf{u}\|_2 > \epsilon,$$

Finalmente, para $\mathbf{u}, \mathbf{v} \in \mathcal{B}(\mathbf{w}; \epsilon)$, concluimos que,

$$\|\nabla\varphi(\mathbf{u}) - \nabla\varphi(\mathbf{v})\|_2 \leq \eta \|\mathbf{u} - \mathbf{v}\|_2,$$

$$\text{donde, } \eta = \frac{2}{\sqrt{\alpha_{\min}}\epsilon} \left(1 + \frac{2\sqrt{2}}{\sqrt{\alpha_{\min}}} \right).$$

★

Corolario 2.2. *Las funciones χ y ψ descritas en (2.11) satisfacen la desigualdades siguientes.*

$$|\chi(a, b) - \chi(a, c)| \leq \eta |b - c| \quad y \quad |\psi(a, b) - \psi(a, c)| \leq \eta |b - c|. \quad (2.24)$$

Demostración. Claramente, a partir de (2.10) y (2.11)

$$\nabla\varphi_\lambda(a, b) - \nabla\varphi_\lambda(a, c) = \begin{pmatrix} \chi(a, b) - \chi(a, c) \\ \psi(a, b) - \psi(a, c) \end{pmatrix}.$$

Por propiedades de la *norma vectorial dos*, tenemos,

$$|\chi(a, b) - \chi(a, c)| \leq \|\nabla\varphi_\lambda(a, b) - \nabla\varphi_\lambda(a, c)\|_2,$$

y por el **Lema 2.5**,

$$\|\nabla\varphi_\lambda(a, b) - \nabla\varphi_\lambda(a, c)\|_2 \leq \eta \left\| \begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} a \\ c \end{pmatrix} \right\|_2 = \eta |b - c|,$$

por lo tanto,

$$|\chi(a, b) - \chi(a, c)| \leq \eta |b - c|.$$

En forma análoga, se demuestra la desigualdad,

$$|\psi(a, b) - \psi(a, c)| \leq \eta |b - c|.$$

★

Capítulo 3

El PCNL y su reformulación mediante la *función de Kanzow*

En este capítulo, reformulamos el Problema de Complementariedad No Lineal PCNL, como un sistema de ecuaciones no lineales no suave usando la función Φ_λ definida en (1.3). A partir de la definición de *jacobiano generalizado* dada en [11] construimos un subconjunto de matrices del *jacobiano generalizado* de Φ_λ en \mathbf{x} y demostramos que este subconjunto, en una solución del sistema $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$, es compacto.

En el planteamiento del PCNL, la tercera condición de (1.1) exige que los vectores \mathbf{x} y $F(\mathbf{x})$ sean ortogonales, por ello, es llamada *condición de complementariedad*. Además, de la misma condición se infiere que todos los términos $x_i F_i(\mathbf{x})$ deben ser cero y dado que el vector $\mathbf{x} \geq 0$ debe estar en una región de \mathbb{R}^n donde $F(\mathbf{x}) \geq 0$, se tiene que el vector cero de \mathbb{R}^n no siempre es una solución del PCNL. Este es el caso de la función F definida por,

$$F(x, y) = \begin{pmatrix} F_1(x, y) \\ F_2(x, y) \end{pmatrix} = \begin{pmatrix} (x - 2)^2 + y^2 - 1 \\ x^2 - y - 2 \end{pmatrix}.$$

El punto $\mathbf{x}_1 = (0, 0)$, quien a simple vista parecería ser solución del PCNL asociado a F , (Figura 3.1) no lo es, ya que $F(\mathbf{x}_1) = (3, -2)$ no satisface la condición de no negatividad.

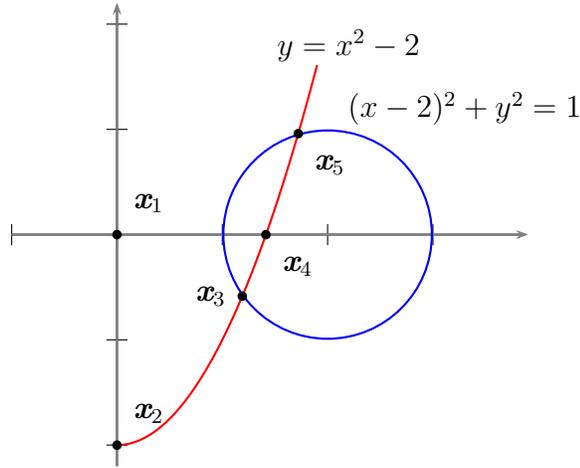


Figura 3.1: Gráficas de $F_1(x, y) = 0$ y $F_2(x, y) = 0$.

En la Figura 3.1 podemos observar que los puntos \mathbf{x}_2 y \mathbf{x}_3 , tampoco satisfacen la condición de no negatividad y el punto $\mathbf{x}_4 = (\sqrt{2}, 0)$ no satisface la condición de complementariedad. Por último, el punto $\mathbf{x}_5 = (1.7205, 0.9601)$ cumple las tres condiciones, luego es una solución del PCNL asociado a F .

A continuación, consideramos la *función de Kanzow* φ_λ definida en (2.1) y reformulamos el PCNL como el sistema de ecuaciones no lineales no diferenciable

$$\Phi_\lambda(\mathbf{x}) = \begin{pmatrix} \varphi_\lambda(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi_\lambda(x_n, F_n(\mathbf{x})) \end{pmatrix} = \mathbf{0}. \quad (3.1)$$

Es importante destacar el hecho de que (3.1) es un sistema de ecuaciones no lineales, *no diferenciable*. Aspecto clave para tener en cuenta, al resolverlo.

Entre los métodos más populares para resolver un sistema de ecuaciones no lineales, **diferenciable** $H(\mathbf{x}) = \mathbf{0}$, se encuentran los métodos *tipo Newton* [1] [6], los cuales requieren calcular, en cada iteración, la *matriz jacobiana* de H . Esto, desde el punto de vista del costo computacional, resulta poco eficiente. Una alternativa menos costosa y muy utilizada, la constituyen los métodos cuasi-Newton [1] [6] que aproximan en cada iteración la *matriz jacobiana*. Entre estos últimos, se encuentran los llamados *métodos secante de cambio mínimo* [14] [33], los cuales forman una familia que se caracteriza porque, en cada iteración, la matriz que aproxima a la matriz jacobiana satisface una

ecuación que se conoce como *ecuación secante* [14] [33] y el cambio que se presenta entre dos aproximaciones consecutivas es el *mínimo posible*. Es esto lo que justifica el calificativo de *métodos secante de cambio mínimo*. Un estudio detallado de estos métodos se presenta en [32] [33] [14]. El precio de usar una aproximación de la matriz jacobiana se ve reflejado en la disminución de la velocidad de convergencia del respectivo algoritmo secante.

Sin embargo, cuando una función **no es diferenciable**, como en el caso de la función Φ_λ , hablar de “la matriz jacobiana” no tiene sentido.

En 1973, *Frank H. Clarke*, en su *tesis doctoral* [9], presentó un concepto que extiende el de matriz jacobiana para algunas funciones no diferenciables, este concepto es el de *jacobiano generalizado* definido para $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, *Lipschitz continua* en un punto \mathbf{x} , como el conjunto,

$$\partial F(\mathbf{x}) = \text{conv} \left\{ \lim_{k \rightarrow \infty} F'(\mathbf{x}_k) \in \mathbb{R}^{n \times n} : \mathbf{x}_k \rightarrow \mathbf{x}, \mathbf{x}_k \in D_F \right\} \quad (3.2)$$

donde D_F es el conjunto de todos los puntos de \mathbb{R}^n en los que la función F es diferenciable y $\text{conv} \{A\}$ representa la envolvente convexa de A . El conjunto $\partial F(\mathbf{x})$ es *no vacío, convexo y compacto* y [10] [11].

En el caso particular en que F es diferenciable en el punto \mathbf{x} , el conjunto $\partial F(\mathbf{x})$ tiene un único elemento: la *matriz jacobiana* de F en \mathbf{x} , $F'(\mathbf{x})$.

Ya que la *función de Kanzow* φ_λ es *Lipschitz continua* (**Lema 2.3**), también lo es la función Φ_λ . Así, el *jacobiano generalizado* de $\Phi_\lambda(\mathbf{x})$ existe.

Con el fin de construir matrices en $\partial \Phi_\lambda(\mathbf{x})$ consideramos una sucesión de vectores $\{\mathbf{y}_k\}$ en \mathbb{R}^n que converge a \mathbf{x} y tal que $\Phi'_\lambda(\mathbf{y}_k)$ existe, luego mostraremos que $\lim_{k \rightarrow \infty} \Phi'_\lambda(\mathbf{y}_k)$ existe.

Para clasificar los índices de las componentes del punto \mathbf{x} definimos el conjunto

$$\beta = \beta(\mathbf{x}) = \{i : x_i = F_i(\mathbf{x}) = 0\}. \quad (3.3)$$

La sucesión que usaremos es la siguiente¹

$$\mathbf{y}_k = \mathbf{x} + \varepsilon_k \mathbf{z} \quad (3.4)$$

donde, $\{\varepsilon_k\}$ es una sucesión de números positivos tales que $\lim_{k \rightarrow \infty} \varepsilon_k = 0$ y el vector \mathbf{z} se elige de forma que $z_i \neq 0$ cuando $i \in \beta$. Evidentemente \mathbf{y}_k converge a \mathbf{x} cuando $k \rightarrow \infty$.

Para analizar la diferenciabilidad de Φ_λ en \mathbf{y}_k consideramos dos casos.

1. Si $i \notin \beta$, es decir $x_i \neq 0$ o $F_i(\mathbf{x}) \neq 0$, como F_i es continua, podemos suponer ε_k tan pequeño que $y_i^k \neq 0$ o $F_i(\mathbf{x}) \neq 0$, por lo cual Φ_λ es diferenciable en \mathbf{y}_k .
2. Si $i \in \beta$, se tiene que $z_i \neq 0$; por lo tanto, $y_i^k \neq 0$, lo cual es suficiente para que Φ_λ sea diferenciable en \mathbf{y}_k .

Dado que Φ_λ es diferenciable en \mathbf{y}_k , existe la matriz jacobiana de Φ_λ en cada \mathbf{y}_k . En efecto,

$$\Phi'_\lambda(\mathbf{y}_k) = \begin{pmatrix} \nabla\varphi_\lambda(y_1^k, F_1(\mathbf{y}_k))^T \\ \vdots \\ \nabla\varphi_\lambda(y_n^k, F_n(\mathbf{y}_k))^T \end{pmatrix},$$

donde la fila i -ésima de $\Phi'_\lambda(\mathbf{y}_k)$ es de la forma

$$(\chi(y_i^k, F_i(\mathbf{y}_k)) - 1) \mathbf{e}_i^T + (\psi(y_i^k, F_i(\mathbf{y}_k)) - 1) \nabla F_i(\mathbf{y}_k)^T$$

con, χ y ψ definidas por (2.11) y $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ son los vectores canónicos de \mathbb{R}^n .

Así,

$$\lim_{k \rightarrow \infty} \Phi'_\lambda(\mathbf{y}_k) = \lim_{k \rightarrow \infty} \begin{pmatrix} \nabla\varphi_\lambda(y_1^k, F_1(\mathbf{y}_k))^T \\ \vdots \\ \nabla\varphi_\lambda(y_n^k, F_n(\mathbf{y}_k))^T \end{pmatrix}. \quad (3.5)$$

¹Consideramos la misma sucesión de [13], donde los autores la usan para desarrollos teórico con la *función de Fischer*.

Para calcular este límite, consideremos dos casos:

1. Si $i \notin \beta$, como toda la fila es continua en los puntos de la sucesión, se tiene que

$$\lim_{k \rightarrow \infty} \nabla \varphi_\lambda(y_i^k, F_i(\mathbf{y}_k))^T = [H]_i$$

donde

$$[H]_i = (\chi(x_i, F_i(\mathbf{x})) - 1) \mathbf{e}_i^T + (\psi(x_i, F_i(\mathbf{x})) - 1) \nabla F_i(\mathbf{x})^T.$$

2. Si $i \in \beta$, de (3.4), tenemos que

$$y_i^k = \varepsilon_k z_i, \quad (3.6)$$

además, por el *Teorema de Taylor* tenemos que

$$F_i(\mathbf{y}_k) = F_i(\mathbf{x} + \varepsilon_k \mathbf{z}) = F_i(\mathbf{x}) + \varepsilon^k \nabla F_i(\boldsymbol{\zeta}^k)^T \mathbf{z} = \varepsilon^k \nabla F_i(\boldsymbol{\zeta}^k)^T \mathbf{z}, \quad (3.7)$$

donde $\boldsymbol{\zeta}^k \rightarrow \mathbf{x}$, cuando $k \rightarrow \infty$. Al sustituir (3.6) y (3.7) en la fila i -ésima de $\Phi'_\lambda(\mathbf{y}_k)$, con $i \in \beta$, se tiene que

$$\lim_{k \rightarrow \infty} \nabla \varphi_\lambda(y_i^k, F_i(\mathbf{y}_k))^T = [H]_i$$

con

$$[H]_i = (\chi(z_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) - 1) \mathbf{e}_i^T + (\psi(z_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) - 1) \nabla F_i(\mathbf{x})^T.$$

Luego, para todo $i = 1, \dots, n$, el límite cuando $k \rightarrow \infty$ de cada fila de la matriz $\Phi'_\lambda(\mathbf{y}_k)$ existe. Por lo tanto, de (3.5),

$$\lim_{k \rightarrow \infty} \Phi'_\lambda(\mathbf{y}_k) = \begin{pmatrix} [H]_1 \\ \vdots \\ [H]_n \end{pmatrix} = H, \quad (3.8)$$

donde

$$[H]_i = \begin{cases} (\chi(x_i, F_i(\mathbf{x})) - 1) \mathbf{e}_i^T + (\psi(x_i, F_i(\mathbf{x})) - 1) \nabla F_i(\mathbf{x})^T, & i \notin \beta \\ (\chi(z_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) - 1) \mathbf{e}_i^T + (\psi(z_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) - 1) \nabla F_i(\mathbf{x})^T, & i \in \beta. \end{cases}$$

Debido a que existe un número no contable de formas de elegir \mathbf{z} , tenemos un conjunto no contable de matrices H en $\partial\Phi_\lambda(\mathbf{x})$. Por otro lado, la anterior construcción de la matriz H , plantea un procedimiento para elegir las matrices del *jacobiano generalizado* de Φ_λ en \mathbf{x} .

Consideremos el caso particular en que \mathbf{x}^* es una solución del sistema de ecuaciones no lineales $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$. Es decir,

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

Cabe mencionar que, si para una solución \mathbf{x}^* del PCNL, existe algún índice i tal que $x_i = F_i(\mathbf{x}^*) = 0$, decimos que \mathbf{x}^* es una *solución degenerada*.

Las matrices (3.8) en \mathbf{x}^* las denotaremos por $H_*(\mathbf{z})$. Teniendo en cuenta lo anterior, podemos concluir que $H_*(\mathbf{z})$ es de la forma,

$$H_* = H_*(\mathbf{z}) = \begin{pmatrix} [H_*(\mathbf{z})]_1 \\ \vdots \\ [H_*(\mathbf{z})]_n \end{pmatrix} = \begin{pmatrix} [H_*]_1 \\ \vdots \\ [H_*]_n \end{pmatrix}, \quad (3.9)$$

con,

$$[H_*(\mathbf{z})]_i = \begin{cases} \left(\frac{\lambda}{2} - 2\right) \mathbf{e}_i^T, & x_i^* = 0, F_i(\mathbf{x}^*) \neq 0 \\ \left(\frac{\lambda}{2} - 2\right) \nabla F_i(\mathbf{x}^*)^T, & x_i^* \neq 0, F_i(\mathbf{x}^*) = 0 \\ (\alpha_i(\mathbf{z}) - 1) \mathbf{e}_i^T + (\rho_i(\mathbf{z}) - 1) \nabla F_i(\mathbf{x}^*)^T, & x_i^* = 0, F_i(\mathbf{x}^*) = 0 \end{cases}$$

donde

$$\alpha_i(\mathbf{z}) = \chi(z_i, \nabla F_i(\mathbf{x}^*)^T \mathbf{z}),$$

$$\rho_i(\mathbf{z}) = \psi(z_i, \nabla F_i(\mathbf{x}^*)^T \mathbf{z}).$$

El conjunto formado por todas las matrices $H_*(\mathbf{z})$ lo llamaremos \mathcal{Z}_* , esto es,

$$\mathcal{Z}_* = \{H_*(\mathbf{z}) : \mathbf{z} \in \mathbb{R}^n \text{ tal que } z_i \neq 0, \text{ si } x_i = F_i(\mathbf{x}^*) = 0\}. \quad (3.10)$$

Claramente, para cada $\mathbf{z} \in \mathbb{R}^n$, existe una matriz $H_*(\mathbf{z})$. Así, el conjunto \mathcal{Z}_* es no contable y además es *compacto*. Para verificar la compacidad, es suficiente con demostrar

que \mathcal{Z}_* es *cerrado*, pues $\partial\Phi(\mathbf{x}^*)$ es *compacto* [11] y $\mathcal{Z}_* \subseteq \partial\Phi(\mathbf{x}^*)$.

Sea $\{H_*^p\}$ una sucesión de matrices en \mathcal{Z}_* . Supongamos que

$$\lim_{p \rightarrow \infty} H_*^p = H_*.$$

Como $H_*^p = H_*^p(\mathbf{z})$, esta sucesión de matrices depende de una sucesión de vectores $\{\mathbf{z}^p\}$, supongamos que $\lim_{p \rightarrow \infty} \mathbf{z}^p = \mathbf{z}$. Debemos probar que H_* pertenece a \mathcal{Z}_* , para ello, enfocaremos este análisis en cada una de las filas de H_* , en efecto,

$$H_* = \lim_{p \rightarrow \infty} H_*^p = \begin{pmatrix} \lim_{p \rightarrow \infty} [H_*^p]_1 \\ \vdots \\ \lim_{p \rightarrow \infty} [H_*^p]_n \end{pmatrix}$$

Definamos los siguientes conjuntos de índices,

$$\beta_0 = \{i : x_i^* = 0 = F_i(\mathbf{x}^*)\}, \quad \beta_1 = \{i : x_i^* = 0 < F_i(\mathbf{x}^*)\}, \quad \beta_2 = \{i : x_i^* > 0 = F_i(\mathbf{x}^*)\}.$$

En los casos en que $i \in \beta_1 \cup \beta_2$, las filas de H_* no dependen de \mathbf{z}^p , se mantienen constantes; es decir,

$$\lim_{p \rightarrow \infty} [H_*^p]_i = [H_*]_i.$$

Por otra parte, si $i \in \beta_0$

$$\begin{aligned} \lim_{p \rightarrow \infty} [H_*^p]_i &= \lim_{p \rightarrow \infty} [(\alpha_i(\mathbf{z}^p) - 1) \mathbf{e}_i^T + (\rho_i(\mathbf{z}^p) - 1) \nabla F_i(\mathbf{x}^*)^T] \\ &= \left(\lim_{p \rightarrow \infty} \alpha_i(\mathbf{z}^p) - 1 \right) \mathbf{e}_i^T + \left(\lim_{p \rightarrow \infty} \rho_i(\mathbf{z}^p) - 1 \right) \nabla F_i(\mathbf{x}^*)^T \\ &= (\alpha_i(\mathbf{z}) - 1) \mathbf{e}_i^T + (\rho_i(\mathbf{z}) - 1) \nabla F_i(\mathbf{x}^*)^T. \end{aligned}$$

Probamos así, que el conjunto \mathcal{Z}_* es *cerrado* y por lo tanto *compacto*.

Capítulo 4

Algoritmo y teoría de convergencia

En la primera parte de este capítulo, proponemos un nuevo algoritmo *cuasi Newton* genérico para resolver el sistema de ecuaciones no lineales, no diferenciable, $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$ descrito en (3.1) y desarrollamos la teoría de convergencia local para el mismo. En la segunda parte, generamos una familia de *métodos secante de cambio mínimo*, siguiendo las reglas de [31] para métodos de este tipo. Demostramos que dicha familia proporciona algoritmos local y superlinealmente convergentes.

Motivados en el hecho que una aproximación *cuasi Newton* es una buena alternativa para resolver un sistema de ecuaciones no lineales cuando la matriz jacobiana (caso diferenciable) no está disponible o su cálculo involucra muchas operaciones, proponemos el algoritmo *cuasi Newton* siguiente para resolver el sistema $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$ y con ello resolver el PCNL.

Algoritmo 4.1. *Dados \mathbf{x}_0 y $\lambda \in (0, 4)$, para $k = 1, 2, \dots$, generamos la iteración siguiente por*

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

donde

$$B_k = \begin{pmatrix} [B_k]_1 \\ \vdots \\ [B_k]_n \end{pmatrix} \quad (4.1)$$

con,

$$[B_k]_i = \begin{cases} (\chi(x_i, F_i(\mathbf{x})) - 1) \mathbf{e}_i^T + (\psi(x_i, F_i(\mathbf{x})) - 1) [A_k]_i, & i \notin \beta \\ (\chi(z_i, [A_k]_i \mathbf{z}_k) - 1) \mathbf{e}_i^T + (\psi(z_i, [A_k]_i \mathbf{z}_k) - 1) [A_k]_i, & i \in \beta. \end{cases}$$

donde $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ es la base canónica de \mathbb{R}^n , la matriz

$$A_k = \begin{pmatrix} [A_k]_1 \\ \vdots \\ [A_k]_n \end{pmatrix}$$

es una aproximación de la matriz jacobiana $F'(\mathbf{x}_k)$, y $\mathbf{z}^k \in \mathbb{R}^n$ es un vector tal que $z_i^k \neq 0$, si $x_i^k = F_i(\mathbf{x}^k) = 0$.

Las hipótesis bajo las cuales desarrollamos la teoría de convergencia local del **Algoritmo 4.1** son las siguientes.

4.1. Hipótesis

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

H2. La matriz jacobiana de F es *Lipschitz continua* en una vecindad de $\mathbf{x}^* \in \mathbb{R}^n$. Es decir, existen constantes $\gamma > 0$ y $\delta > 0$ tales que

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

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. Las matrices de \mathcal{Z}_* son no singulares.

Por **H3** y la compacidad de \mathcal{Z}_* tenemos que existe una constante μ tal que para toda $H_*(\mathbf{z}) \in \mathcal{Z}_*$,

$$\|H_*^{-1}(\mathbf{z})\| \leq \mu. \quad (4.2)$$

4.2. Resultados de convergencia

A continuación presentamos dos lemas útiles en la demostración del **Teorema 4.1**.

Lema 4.1. Sean $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F \in C^1$, cuya matriz jacobiana satisface la hipótesis **H2**, $\epsilon > 0$ y $\delta > 0$ dados, H y B definidas en (3.8) y (4.1), respectivamente. Entonces, para cada $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*; \epsilon)$ y para cada $A \in \mathcal{B}(F'(\mathbf{x}^*); \delta)$ existe una constante $\theta > 0$ tal que

$$\|H - B\|_\infty \leq \theta. \quad (4.3)$$

Demostración. Consideremos las matrices H y B definidas en (3.8) y (4.1), respectivamente. Sean

$$\begin{aligned} a_i &= \chi(x_i, F_i(\mathbf{x})), & b_i &= \psi(x_i, F_i(\mathbf{x})), \\ c_i &= \chi(z_i, \nabla F_i(\mathbf{x})^T \mathbf{z}), & d_i &= \psi(z_i, \nabla F_i(\mathbf{x})^T \mathbf{z}), \\ \tilde{c}_i &= \chi(z_i, [A]_i \mathbf{z}) & y & \quad \tilde{d}_i = \psi(z_i, [A]_i \mathbf{z}). \end{aligned}$$

De la definición de *norma matricial infinito* [44], tenemos que:

$$\|H - B\|_\infty = \max_{1 \leq i \leq n} \{ \| [H]_i - [B]_i \|_1 \} \quad (4.4)$$

por lo cual, nos interesa analizar cada una de las filas de la matriz $H - B$.

De (3.8) y (4.1), obtenemos

$$[H - B]_i = \begin{cases} (b_i - 1) \left(\nabla F_i(\mathbf{x})^T - [A]_i \right), & i \notin \beta, \\ (c_i - \tilde{c}_i) \mathbf{e}_i^T + (d_i - 1) \nabla F_i(\mathbf{x})^T - (\tilde{d}_i - 1) [A]_i, & i \in \beta. \end{cases}$$

Supongamos que el máximo (4.4) se alcanza en la fila j . Así, debemos considerar dos casos.

1. Si $j \notin \beta$, donde β esta definido en (3.3), entonces

$$\begin{aligned}
\|H - B\|_\infty &= \left\| [H]_j - [B]_j \right\|_1 \\
&= \left\| (b_j - 1) \left(\nabla F_j(\mathbf{x})^T - [A]_j \right) \right\|_1 \\
&\leq |b_j - 1| \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 \\
&\leq (|b_j| + 1) \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 \\
&\leq (|\psi(x_j, F_j(\mathbf{x}))| + 1) \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1. \tag{4.5}
\end{aligned}$$

Luego, de (4.5) y (2.19), obtenemos

$$\|H - B\|_\infty \leq (\sqrt{2} + 1) \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1. \tag{4.6}$$

Por otra parte, el factor $\left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1$ lo acotamos usando la *norma vectorial infinito*¹, luego utilizamos la hipótesis **H2** y el hecho que $A \in \mathcal{B}(F'(\mathbf{x}^*); \delta)$, con lo cual,

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

de donde obtenemos la cota,

$$\left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 \leq n(\gamma\epsilon + \delta), \tag{4.7}$$

¹Para todo $\mathbf{y} \in \mathbb{R}^n$, se tiene que $\|\mathbf{y}\|_1 \leq n \|\mathbf{y}\|_\infty$ [44].

por lo tanto, de (4.6) y (4.7)

$$\|H - B\|_\infty \leq (\sqrt{2} + 1) n(\gamma\epsilon + \delta).$$

2. Si $j \in \beta$,

$$\begin{aligned} \|H - B\|_\infty &= \left\| [H]_j - [B]_j \right\|_1 \\ &= \left\| (c_j - \tilde{c}_j) \mathbf{e}_j^T + (d_j - 1) \nabla F_j(\mathbf{x})^T - (\tilde{d}_j - 1) [A]_j \right\|_1 \\ &= \left\| (c_j - \tilde{c}_j) \mathbf{e}_j^T \right\| + \left\| d_j \nabla F_j(\mathbf{x})^T - \tilde{d}_j [A]_j \right\|_1 + \left\| \nabla F_j(\mathbf{x})^T - [A]_j \right\|_1 \end{aligned} \quad (4.8)$$

A continuación, acotamos los dos primeros sumandos del lado derecho de (4.8).

$$\begin{aligned} \left\| (c_j - \tilde{c}_j) \mathbf{e}_j^T \right\|_1 &= |c_j - \tilde{c}_j| \left\| \mathbf{e}_j^T \right\|_1 = |c_j - \tilde{c}_j| \\ &\leq \left| \chi(z_j, \nabla F_j(\mathbf{x})^T \mathbf{z}) - \chi(z_j, [A]_j \mathbf{z}) \right| \\ &\leq \eta \left| \nabla F_j(\mathbf{x})^T \mathbf{z} - [A]_j \mathbf{z} \right| \\ &\leq \eta \left| \left(\nabla F_j(\mathbf{x})^T - [A]_j \right) \mathbf{z} \right| \\ &\leq \eta \delta \|\mathbf{z}\|_\infty, \end{aligned} \quad (4.9)$$

donde la última desigualdad se tiene por el **Corolario 2.2**.

Además,

$$\begin{aligned}
\left\| d_j \nabla F_j(\mathbf{x})^T - \tilde{d}_j [A]_j \right\|_1 &= \left\| d_j \nabla F_j(\mathbf{x})^T - d_j \nabla F_j(\mathbf{x}^*)^T + d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1 \\
&\leq |d_j| \left\| \nabla F_j(\mathbf{x}) - \nabla F_j(\mathbf{x}^*) \right\|_1 + \left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1 \\
&\leq n |d_j| \left\| \nabla F_j(\mathbf{x}) - \nabla F_j(\mathbf{x}^*) \right\|_\infty + \left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1 \\
&\leq n |d_j| \gamma \epsilon + \left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1 \\
&= n \left| \psi(z_j, \nabla F_j(\mathbf{x})^T \mathbf{z}) \right| \gamma \epsilon + \left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1, \\
&\leq n \sqrt{2} \gamma \epsilon + \left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1. \tag{4.10}
\end{aligned}$$

Con el fin de acotar el segundo sumando en la desigualdad (4.10), sumamos y restamos la expresión $\tilde{d}_j \nabla F_j(\mathbf{x}^*)^T$, usamos (2.24) y la cota para $\left| \tilde{d}_j \right|$ dada por (2.19), teniendo en cuenta que $d_j = \psi(z_j, \nabla F_j(\mathbf{x})^T \mathbf{z})$ y $\tilde{d}_j = \psi(z_j, [A]_j \mathbf{z})$. Así,

$$\begin{aligned}
\left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1 &= \left\| (d_j - \tilde{d}_j) \nabla F_j(\mathbf{x}^*)^T + \tilde{d}_j \left(\nabla F_j(\mathbf{x}^*)^T - [A]_j \right) \right\|_1 \\
&\leq |d_j - \tilde{d}_j| \left\| \nabla F_j(\mathbf{x}^*) \right\|_1 + \left| \tilde{d}_j \right| \left\| \nabla F_j(\mathbf{x}^*)^T - [A]_j \right\|_1 \\
&\leq n |d_j - \tilde{d}_j| \left\| \nabla F_j(\mathbf{x}^*) \right\|_\infty + n \left| \tilde{d}_j \right| \left\| \nabla F_j(\mathbf{x}^*)^T - [A]_j \right\|_\infty.
\end{aligned}$$

Por el **Corolario 2.2**, tenemos

$$\begin{aligned}
\left\| d_j \nabla F_j(\mathbf{x}^*)^T - \tilde{d}_j [A]_j \right\|_1 &\leq n \left(\eta \left| \nabla F_j(\mathbf{x})^T \mathbf{z} - [A]_j^T \mathbf{z} \right| \right) \left\| \nabla F_j(\mathbf{x}^*) \right\|_\infty + n \left| \tilde{d}_j \right| \delta \\
&\leq n \left(\eta \left| \left(\nabla F_j(\mathbf{x})^T - [A]_j \right) \mathbf{z} \right| \right) \left\| \nabla F_j(\mathbf{x}^*) \right\|_\infty + \sqrt{2} \delta \\
&\leq n \delta \left(\eta \left\| \mathbf{z} \right\|_\infty \left\| \nabla F_j(\mathbf{x}^*) \right\|_\infty + \sqrt{2} \right).
\end{aligned}$$

Reemplazando esta cota en (4.10),

$$\left\| d_j \nabla F_j(\mathbf{x})^T - \tilde{d}_j [A]_j \right\|_1 \leq n \sqrt{2} \gamma \epsilon + n \delta (\eta \|\mathbf{z}\|_\infty \|\nabla F_j(\mathbf{x}^*)\|_\infty + \sqrt{2}). \quad (4.11)$$

Finalmente, de (4.7), (4.9) y (4.11) obtenemos una cota superior para $\|H - B\|_\infty$.

$$\|H - B\|_\infty \leq \delta (\eta \|\mathbf{z}\|_\infty + n \eta \|\mathbf{z}\|_\infty \|\nabla F_j(\mathbf{x}^*)\|_\infty + \sqrt{2}n + n) + \epsilon n \gamma (\sqrt{2} + 1).$$

Por lo tanto,

$$\|H - B\|_\infty \leq \theta \quad (4.12)$$

donde

$$\theta = \delta \tau + \epsilon \omega, \quad (4.13)$$

con

$$\tau = \eta \|\mathbf{z}\|_\infty + n \eta \|\mathbf{z}\|_\infty \|\nabla F_j(\mathbf{x}^*)\|_\infty + \sqrt{2}n + n, \quad (4.14)$$

y

$$\omega = n \gamma (\sqrt{2} + 1) \quad (4.15)$$

pues es mayor que la cota resultante en el caso $j \notin \beta$. ★

Lema 4.2. Sean $r \in (0, 1)$ y B la matriz definida por (4.1). Existen constantes positivas ϵ_0 y δ_0 tales que,

$$\text{si } \|\mathbf{x} - \mathbf{x}^*\|_\infty \leq \epsilon_0 \quad \text{y} \quad \|A - F'(\mathbf{x}^*)\|_\infty \leq \delta_0,$$

la función \mathcal{Q} definida por

$$\mathcal{Q}(\mathbf{x}, A) = \mathbf{x} - B^{-1} \Phi_\lambda(\mathbf{x}), \quad (4.16)$$

está bien definida y

$$\|\mathcal{Q}(\mathbf{x}, A) - \mathbf{x}^*\|_\infty \leq r \|\mathbf{x} - \mathbf{x}^*\|_\infty. \quad (4.17)$$

Demostración. Sean $r \in (0, 1)$, $0 < \hat{\epsilon} < \frac{r}{8\mu(\omega + \sqrt{2}n\eta)}$ y $0 < \delta_0 < \frac{r}{8\mu\tau}$,

donde ω es la constante (4.15), η es la constante de Lipschitz de $\nabla \varphi_\lambda$ dada en (2.20), τ está definida en (4.14) y μ está definida en (4.2).

Supongamos que $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*; \hat{\epsilon})$, $A \in \mathcal{B}(F'(\mathbf{x}^*); \delta_0)$, B la matriz asociada a A por la regla (4.1), H_* la matriz asociada a $F'(\mathbf{x}^*)$ por la misma regla y H definida en (3.8).

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

$$\|B - H_*\|_\infty \leq \|B - H\|_\infty + \|H - H_*\|_\infty. \quad (4.18)$$

El primer sumando del lado derecho de (4.18) está acotado por (4.12), así

$$\|B - H\|_\infty \leq \delta_0 \tau + \hat{\epsilon} \omega < \frac{r}{8\mu} + \frac{\omega r}{8\mu(\omega + \sqrt{2n\eta})}. \quad (4.19)$$

Acotemos el segundo sumando del lado derecho de (4.18), para ello, utilicemos la continuidad de F , la *continuidad Lipschitz* de $\nabla\varphi_\lambda$ y la definición de *norma matricial infinito*.

Por la continuidad de F , tenemos que para $\hat{\epsilon}$ existe $\hat{\delta} > 0$ tal que,

$$\text{si } \|\mathbf{x} - \mathbf{x}^*\|_\infty < \hat{\delta} \quad \text{entonces} \quad |F_j(\mathbf{x}) - F_j(\mathbf{x}^*)| < \hat{\epsilon}.$$

Sea $\tilde{\epsilon} = \min\{\hat{\epsilon}, \hat{\delta}\}$.

$$\text{Si } \|\mathbf{x} - \mathbf{x}^*\|_\infty < \tilde{\epsilon} \quad \text{entonces} \quad |F_j(\mathbf{x}) - F_j(\mathbf{x}^*)| < \hat{\epsilon},$$

y por otra parte,

$$\begin{aligned} \|H - H_*\|_\infty &= \left\| [H]_j - [H_*]_j \right\|_1 \\ &\leq n \lim_{k \rightarrow \infty} \left\| \nabla\varphi_\lambda(y_j^k, F_j(\mathbf{y}^k)) - \nabla\varphi_\lambda(\bar{y}_j^k, F_j(\bar{\mathbf{y}}^k)) \right\|_\infty \\ &\leq n \lim_{k \rightarrow \infty} \left\| \nabla\varphi_\lambda(y_j^k, F_j(\mathbf{y}^k)) - \nabla\varphi_\lambda(\bar{y}_j^k, F_j(\bar{\mathbf{y}}^k)) \right\|_2 \end{aligned}$$

Como el gradiente de φ_λ es *Lipschitz continua* (2.20) tenemos

$$\begin{aligned} \|H - H_*\|_\infty &= n \lim_{k \rightarrow \infty} \eta \left\| \begin{pmatrix} y_j^k - \bar{y}_j^k \\ F_j(\mathbf{y}^k) - F_j(\bar{\mathbf{y}}^k) \end{pmatrix} \right\|_2 \\ &\leq \sqrt{2} n \eta \left\| \begin{pmatrix} x_j - x_j^* \\ F_j(\mathbf{x}) - F_j(\mathbf{x}^*) \end{pmatrix} \right\|_\infty \\ &= \sqrt{2} n \eta \max \{ |x_j - x_j^*|, |F_j(\mathbf{x}) - F_j(\mathbf{x}^*)| \}. \end{aligned}$$

Consideremos las dos posibilidades para este máximo.

1. Si $\max \{ |x_j - x_j^*|, |F_j(\mathbf{x}) - F_j(\mathbf{x}^*)| \} = |x_j - x_j^*| \leq \|\mathbf{x} - \mathbf{x}^*\|_\infty \leq \tilde{\epsilon} < \hat{\epsilon}$.
2. Si $\max \{ |x_j - x_j^*|, |F_j(\mathbf{x}) - F_j(\mathbf{x}^*)| \} = |F_j(\mathbf{x}^k) - F_j(\mathbf{x}^*)| < \hat{\epsilon}$.

de lo anterior

$$\|H - H_*\|_\infty < \sqrt{2} n \eta \hat{\epsilon} < \frac{\sqrt{2} n \eta r}{8 \mu (\omega + \sqrt{2} n \eta)}. \quad (4.20)$$

Sustituyendo (4.19) y (4.20) en la desigualdad (4.18)

$$\|B - H_*\|_\infty < \frac{r}{8 \mu} + \frac{\omega r}{8 \mu (\omega + \sqrt{2} n \eta)} + \frac{\sqrt{2} n \eta r}{8 \mu (\omega + \sqrt{2} n \eta)} = \frac{r}{4 \mu}.$$

Por lo tanto, de (4.2) se tiene

$$\|B - H_*\|_\infty < \frac{r}{4 \mu} < \frac{1}{4 \|H_*^{-1}\|_\infty}, \quad (4.21)$$

luego,

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

de donde $\|H_*^{-1}B - I_n\|_\infty < 1$. Por el *lema de Banach*² existe B^{-1} y por lo tanto la función \mathcal{Q} está bien definida. Además,

$$\|B^{-1}\|_\infty \leq \frac{\|H_*^{-1}\|_\infty}{1 - \|H_*^{-1}B - I_n\|_\infty} \leq \frac{\mu}{1 - \frac{1}{4}} = \frac{4}{3}\mu.$$

La segunda parte de la prueba consiste en probar (4.17). Para ello, sustraemos \mathbf{x}^* en (4.16), aplicamos $\|\cdot\|_\infty$ y realizamos algunas manipulaciones algebraicas.

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

de donde obtenemos la desigualdad,

$$\|\mathcal{Q}(\mathbf{x}, A) - \mathbf{x}^*\|_\infty \leq \frac{4}{3}\mu \left[\frac{r}{4\mu} + \frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}^*) + H_*(\mathbf{x} - \mathbf{x}^*)\|_\infty}{\|\mathbf{x} - \mathbf{x}^*\|_\infty} \right] \|\mathbf{x} - \mathbf{x}^*\|_\infty. \quad (4.22)$$

Por otra parte,

$$\frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}^*) + H_*(\mathbf{x} - \mathbf{x}^*)\|_\infty}{\|\mathbf{x} - \mathbf{x}^*\|_\infty} \leq \frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}^*) + H(\mathbf{x} - \mathbf{x}^*)\|_\infty}{\|\mathbf{x} - \mathbf{x}^*\|_\infty} + \|H_* - H\|_\infty, \quad (4.23)$$

²Sean $\|\cdot\|$ una norma matricial inducida en $\mathbb{R}^{n \times n}$, A y $C \in \mathbb{R}^{n \times n}$. Si C es no singular y $\|I_n - C^{-1}A\| < 1$, entonces A es no singular y $\|A^{-1}\| \leq \frac{\|C^{-1}\|}{1 - \|I_n - C^{-1}A\|}$.

donde $H \in \partial\Phi_\lambda(\mathbf{x})$.

En [23], *Kanzow y Kleinmichael* prueban que $\Phi_\lambda(\mathbf{x})$ es semisuave, es decir

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}^*) + H(\mathbf{x} - \mathbf{x}^*)\|_\infty}{\|\mathbf{x} - \mathbf{x}^*\|_\infty} = 0. \quad (4.24)$$

Así, para cualquier $\rho > 0$ existe $\epsilon_2 > 0$ tal que si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \epsilon_2$ 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 $\rho = \frac{\omega r}{8\mu(\omega + \sqrt{2}n\eta)}$, existe $\epsilon_r > 0$ tal que, si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \epsilon_r$ entonces

$$\frac{\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{x}^*) + H(\mathbf{x} - \mathbf{x}^*)\|_\infty}{\|\mathbf{x} - \mathbf{x}^*\|_\infty} < \frac{\omega r}{8\mu(\omega + \sqrt{2}n\eta)}. \quad (4.25)$$

Finalmente, sea $\epsilon_0 = \min\{\tilde{\epsilon}, \epsilon_r\}$. Si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \epsilon_0$ y $\|A - F'(\mathbf{x}^*)\|_\infty < \delta_0$ entonces, de (4.25), (4.20), (4.23) y (4.22),

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

completando así la demostración del **Lema 4.2**. ★

A continuación, presentamos un teorema análogo al *teorema de las dos vecindades* del caso diferenciable [14], con el cual demostramos *convergencia lineal* del algoritmo propuesto. El nombre de *dos vecindades* se debe a que, en las hipótesis del teorema, se requieren dos vecindades, una de la solución y en la cual debe estar el punto inicial, y otra de la matriz jacobiana de F en la solución y en la cual debe estar su aproximación inicial.

Teorema 4.1. Dado $r \in (0, 1)$, existen constantes positivas ϵ_1 y δ_1 tales que

$$\text{si } \|\mathbf{x}_0 - \mathbf{x}^*\|_\infty \leq \epsilon_1 \quad \text{y} \quad \|A_0 - F'(\mathbf{x}^*)\|_\infty \leq \delta_1,$$

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

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

con B_k la matriz cuyas filas están definidas por (4.1), está bien definida, converge a \mathbf{x}^* y satisface

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|_\infty \leq r \|\mathbf{x}_k - \mathbf{x}^*\|_\infty, \quad (4.26)$$

para todo $k = 0, 1, 2, \dots$.

Demostración. Consideremos la función \mathcal{Q} definida en (4.16). Así, para todo $k = 0, 1, \dots$

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

con B_k definida por (4.1).

Dado $r \in (0, 1)$, sean $\epsilon_1 \in (0, \epsilon_0)$ y $\delta_1 \in (0, \delta_0)$ donde ϵ_0 y δ_0 son las constantes positivas del **Lema 4.2**.

Realizaremos la demostración por *inducción matemática* sobre k .

- Para $k = 0$.

$$\text{si } \|\mathbf{x}_0 - \mathbf{x}^*\|_\infty \leq \epsilon_1 \leq \epsilon_0 \quad \text{y} \quad \|A_0 - F'(\mathbf{x}^*)\|_\infty \leq \delta_1 \leq \delta_0,$$

$\mathbf{x}_1 = \mathcal{Q}(\mathbf{x}_0, A_0)$ está bien definido y satisface

$$\|\mathbf{x}_1 - \mathbf{x}^*\|_\infty \leq r \|\mathbf{x}_0 - \mathbf{x}^*\|_\infty. \quad (4.27)$$

- Supóngamos la *hipótesis inductiva* para $k = m - 1$, esto es,

$$\text{si } \|\mathbf{x}_0 - \mathbf{x}^*\|_\infty \leq \epsilon_1 \quad \text{y} \quad \|A_0 - F'(\mathbf{x}^*)\|_\infty \leq \delta_1,$$

entonces

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

está bien definido y

$$\|\mathbf{x}_m - \mathbf{x}^*\|_\infty = \|\mathcal{Q}(\mathbf{x}_{m-1}, A_{m-1}) - \mathbf{x}^*\|_\infty \leq r \|\mathbf{x}_{m-1} - \mathbf{x}^*\|_\infty. \quad (4.28)$$

Dado que $\|\mathbf{x}_0 - \mathbf{x}^*\|_\infty < \epsilon_1$ entonces de (4.28)

$$\begin{aligned} \|\mathbf{x}_m - \mathbf{x}^*\|_\infty &\leq r \|\mathbf{x}_{m-1} - \mathbf{x}^*\|_\infty \\ &\leq r^m \|\mathbf{x}_0 - \mathbf{x}^*\|_\infty \\ &\leq r^m \epsilon_1 \\ &< \epsilon_0, \end{aligned}$$

Así, por el **Lema 4.2**, \mathbf{x}_{m+1} está bien definido y satisface

$$\|\mathbf{x}_{m+1} - \mathbf{x}^*\|_\infty \leq r \|\mathbf{x}_m - \mathbf{x}^*\|_\infty. \quad (4.29)$$

Por lo tanto, concluimos que (4.26) se cumple para todo $k = 0, 1, \dots$ ★

Observe que en la demostración del **Teorema 4.1** usamos la *norma infinito*. Pero, si $\mathbf{e}_k = \|\mathbf{x}_k - \mathbf{x}^*\|$ es el error en cualquier otra norma, entonces $\mathbf{e}_k \leq \alpha r^k \mathbf{e}_0$, donde α es una constante positiva que no depende de k y r es la constante del **Teorema 4.1**. Así, tendremos convergencia *r lineal* en cualquier otra norma.

Entre los teoremas estándares de la teoría *cuasi Newton* para sistemas de ecuaciones no lineales diferenciables está el teorema conocido como *condición de Dennis-Moré* [15], el cual da una condición suficiente para *convergencia superlineal*. El teorema siguiente es análogo al teorema mencionado y nos permitirá, en la sección siguiente, demostrar *convergencia superlineal* del **Algoritmo 4.1**.

En la demostración del **Teorema 4.2** usamos $\|\cdot\| = \|\cdot\|_\infty$, pero, recordemos que los resultados de *convergencia superlineal* son independientes de la norma.

Teorema 4.2. *Supóngamos que son válidas las hipótesis **H1** a **H3** y que la sucesión $\{\mathbf{x}_k\}$ generada por*

$$\mathbf{x}_{k+1} = \mathbf{x}_k - B_k^{-1} \Phi_\lambda(\mathbf{x}_k), \quad (4.30)$$

satisface que $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^$, con B_k la matriz cuyas filas están definidas por (4.1) y H_* la matriz definida por (3.9). Si*

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - H_*) \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = 0 \quad (4.31)$$

donde $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ entonces, la sucesión $\{\mathbf{x}_k\}$ converge superlinealmente a \mathbf{x}^ .*

Demostración. Como se mencionó anteriormente la Φ_λ es *semisuave* en \mathbf{x}^* , así

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

donde $H_* \in \partial\Phi_\lambda(\mathbf{x}^*)$. Como $\Phi_\lambda(\mathbf{x}^*) = 0$, tenemos

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_k) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0. \quad (4.32)$$

Por otra parte

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq \|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|, \quad (4.33)$$

además,

$$\left| \|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \right| \leq \|\Phi_\lambda(\mathbf{x}_k) - H_*(\mathbf{x}_k - \mathbf{x}^*)\|. \quad (4.34)$$

Al reemplazar (4.33) en (4.32) y utilizar (4.34) tenemos,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{|\|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\||}{\|H_*^{-1}\| \|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} = 0, \quad (4.35)$$

luego,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{|\|\Phi_\lambda(\mathbf{x}_k)\| - \|H_*(\mathbf{x}_k - \mathbf{x}^*)\||}{\|H_*(\mathbf{x}_k - \mathbf{x}^*)\|} = 0. \quad (4.36)$$

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 (4.33) obtenemos

$$\|\Phi_\lambda(\mathbf{x}_k)\| > \frac{1}{2} \|H_*(\mathbf{x}_k - \mathbf{x}^*)\| \geq \frac{1}{2\|H_*^{-1}\|} \|\mathbf{x}_k - \mathbf{x}^*\|. \quad (4.37)$$

Por otra parte, de (4.30) tenemos la ecuación

$$0 = B_k \mathbf{s}_k + \Phi_\lambda(\mathbf{x}_k), \quad (4.38)$$

donde $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$. Sumamos y restamos la expresión $H_* \mathbf{s}_k - \Phi_\lambda(\mathbf{x}_{k+1})$ en (4.38), con lo cual

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

Aplicando norma y luego la desigualdad triangular obtenemos

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

de donde

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

el primer sumando del término derecho converge a 0, por la hipótesis (4.31). El segundo sumando converge a 0, por la *semisuavidad* de Φ_λ . Por lo tanto,

$$\lim_{\mathbf{x}_k \rightarrow \mathbf{x}^*} \frac{\|\Phi_\lambda(\mathbf{x}_{k+1})\|}{\|\mathbf{s}_k\|} = 0. \quad (4.39)$$

De la expresión (4.39), (4.37) y (4.2)

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

Por lo tanto,

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

es decir, si se satisface (4.31) la sucesión $\{\mathbf{x}_k\}$ converge superlinealmente a \mathbf{x}^* . ★

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

En esta sección generamos una familia de *métodos secante de cambio mínimo* para el sistema de ecuaciones $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$ siguiendo las reglas generales para métodos de este tipo desarrollados por *Martínez* en [31] y demostramos además, que dicha familia proporciona *algoritmos local y superlinealmente convergentes*.

Los métodos *cuasi Newton* difieren entre sí, por la forma de actualizar la matriz A_k en cada iteración. Entre los algoritmos *cuasi Newton* “prácticos” se encuentran los llamados *métodos secantes de cambio mínimo*, en los cuales la actualización de A_k , la matriz A_{k+1} , debe satisfacer la ecuación

$$A_{k+1}(\mathbf{x}_{k+1} - \mathbf{x}_k) = F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k) \quad (4.40)$$

conocida como *ecuación secante* [14] y su cambio (medido en alguna norma) respecto a A_k debe ser *mínimo*. Es decir,

$$\|A_{k+1} - A_k\| = \min_{A \in V} \|A - A_k\|, \quad (4.41)$$

donde V , es un conjunto de matrices $A \in S \subseteq \mathbb{R}^{n \times n}$ que satisfacen la *ecuación secante* (4.40), donde S es un subespacio vectorial de $\mathbb{R}^{n \times n}$ [16].

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* [14] [33] [6], el cual garantiza que las matrices de la sucesión permanecen en una vecindad de $F'(x^*)$. Esto es esencial para demostrar convergencia *local lineal*.

Por lo anterior, en cada iteración de un algoritmo *secante de cambio mínimo*, el conjunto V está determinado por los vectores \mathbf{x}_k y \mathbf{x}_{k+1} ,

$$V = V(\mathbf{x}_k, \mathbf{x}_{k+1}) = \{A \in S \subseteq \mathbb{R}^{n \times n} : A(\mathbf{x}_{k+1} - \mathbf{x}_k) = F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k)\}. \quad (4.42)$$

Dado que necesitamos la matriz en V “más cercana” a A_k , es natural pensar en la *proyección ortogonal* de esta matriz sobre V , $P_V(A_k) = P_{\mathbf{x}_k, \mathbf{x}_{k+1}}(A_k)$. Teniendo en cuenta que

$$\|P_V(A_k) - A_k\| = \inf_{A \in V} \|A - A_k\|, \quad (4.43)$$

y que V es un *conjunto cerrado* podemos garantizar que $P_V(A_k) \in V$. Esta proyección es única porque V es un *conjunto convexo* [26]. Por lo tanto,

$$\|P_V(A_k) - A_k\| = \min_{A \in V} \|A - A_k\|. \quad (4.44)$$

Así, de (4.41) y (4.44) concluimos que

$$A_{k+1} = P_V(A_k).$$

Diferentes *actualizaciones secante de cambio mínimo* se obtienen al variar la norma matricial en $\mathbb{R}^{n \times n}$ o al variar al subespacio S . Es así como se genera una *familia de métodos secante de cambio mínimo*. Por ejemplo, denotemos

$$\mathbf{y}_k = F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k) \quad \text{y} \quad \mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k.$$

- Actualización “buena” de Broyden [5].

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k}, \quad (4.45)$$

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

- Actualización “mala” de Broyden [4].

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{y}_k^T A_k}{\mathbf{y}_k^T \mathbf{y}_k}. \quad (4.46)$$

resulta al elegir la *norma de Frobenius* y

$$V = \{A \in \mathbb{R}^{n \times n} : A^{-1} (\mathbf{x}_{k+1} - \mathbf{x}_k) = F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k)\}.$$

- Actualización de Schubert [12].

$$A_{k+1} = A_k + \frac{F(x_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k}. \quad (4.47)$$

- Actualización de Schubert esparsa [12].

$$A_{k+1} = A_k + \frac{(\mathbf{y}_k - A_k \mathbf{s}_k) \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{s}_k}, \quad (4.48)$$

resulta al elegir la *norma de Frobenius* y S como el conjunto de las matrices esparsas.

Martínez en [31] [32] [33] generaliza la *teoría de convergencia local para los métodos secantes de cambio mínimo* para resolver sistemas de ecuaciones no lineales diferenciables y problemas de optimización. Por ello en sus hipótesis hace uso de dos normas. En nuestro caso, sólo usaremos una norma dado que nuestro interés es la solución de sistemas de ecuaciones no lineales.

Algoritmo 4.2. Supongamos \mathbf{x}_0 y A_0 arbitrarios. Para $k = 0, 1, \dots$, las actualizaciones \mathbf{x}_{k+1} y A_{k+1} son generadas como sigue

$$B_k = D_a + D_b A_k \quad (4.49)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k - B_k^{-1} \Phi(\mathbf{x}_k) \quad (4.50)$$

$$A_{k+1} = P_{\mathbf{x}_k, \mathbf{x}_{k+1}}(A_k) \quad (4.51)$$

donde $D_a = \text{diag}(a_1, \dots, a_n)$ y $D_b = \text{diag}(b_1, \dots, b_n)$ con

$$a_i = \chi(x_i^k, F_i(\mathbf{x}^k)) - 1 \quad y \quad b_i = \psi(x_i^k, F_i(\mathbf{x}^k)) - 1.$$

Con el fin de desarrollar la teoría de convergencia de la familia de *métodos secante de cambio mínimo* generados por el **Algoritmo 4.2** supondremos una hipótesis adicional a **H1**, **H2** y **H3**.

H4. Supongamos que para todo \mathbf{x}, \mathbf{z} en una vecindad de \mathbf{x}^* , existen $A \in V(\mathbf{x}, \mathbf{z})$ y $\alpha_1 > 0$ tal que,

$$\|A - F'(\mathbf{x}^*)\| \leq \alpha_1 \sigma(\mathbf{x}, \mathbf{z}), \quad (4.52)$$

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

Esta hipótesis es análoga a la hipótesis adicional en [28].

4.4. Resultados de convergencia adicionales

En el siguiente lema demostramos que una matriz generada utilizando la regla (4.51) para actualizar la matriz A_k puede deteriorarse pero, en forma controlada.

Lema 4.3. Supongamos las hipótesis **H1** a **H4**. Sean A_+ la proyección ortogonal de A sobre el conjunto $V(\mathbf{x}, \mathbf{z})$ y \hat{A} la proyección ortogonal de $F'(\mathbf{x})$ sobre $V(\mathbf{x}, \mathbf{z})$ entonces,

$$\|A_+ - F'(\mathbf{x})\| \leq \|A - F'(\mathbf{x})\| + \alpha_2 \sigma(\mathbf{x}, \mathbf{z}) \quad (4.53)$$

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

Demostración. Por la desigualdad triangular tenemos,

$$\begin{aligned} \|A_+ - F'(\mathbf{x})\| &\leq \|A_+ - \widehat{A}\| + \|\widehat{A} - F'(\mathbf{x})\| \\ &\leq \|A - \widehat{A}\| + \|\widehat{A} - F'(\mathbf{x})\| \\ &\leq \|A - F'(\mathbf{x})\| + \|F'(\mathbf{x}) - \widehat{A}\| + \|\widehat{A} - F'(\mathbf{x})\| \end{aligned} \quad (4.54)$$

Pero $\|\widehat{A} - F'(\mathbf{x})\| = \min_{W \in V(\mathbf{x}, \mathbf{z})} \|W - F'(\mathbf{x})\|$ y, en particular, por **H4** existe $U \in V(\mathbf{x}, \mathbf{z})$ tal que

$$\|U - F'(\mathbf{x})\| \leq \alpha_1 \sigma(\mathbf{x}, \mathbf{z}).$$

Entonces, de (4.54) tenemos que

$$\begin{aligned} \|A_+ - F'(\mathbf{x})\| &\leq \|A - F'(\mathbf{x})\| + 2 \|\widehat{A} - F'(\mathbf{x})\| \\ &\leq \|A - F'(\mathbf{x})\| + 2 \|U - F'(\mathbf{x})\| \\ &\leq \|A - F'(\mathbf{x})\| + \alpha_2 \sigma(\mathbf{x}, \mathbf{z}) \end{aligned}$$

con, $\alpha_2 = 2\alpha_1$. ★

Lema 4.4. *Supongamos las hipótesis **H1** a **H4**. Existe $c > 0$ tal que*

$$\|P_{\mathbf{x}, \mathbf{y}}(A) - F'(\mathbf{x}^*)\| \leq \|A - F'(\mathbf{x}^*)\| + c \|\mathbf{x} - \mathbf{x}^*\| \quad (4.55)$$

siempre que, \mathbf{x} y \mathbf{y} pertenezcan a una vecindad de \mathbf{x}^ , con $\|\mathbf{y} - \mathbf{x}^*\| \leq \|\mathbf{x} - \mathbf{x}^*\|$ y la matriz A en una vecindad de $F'(\mathbf{x}^*)$.*

Demostración. ver [31]. ★

Los dos lemas anteriores, así como las hipótesis **H1** a **H4** son centrales para garantizar el siguiente resultado.

Teorema 4.3. *Supongamos las hipótesis **H1** a **H4** y que la sucesión $\{A_k\}$ está definida por (4.2). Dado $r \in (0, 1)$, existen constantes positivas $\bar{\epsilon}$ y $\bar{\delta}$ tales que*

$$\text{si } \|\mathbf{x}_0 - \mathbf{x}^*\| \leq \bar{\epsilon} \quad \text{y} \quad \|A_0 - F'(\mathbf{x}^*)\| \leq \bar{\delta}$$

cualquier 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,

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq r \|\mathbf{x}_k - \mathbf{x}^*\|. \quad (4.56)$$

Demostración. Dado $\bar{\delta} \in (0, \delta_1)$, podemos escoger $\bar{\epsilon} \in (0, \epsilon_1)$ con

$$\bar{\epsilon} < \frac{(\delta_1 - \bar{\delta})(1 - r)}{c}, \quad (4.57)$$

donde δ_1 y ϵ_1 son las constantes positivas del **Lema 4.2**, y la constante c es definida en (4.60).

Algunas consideraciones sobre la forma de elegir la constante $\bar{\epsilon}$, son las siguientes.

1. Si $\frac{(\delta_1 - \bar{\delta})(1 - r)}{c} < \epsilon_1$, la constante $\bar{\delta}$ también dependerá de ϵ_1 , ya que,

$$\bar{\delta} > \delta_1 - \frac{c\epsilon_1}{1 - r}.$$

2. Si $\frac{(\delta_1 - \bar{\delta})(1 - r)}{c} > \epsilon_1$, podemos elegir cualquier valor en $(0, \epsilon_1)$ como $\bar{\epsilon}$.

Así, en cualquiera de los casos, se puede elegir $\bar{\epsilon}$ en el conjunto

$$(0, \epsilon_1) \cap \left(0, \frac{(\delta_1 - \bar{\delta})(1-r)}{c}\right).$$

Luego, a partir de la desigualdad (4.57) obtenemos

$$\bar{\delta} + c \frac{\bar{\epsilon}}{1-r} < \bar{\delta} + c \frac{(\delta_1 - \bar{\delta})(1-r)}{c(1-r)} < \delta_1. \quad (4.58)$$

Realizaremos la demostración por inducción sobre k .

1. Para $k = 0$,

$$\text{si } \|\mathbf{x}_0 - \mathbf{x}^*\| \leq \bar{\epsilon} < \epsilon_1 \quad \text{y} \quad \|A_0 - F'(\mathbf{x}^*)\| \leq \bar{\delta} < \delta_1$$

entonces por el **Lema 4.2**, $\mathbf{x}_1 = \mathbf{x}_0 - B_0^{-1}\Phi(\mathbf{x}_0)$ está bien definido y satisface

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

así, (4.56) es verdadera para $k = 0$.

2. Supongamos (4.56) verdadera para $k = m - 1$, esto es,

$$\text{si } \|\mathbf{x}_{m-1} - \mathbf{x}^*\| \leq \bar{\epsilon} \quad \text{y} \quad \|A_{m-1} - F'(\mathbf{x}^*)\| \leq \bar{\delta},$$

$\mathbf{x}_m = \mathbf{x}_{m-1} - B_0^{-1}\Phi(\mathbf{x}_{m-1})$ está bien definido y satisface

$$\|\mathbf{x}_m - \mathbf{x}^*\| \leq r \|\mathbf{x}_{m-1} - \mathbf{x}^*\|. \quad (4.59)$$

3. De (4.59) obtenemos

$$\begin{aligned} \|\mathbf{x}_m - \mathbf{x}^*\| &\leq r \|\mathbf{x}_{m-1} - \mathbf{x}^*\| \\ &\leq r^m \|\mathbf{x}_0 - \mathbf{x}^*\| \\ &< r^m \bar{\epsilon} \\ &< \bar{\epsilon} < \epsilon_1. \end{aligned}$$

Por (4.55),

$$\begin{aligned}
\|A_m - F'(\mathbf{x}^*)\| &\leq \|A_{m-1} - F'(\mathbf{x}^*)\| + c_{m-1} \|\mathbf{x}_{m-1} - \mathbf{x}^*\| \\
&\leq \bar{\delta} + c\bar{\epsilon} \sum_{j=0}^{m-1} r^j \\
&< \bar{\delta} + c\bar{\epsilon} \sum_{j=0}^{\infty} r^j \\
&< \bar{\delta} + c \frac{\bar{\epsilon}}{1-r} < \delta_1,
\end{aligned}$$

donde

$$c = \max_{0 \leq j \leq m-1} c_j \quad (4.60)$$

con las constantes c_i dadas en lema anterior y la última desigualdad se obtiene de (4.58). Así, por **Lema 4.2**, \mathbf{x}_{m+1} está bien definido y satisface

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

Lema 4.5. *Supongamos las hipótesis **H1** a **H4** y que la sucesión $\{A_k\}$ está definida por (4.2). Existen constantes positivas $\bar{\epsilon}$ y $\bar{\delta}$ tales que,*

$$si \quad \|\mathbf{x}_0 - \mathbf{x}^*\| \leq \bar{\epsilon} \quad y \quad \|A_k - F'(\mathbf{x}^*)\| \leq \bar{\delta}$$

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

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

con B_k definida por (4.1), entonces

$$\lim_{k \rightarrow \infty} \|B_{k+1} - B_k\| = 0. \quad (4.61)$$

Demostración. Sean $H_* \in \partial\Phi_\lambda(\mathbf{x}^*)$, $r \in (0, 1)$, $\bar{\epsilon} < \frac{r}{8\mu(\omega + \sqrt{2}n\eta)}$ y $\bar{\delta} < \frac{r}{8\mu\tau}$,

donde ω es la constante (4.15), η es la constante de *Lipschitz* de $\nabla\varphi_\lambda$ dada en (2.20), τ está definida en (4.14) y μ está definida en (4.2).

Por la desigualdad triangular

$$\|B_{k+1} - B_k\| \leq \|B_{k+1} - H_*\| + \|B_k - H_*\| \quad (4.62)$$

y de la desigualdad (4.21)

$$\|B_k - H_*\|_\infty \leq \frac{r}{4\mu} \quad \text{y} \quad \|B_{k+1} - H_*\|_\infty \leq \frac{r}{4\mu}. \quad (4.63)$$

Usando (4.63) en (4.62) tenemos,

$$\|B_{k+1} - B_k\|_\infty \leq \frac{r}{2\mu},$$

de lo cual sigue (4.61). ★

Con el resultado del **Lema 4.5** derivamos una condición suficiente para convergencia superlineal, como lo afirma el siguiente teorema.

Teorema 4.4. *Supongamos las hipótesis **H1** a **H4** y que las sucesiones $\{\mathbf{x}_k\}$, $\{A_k\}$ son generadas por el **Algoritmo 4.2** y $\lim_{k \rightarrow \infty} \mathbf{x}_k = \mathbf{x}^*$. Si*

$$\lim_{k \rightarrow \infty} \frac{\|(B_{k+1} - H_*) \mathbf{s}_k\|}{\|\mathbf{s}_k\|} = 0 \quad (4.64)$$

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

Demostración. Es una aplicación directa de los **Teoremas 4.2** y **4.3**. En efecto,

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

Por el lema anterior,

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

y el segundo sumando del término derecho es cero por la hipótesis (4.64). Luego,

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

que es la condición tipo *Dennis-Moré* del **Teorema 4.2**. Por lo tanto, concluimos que la sucesión $\{\mathbf{x}_k\}$ converge superlinealmente. ★

Por los resultados demostrados en esa sección tenemos que bajo las hipótesis **H1** a **H4**, la familia de métodos secantes de cambio mínimo generada por (4.40) proporciona algoritmos local y superlinealmente convergentes, lo cual se resume en el siguiente teorema.

Teorema 4.5. *Supongamos las hipótesis **H1** a **H4** y que las sucesiones $\{\mathbf{x}_k\}$ y $\{A_k\}$ son generadas por (4.2). Dado $r \in (0, 1)$. Existen constantes positivas $\bar{\epsilon}$ y $\bar{\delta}$ tales que,*

$$\text{si } \|\mathbf{x}_0 - \mathbf{x}^*\| \leq \bar{\epsilon} \quad \text{y} \quad \|A - F'(\mathbf{x}^*)\| \leq \bar{\delta}$$

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

$$\lim_{k \rightarrow \infty} \frac{\|(B_{k+1} - H_*) s_k\|}{\|s_k\|} = 0. \quad (4.65)$$

entonces la sucesión $\{\mathbf{x}_k\}$ converge superlinealmente a \mathbf{x}^* .

Demostración. Es una aplicación directa de los **Teoremas 4.2**, **4.3** y **4.4**. ★

Capítulo 5

Pruebas numéricas

En este capítulo, analizamos numéricamente el comportamiento local de la familia de *métodos secantes de cambio mínimo* propuesta en el Capítulo 3. Para ello, comparamos nuestros algoritmos con el *método de Newton generalizado* propuesto en [40].

Para mayor claridad en la lectura del capítulo, incluimos a continuación los algoritmos mencionados en el párrafo anterior.

El **Algoritmo 5.1** es de tipo *Newton generalizado*. Este usa en cada iteración la matriz H_k definida en (3.8), la que a su vez usa la matriz jacobiana de F .

Algoritmo 5.1. *Dados \mathbf{x}_0 y $\lambda \in (0, 4)$, para $k = 1, 2, \dots$*

Mientras $\|\Phi(\mathbf{x}_k)\| \geq \sqrt{n} 10^{-5}$ y $k < N$

Calculamos $F'(\mathbf{x}_k)$.

Calculamos H_k mediante (3.8).

Calculamos la iteración siguiente por

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

$k \leftarrow k + 1$

fin.

El **Algoritmo 5.2** se basa en el **Algoritmo 4.2** que propusimos en el Capítulo 4 y el cual es de tipo *secante de cambio mínimo*. Usamos las fórmulas de *Broyden* (4.45), (4.46), la fórmula de *Schubert* (4.47) y *Schubert esparsa* (4.3) para actualizar en cada iteración la matriz A_k , con lo cual tenemos cuatro versiones del **Algoritmo 5.2**.

Algoritmo 5.2. *Dados \mathbf{x}_0 , A_0 y $\lambda \in (0, 4)$, para $k = 1, 2, \dots$*

Mientras $\|\Phi(\mathbf{x}_k)\| \geq \sqrt{n} 10^{-5}$ *y* $k < N$

Calculamos B_k mediante (4.1).

Calculamos la iteración siguiente:

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

Actualizamos A_k , usando (4.45), (4.46), (4.47) o (4.3).

$$k \leftarrow k + 1$$

fin.

Para escribir los códigos de los algoritmos y de las *funciones de prueba* utilizamos el *software* MATLAB[®]. Realizamos los experimentos numéricos en un computador con procesador AMD Sempron (tm) de 2.21 GHz.

Utilizamos ocho *problemas de prueba* para algoritmos de complementariedad no lineal, cuatro de los cuales los elegimos de una lista propuesta en [37] y que son considerados problemas “difíciles”, en lo que respecta a convergencia. Estos problemas son *Kojima-Shindo* (aplicación en economía al problema de *equilibrio económico* [24]), *Kojima-Josephy*, *Nash-Cornout* (aplicación en *teoría de juegos* [21]) y *Mathiesen Modificado* (aplicación al problema de *equilibrio económico walrasiano* [35]). Para cada uno de estos problemas describimos a continuación la función que los define, así como el punto inicial que utilizamos (\mathbf{x}_0) y la o las soluciones encontradas (\mathbf{x}^*).

1. **Problema Kojima-Shindo.** Sea $F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definido por

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

$$\mathbf{x}_0 = (1 \ 1 \ 1 \ 1)^T \text{ y } \mathbf{x}^* = (\sqrt{6}/2 \ 0 \ 0 \ 1/2)^T.$$

2. **Problema Kojima-Josephy** Sea $F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definida por

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

$$\mathbf{x}_0 = (1 \ 1 \ 1 \ 1)^T \text{ y } \mathbf{x}^* = (\sqrt{6}/2 \ 0 \ 0 \ 1/2)^T.$$

3. **Problema Nash-Cournot** Sea $F : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ definida por

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

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

Punto inicial $\mathbf{x}_0 = (10 \ 10 \ 10 \ 10 \ 10)^T$ y la solución encontrada fue $\mathbf{x}^* = (15.4293076 \ 12.4985817 \ 9.6634730 \ 7.1650935 \ 5.1325661)^T$.

4. **Problema Mathiesen-Modificado.** Sea $F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definida por

$$F(\mathbf{x}) = \begin{pmatrix} -x_2 + x_3 + x_4 \\ x_1 - \frac{4.5x_3 + 2.7x_4}{x_2 + 1} \\ 5 - x_1 - \frac{0.5x_3 + 0.3x_4}{x_3 + 1} \\ 3 - x_1 \end{pmatrix},$$

Este problema tiene infinitas soluciones de la forma

$$\mathbf{x}^* = (a \ 0 \ 0 \ 0)^T,$$

con $a \in [0, 3]$.

A partir del punto inicial $\mathbf{x}_0 = (1 \ 1 \ 1 \ 1)^T$ encontramos las soluciones siguientes. Con el *método de Newton generalizado*,

$$\mathbf{x}^* = (0 \ 0 \ 0 \ 0)^T.$$

Con el **Algoritmo 4.2** y las diferentes actualizaciones,

$$\begin{aligned} \mathbf{x}^* &= (0.8268872 \ 0 \ 0 \ 0)^T && \text{actualización de Broyden "bueno"}. \\ \mathbf{x}^* &= (0.8501340 \ 0 \ 0 \ 0)^T && \text{actualización de Broyden "malo"}. \\ \mathbf{x}^* &= (1.0983616 \ 0 \ 0 \ 0)^T && \text{actualización de Schubert}. \\ \mathbf{x}^* &= (0.3069555 \ 0 \ 0 \ 0)^T && \text{actualización de Schubert esparsa}. \end{aligned}$$

Los cuatro problemas restantes, los generamos como en [20], para lo cual definimos la función $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ por

$$F_i(\mathbf{x}) = \begin{cases} h_i(\mathbf{x}) - h_i(\mathbf{x}^*), & \text{si } i \text{ es impar o si } i > \frac{n}{2}, \\ h_i(\mathbf{x}) - h_i(\mathbf{x}^*) + 1, & \text{si } i \text{ es par,} \end{cases}$$

donde $\mathbf{x}^* = (1 \ 0 \ 1 \dots 1)^T$ y las h_i con $i = 1, \dots, n$ son algunas de las funciones propuestas en [29], a saber *Sistema Trigonométrico*, *Trigonométrico Exponencial*, *Tri-diagonal* y *Rosenbrock*, los cuales describimos a continuación.

1. **Problema Sistema Trigonométrico.** Sea $h_i : \mathbb{R}^{10} \rightarrow \mathbb{R}$ definida por

$$h_i(\mathbf{x}) = \sum_{j=1}^{10} \left(5 - (i+1)(1 - \cos x_j) - \sin x_j - \sum_{m=5i+1}^{5i+5} \cos x_j \right),$$

donde $i = \text{div}(k-1, 5)$. El punto inicial fue $\mathbf{x}_0 = (0.9 \ 0.1 \ 0.9 \ 0.1 \dots 0.1 \ 0.9 \ 0.1)^T$ y la solución es $\mathbf{x}^* = (1 \ 0 \ 1 \ 0 \dots 0 \ 1 \ 0)^T$.

2. **Problema Trigonométrico Exponencial.** Sea $h_i : \mathbb{R}^{100} \rightarrow \mathbb{R}$ definida por:

$$h_i(\mathbf{x}) = 3x_1^2 + 2x_2 + \sin(x_1 - x_2) \sin(x_1 + x_2) + S + 4x_{100} - x_{99}e^{x_{99}+x_{100}} - 8,$$

donde

$$S = \sum_{j=2}^{99} (3x_j^2 + 2x_{j+1} - 5 + \sin(x_j - x_{j+1}) \sin(x_j + x_{j+1}) + 4x_j - x_{j-1}e^{x_{j-1}+x_j} - 3).$$

Este problema tiene solución $\mathbf{x}^* = (1 \ 0 \ 1 \ 0 \ \dots \ 0 \ 1 \ 0)^T$. El punto inicial utilizado fue $\mathbf{x}_0 = (0.9 \ 0.1 \ 0.9 \ 0.1 \ \dots \ 0.1 \ 0.9 \ 0.1)^T$.

3. **Problema Tridiagonal**, Sea $h : \mathbb{R}^{20} \rightarrow \mathbb{R}$ definida por

$$h_i(\mathbf{x}) = \sum_{j=1}^{20} (x_j - 1)^2 - \sum_{j=2}^{20} x_j x_{j-1},$$

Este problema tiene solución $\mathbf{x}^* = (1 \ 0 \ 1 \ 0 \ \dots \ 0 \ 1 \ 0)^T$. El punto inicial utilizado fue $\mathbf{x}_0 = (0.9 \ 0.1 \ 0.9 \ 0.1 \ \dots \ 0.1 \ 0.9 \ 0.1)^T$.

4. **Problema Rosenbrock**. Sea $h : \mathbb{R}^{20} \rightarrow \mathbb{R}$ definida por

$$h_i(\mathbf{x}) = \sum_{j=1}^{19} \left[100 (x_j^2 - x_{j+1})^2 + (x_j - 1)^2 \right],$$

Este problema tiene solución $\mathbf{x}^* = (1 \ 0 \ 1 \ 0 \ \dots \ 0 \ 1 \ 0)^T$. El punto inicial utilizado es $\mathbf{x}_0 = (0.9 \ 0.1 \ 0.9 \ 0.1 \ \dots \ 0.1 \ 0.9 \ 0.1)^T$.

Observemos que \mathbf{x}^* es una solución degenerada de los PCNL así generados, ya que

$$x_i^* = 0 = F_i(\mathbf{x}^*), \quad \text{para } i > \frac{n}{2} \text{ y par,}$$

con lo cual, Φ_λ , no es diferenciable en \mathbf{x}^* . Esto nos motivó a elegir dichos problemas para analizar la eficiencia de nuestros algoritmos.

Usamos los mismos *criterios de convergencia y divergencia* propuestos en [39]. Exactamente, declaramos *convergencia* si $\|\Phi_\lambda(x_k)\|_2 < \sqrt{n} 10^{-5}$ y *divergencia* si el número de iteraciones excedió 500 o si $\|\Phi_\lambda(x_k)\|_\infty > 10^{20}$. Los *puntos iniciales* utilizados en las pruebas numéricas son los propuestos en [37] y [29], respectivamente.

Para las pruebas numéricas procedimos de la siguiente forma: variamos λ en el intervalo $(0, 4)$ desde $\lambda = 10^{-3}$ hasta $\lambda = 3.999$ con incrementos de 10^{-3} .

1. Para cada valor de λ , registramos si hubo o no convergencia de los **Algoritmos 5.1 y 5.2**, éste último con las cuatro fórmulas de actualización mencionadas anteriormente.

2. De todos los valores de λ utilizados, aquél con el cual el *método de Newton generalizado* converge en menos iteraciones lo llamamos λ_{\min} y usamos este valor para el parámetro λ en los **Algoritmos 5.1 y 5.2**.

Los resultados de pruebas del numeral 1) para cada una de las ocho funciones utilizadas, los presentamos en diagramas de barras con la siguiente convención de colores.

	<i>M. Newton</i>		<i>M. Broyden bueno</i>		<i>M. Broyden malo</i>
	<i>M. Schubert</i>		<i>M. Schubert esparsa</i>		

El *eje horizontal* indica la variación de λ . Para mayor claridad en la visualización de los diagramas, sólo resaltamos los valores inicial y final de λ a saber ($\lambda = 10^{-1}$) y ($\lambda = 3.9$), así como el primer valor para el cual el *método de Newton generalizado* converge. El *eje vertical* indica número de iteraciones.

Las Figuras 5.1, 5.2 y 5.3 nos permiten observar que en cada caso, existe un intervalo de valores de λ en el que el *método de Newton generalizado*, con las funciones *Kojima-Shindo*, *Kojima-Josephy* y *sistema trigonométrico*, no converge debido al mal condicionamiento de la matriz H_k , mientras que en dichos intervalos el **Algoritmo 5.2** converge para algunas de las actualizaciones utilizadas.

Algo importante de destacar relacionado con el problema *sistema trigonométrico* (Figura 5.3) es que los **Algoritmos 5.1 y 5.2** convergen (para la mayoría de los valores de $\lambda \in (1.6, 3.9)$), a diferencia de los algoritmos locales que usan solamente la *función de Fisher* o la *función mínimo* como fue reportado en [39].

En las figuras restantes, observamos que en general los dos algoritmos utilizados convergen. El **Algoritmo 5.2** tuvo un *pobre* desempeño numérico para las funciones *Nash-Cornout* (Figura 5.4) y *Ronsenbrock* (Figura 5.8) con las fórmulas de actualización (4.45) y (4.46), respectivamente.

Los resultados de las pruebas numéricas descritas en el numeral 2) los resumimos en la **Tabla 5.1**. En esta tabla reportamos la siguiente información.

- Problema** nombre del problema de prueba.
n: dimensión del problema.
 λ_{min} : valor de λ para el **Algoritmo 5.1** converge en menos iteraciones.
NG: número de iteraciones con el **Algoritmo 5.1**.
BB: número de iteraciones con el **Algoritmo 5.2** y la fórmula (4.45).
BM: número de iteraciones con el **Algoritmo 5.2** y la fórmula (4.46).
SN: número de iteraciones con el **Algoritmo 5.2** y la fórmula (4.47).
SE: número de iteraciones con el **Algoritmo 5.2** y la fórmula (4.3).
 - : hubo divergencia.

Problema	n	λ_{\min}	NG	SE	BM	SN	BB
<i>Kojima-Shindo</i>	4	2.7	7	17	14	-	16
<i>Kojima-Josephy</i>	4	3.860	9	14	9	-	12
<i>Nash-Cornout</i>	5	1.540	6	8	7	135	7
<i>Mathiesen Modificado</i>	4	0.010	3	3	3	3	3
<i>Sistema Trigonom.</i>	10	1.930	5	18	24	-	-
<i>Trigonométrico exp.</i>	100	0.010	3	6	6	8	4
<i>Tridiagonal</i>	20	0.010	3	5	5	8	4
<i>Rosembrock</i>	20	0.010	3	3	3	3	3

Tabla 5.1: Comportamiento local de los **Algoritmos** 5.1 y 5.2.

En la **Tabla 5.1**, observamos que, para estas pruebas numéricas preliminares, el **Algoritmo 5.2** que proponemos para resolver el problema PCNL presenta un buen comportamiento local. En particular, destacamos el caso del problema *Mathiesen modificado*, en el cual cada método converge en el mismo número de iteraciones pero, a soluciones distintas del problema.

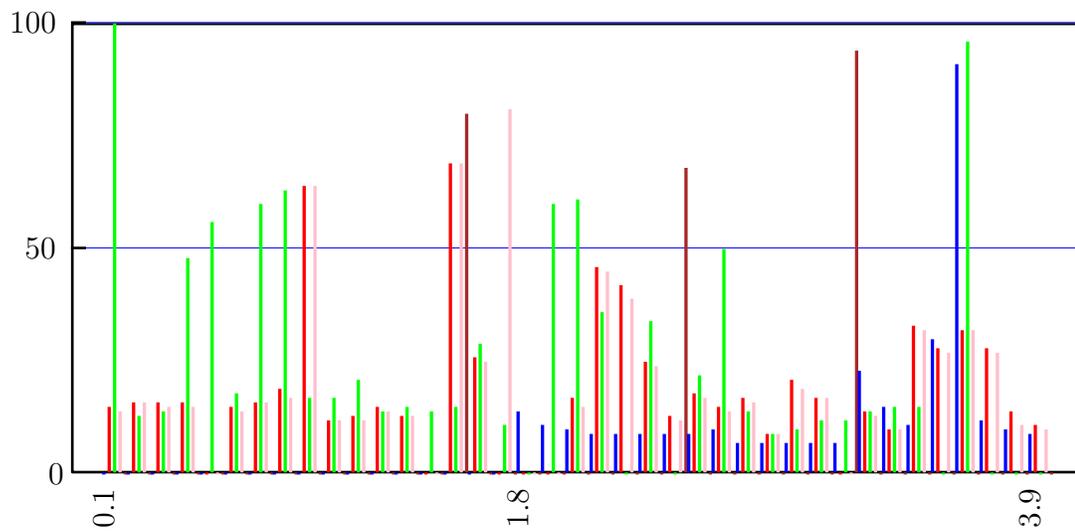


Figura 5.1: *Problema Kojima-Shindo*.

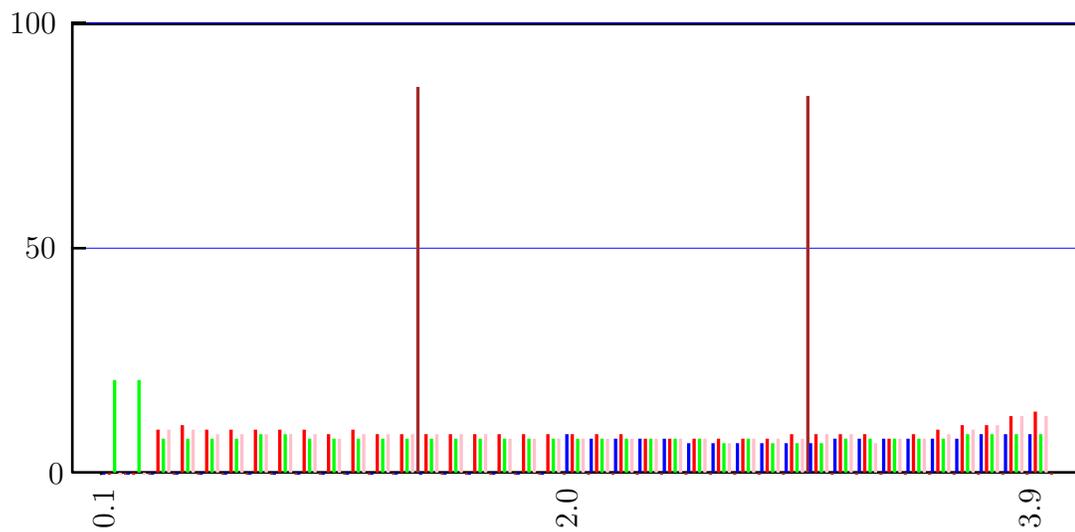


Figura 5.2: *Problema Kojima-Josephy*.

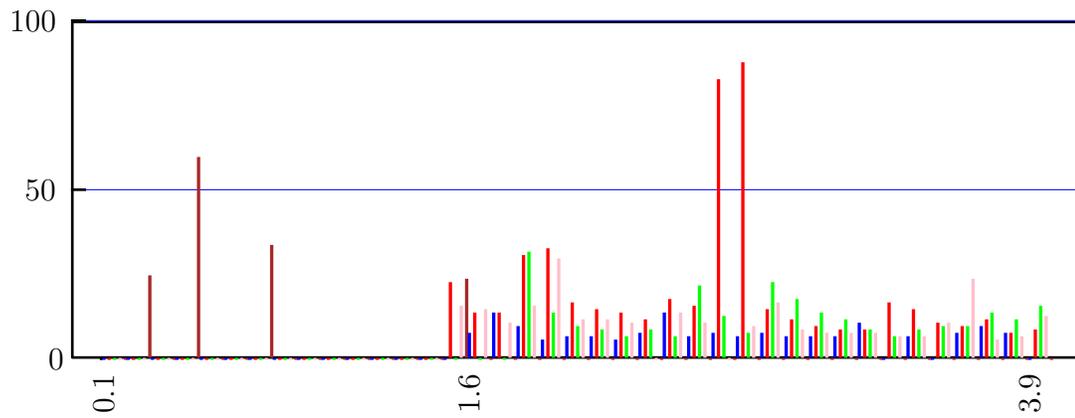


Figura 5.3: *Problema Sistema Trigonometrico.*

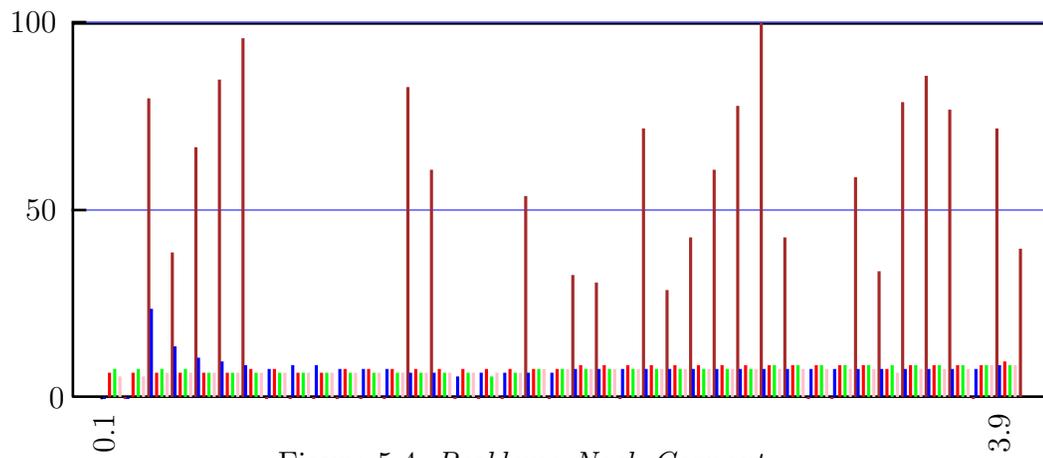


Figura 5.4: *Problema Nash-Cornout.*

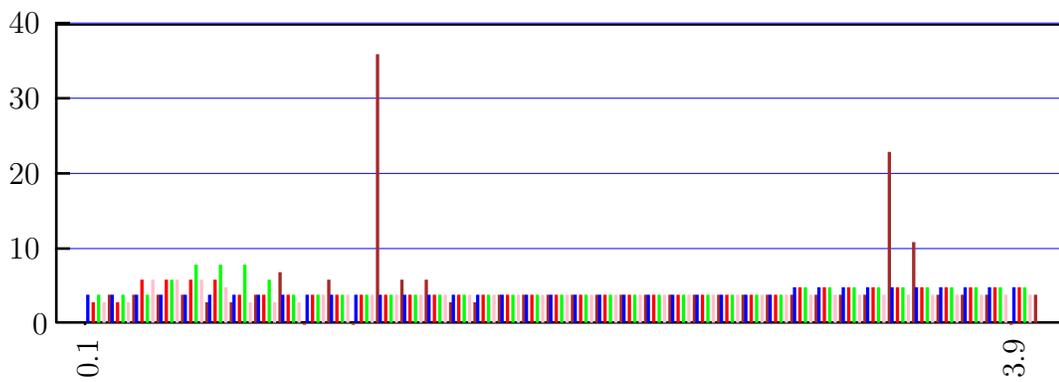


Figura 5.5: *Problema Mathiesen Modificado.*

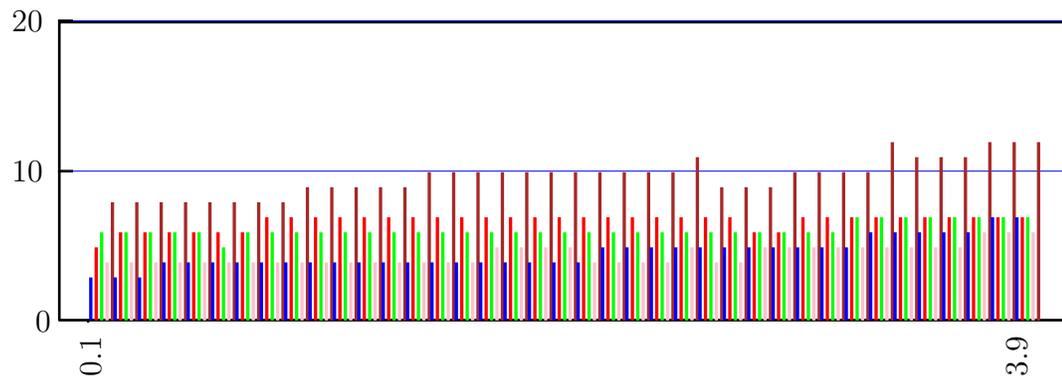


Figura 5.6: *Problema Sistema Tridiagonal.*

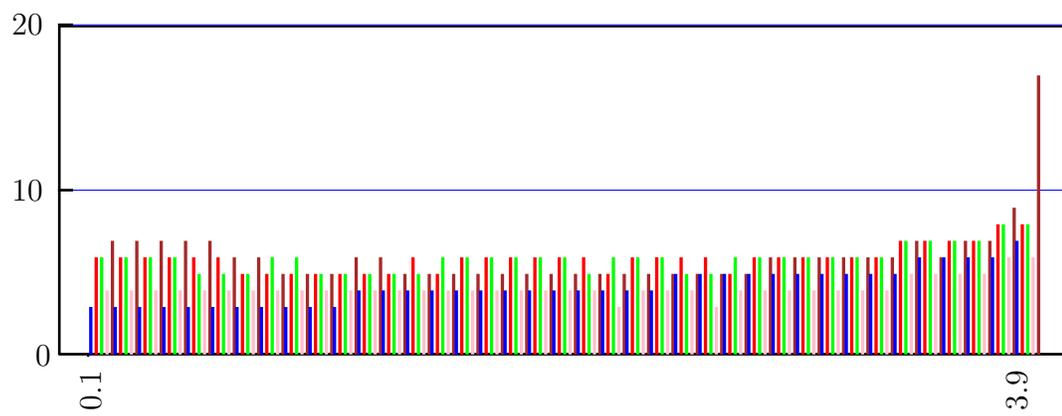


Figura 5.7: *Problema Sistema Trigonómico Exponencial.*

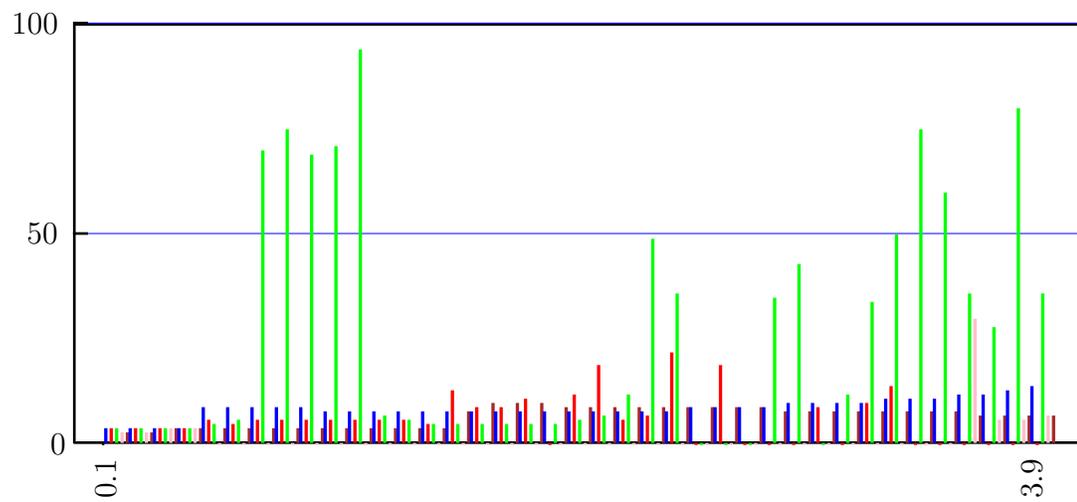


Figura 5.8: *Problema Rosenbrock.*

Capítulo 6

Comentarios finales

La técnica de reformular el Problema de Complementariedad No Lineal como un sistema de ecuaciones no lineales es muy importante para resolver ese tipo de problemas porque de esta forma, el trabajo principal en cada iteración de un algoritmo tipo Newton para resolver dichos sistemas consiste en hacer una única evaluación de la función y en encontrar la solución de un único sistema de ecuaciones lineales.

En este trabajo propusimos un método *cuasi Newton* que puede ser útil cuando la matriz jacobiana de la función que define el problema es difícil o costosa (en término de número de operaciones) de calcular.

Un estudio detallado de la función de complementariedad que usamos para reformular el PCNL y la cual denominamos, *la función de Kanzow* nos permitió encontrar cotas que fueron fundamentales en el desarrollo de la teoría de convergencia del método propuesto.

Además, generamos una *familia de métodos secantes de cambio mínimo* que bajo ciertas hipótesis, convergen local y superlinealmente a la solución del problema.

Experimentos numéricos preliminares muestran un buen desempeño local de los algoritmos propuestos. Pero, consideramos que es necesario realizar más pruebas numéricas.

Finalmente, pensamos que es necesario incorporar estrategias de globalización al algoritmo propuesto y hacer el análisis teórico y numérico del algoritmo globalizado.

Bibliografía

- [1] Jozsef. Abaffy, Charles Broyden, and Emilio Spedicato, *A class of direct methods for linear systems*, Numerische Mathematik **45** (1984), 361–376.
- [2] Mihai Anitescu, James F. Cremer, and Florian A. Potra, *On the existence of solutions to complementarity formulations of contact problems with friction*, In: Complementarity and Variational Problems: State of the art, SIAM Publications, 1997, pp. 12–21.
- [3] Tom M. Apostol, *Análisis matemático*, Reverté, 1976.
- [4] Peter N. Brown and Youcef Saad, *Convergence theory of nonlinear newton–krylov algorithms*, SIAM Journal on Optimization **4** (1994), no. 2, 297–330.
- [5] Charles G. Broyden, *A Class of Methods for Solving Nonlinear Simultaneous Equations*, Mathematics of Computation **19** (1965), no. 92, 577–593.
- [6] Charles G. Broyden, Jhon. E. Dennis, and Jorge J. Moré, *On the local and superlinear convergence of quasi-newton methods*, PB 211 923, Department of Computer Science, Cornell University, 1972.
- [7] Sandra Buhmiller and Nataža Krejić, *A new smoothing quasi-newton method for nonlinear complementarity problems*, Journal of Computational and Applied Mathematics **211** (2008), no. 2, 141 – 155.
- [8] Antony Chen, Jun-Seok Oh, Dongjoo Park, and Will Recker, *Solving the bicriteria traffic equilibrium problem with variable demand and nonlinear path costs*, Applied Mathematics and Computation **217** (2010), no. 7, 3020–3031.
- [9] Frank H. Clarke, *Necessary Conditions for Nonsmooth Problems in Optimal Control and the Calculus of Variations*, Ph.D. thesis, University of Washington, 1973.

-
- [10] Frank H. Clarke, *Generalized gradients and applications*, Transactions of the American Society **205** (1975), 247–262.
- [11] Frank H. Clarke, *Optimization and nonsmooth analysis*, Society for Industrial and Applied Mathematics, 1990.
- [12] Thomas F. Coleman, Burton S. Garbow, and Jorge J. More, *Software for estimating sparse jacobian matrices*, ACM Trans. Math. Softw. **10** (1984), no. 3, 329–345.
- [13] Tecla De Luca, Francisco Facchinei, and Christian Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Mathematical Programming **75** (1996), 407–439.
- [14] John E. Dennis and Robert B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Society for Industrial and Applied Mathematics, 1996.
- [15] John E. Dennis, Jr. and Jorge J. More, *A characterization of superlinear convergence and its application to quasi-newton methods*, Mathematics of Computation **28** (1974), no. 126, 549–560.
- [16] Jr. Dennis, J. E. and R. B. Schnabel, *Least change secant updates for quasi-newton methods*, SIAM Review **21** (1979), no. 4, pp. 443–459 (English).
- [17] Jean Dieudonné, *Referencia*, vol. 1, p. 92, Academic Press New York-London., 1969.
- [18] Michael C. Ferris and Jong-Shi Pang, *Engineering and economic applications of complementarity problems*, SIAM Review **39** (1997), 669–713.
- [19] Andreas Fischer and Christian Kanzow, *On finite termination of an iterative method for linear complementarity problems*, Math. Program. **74** (1996), no. 3, 279–292.
- [20] Márcia A Gomes-Ruggiero, JM Martinez, and Sandra Augusta Santos, *Solving nonsmooth equations by means of quasi-newton methods with globalization*, Recent Advances in Nonsmooth Optimization, World Scientific Publisher, Singapore (1995), 121–140.
- [21] Patrick T. Harker and Jong-Shi Pang, *Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications*, Mathematical Programming **48** (1990), 161–220.
- [22] Alston S. Householder, *Referencia*, p. 83, Dover Publications, INC., 1964.

- [23] Christian Kanzow and Helmut Kleinmichel, *A new class of semismooth newton-type methods for nonlinear complementarity problems*, Comput. Optim. Appl. **11** (1998), no. 3, 227–251.
- [24] Masakazu Kojima and Susumu Shindoh, *Extensions of newton and quasi-newton methods to systems of pc 1 equations*, Research reports on information sciences, Inst. of Technology, Department of Information Sciences, 1986.
- [25] Michael Kostreva, *Elasto-hydrodynamic lubrication: A non-linear complementarity problem*, International Journal for Numerical Methods in Fluids **4** (1984), no. 4, 377–397.
- [26] Erwin Kreyszig, *Referencia*, p. 144, Wiley, 1978.
- [27] Dong-Hui Li and Masao Fukushima, *Globally convergent broyden-like methods for semismooth equations and applications to vip, ncp and mcp*, Annals of Operations Research **103** (2001), no. 1-4, 71–97.
- [28] Vera L. R. Lopes, Jose M. Martínez, and Rosana Pérez, *On the local convergence of quasi-newton methods for nonlinear complementary problems*, Applied Numerical Mathematics **30** (1999), 3–22.
- [29] Ladislav Lukšan, *Inexact trust region method for large sparse systems of nonlinear equations*, Journal of Optimization Theory and Applications **81** (1994), 569–590.
- [30] Changfeng Ma, *A new smoothing quasi-newton method for nonlinear complementarity problems*, Applied Mathematics and Computation **171** (2005), no. 2, 807–823.
- [31] José M. Martínez, *Local convergence theory of inexact newton methods based on structured least change updates*, Mathematics of computation **55** (1990), 143–168.
- [32] José M. Martínez, *On the relation between two local convergence theories of least change secant update methods*, Mathematics of Computation **59** (1992), no. 200, pp. 457–481.
- [33] ———, *Practical quasi-newton methods for solving nonlinear systems*, Journal of Computational and Applied Mathematics **124** (2000), no. 1, 97 – 121.
- [34] Lars Mathiesen, *Computation of economic equilibria by a sequence of linear complementarity problems*, Economic Equilibrium: Model Formulation and Solution (Alan S. Manne, ed.), Mathematical Programming Studies, vol. 23, Springer Berlin Heidelberg, 1985, pp. 144–162.

- [35] ———, *An algorithm based on a sequence of linear complementarity problems applied to a walrasian equilibrium model: An example*, *Mathematical Programming* **37** (1987), 1–18.
- [36] Carl D. Meyer (ed.), *Matrix analysis and applied linear algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.
- [37] Jong-Shi Pang and Liqun Qi, *Nonsmooth equations: Motivation and algorithms*, *SIAM Journal on Optimization* **3** (1993), no. 3, 443–465.
- [38] Rosana Pérez and Vera L. R. Lopes, *Recent applications and numerical implementation of quasi-newton methods for solving nonlinear systems of equations*, *Numerical Algorithms* **35** (2004), 261–285.
- [39] Rosana Pérez M., *Algoritmos para complementariedade não lineal e problemas relacionados*, Ph.D. thesis, IMECC, UNICAM, Campinas, Brazil, Dezembro 1997.
- [40] Liqun Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, *Math. Oper. Res.* **18** (1993), no. 1, 227–244.
- [41] Thomas F. Rutherford, *Applied general equilibrium modeling with mpsge as a gams subsystem: An overview of the modeling framework and syntax*, *Computational Economics* **14** (1999), 1–46.
- [42] Nicol N. Schraudolph, Jin Yu, and Simon Günter, *A stochastic quasi-Newton method for online convex optimization*, Proc. 11th Intl. Conf. Artificial Intelligence and Statistics (AISTATS), Society for Artificial Intelligence and Statistics, March 2007.
- [43] Andrew Sherman, *On newton-iterative methods for the solution of systems of nonlinear equations*, *SIAM Journal on Numerical Analysis* **15** (1978), no. 4, 755–711.
- [44] David S. Watkins, *Referencia*, pp. 32–33, John Wiley and Sons, Inc., 2005.
- [45] ———, *Referencia*, p. 115, John Wiley and Sons, Inc., 2005.
- [46] ———, *Referencia*, p. 34, John Wiley and Sons, Inc., 2005.
- [47] Longquan Yong, *Nonlinear complementarity problem and solution methods*, Proceedings of the 2010 international conference on Artificial intelligence and computational intelligence: Part I, Springer-Verlag, 2010, pp. 461–469.