

Un algoritmo global con jacobiano suavizado para complementariedad no lineal



WILMER SÁNCHEZ GRUESO

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Matemáticas
Maestría en Ciencias Matemáticas
Popayán
2020

Un algoritmo global con jacobiano suavizado para complementariedad no lineal

Wilmer Sánchez Grueso

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. Hector 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

2020

Nota de aceptación

Director _____

Dra. Rosana Pérez Mera

Jurado *Anibal Sosa*

Dr. Uram Anibal Sosa Aguirre

Jurado _____

Dr. Carlos Andrés Arias Torres

Popayán, 22 de enero de 2020.

Agradecimientos

A mi Madre, quien con su amor, paciencia y esfuerzo me ha dado la fortaleza para realizar cada una de las metas que me he propuesto.

A mi Padre, por su dedicación, sabio consejo y su apoyo incondicional. Su tenacidad y lucha interminable han hecho de él un ejemplo a seguir por mi.

A mi Hermano, por ser mi amigo incondicional, mostrarme el camino en los momentos más difíciles. Sus conocimientos y consejos han motivado gran parte de este trabajo.

Especialmente a mi directora y amiga Dra. *Rosana Pérez Mera*, por darme la oportunidad de realizar este gran trabajo de investigación, por brindarme su dedicación incondicional, acertada orientación, conocimientos y su experiencia investigativa. Agradezco profundamente por llenar mi vida con sus valiosos consejos.

A mi codirector, *Hector J Martínez* por acompañarme en este trabajo y en esta carrera investigativa que apenas comienza.

A los Doctores *Uram Anibal Sosa Aguirre* y *Carlos Andrés Arias Torres*, quienes con su gran nivel académico e investigativo me brindaron sugerencias y recomendaciones permitiendo que este trabajo llegará a buen término.

Al Magíster *Favián Enrique Arenas Aparicio* por brindarme sus conocimientos y encaminarme a esta hermosa rama de las matemáticas, la optimización.

A mis compañeros y colegas de estudio, pues cuando compartes mucho tiempo con un grupo de personas, automáticamente se convierten en tu familia.

A todas y cada una de las personas que me acompañaron en este recorrido, muchas de las cuales han sido un soporte muy fuerte en la realización de este trabajo.

Wilmer Sánchez Grueso.

Índice general

Índice general	3
1. Introducción	5
2. La estrategia de suavización del jacobiano aplicada a $\Phi_\lambda(x) = 0$.	10
3. Preliminares	21
4. Algoritmo propuesto	30
5. Convergencia global	35
6. Convergencia local	48
7. Pruebas numéricas	54
7.1. Problemas de prueba, puntos iniciales y valores de los parámetros	55
7.2. Pruebas numéricas	60

ÍNDICE GENERAL	4
8. Comentarios finales	72
Bibliografía	74

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(F)) 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, decimos que un vector es no negativo, si todas sus componentes lo son. La tercera condición en (1.1) exige que los vectores \mathbf{x} y $F(\mathbf{x})$ sean ortogonales, por ello, es llamada condición de complementariedad.

El Problema de Complementariedad no Lineal surge en diversas aplicaciones de Física, Ingeniería y Economía [2] [21] [9] [14]. Su importancia se debe a que el concepto de complementariedad es sinónimo de la noción de sistema en equilibrio.

Una técnica, quizá la más popular, para resolver el problema consiste en reformularlo como un sistema de ecuaciones no lineales mediante una función de complementariedad $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

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

Ejemplos de este tipo de funciones son las siguientes:

- *Función mínimo* [28].

$$\varphi(a, b) = \min \{a, b\}.$$

- *Función de Fischer* [15].

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

- *Familia uniparamétrica de funciones de complementariedad* [19].

$$\varphi_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} - a - b, \quad \lambda \in (0, 4). \quad (1.3)$$

- *Clase de funciones de complementariedad* [18]

$$\varphi_\nu(a, b) = a + b - \sqrt{\nu(a - b)^2 + (1 - \nu)(a^2 + b^2)}, \quad \nu \in [0, 1]. \quad (1.4)$$

Observemos que φ_λ se reduce a la *función de Fischer* cuando $\lambda = 2$ y tiende a un múltiplo de la *función mínimo* cuando λ tiende a cero. Por otra parte, φ_ν se reduce a la *función de Fischer* cuando $\nu = 0$. Vale la pena mencionar que a pesar que existen numerosas funciones de complementariedad las funciones *mínimo* y *Fischer* han sido las más utilizadas en la reformulación del problema de complementariedad no lineal [19].

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

$$\Phi(\mathbf{x}) = \begin{pmatrix} \varphi(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi(x_n, F_n(\mathbf{x})) \end{pmatrix} = \mathbf{0}, \quad (1.5)$$

el cual, debido a la falta de suavidad de φ , es un sistema no diferenciable. De la definición de función de complementariedad dada en (1.2), se deduce que un vector \mathbf{x}_* resuelve el sistema (1.5) si, y solo si, \mathbf{x}_* resuelve el PCNL(F).

Para resolver (1.5) y con ello, resolver el PCNL(F), se han propuesto algoritmos tipo *Newton* no suaves [31] [34], también llamados de *Newton* generalizados; tipo cuasi *Newton* generalizados [6] [22] [23], entre otros [1] [7] [13] [24] [29] [33]. Para globalizar estos métodos, se resuelve el problema de minimización

$$\begin{aligned} & \text{Minimizar } \Psi(\mathbf{x}), \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned} \quad (1.6)$$

donde la función de mérito natural [27] $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}$ es definida por

$$\Psi(\mathbf{x}) = \frac{1}{2} \|\Phi(\mathbf{x})\|_2^2.$$

Por otra parte, en los métodos tipo *Newton* generalizados es necesario acudir al concepto de *jacobiano generalizado* [11], definido para $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$, *Lipschitz* continua, en un punto \mathbf{x} , como el conjunto

$$\partial G(\mathbf{x}) = \text{conv} \left\{ \lim_{k \rightarrow \infty} G'(\mathbf{x}_k) \in \mathbb{R}^{n \times n}: \lim_{k \rightarrow \infty} \mathbf{x}_k \rightarrow \mathbf{x}, \mathbf{x}_k \in D_G \right\}, \quad (1.7)$$

donde D_G es el conjunto de todos los puntos de \mathbb{R}^n donde G es diferenciable y $\text{conv} \{A\}$ representa la envolvente convexa de A . Este conjunto es no vacío, convexo y compacto [12]. Usualmente, el conjunto (1.7) es difícil de calcular, en especial cuando $n > 1$. Por ello, una alternativa para su implementación práctica es utilizar la subestimación propuesta en [12]:

$$\partial G(\mathbf{x})^T \subseteq \partial G_1(\mathbf{x}) \times \cdots \times \partial G_n(\mathbf{x}),$$

donde el lado derecho, denotado en [20] por $\partial_C G(\mathbf{x})^T$ y llamado el *C-subdiferencial* de G en \mathbf{x} , es el conjunto de matrices en $\mathbb{R}^{n \times n}$ con la i -ésima columna dada por el gradiente generalizado de la i -ésima componente de la función G .

Otra estrategia para resolver (1.5) es la de *suavización del jacobiano* propuesta en [10] y llamada así, en [20]. La idea general de los métodos que usan esta estrategia, es aproximar la función Φ por una función suave $\Phi_\mu: \mathbb{R}^n \rightarrow \mathbb{R}^n$, donde $\mu > 0$ denota el parámetro de suavización, para posteriormente resolver una sucesión de sistemas de ecuaciones no lineales suavizados,

$$\Phi_\mu(\mathbf{x}) = \begin{pmatrix} \varphi_\mu(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi_\mu(x_n, F_n(\mathbf{x})) \end{pmatrix} = \mathbf{0}, \quad (1.8)$$

con μ tendiendo a cero y φ_μ una suavización de la función de complementariedad φ usada en la reformulación (1.5). El sistema (1.8) se resuelve, en cada iteración, mediante la solución del sistema de ecuaciones lineales (ecuación de *Newton* híbrida),

$$\Phi'_\mu(\mathbf{x}_k) \mathbf{s}_k = -\Phi(\mathbf{x}_k), \quad (1.9)$$

donde $\Phi'_\mu(\mathbf{x}_k)$ es la matriz jacobiana de la función Φ_μ en \mathbf{x}_k . Asociada al sistema suavizado (1.8). En [20], los autores consideran la función de mérito relacionada con la

función Ψ previamente definida

$$\Psi_\mu(\mathbf{x}) = \frac{1}{2} \|\Phi_\mu(\mathbf{x})\|_2^2,$$

en la globalización de su algoritmo.

En [20], los autores presentan un nuevo algoritmo basado en la estrategia de suavización del jacobiano para resolver problemas de complementariedad no lineal, mediante su reformulación como un sistema de ecuaciones no lineales usando la función de *Fischer-Burmeister* y muestran un buen desempeño numérico de su algoritmo. Motivados por los resultados obtenidos en [20] con la función de *Fischer* y teniendo en cuenta que, esta función es un caso particular de la familia (1.3), cuando $\lambda = 2$, en el presente trabajo de investigación, usamos dicha familia de funciones de complementariedad en conexión con la estrategia del jacobiano suavizado para proponer un nuevo algoritmo para resolver el PCNL(F) mediante su reformulación como un sistema no lineal no diferenciable. Este algoritmo, puede verse como una generalización del propuesto en [20] para cualquier miembro de la familia φ_λ con λ en $(0, 4)$.

Por otra parte, nos gustaría mencionar que finalizando nuestra investigación, conocimos de la publicación de un nuevo algoritmo tipo *Newton* suavizado para resolver problemas de complementariedad no lineal [36], el cual considera la familia de funciones de complementariedad (1.4) y una suavización de la misma, propuesta en [35]. Este algoritmo usa una técnica tipo *Levenberg-Marquard* (usada para resolver problemas de mínimos cuadrados no lineales [13]) para aproximar la matriz hessiana de la función de mérito, respectiva. Con ello, también controlan la no singularidad de dicha matriz. Teniendo en cuenta lo novedoso de su estrategia, que también usan una familia de funciones de complementariedad (diferente de la que usamos nosotros, pero familia, al fin) y los buenos resultados del desempeño de su algoritmo, decidimos incluir este algoritmo en la comparación numérica que realizamos en nuestra investigación. Los resultados numéricos muestran que nuestro algoritmo, en general, presenta un mejor desempeño (medido con la norma 2).

El presente documento está organizado de la siguiente forma.

En el **Capítulo 2**, presentamos detalladamente la estrategia de suavización del jacobiano aplicada a la función Φ_λ , describimos explícitamente la matriz jacobiana de su suavización, que corresponde a la matriz de coeficientes de los sistemas lineales (1.9) y encontramos una cota superior del parámetro μ que será de gran importancia en la propuesta algorítmica.

En el **Capítulo 3**, presentamos algunos resultados teóricos preliminares que serán utilizados en el desarrollo de la teoría de convergencia del algoritmo propuesto.

En el **Capítulo 4**, proponemos un algoritmo con jacobiano suavizado para resolver problemas de complementariedad no lineal que generaliza el algoritmo presentado en [20] a todos los miembros de la familia φ_λ , con $\lambda \in (0, 4)$. Además, demostramos que el algoritmo está bien definido.

En el **Capítulo 5**, desarrollamos la teoría de convergencia global del método propuesto.

En el **Capítulo 6**, bajo ciertas hipótesis, demostramos que el **Algoritmo 4.1** converge local y, q -súperlinealmente o q -cuadráticamente a una solución del problema de complementariedad.

En el **Capítulo 7**, analizamos el desempeño numérico del algoritmo propuesto. Para ello, exploramos su comportamiento cuando variamos el parámetro λ y comparamos los resultados con los obtenidos en [4] [20] [36], para otras propuestas algorítmicas. Además, mostramos la bondad de tener un algoritmo global frente a uno local.

Finalmente, el **Capítulo 8**, contiene algunos comentarios finales sobre lo realizado y trabajos futuros.

Capítulo 2

La estrategia de suavización del jacobiano aplicada a $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$.

En este capítulo describimos detalladamente, la estrategia del jacobiano suavizado para resolver el problema de complementariedad no lineal usando su reformulación (1.5), con φ definida por (1.3) y su suavización definida recientemente en [3]. Complementamos esta descripción con la deducción de la forma explícita de la matriz jacobiana de la suavización y presentamos algunos resultados teóricos útiles en desarrollos posteriores.

En primer lugar, consideramos la reformulación (1.5) del PCNL(F) como el sistema de ecuaciones no lineales, no diferenciable,

$$\Phi_\lambda(\mathbf{x}) = \begin{pmatrix} \varphi_\lambda(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi_\lambda(x_n, F_n(\mathbf{x})) \end{pmatrix} = \mathbf{0}, \quad (2.1)$$

donde φ_λ es la familia de funciones de complementariedad (1.3).

Recordemos que, un método de *Newton* generalizado para resolver el sistema (2.1), debe resolver, en cada iteración, el sistema de ecuaciones lineales

$$H_k \mathbf{s}_k = -\Phi_\lambda(\mathbf{x}_k), \quad (2.2)$$

donde $H_k \in \partial\Phi_\lambda(\mathbf{x}_k)$ o $H_k \in \partial_C\Phi_\lambda(\mathbf{x}_k)$. Vale la pena mencionar que, en este documento usaremos matrices $H_k \in \partial_C\Phi_\lambda(\mathbf{x}_k)$.

Con el propósito de definir un método de jacobiano suavizado para (2.1), seguimos la idea básica de la estrategia del jacobiano suavizado dada en [20] y consideramos la suavización de φ_λ propuesta en [3] definida por

$$\varphi_{\lambda\mu}(a, b) = \sqrt{(a-b)^2 + \lambda ab + (4-\lambda)\mu} - a - b, \quad (2.3)$$

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

Para simplificar notación en cálculos posteriores, usaremos la notación propuesta en [3], para el primer sumando del lado derecho en (2.3), es decir,

$$G_{\lambda\mu}(a, b) = \sqrt{(a-b)^2 + \lambda ab + (4-\lambda)\mu}. \quad (2.4)$$

Observemos que el problema de la no diferenciabilidad de φ_λ se soluciona con su suavización $\varphi_{\lambda\mu}$, la cual es una función continuamente diferenciable para todo $(a, b) \in \mathbb{R}^2$. Además, es claro que a medida que μ tiende a cero, $\varphi_{\lambda\mu}$, se acerca a φ_λ .

Como es de esperarse, la distancia entre la función de complementariedad φ_λ y su suavización $\varphi_{\lambda\mu}$ está acotada superiormente por una constante que depende de los parámetros λ y μ . Esto, es un caso particular de la siguiente proposición, la cual será de utilidad en la pruebas de convergencia que presentamos en un capítulo posterior.

Proposición 2.1. *La función $\varphi_{\lambda\mu}$ satisface la desigualdad*

$$|\varphi_{\lambda\mu_1}(a, b) - \varphi_{\lambda\mu_2}(a, b)| \leq \sqrt{4-\lambda} |\sqrt{\mu_1} - \sqrt{\mu_2}|,$$

para todo $(a, b) \in \mathbb{R}^2$ y todo $\mu_1, \mu_2 \geq 0$. En particular, tenemos

$$|\varphi_{\lambda\mu} - \varphi_\lambda| \leq \sqrt{4-\lambda} \sqrt{\mu},$$

para todo $(a, b) \in \mathbb{R}^2$, $\lambda \in (0, 4)$ y todo $\mu \geq 0$.

Demostración. Sean $(a, b) \in \mathbb{R}^2$, μ_1 y μ_2 no negativos, tales que $\mu_1 \neq \mu_2$.

$$\begin{aligned} |\varphi_{\lambda\mu_1}(a, b) - \varphi_{\lambda\mu_2}(a, b)| &= |G_{\lambda\mu_1}(a, b) - G_{\lambda\mu_2}(a, b)| \\ &= \left| \frac{G_{\lambda\mu_1}^2(a, b) - G_{\lambda\mu_2}^2(a, b)}{G_{\lambda\mu_1}(a, b) + G_{\lambda\mu_2}(a, b)} \right| \\ &= \left| \frac{(\mu_1 - \mu_2)(4 - \lambda)}{G_{\lambda\mu_1}(a, b) + G_{\lambda\mu_2}(a, b)} \right| \end{aligned}$$

Ahora, observemos que $G_{\lambda\mu}(a, b) \geq \sqrt{\mu(4 - \lambda)}$, luego

$$\begin{aligned} \frac{1}{G_{\lambda\mu_1}(a, b)} &\leq \frac{1}{\sqrt{\mu_1}\sqrt{4 - \lambda}} \\ |\varphi_{\lambda\mu_1}(a, b) - \varphi_{\lambda\mu_2}(a, b)| &\leq \frac{|(\mu_1 - \mu_2)(4 - \lambda)|}{(\sqrt{\mu_1} + \sqrt{\mu_2})\sqrt{4 - \lambda}} \end{aligned}$$

de ahí que

$$|\varphi_{\lambda\mu_1}(a, b) - \varphi_{\lambda\mu_2}(a, b)| \leq |\sqrt{\mu_1} - \sqrt{\mu_2}|\sqrt{4 - \lambda}.$$

En particular, cuando $\mu_1 = \mu$ y $\mu_2 = 0$, se obtiene

$$|\varphi_{\lambda\mu} - \varphi_{\lambda}| \leq \sqrt{4 - \lambda}\sqrt{\mu}. \quad \square$$

A partir de (2.3), definimos la función $\Phi_{\lambda\mu}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ de la siguiente forma,

$$\Phi_{\lambda\mu}(\mathbf{x}) = \begin{pmatrix} \varphi_{\lambda\mu}(x_1, F_1(\mathbf{x})) \\ \varphi_{\lambda\mu}(x_2, F_2(\mathbf{x})) \\ \vdots \\ \varphi_{\lambda\mu}(x_n, F_n(\mathbf{x})) \end{pmatrix}. \quad (2.5)$$

Como consecuencia de la **Proposición 2.1** tenemos la siguiente proposición que, en particular, da una cota superior para la distancia entre Φ_{λ} y su aproximación $\Phi_{\lambda\mu}$.

Proposición 2.2. *Las siguientes desigualdades se tienen para la función $\Phi_{\lambda\mu}$*

$$i) \quad \|\Phi_{\lambda\mu_1} - \Phi_{\lambda\mu_2}\| \leq \kappa|\sqrt{\mu_1} - \sqrt{\mu_2}|.$$

$$2i) \quad \|\Phi_{\lambda\mu} - \Phi_\lambda\| \leq \kappa\sqrt{\mu}.$$

para todo $\mathbf{x} \in \mathbb{R}^n$, y $\mu, \mu_1, \mu_2 \geq 0$, donde $\kappa = \sqrt{n(4 - \lambda)}$.

Demostración.

$$\|\Phi_{\lambda\mu_1}(\mathbf{x}) - \Phi_{\lambda\mu_2}(\mathbf{x})\| = \left\| \begin{pmatrix} \varphi_{\lambda\mu_1}(x_1, F_1(\mathbf{x})) \\ \varphi_{\lambda\mu_1}(x_2, F_2(\mathbf{x})) \\ \vdots \\ \varphi_{\lambda\mu_1}(x_n, F_n(\mathbf{x})) \end{pmatrix} - \begin{pmatrix} \varphi_{\lambda\mu_2}(x_1, F_1(\mathbf{x})) \\ \varphi_{\lambda\mu_2}(x_2, F_2(\mathbf{x})) \\ \vdots \\ \varphi_{\lambda\mu_2}(x_n, F_n(\mathbf{x})) \end{pmatrix} \right\|.$$

Usando la definición de la norma dos y la **Proposición 2.1**, tenemos que,

$$\begin{aligned} \|\Phi_{\lambda\mu_1}(\mathbf{x}) - \Phi_{\lambda\mu_2}(\mathbf{x})\| &= \sqrt{\sum_{i=1}^n [\varphi_{\lambda\mu_1}(x_i, F_i(\mathbf{x})) - \varphi_{\lambda\mu_2}(x_i, F_i(\mathbf{x}))]^2} \\ &\leq \sqrt{\sum_{i=1}^n [(\sqrt{\mu_1} - \sqrt{\mu_2})\sqrt{4 - \lambda}]^2} = \sqrt{n} [(\sqrt{\mu_1} - \sqrt{\mu_2})\sqrt{4 - \lambda}]^2 \\ &= \sqrt{n(4 - \lambda)} |\sqrt{\mu_1} - \sqrt{\mu_2}|. \end{aligned}$$

La segunda parte de la proposición se obtiene eligiendo $\mu_1 = \mu$ y $\mu_2 = 0$. \square

Con la información anterior, tenemos que, una iteración del método con jacobiano suavizado para resolver (2.1), es de la forma:

$$\begin{aligned} \Phi'_{\lambda\mu}(\mathbf{x}_k) \mathbf{s}_k &= -\Phi_\lambda(\mathbf{x}_k) \\ \mathbf{x}_{k+1} &= \mathbf{s}_k + \mathbf{x}_k, \end{aligned} \tag{2.6}$$

donde $\Phi'_{\lambda\mu}(\mathbf{x}_k)$ es la matriz jacobiana de $\Phi_{\lambda\mu}$ en \mathbf{x}_k .

De (2.2) y (2.6), podemos inferir que un método de suavización del jacobiano trata de resolver directamente la reformulación $\Phi_\lambda(\mathbf{x}) = 0$, reemplazando la matriz H_k en el conjunto $\partial_C \Phi_\lambda(\mathbf{x})$, por una aproximación $\Phi'_{\lambda\mu_k}(\mathbf{x}_k)$ de dicha matriz. Desde este punto de vista, podemos interpretar los métodos de jacobiano suavizado como métodos tipo, casi-Newton generalizados.

La matriz jacobiana $\Phi'_{\lambda\mu}(\mathbf{x})$ juega un papel central en el método de jacobiano suavizado que propondremos. Una forma general de la matriz jacobiana de $\Phi_{\lambda\mu}$ está dada por:

$$\Phi'_{\lambda\mu}(\mathbf{x}) = \begin{pmatrix} \nabla\varphi_{\lambda\mu}(x_1, F_1(\mathbf{x}))^T \\ \nabla\varphi_{\lambda\mu}(x_2, F_2(\mathbf{x}))^T \\ \vdots \\ \nabla\varphi_{\lambda\mu}(x_n, F_n(\mathbf{x}))^T \end{pmatrix}. \quad (2.7)$$

Con el fin de describir explícitamente cada una de las filas de $\Phi'_{\lambda\mu}(\mathbf{x})$, usamos la función definida en (2.4) para la cual, su gradiente esta dado por,

$$\nabla G_{\lambda\mu}(a, b) = \begin{pmatrix} \frac{2(a-b) + \lambda b}{2G_{\lambda\mu}(a, b)} \\ -2(a-b) + \lambda a \\ \frac{-2(a-b) + \lambda a}{2G_{\lambda\mu}(a, b)} \end{pmatrix} := \begin{pmatrix} \alpha_{\lambda\mu}(a, b) \\ \beta_{\lambda\mu}(a, b) \end{pmatrix}. \quad (2.8)$$

De (2.3) y (2.8) tenemos que, para cualquier $(x, y) \in \mathbb{R}^2$,

$$\nabla\varphi_{\lambda\mu}(a, b) = \nabla G_{\lambda\mu}(a, b) - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_{\lambda\mu}(a, b) - 1 \\ \beta_{\lambda\mu}(a, b) - 1 \end{pmatrix}. \quad (2.9)$$

Para mayor claridad en la lectura de este documento definimos las variables auxiliares u y v por, $u = x_1$ y $v = F_1(\mathbf{x})$. Así, usando la regla de la cadena tenemos que,

$$\nabla\varphi_{\lambda\mu}(x_1, F_1(\mathbf{x})) = \nabla\varphi_{\lambda\mu}(u, v) = \begin{pmatrix} \frac{\partial\varphi_{\lambda\mu}}{\partial u} \frac{\partial u}{\partial x_1} + \frac{\partial\varphi_{\lambda\mu}}{\partial v} \frac{\partial v}{\partial x_1} \\ \frac{\partial\varphi_{\lambda\mu}}{\partial u} \frac{\partial u}{\partial x_2} + \frac{\partial\varphi_{\lambda\mu}}{\partial v} \frac{\partial v}{\partial x_2} \\ \vdots \\ \frac{\partial\varphi_{\lambda\mu}}{\partial u} \frac{\partial u}{\partial x_n} + \frac{\partial\varphi_{\lambda\mu}}{\partial v} \frac{\partial v}{\partial x_n} \end{pmatrix},$$

con lo cual,

$$\begin{aligned} \nabla\varphi_{\lambda\mu}(u, v) &= \begin{pmatrix} (\alpha_{\lambda\mu}(u, v) - 1) 1 + (\beta_{\lambda\mu}(u, v) - 1) \frac{\partial F_1(\mathbf{x})}{\partial x_1} \\ (\alpha_{\lambda\mu}(u, v) - 1) 0 + (\beta_{\lambda\mu}(u, v) - 1) \frac{\partial F_1(\mathbf{x})}{\partial x_2} \\ \vdots \\ (\alpha_{\lambda\mu}(u, v) - 1) 0 + (\beta_{\lambda\mu}(u, v) - 1) \frac{\partial F_1(\mathbf{x})}{\partial x_n} \end{pmatrix} \\ &= (\alpha_{\lambda\mu}(u, v) - 1) \mathbf{e}_1 + (\beta_{\lambda\mu}(u, v) - 1) \nabla F_1(\mathbf{x}). \end{aligned}$$

Luego, la primera fila de la matriz (2.7) es dada por

$$\nabla\varphi_{\lambda\mu}(x_1, F_1(\mathbf{x}))^T = (\alpha_{\lambda\mu}(x_1, F_1(\mathbf{x})) - 1) \mathbf{e}_1^T + (\beta_{\lambda\mu}(x_1, F_1(\mathbf{x})) - 1) \nabla F_1(\mathbf{x})^T,$$

y de forma análoga, la fila i -ésima de la matriz $\Phi'_{\lambda\mu}(\mathbf{x})$ está dada por,

$$\nabla\varphi_{\lambda\mu}(x_i, F_i(\mathbf{x}))^T = (\alpha_{\lambda\mu}(x_i, F_i(\mathbf{x})) - 1) \mathbf{e}_i^T + (\beta_{\lambda\mu}(x_i, F_i(\mathbf{x})) - 1) \nabla F_i(\mathbf{x})^T.$$

Por lo tanto, una forma matricial de la expresión $\Phi'_{\lambda\mu}(\mathbf{x})$ es:

$$\Phi'_{\lambda\mu}(\mathbf{x}) = \begin{pmatrix} (\alpha_{\lambda\mu}(x_1, F_1(\mathbf{x})) - 1) \mathbf{e}_1^T + (\beta_{\lambda\mu}(x_1, F_1(\mathbf{x})) - 1) \nabla F_1(\mathbf{x})^T \\ (\alpha_{\lambda\mu}(x_2, F_2(\mathbf{x})) - 1) \mathbf{e}_2^T + (\beta_{\lambda\mu}(x_2, F_2(\mathbf{x})) - 1) \nabla F_2(\mathbf{x})^T \\ \vdots \\ (\alpha_{\lambda\mu}(x_n, F_n(\mathbf{x})) - 1) \mathbf{e}_n^T + (\beta_{\lambda\mu}(x_n, F_n(\mathbf{x})) - 1) \nabla F_n(\mathbf{x})^T \end{pmatrix},$$

en forma compacta,

$$\Phi'_{\lambda\mu}(\mathbf{x}) = \text{diag}(A_1(\mathbf{x}), \dots, A_n(\mathbf{x})) + \text{diag}(B_1(\mathbf{x}), \dots, B_n(\mathbf{x})) F'(\mathbf{x}), \quad (2.10)$$

donde para todo, $i = 1, \dots, n$,

$$A_i(\mathbf{x}) = \alpha_{\lambda\mu}(x_i, F_i(\mathbf{x})) - 1,$$

$$B_i(\mathbf{x}) = \beta_{\lambda\mu}(x_i, F_i(\mathbf{x})) - 1.$$

La siguiente proposición [3], garantiza que cuando el parámetro μ tiende a cero entonces la distancia entre $\Phi'_{\lambda\mu}(\mathbf{x})$ y el conjunto $\partial_C\Phi_\lambda(\mathbf{x})$, también tiende a cero, lo cual hace que tenga sentido reemplazar la iteración de *Newton* generalizada (2.2) por (2.6).

Proposición 2.3. *Sean $\mathbf{x} \in \mathbb{R}^n$ y $\mu > 0$, arbitrario pero fijo. Entonces*

$$\lim_{\mu \rightarrow 0} \text{dist}(\Phi'_{\lambda\mu}(\mathbf{x}), \partial_C\Phi_\lambda(\mathbf{x})) = 0.$$

Como consecuencia de esta proposición, para cada $\delta > 0$ existe $\bar{\mu} = \bar{\mu}(\mathbf{x}, \delta) > 0$ tal que,

$$\text{dist}(\Phi'_{\lambda\mu}(\mathbf{x}), \partial_C\Phi_\lambda(\mathbf{x})) \leq \delta,$$

para todo $0 < \mu \leq \bar{\mu}$.

Para obtener una expresión explícita para $\bar{\mu}$, de gran importancia en el algoritmo que propondremos, ya que da una cota superior para μ , procedemos en forma análoga a lo hecho en [20], como se verifica en la siguiente proposición.

Proposición 2.4. *Sea $\mathbf{x} \in \mathbb{R}^n$ arbitrario. Suponga que \mathbf{x} no es una solución del problema de complementariedad no lineal y*

$$\gamma(\mathbf{x}) = \frac{1}{2} \max_{i \notin \beta(\mathbf{x})} \{ \| [2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})] \mathbf{e}_i + [-2(x_i - F_i(\mathbf{x})) + \lambda x_i] \nabla F_i(\mathbf{x}) \| \}$$

y

$$\alpha(\mathbf{x}) = \min_{i \notin \beta(\mathbf{x})} \{ (x_i - F_i(\mathbf{x}))^2 + \lambda x_i F_i(\mathbf{x}) \} > 0,$$

con definido por $\beta(\mathbf{x}) = \{i: \mathbf{x}_i = 0 = F_i(\mathbf{x})\}$. Sea $\delta > 0$, defina

$$\bar{\mu}(\mathbf{x}, \delta, \lambda) := \begin{cases} 1 & \text{si } \frac{n\gamma(\mathbf{x})^2}{\delta^2} - \alpha(\mathbf{x}) \leq 0 \\ \frac{\alpha(\mathbf{x})^2}{4 - \lambda} \left(\frac{\delta^2}{n\gamma(\mathbf{x})^2 - \delta^2\alpha(\mathbf{x})} \right) & \text{otro caso.} \end{cases}$$

Entonces

$$\text{dist}_F(\Phi'_{\lambda\mu}(\mathbf{x}), \partial_C\Phi_\lambda(\mathbf{x})) \leq \delta,$$

para todo μ tal que $0 < \mu < \bar{\mu}(\mathbf{x}, \delta, \lambda)$.

Demostración. Dado que \mathbf{x} no es solución del PCNL(F), $\{1, \dots, n\} \setminus \beta(\mathbf{x}) \neq \emptyset$, con lo cual, $\alpha(\mathbf{x}) > 0$. Además

$$\begin{aligned} dist_F(\Phi'_{\lambda\mu}(\mathbf{x}), \partial_C \Phi_\lambda(\mathbf{x})) &= dist_F(\nabla \Phi_{\lambda\mu}(\mathbf{x}), \partial_C \Phi(\mathbf{x})^T) \\ &= \sqrt{\sum_{i=1}^n [dist_2([\nabla \Phi_{\lambda\mu}(\mathbf{x})]_i, \partial \Phi_{\lambda,i}(\mathbf{x}))]^2}. \end{aligned} \quad (2.11)$$

Por la **Proposición 3.2**, es suficiente considerar la distancia entre las i -ésimas columnas de $\nabla \Phi_{\lambda\mu}(\mathbf{x})$ y $\partial_C \Phi_\lambda^T(\mathbf{x})$. De la forma matricial en (2.7) y de la **Proposición 3.6** sabemos que las columnas están dadas respectivamente por

$$[\nabla \Phi_{\lambda\mu}(\mathbf{x})]_i = (\alpha_{\lambda\mu}(x_i, F_i(\mathbf{x})) - 1)\mathbf{e}_i + (\beta_{\lambda\mu}(x_i, F_i(\mathbf{x})) - 1)\nabla F_i(\mathbf{x})$$

y

$$\partial \Phi_{\lambda,i}(\mathbf{x}) = \begin{cases} (\alpha_\lambda(x_i, F_i(\mathbf{x})) - 1)\mathbf{e}_i + (\beta_\lambda(x_i, F_i(\mathbf{x})) - 1)\nabla F_i(\mathbf{x}) & \text{si } i \in \beta(\mathbf{x}) \\ (\sigma_i - 1)\mathbf{e}_i + (\rho_i - 1)\nabla F_i(\mathbf{x}) & \text{si } i \notin \beta(\mathbf{x}) \end{cases}$$

donde $(\sigma_i, \rho_i) \in \mathbb{R}^2$ y $\|(\sigma_i, \rho_i)\| \leq \sqrt{c_\lambda}$, $c_\lambda \in (0, 2)$.

Para el análisis de la distancia entre las columnas, es necesario considerar dos casos

Caso 1: Si $i \in \beta(\mathbf{x})$ entonces $(x_i, F_i(\mathbf{x})) = (0, 0)$,

$$[\nabla \Phi_{\lambda\mu}(\mathbf{x})]_i = -\mathbf{e}_i - \nabla F_i(\mathbf{x}),$$

que coincide con $\partial \Phi_{\lambda,i}(\mathbf{x})$ para $(\sigma_i, \rho_i) = (0, 0)$, por lo tanto,

$$dist_2([\nabla \Phi_{\lambda\mu}(\mathbf{x})]_i, \partial \Phi_{\lambda,i}(\mathbf{x})) = 0 \quad (2.12)$$

y cualquier $\delta > 0$ sirve.

Caso 2: Si $i \notin \beta(\mathbf{x})$ entonces el Jacobiano generalizado está formado por un solo elemento, la matriz jacobiana de $\Phi_\lambda(\mathbf{x})$. La distancia $D = dist_2([\nabla \Phi'_{\lambda\mu}(\mathbf{x})]_i, \partial \Phi_{\lambda,i}(\mathbf{x}))$ esta dada por:

$$\begin{aligned}
D &= \left\| [\nabla\Phi_{\lambda\mu}(\mathbf{x})]_i - \nabla\Phi_{\lambda,i}(\mathbf{x}) \right\| \\
&= \left\| \left(\frac{2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})}{2G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} - 1 \right) \mathbf{e}_i + \left(\frac{-2(x_i - F_i(\mathbf{x})) + \lambda x_i}{2G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} - 1 \right) \nabla F_i(\mathbf{x}) \right. \\
&\quad \left. - \left(\frac{2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})}{2G_{\lambda}(x_i, F_i(\mathbf{x}))} - 1 \right) \mathbf{e}_i - \left(\frac{-2(x_i - F_i(\mathbf{x})) + \lambda x_i}{2G_{\lambda}(x_i, F_i(\mathbf{x}))} - 1 \right) \nabla F_i(\mathbf{x}) \right\| \\
&= \left\| \frac{1}{2}(2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})) \left(\frac{1}{G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} - \frac{1}{G_{\lambda}(x_i, F_i(\mathbf{x}))} \right) \mathbf{e}_i \right. \\
&\quad \left. + \frac{1}{2}(-2(x_i - F_i(\mathbf{x})) + \lambda x_i) \left(\frac{1}{G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} - \frac{1}{G_{\lambda}(x_i, F_i(\mathbf{x}))} \right) \nabla F_i(\mathbf{x}) \right\| \\
&= \frac{1}{2} \left(\frac{1}{G_{\lambda}(x_i, F_i(\mathbf{x}))} - \frac{1}{G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} \right) \left\| (2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})) \mathbf{e}_i \right. \\
&\quad \left. + (-2(x_i - F_i(\mathbf{x})) + \lambda x_i) \nabla F_i(\mathbf{x}) \right\|
\end{aligned}$$

Usando la definición de $\gamma(\mathbf{x})$, tenemos que

$$D = \text{dist}_2([\nabla\Phi'_{\lambda\mu}(\mathbf{x})]_i, \partial\Phi_{\lambda,i}(\mathbf{x})) \leq \left(\frac{1}{G_{\lambda}(x_i, F_i(\mathbf{x}))} - \frac{1}{G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} \right) \gamma(\mathbf{x})$$

y por la definición $\alpha(\mathbf{x})$ tenemos que, $\alpha(\mathbf{x}) \leq (x_i - F_i(\mathbf{x}))^2 + \lambda x_i F_i(\mathbf{x})$, para todo i . Además, por el **Lema 3.1** tenemos que

$$h(\alpha(\mathbf{x})) \geq h((x_i - F_i(\mathbf{x}))^2 + \lambda x_i F_i(\mathbf{x})),$$

es decir,

$$\frac{1}{G_{\lambda}(x_i, F_i(\mathbf{x}))} - \frac{1}{G_{\lambda\mu}(x_i, F_i(\mathbf{x}))} \leq \frac{1}{\sqrt{\alpha(\mathbf{x})}} - \frac{1}{\sqrt{\alpha(\mathbf{x}) + (4 - \lambda)\mu}},$$

por lo tanto,

$$\begin{aligned}
\text{dist}_2([\nabla\Phi'_{\lambda\mu}(\mathbf{x})]_i, \partial\Phi_{\lambda,i}(\mathbf{x})) &\leq \left(\frac{1}{\sqrt{\alpha(\mathbf{x})}} - \frac{1}{\sqrt{\alpha(\mathbf{x}) + (4 - \lambda)\mu}} \right) \gamma(\mathbf{x}) \\
&\leq \left(\frac{\sqrt{(4 - \lambda)\mu}}{\sqrt{\alpha(\mathbf{x})}\sqrt{\alpha(\mathbf{x}) + (4 - \lambda)\mu}} \right) \gamma(\mathbf{x}).
\end{aligned}$$

Ahora demostremos que,

$$\left(\frac{\sqrt{(4-\lambda)\mu}}{\sqrt{\alpha(\mathbf{x})}\sqrt{\alpha(\mathbf{x})+(4-\lambda)\mu}} \right) \gamma(\mathbf{x}) \leq \frac{\delta}{\sqrt{n}}, \quad (2.13)$$

para todo $0 < \mu < \bar{\mu}(\mathbf{x}, \delta, \lambda)$.

Si $\gamma(\mathbf{x}) = 0$ la desigualdad se satisface trivialmente para todo $\mu > 0$. Supongamos que, $\gamma(\mathbf{x}) > 0$ y consideremos la expresión equivalente a (2.13)

$$\alpha(\mathbf{x})^2 \geq (4-\lambda)\mu \left(\frac{n\gamma(\mathbf{x})^2}{\delta^2} - \alpha(\mathbf{x}) \right).$$

Si $\frac{n\gamma(\mathbf{x})^2}{\delta^2} - \alpha(\mathbf{x}) \leq 0$ entonces la desigualdad se satisface para todo $\mu > 0$, en particular para $\mu \in (0, 1]$. En otro caso, se tiene que

$$\mu \leq \frac{\alpha(\mathbf{x})^2}{4-\lambda} \left(\frac{\delta^2}{n\gamma(\mathbf{x})^2 - \delta^2\alpha(\mathbf{x})} \right) = \bar{\mu}(\mathbf{x}, \delta, \lambda)$$

por tanto, para todo $0 < \mu < \bar{\mu}(\mathbf{x}, \delta, \lambda)$,

$$dist_2([\nabla\Phi'_{\lambda\mu}(\mathbf{x})]_i, \partial\Phi_{\lambda,i}(\mathbf{x})) \leq \frac{\delta}{\sqrt{n}}, \quad (2.14)$$

por (2.11), (2.12) y (2.14) concluimos que,

$$dist_F(\Phi'_{\lambda\mu}(\mathbf{x}), \partial_C\Phi_\lambda(\mathbf{x})) \leq \sqrt{\sum_{i=1}^n \frac{\delta^2}{n}} = \delta,$$

para todo $0 < \mu < \bar{\mu}(\mathbf{x}, \delta, \lambda)$. □

Dado que nuestra propuesta es un algoritmo global para resolver el PCNL(F), indirectamente a través de su reformulación (2.1), consideramos la función de mérito natural para el problema (2.1), $\Psi_\lambda: \mathbb{R}^n \rightarrow \mathbb{R}$, definida por,

$$\Psi_\lambda(\mathbf{x}) = \frac{1}{2}\Phi_\lambda(\mathbf{x})^T\Phi_\lambda(\mathbf{x}). \quad (2.15)$$

Así, la idea es resolver el PCNL(F) minimizando la función de mérito Ψ_λ . Un problema a tener en cuenta al considerar ese método de minimización es que la dirección encontrada

en (2.6) no necesariamente es una dirección de descenso para Ψ_λ en el punto \mathbf{x}_k . Siguiendo lo hecho en [20], una alternativa, es usar dicha dirección para reducir la función de mérito relacionada

$$\Psi_{\lambda\mu}(\mathbf{x}) = \frac{1}{2}\Phi_{\lambda\mu}(\mathbf{x})^T\Phi_{\lambda\mu}(\mathbf{x}). \quad (2.16)$$

Además, para garantizar que el algoritmo este bien definido para un PCNL(F) arbitrario, usamos un paso del gradiente para la función de mérito Ψ_λ cuando (2.6) no tenga solución o dé una dirección con poco descenso para $\Psi_{\lambda\mu}$.

Preliminares

En este capítulo presentamos algunas definiciones, proposiciones y lemas relacionados con la teoría de complementariedad no lineal que serán de gran utilidad para el desarrollo de la teoría de convergencia de nuestra propuesta algorítmica.

Para claridad en su lectura organizamos la presentación de los resultados de la siguiente forma: primero definiciones, luego proposiciones y lemas.

Definición 3.1. Sea $A \in \mathbb{R}^{n \times n}$. La norma de Frobenius de A se define por,

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{traza}(A^T A)}.$$

Definición 3.2. Sea $A \in \mathbb{R}^{n \times n}$ y \mathcal{M} es un subconjunto no vacío de matrices de tamaño $n \times n$. Se define la distancia entre A y \mathcal{M} como,

$$\text{dist}(A, \mathcal{M}) = \inf_{B \in \mathcal{M}} \{\|A - B\|\},$$

donde $\|\cdot\|$ es alguna norma matricial.

Definición 3.3. Sea $\{t_k\} \subseteq \mathbb{R}$, si existe un número real U que satisface las siguientes condiciones:

i) Para cada $\epsilon > 0$ existe un entero N tal que $k > N$ implica

$$t_k < U + \epsilon.$$

2i) Dado $\epsilon > 0$ y $m > 0$, existe un entero $k > m$ tal que

$$t_k > U - \epsilon$$

entonces U se llama límite superior de $\{t_k\}$ y se escribe

$$U = \limsup_{k \rightarrow \infty} t_k.$$

Asociados a la solución, \mathbf{x}^* , del PCNL(F) tenemos los siguientes conjuntos de índices

$$\begin{aligned} \hat{\alpha} &:= \{i: \mathbf{x}_i^* > 0 = F_i(\mathbf{x}^*)\}, \\ \hat{\beta} &:= \{i: \mathbf{x}_i^* = 0 = F_i(\mathbf{x}^*)\}, \\ \hat{\gamma} &:= \{i: \mathbf{x}_i^* < 0 = F_i(\mathbf{x}^*)\}. \end{aligned} \tag{3.1}$$

Cuando $\beta \neq \emptyset$ entonces la solución \mathbf{x}^* se denomina *solución degenerada*.

Definición 3.4. Sea \mathbf{x}^* una solución del PCNL(F).

1. Si todas la matrices $H \in \partial_B \Phi_\lambda(\mathbf{x}^*)$ son no singulares, \mathbf{x}^* es llamada una solución *BD-regular*.

2. Si la submatriz¹ $F'(\mathbf{x}^*)_{\hat{\alpha}\hat{\alpha}}$ es no singular y el complemento de Schur

$$F'(\mathbf{x}^*)_{\hat{\beta}\hat{\beta}} - F'(\mathbf{x}^*)_{\hat{\beta}\hat{\alpha}} F'(\mathbf{x}^*)_{\hat{\alpha}\hat{\alpha}}^{-1} F'(\mathbf{x}^*)_{\hat{\alpha}\hat{\beta}}$$

es una *P-matriz*², \mathbf{x}^* es llamada una solución *R-regular*.

¹Dada una matriz $A = (a_{ij})$ de tamaño $m \times n$ y los conjuntos de índices η y τ , la matriz $A_{\eta\tau}$ es aquella con componentes a_{ij} tales que $i \in \eta$ y $j \in \tau$.

²Una matriz $M \in \mathbb{R}^{n \times m}$ es llamada una *P-matriz*, si para cualquier vector no nulo $y \in \mathbb{R}^n$, existe un índice $i_0 = i_0(\mathbf{y}) \in \{1, 2, \dots, n\}$ tal que $y_{i_0} [M\mathbf{y}]_{i_0}$.

Definición 3.5. Dadas dos sucesiones $\{\alpha_k\}$ y $\{\beta_k\}$ tales que $\beta_k \geq 0$ para todo k , escribimos

$$\alpha_k = O(\beta_k),$$

si existe una constante $M > 0$ tal que $|\alpha_k| \leq M\beta_k$ para todo k . Escribimos

$$\alpha_k = o(\beta_k) \quad \text{cuando} \quad k \rightarrow \infty,$$

si

$$\lim_{k \rightarrow \infty} \frac{\alpha_k}{\beta_k} = 0.$$

El siguiente resultado garantiza, bajo ciertas hipótesis, la suavidad y semisuavidad fuerte de Φ_λ y cuya demostración es básicamente la presentada en [19] usando elementos de C -subdiferenciabilidad [30].

Proposición 3.1. Suponga que $\{\mathbf{x}^k\} \subseteq \mathbb{R}^n$ es una sucesión que converge a $\mathbf{x}^* \in \mathbb{R}^n$. Entonces se cumple lo siguiente:

1. La función Φ_λ es semisuave, así que

$$\|\Phi_\lambda(\mathbf{x}^k) - \Phi_\lambda(\mathbf{x}^*) - H_k(\mathbf{x}^k - \mathbf{x}^*)\| = o(\|\mathbf{x}^k - \mathbf{x}^*\|),$$

para cualquier $H_k \in \partial_C \Phi_\lambda(\mathbf{x}^k)$.

2. Si F es continuamente diferenciable con jacobiano localmente Lipchitz, entonces Φ_λ es fuertemente semisuave así que

$$\|\Phi_\lambda(\mathbf{x}^k) - \Phi_\lambda(\mathbf{x}^*) - H_k(\mathbf{x}^k - \mathbf{x}^*)\| = O(\|\mathbf{x}^k - \mathbf{x}^*\|^2),$$

para cualquier $H_k \in \partial_C \Phi_\lambda(\mathbf{x}^k)$.

Proposición 3.2. [20] Sea $\mathbf{x} \in \mathbb{R}^n$ y $\mu > 0$ arbitrario pero fijo. Entonces

$$[\text{dist}_F(\nabla \Phi_{\lambda\mu}(\mathbf{x}), \partial_C \Phi(\mathbf{x})^T)]^2 = \sum_{i=1}^n [\text{dist}_2([\nabla \Phi_{\lambda\mu}(\mathbf{x})]_i, \partial \Phi_{\lambda,i}(\mathbf{x}))]^2.$$

La siguiente proposición tomada de [26], es un resultado general para cualquier tipo de sucesión.

Proposición 3.3. *Supongamos que $\mathbf{x}^* \in \mathbb{R}^n$ es un punto de acumulación aislado³ de la sucesión $\{\mathbf{x}^k\} \subseteq \mathbb{R}^n$ tal que $\{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|\}_L \rightarrow 0$ para cualquier subsucesión $\{\mathbf{x}^k\}_L$ converge a \mathbf{x}^* . Entonces la sucesión $\{\mathbf{x}^k\}$ converge a \mathbf{x}^**

Proposición 3.4. [19] *Sea $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ localmente Lipschitz y $\mathbf{x}^* \in \mathbb{R}^n$ con $G(\mathbf{x}^*) = \mathbf{0}$ tal que todas las matrices en $\partial G(\mathbf{x}^*)$ son no singulares y suponga que hay dos sucesiones $\{\mathbf{x}^k\} \subseteq \mathbb{R}^n$ y $\{\mathbf{d}^k\} \subseteq \mathbb{R}^n$ con*

$$\lim_{k \rightarrow \infty} \mathbf{x}^k = \mathbf{x}^* \quad \text{y} \quad \|\mathbf{x}^k + \mathbf{d}^k - \mathbf{x}^*\| = o(\|\mathbf{x}^k - \mathbf{x}^*\|).$$

Entonces

$$\|G(\mathbf{x}^k + \mathbf{d}^k)\| = o(\|G(\mathbf{x}^k)\|).$$

La siguiente proposición será útil posteriormente, para mostrar que el algoritmo propuesto en este trabajo de investigación, está bien definido.

Proposición 3.5. *Si \mathbf{x} y \mathbf{z} son vectores de \mathbb{R}^n tales que*

$$\|\mathbf{x} - \mathbf{z}\| \leq \alpha \|\mathbf{x}\|, \quad \alpha \in (0, 1).$$

entonces

$$\mathbf{x}^T(\mathbf{x} - \mathbf{z}) \leq \alpha \|\mathbf{x}\|^2.$$

Demostración. Usando la desigualdad de *Cauchy-Schwarz* se tiene que:

$$\mathbf{x}^T(\mathbf{x} - \mathbf{z}) \leq |\mathbf{x}^T(\mathbf{x} - \mathbf{z})| \leq \|\mathbf{x}\| \|\mathbf{x} - \mathbf{z}\| \leq \alpha \|\mathbf{x}\|^2. \quad \square$$

La siguiente proposición tomada de [19] caracteriza el *C-subdiferencial* de Φ_λ en \mathbf{x} .

Proposición 3.6. *Sea $\mathbf{x} \in \mathbb{R}^n$ arbitrario, entonces*

$$\partial_C \Phi_\lambda(\mathbf{x}) = D_a(\mathbf{x}) + D_b(\mathbf{x})F'(\mathbf{x}) \tag{3.2}$$

donde $D_a = \text{diag}(a_1(\mathbf{x}), \dots, a_n(\mathbf{x}))$, $D_b = \text{diag}(b_1(\mathbf{x}), \dots, b_n(\mathbf{x}))$ son matrices diagonales en $\mathbb{R}^{n \times n}$.

³Si Ω es el conjunto de puntos de acumulación de $\{\mathbf{x}^k\}$, se dice que $\mathbf{x}^* \in \Omega$ es un punto de acumulación aislado si existe $\delta > 0$ tal que $\mathcal{B}(\mathbf{x}^*; \delta) = \{\mathbf{x}^*\}$

- Si $(\mathbf{x}_i, F(\mathbf{x}_i)) \neq (0, 0)$ entonces

$$a_i(\mathbf{x}) = \frac{2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})}{2\sqrt{(x_i - F_i(\mathbf{x}))^2 + \lambda x_i F_i(\mathbf{x})}} - 1 := \alpha_\lambda(x_i, F_i(\mathbf{x})) - 1$$

$$b_i(\mathbf{x}) = \frac{-2(x_i - F_i(\mathbf{x})) + \lambda x_i}{2\sqrt{(x_i - F_i(\mathbf{x}))^2 + \lambda x_i F_i(\mathbf{x})}} - 1 := \beta_\lambda(x_i, F_i(\mathbf{x})) - 1.$$

- Si $(\mathbf{x}_i, F(\mathbf{x}_i)) = (0, 0)$ entonces

$$\begin{aligned} a_i(\mathbf{x}) &= \sigma_i - 1 \\ b_i(\mathbf{x}) &= \rho_i - 1, \end{aligned}$$

para cualquier $(\sigma_i, \rho_i) \in \mathbb{R}^2$ tal que $\|(\sigma_i, \rho_i)\| \leq \sqrt{c_\lambda}$, $c_\lambda \in (0, 2)$.

La demostración de este teorema se sigue de la **Proposición 2.5** en [19] y la definición de *C-subdiferencial*.

En [19], los autores demuestran la siguiente proposición que garantiza la diferenciabilidad de la función de mérito Ψ_λ y da una expresión para su gradiente.

Proposición 3.7. *La función de mérito Ψ_λ es continuamente diferenciable y*

$$\nabla \Psi_\lambda(\mathbf{x}) = V^T \Phi_\lambda(\mathbf{x}),$$

para cualquier $V \in \partial_C \Phi_\lambda(\mathbf{x})$.

El siguiente lema técnico, cuya demostración aparece en [20], se usará posteriormente para encontrar cotas que garanticen la convergencia del algoritmo que se propondrá en este trabajo.

Lema 3.1. *Sean $\mu > 0$ y $\lambda \in (0, 4)$. La función $h: (0, \infty) \rightarrow \mathbb{R}$ definida por*

$$h(t) = \frac{1}{\sqrt{t}} - \frac{1}{\sqrt{t + (4 - \lambda)\mu}},$$

es estrictamente decreciente para todo $t > 0$.

Finalizamos los preliminares con un lema en el cual, consideramos la función de mérito $\Psi_{\lambda\mu}$ que definimos en el **Capítulo 2**. A pesar de que la demostración es similar al **Lema 2.6** en [20] la incluimos aquí, pues es un resultado nuevo de nuestra investigación.

Lema 3.2. *Sea $\{\mathbf{x}^k\} \subseteq \mathbb{R}^n$ y $\{\mu_k\} \subseteq \mathbb{R}$ dos sucesiones tales que $\{\mathbf{x}^k\} \rightarrow \mathbf{x}^*$ para algún $\mathbf{x}^* \in \mathbb{R}^n$ y $\{\mu_k\} \rightarrow 0$. Entonces*

$$\lim_{k \rightarrow \infty} \nabla \Psi_{\lambda\mu_k}(\mathbf{x}^k) = \nabla \Psi_{\lambda}(\mathbf{x}^*) \quad (3.3)$$

y

$$\lim_{\mu \rightarrow 0} \Phi'_{\lambda\mu}(\mathbf{x}^k) \Phi_{\lambda}(\mathbf{x}^k) = \nabla \Psi_{\lambda}(\mathbf{x}^*). \quad (3.4)$$

Demostración. Observe que la función $\Psi_{\lambda\mu}(\mathbf{x})$ es continuamente diferenciable para todo $\mu > 0$ y se tiene que

$$\begin{aligned} \nabla \Psi_{\lambda\mu_k}(\mathbf{x}^k) &= \nabla \left(\frac{1}{2} \Phi_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_{\lambda\mu_k}(\mathbf{x}^k) \right) \\ &= \frac{1}{2} (\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_{\lambda\mu_k}(\mathbf{x}^k) + \Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_{\lambda\mu_k}(\mathbf{x}^k)) \\ &= \Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_{\lambda\mu_k}(\mathbf{x}^k) \\ &= \sum_{i=1}^n \varphi_{\lambda\mu_k}(x_i^k, F_i(\mathbf{x}^k)) [\nabla \Phi_{\lambda\mu_k}(\mathbf{x}^k)]_i, \end{aligned}$$

donde $[\nabla \Phi_{\lambda\mu_k}(\mathbf{x}^k)]_i$ denota la i -ésima fila de la matriz jacobiana de $\Phi_{\lambda\mu_k}(\mathbf{x}^k)$. Por otro lado, para un elemento arbitrario $V \in \partial_C \Phi_{\lambda}(\mathbf{x}^*)$, de la **Proposición 3.7** se sigue que

$$\nabla \Psi_{\lambda}(\mathbf{x}^*) = V^T \Phi_{\lambda}(\mathbf{x}^*) = \sum_{i=1}^n \varphi_{\lambda}(x_i^k, F_i(\mathbf{x}^k)) [V^T]_i,$$

donde $[V^T]_i$ denota la i -ésima fila de la matriz V^T . Para calcular el límite (3.3) debemos considerar dos casos, para lo cual tenemos en cuenta el conjunto $\hat{\beta}$ definido en (3.1).

Caso 1. Si $i \notin \hat{\beta}$ entonces φ_{λ} es continuamente diferenciable en $(x_i^*, F_i(\mathbf{x}^*))$, y por tanto, la i -ésima columna V^T consta de un solo elemento, $[\nabla \Phi_{\lambda}(\mathbf{x}^*)]_i$, esto se sigue de la **Proposición 3.6**.

Ahora, dado que φ_{λ} es continua y F es continuamente diferenciable en \mathbf{x}^* , tenemos lo

siguiente

$$\begin{aligned}
\lim_{k \rightarrow \infty} \nabla \Psi_{\lambda\mu_k}(\mathbf{x}^k) &= \lim_{k \rightarrow \infty} \varphi_{\lambda\mu_k}(x_i^k, F_i(\mathbf{x}^k)) [\nabla \Phi_{\lambda\mu_k}(\mathbf{x}^k)]_i \\
&= \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) \lim_{k \rightarrow \infty} [\nabla \varphi_{\lambda\mu_k}(x_i^k, F_i(\mathbf{x}^k))]_i \\
&= \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) \lim_{k \rightarrow \infty} [(\alpha_{\lambda\mu_k}(x_i, F_i(\mathbf{x}^k)) - 1)\mathbf{e}_i + (\beta_{\lambda\mu_k}(x_i^k, F_i(\mathbf{x}^k)) - 1)\nabla F_i(\mathbf{x}^k)] \\
&= \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) [(\alpha_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) - 1)\mathbf{e}_i + (\beta_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) - 1)\nabla F_i(\mathbf{x}^*)] \\
&= \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) \nabla \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) \\
&= \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) [\nabla \Phi_{\lambda}(\mathbf{x}^*)]_i \\
&= \varphi_{\lambda}(x_i^*, F_i(\mathbf{x}^*)) [V^T]_i.
\end{aligned}$$

Por tanto,

$$\lim_{k \rightarrow \infty} \nabla \Psi_{\lambda\mu_k}(\mathbf{x}^k) = \nabla \Psi_{\lambda}(\mathbf{x}^*).$$

Caso 2. Si $i \in \hat{\beta}$. Probemos que

$$\frac{\partial \varphi_{\lambda\mu}}{\partial a}(a, b) \in (-2, 0) \quad \text{y} \quad \frac{\partial \varphi_{\lambda\mu}}{\partial b}(a, b) \in (-2, 0).$$

En efecto, partimos de una desigualdad trivial y mediante algunas manipulaciones algébricas tenemos:

$$\begin{aligned}
-\lambda b^2 &< 4\mu \\
-\lambda b^2(4 - \lambda) &< 4\mu(4 - \lambda) \\
4(a - b)^2 + 4\lambda ab - \lambda b^2(4 - \lambda) &< 4((a - b)^2 + \lambda ab + \mu(4 - \lambda)) \\
(2(a - b) + \lambda b)^2 &< 4G_{\lambda\mu}^2(a, b) \\
|2(a - b) + \lambda b| &< 2G_{\lambda\mu}(a, b),
\end{aligned}$$

donde $G_{\lambda\mu}(x, y)$ esta dada en (2.4) y así,

$$-2 < \frac{2(a-b) + \lambda b}{2G_{\lambda\mu}(a,b)} - 1 < 0$$

$$-2 < \frac{\partial\varphi(a,b)}{\partial a} < 0,$$

para todo (a, b) en \mathbb{R}^2 , $\lambda \in (0, 4)$ y $\mu > 0$. De forma análoga se prueba que la derivada parcial de $\varphi_{\lambda\mu}$ respecto a la segunda variable pertenece al intervalo $(-2, 0)$. Luego, la sucesión $\{[\nabla\Phi_{\lambda\mu_k}(\mathbf{x}^k)]_i\}$ es acotada para $k \rightarrow \infty$. Por tanto,

$$\lim_{k \rightarrow \infty} \varphi_{\lambda\mu_k}(x_i^k, F_i(\mathbf{x}^k)) = \varphi_\lambda(x_i^*, F_i(\mathbf{x}^*)) = 0,$$

Con lo cual,

$$\lim_{k \rightarrow \infty} \varphi_{\lambda\mu_k}(x_i^k, F_i(\mathbf{x}^k)) [\nabla\Phi_{\lambda\mu_k}(\mathbf{x}^k)]_i = 0.$$

También tenemos que $\varphi_\lambda(x_i^*, F_i(\mathbf{x}^*))V_i^T = 0$, para todo $i \in \hat{\beta}$, luego (3.3) se sigue de los **casos 1** y **2**. Para demostrar (3.4) multiplicamos (3.3) por $\Psi_\lambda(\mathbf{x}^k)$ y repetimos un procedimiento similar al anterior. \square

A continuación mostramos un lema técnico, el cual se utilizará en el análisis de la convergencia global

Lema 3.3. [20] *Sea $\{\mathbf{x}^k\}, \{\mathbf{d}^k\} \subseteq \mathbb{R}^n$ y $\{t_k\} \subseteq \mathbb{R}$ sucesiones con $\mathbf{x}^{k+1} := \mathbf{x}^k + t_k\mathbf{d}^k$ tal que $\mathbf{x}^k \rightarrow \mathbf{x}^*$, $\mathbf{d}^k \rightarrow \mathbf{d}^*$ y $\{t_k\} \rightarrow 0$ para ciertos vectores $\mathbf{x}^*, \mathbf{d}^* \in \mathbb{R}^n$. Además $\{\mu_k\} \subseteq \mathbb{R}$ una sucesión tal que $\{\mu_k\} \rightarrow 0$. Entonces*

$$\lim_{k \rightarrow \infty} \frac{\Psi_{\lambda\mu_k}(\mathbf{x}^k + t_k\mathbf{d}^k) - \Psi_{\lambda\mu_k}(\mathbf{x}^k)}{t_k} = \nabla\Psi_\lambda(\mathbf{x}^*)^T \mathbf{d}^*.$$

El resultado que sigue, demostrado en [10], se usará posteriormente para probar la convergencia del algoritmo propuesto.

Lema 3.4. *Si existe un escalar*

$$\omega \in \left[\frac{1}{2} - \frac{(1 - \alpha - 2\sigma)^2}{2(2 + \alpha)^2}, \frac{1}{2} \right]$$

tal que

$$\Psi_\lambda(\mathbf{y}) \leq \Psi_\lambda(\mathbf{x}^k) - 2\omega\Psi_\lambda(\mathbf{x}^k), \quad (3.5)$$

para algún $k \in K$ y $\mathbf{y} \in \mathbb{R}^n$ entonces se tiene

$$\Psi_{\lambda\mu_k}(\mathbf{y}) \leq \Psi_{\lambda\mu_k}(\mathbf{x}^k) - 2\sigma\Psi_\lambda(\mathbf{x}^k), \quad (3.6)$$

donde μ_k es el parámetro de suavización en el k -ésimo paso.

Capítulo 4

Algoritmo propuesto

En este capítulo proponemos un nuevo algoritmo para resolver el problema de complementariedad no lineal. Básicamente, la propuesta es una generalización del método de jacobiano suavizado propuesto en [20], y su iteración básica está dada por (2.6).

Con el fin de garantizar que el algoritmo esté bien definido y sea aplicable a cualquier problema de complementariedad no lineal, se introduce un paso del gradiente de la función de mérito Ψ_λ cuando el sistema lineal (3.7) no tiene solución o esta sea una dirección de poco decrecimiento de Ψ_λ . Con la introducción de este paso cobra importancia la **Proposición 3.7** que garantiza la diferenciabilidad de Ψ_λ y da una expresión para su gradiente.

El algoritmo propuesto es el siguiente.

Algoritmo 4.1. (*Método con jacobiano suavizado*)

P0. Valores iniciales.

$\mathbf{x}^0 \in \mathbb{R}^n$; θ, α, η y $\rho \in (0, 1)$, $\gamma > 0$, $\sigma \in (0, \frac{1-\alpha}{2})$; $p > 2$ y $\epsilon \geq 0$. Sean $\beta_0 = \|\Phi_\lambda(\mathbf{x}^0)\|$, $\kappa = \sqrt{(4-\lambda)\eta}$, $\mu_0 = (\frac{\alpha}{2\kappa}\beta_0)^2$ y $k = 0$.

P1. Criterio de parada.

Si $\|\nabla\Psi_\lambda(\mathbf{x}^k)\| \leq \epsilon$ entonces el algoritmo se detiene.

P2. Cálculo de la dirección.

Encontrar $\mathbf{d}^k \in \mathbb{R}^n$ resolviendo el sistema de ecuaciones lineales,

$$\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\mathbf{d}^k = -\Phi_\lambda(\mathbf{x}^k). \quad (4.1)$$

Si (4.1) no tiene solución o no se cumple la condición

$$\Phi_\lambda(\mathbf{x}^k)^T \Phi'_{\lambda\mu_k}(\mathbf{x}^k)\mathbf{d}^k \leq -\rho\|\mathbf{d}^k\|^p \quad (4.2)$$

entonces

$$\mathbf{d}^k = -\nabla\Psi_\lambda(\mathbf{x}^k). \quad (4.3)$$

P3. Búsqueda lineal.

Encuentre el menor $m_k \in \{0, 1, 2, \dots\}$ tal que

$$\Psi_{\lambda\mu_k}(\mathbf{x}^k + \theta^{m_k}\mathbf{d}^k) \leq \Psi_{\lambda\mu_k}(\mathbf{x}^k) - 2\sigma\theta^{m_k}\Psi_\lambda(\mathbf{x}^k) \quad (4.4)$$

siempre que \mathbf{d}^k este dado por (4.1) y tal que

$$\Psi_\lambda(\mathbf{x}^k + \theta^{m_k}\mathbf{d}^k) \leq \Psi_\lambda(\mathbf{x}^k) - \sigma\theta^{m_k}\|\mathbf{d}^k\|^2 \quad (4.5)$$

siempre que \mathbf{d}^k este dado por (4.3).

Defina

$$t_k = \theta^{m_k} \quad \text{y} \quad \mathbf{x}^{k+1} = \mathbf{x}^k + t_k\mathbf{d}^k.$$

P4. Actualización de parámetros.

Si

$$\|\Phi_\lambda(\mathbf{x}^{k+1})\| \leq \max \left\{ \eta\beta_k, \frac{1}{\alpha}\|\Phi_\lambda(\mathbf{x}^{k+1}) - \Phi_{\lambda\mu_k}(\mathbf{x}^{k+1})\| \right\} \quad (4.6)$$

entonces

$$\beta_{k+1} = \|\Phi_\lambda(\mathbf{x}^{k+1})\|$$

y escoja μ_{k+1} tal que

$$0 < \mu_{k+1} \leq \min \left\{ \left(\frac{\alpha}{2\kappa}\beta_{k+1} \right)^2, \frac{\mu_k}{4}, \bar{\mu}(\mathbf{x}^{k+1}, \gamma\beta_{k+1}) \right\}. \quad (4.7)$$

Si (4.6) no se cumple y $\mathbf{d}_k = -\nabla\Psi_\lambda(\mathbf{x}^k)$ entonces

$$\beta_{k+1} = \beta_k.$$

y escoja μ_{k+1} tal que

$$0 < \mu_{k+1} \leq \min \left\{ \left(\frac{\alpha}{2\kappa} \|\Phi_\lambda(\mathbf{x}^{k+1})\| \right)^2, \left(\frac{\|\Phi_\lambda(\mathbf{x}^k)\| - \|\Phi_\lambda(\mathbf{x}^{k+1})\|}{2\kappa} \right)^2, \frac{\mu_k}{4} \right\}. \quad (4.8)$$

Si ninguna de las condiciones anteriores se satisface entonces

$$\beta_{k+1} = \beta_k, \quad y \quad \mu_{k+1} = \mu_k.$$

P5. Actualización

Actualizar $k = k + 1$ y regresar a **P1**.

Antes de iniciar las pruebas de convergencia y demás detalles técnicos, se hacen algunos comentarios sobre los pasos del algoritmo. En los valores iniciales (**P0**) se introducen los parámetros constantes e inicializamos las variables. En el criterio de parada (**P1**) es natural pensar que el algoritmo se detenga cuando el gradiente de la función de mérito se haga muy pequeño. Sin embargo en la implementación se adicionarán otros criterios clásicos de parada, como el máximo número de iteraciones permitidas y un criterio de aceptación de la dirección para evitar que el algoritmo se estanque en la búsqueda de un tamaño de paso adecuado.

El cálculo de la dirección de descenso en (**P2**) es quizá el paso central en el algoritmo. Aquí se encuentra la dirección \mathbf{d}_k mediante la ecuación mixta de *Newton* (4.1) y se introducen condiciones de tal forma que se garantice encontrar una dirección de descenso. En caso de que el sistema (4.1) no tenga solución o que la dirección no cumpla el criterio de descenso (4.2), se utiliza la dirección menos el gradiente de la función de mérito (4.3), con ello, además de garantizar una dirección de decrecimiento de Ψ_λ también le da el carácter global al algoritmo.

Una vez encontrada la dirección de descenso se procede a la globalización del algoritmo. Para ello, la estrategia usada es una búsqueda lineal (**P3**), que claramente dependerá de la dirección encontrada en el paso anterior. Si la dirección de descenso se obtiene con la ecuación de *Newton* (4.1) entonces se hace la búsqueda lineal mediante la condición (4.4), la cual es usada por [10] como una estrategia de globalización. Por otra parte, si

al dirección de descenso es la del gradiente (4.3) entonces la búsqueda lineal dada por (4.5), es de tipo *Armijo* [13].

La actualización de parámetros (**P4**) más precisamente de μ_k , se inicia con la condición (4.6) utilizada en [10]. Si esta es satisfecha entonces se actualiza el parámetro μ_k por (4.7). Esta condición garantiza primero, que la distancia entre el *C-subdiferencial* y el jacobiano suavizado es pequeña, con el fin de asegurar convergencia a la solución; segundo, hacer que μ_k tienda a cero. Por otra parte, si (4.6) no se cumple entonces se actualiza μ_k mediante (4.8). Así mismo las condiciones (4.8) y (4.6) son esenciales para garantizar que el algoritmo esta bien definido y su convergencia global.

Para el análisis posterior de la convergencia del algoritmo es necesario definir el siguiente conjunto, el cual está motivado en la condición (4.6),

$$K = \{0\} \cup \left\{ k \in \mathbb{N} : \|\Phi_\lambda(\mathbf{x}^k)\| \leq \max \left\{ \eta \beta_{k-1}, \frac{1}{\alpha} \|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_{k-1}}(\mathbf{x}^k)\| \right\} \right\}. \quad (4.9)$$

Este conjunto lo podemos expresar como la unión de tres conjuntos disjuntos

$$K = \{0\} \cup K_1 \cup K_2,$$

donde

$$\begin{aligned} K_1 &= \left\{ k \in K : \eta \beta_{k-1} \geq \frac{1}{\alpha} \|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_{k-1}}(\mathbf{x}^k)\| \right\} \\ K_2 &= \left\{ k \in K : \eta \beta_{k-1} < \frac{1}{\alpha} \|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_{k-1}}(\mathbf{x}^k)\| \right\}. \end{aligned}$$

La siguiente proposición será de gran utilidad para probar que nuestro algoritmo está bien definido, su demostración es similar a la dada en [20].

Proposición 4.1. *Las siguientes afirmaciones son válidas*

a) Para todo $k \geq 0$,

$$\|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_k}(\mathbf{x}^k)\| \leq \alpha \|\Phi_\lambda(\mathbf{x}^k)\|, \quad (4.10)$$

b) Para todo $k \in K$ con $k \geq 1$,

$$\text{dist}_F(\Phi'_{\lambda\mu_k}(\mathbf{x}^k), \partial_C \Phi_\lambda(\mathbf{x}^k)) \leq \gamma \|\Phi_\lambda(\mathbf{x}^k)\| \quad (4.11)$$

A continuación se demuestra que el algoritmo propuesto está bien definido, es decir termina en un número finito de pasos.

Proposición 4.2. *El Algoritmo 4.1 está bien definido.*

Demostración. La prueba es esencialmente la seguida en [10]. Para demostrar que el algoritmo esté bien definido se reduce a probar en **P3**, que m_k es finito para todo $k \in \mathbb{N}$, más precisamente en las ecuaciones (4.4) y (4.5). En efecto, si la dirección de descenso esta dada por (4.3) entonces la condición tipo *Armijo* garantiza que m_k es finito.

Por otra parte, supongamos que la dirección es dada por (4.1). Tendiendo en cuenta que $\Psi_{\lambda\mu_k}$ es continuamente diferenciable y su gradiente es dado por

$$\nabla\Psi_{\lambda\mu_k}(\mathbf{x}^k) = \Phi'_{\lambda\mu_k}(\mathbf{x}^k)\Phi_{\lambda\mu}(\mathbf{x}^k),$$

Por el Teorema de *Taylor* tenemos que,

$$\begin{aligned} \Psi_{\lambda\mu_k}(\mathbf{x}^{k+1}) - \Psi_{\lambda\mu_k}(\mathbf{x}^k) &= t\nabla\Psi_{\lambda\mu_k}(\mathbf{x}^k)^T \mathbf{d}^k + o(t) \\ &= t(\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\Phi_{\lambda\mu}(\mathbf{x}^k))^T \mathbf{d}^k + o(t) \\ &= t\Phi_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi'_{\lambda\mu_k}(\mathbf{x}^k) \mathbf{d}^k + o(t). \end{aligned} \quad (4.12)$$

Usando la dirección de *Newton* (4.1) en 4.12,

$$\begin{aligned} t\Phi_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi'_{\lambda\mu_k}(\mathbf{x}^k) \mathbf{d}^k + o(t) &= -t\Phi_{\lambda}(\mathbf{x}^k)^T \Phi_{\lambda\mu_k}(\mathbf{x}^k) + o(t) \\ &= -t\Phi_{\lambda}(\mathbf{x}^k)^T (\Phi_{\lambda\mu_k}(\mathbf{x}^k) - \Phi_{\lambda}(\mathbf{x}^k) + \Phi_{\lambda}(\mathbf{x}^k)) + o(t) \\ &= -t\Phi_{\lambda}(\mathbf{x}^k)^T \Phi_{\lambda}(\mathbf{x}^k) + t\Phi_{\lambda}(\mathbf{x}^k)^T (\Phi_{\lambda\mu_k}(\mathbf{x}^k) - \Phi_{\lambda}(\mathbf{x}^k)) + o(t) \\ &\leq -2t\Psi_{\lambda}(\mathbf{x}^k) + 2t\alpha\Psi_{\lambda}(\mathbf{x}^k) + o(t) \\ &= -2t(1 - \alpha)\Psi_{\lambda}(\mathbf{x}^k) + o(t), \end{aligned} \quad (4.13)$$

donde la última desigualdad se deduce de las **Proposiciones 3.5** y **4.1**. Por otro lado, se tiene que

$$\sigma < \frac{1}{2}(1 - \alpha) < 1 - \alpha. \quad (4.14)$$

Por tanto, de (4.12), (4.13) y (4.14),

$$\Psi_{\lambda\mu_k}(\mathbf{x}^{k+1}) - \Psi_{\lambda\mu_k}(\mathbf{x}^k) \leq -2t\sigma\Psi_{\lambda}(\mathbf{x}^k) + o(t).$$

Esto permite concluir que existe un entero no negativo m_k tal que (4.4) es satisfecha. \square

Convergencia global

En este capítulo presentamos los resultados de convergencia del algoritmo propuesto, básicamente se muestra que todo punto de acumulación de la sucesión generada por el **Algoritmo 4.1** es un punto estacionario de la función Ψ_λ . Estos resultados generalizan a los presentados en [20] para el algoritmo que usa la función de *Fischer-Burmeister*, los cuales a su vez se basan en la teoría desarrollada por [10] para métodos tipo *Newton* suavizados.

Iniciamos el estudio de la teoría de convergencia con dos lemas cuyas demostraciones se basa en las reglas de actualización del **Algoritmo 4.1**.

Lema 5.1. *Supongamos que una sucesión generada por el **Algoritmo 4.1** tiene un punto de acumulación \mathbf{x}^* , el cual es solución del PCNL(F). Entonces el conjunto de índices K definido en (4.9) es infinito y la sucesión $\{\mu_k\}$ converge a cero.*

Demostración. La demostración se sigue por contradicción suponiendo que el conjunto K es finito, de esto y la regla de actualización (4.6), existe un entero k_0 , tal que $\beta_k = \beta_{k_0}$, además

$$\|\Phi_\lambda(\mathbf{x}^{k+1})\| > \eta\beta_{k_0} \quad \text{para todo } k \geq k_0.$$

lo anterior contradice el hecho de que \mathbf{x}^* sea solución del PCNL(F). \square

Lema 5.2. *Las siguientes afirmaciones se cumplen*

a) Si \mathbf{d}^k es dado por (4.1) entonces

$$\|\Phi_{\lambda\mu_k}(\mathbf{x}^{k+1})\| < \|\Phi_{\lambda\mu_k}(\mathbf{x}^k)\|.$$

b) Si $\mathbf{d}^k = -\nabla\Psi_\lambda(\mathbf{x}^k)$ y μ_k es actualizado por (4.8), entonces

$$\|\Phi_{\lambda\mu_{k+1}}(\mathbf{x}^{k+1})\| \leq \|\Phi_{\lambda\mu_{k+1}}(\mathbf{x}^k)\|.$$

Demostración. La parte (a) se verifica utilizando la regla de actualización (4.4). La parte (b) se cumple si (4.6) es satisfecha, además se hace uso del **Corolario 2.2** y algunas manipulaciones algebraicas tales como sumar y restar términos adecuados. \square

El siguiente corolario es una consecuencia importante del resultado anterior

Corolario 5.1. Si $k \notin K$, entonces

$$\|\Phi_{\lambda\mu_k}(\mathbf{x}^k)\| \leq \|\Phi_{\lambda\mu_k}(\mathbf{x}^{k-1})\|.$$

Los dos resultados siguientes son lemas técnicos que proporcionan algunas cotas que usaremos en la demostración de la **Proposición 5.1**.

Lema 5.3. Suponga que el conjunto K definido en (4.9) está ordenado de la siguiente manera $k_0 = 0 < k_1 < k_2 < \dots$. Sean $k \in \mathbb{N}$, arbitrario pero fijo y k_j el entero más grande en K tales que $k_j \leq k$ entonces

$$\|\Phi_\lambda(\mathbf{x}^k)\| \leq \beta_{k_j} + 2\kappa\sqrt{\mu_{k_j}}. \quad (5.1)$$

Demostración. Por la hipótesis del lema y usando reglas de actualización de los parámetros dadas en el paso **P4** del algoritmo, tenemos las siguientes desigualdades:

$$\begin{aligned} \mu_k &\leq \mu_{k_j} \\ \beta_k &= \beta_{k_j}. \end{aligned} \quad (5.2)$$

Si $k_j = k$, entonces de la condición (4.6) en el paso **P4** del algoritmo se tiene que $\beta_{k_j} = \|\Phi_\lambda(\mathbf{x}^{k_j})\|$ y trivialmente se cumple que

$$\|\Phi_\lambda(\mathbf{x}^k)\| \leq \beta_{k_j} + 2\kappa\sqrt{\mu_{k_j}}.$$

Si $k_j < k$ es decir, $k \notin K$ y por **Corolario 5.1** tenemos que

$$\|\Phi_{\lambda\mu_l}(\mathbf{x}^l)\| \leq \|\Phi_{\lambda\mu_l}(\mathbf{x}^{l-1})\|,$$

para todo, $k_j < l < k_{j+1}$. Dado $k < k_{j+1}$ se tiene que

$$\|\Phi_{\lambda\mu_l}(\mathbf{x}^l)\| \leq \|\Phi_{\lambda\mu_l}(\mathbf{x}^{l-1})\|,$$

para todo, $k_j < l \leq k$, equivalentemente

$$\|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^{l+1})\| \leq \|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^l)\|,$$

para todo l tal que $k_j \leq l \leq k-1$, Además, para estos índices l , por el **Corolario 5.1**, tenemos que,

$$\|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^{l+1})\| + \kappa\sqrt{\mu_{l+1}} \leq \|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^l)\| + \kappa\sqrt{\mu_{l+1}}.$$

Mediante operaciones algebraicas y utilizando la **Proposición 2.2**, tenemos

$$\begin{aligned} \|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^l)\| + \kappa\sqrt{\mu_{l+1}} &\leq \|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^l) - \Phi_{\lambda\mu_l}(\mathbf{x}^l) + \Phi_{\lambda\mu_l}(\mathbf{x}^l)\| + \kappa\sqrt{\mu_{l+1}} \\ &\leq \|\Phi_{\lambda\mu_{l+1}}(\mathbf{x}^l) - \Phi_{\lambda\mu_l}(\mathbf{x}^l)\| + \|\Phi_{\lambda\mu_l}(\mathbf{x}^l)\| + \kappa\sqrt{\mu_{l+1}} \\ &\leq \|\Phi_{\lambda\mu_l}(\mathbf{x}^l)\| + \kappa|\sqrt{\mu_{l+1}} - \sqrt{\mu_l}| + \kappa\sqrt{\mu_{l+1}} \\ &\leq \|\Phi_{\lambda\mu_l}(\mathbf{x}^l)\| + \kappa(\sqrt{\mu_l} - \sqrt{\mu_{l+1}}) + \kappa\sqrt{\mu_{l+1}} \\ &\leq \|\Phi_{\lambda\mu_l}(\mathbf{x}^l)\| + \kappa\sqrt{\mu_l}. \end{aligned}$$

Usando repetidamente la desigualdad anterior y por el **Corolario 2.2**, se sigue que

$$\begin{aligned}
\|\Phi_\lambda(\mathbf{x}^k)\| &\leq \|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_k}(\mathbf{x}^k)\| + \|\Phi_{\lambda\mu_k}(\mathbf{x}^k)\| \\
&\leq \|\Phi_{\lambda\mu_k}(\mathbf{x}^k)\| + \kappa\sqrt{\mu_k} \\
&\leq \|\Phi_{\lambda\mu_{k-1}}(\mathbf{x}^{k-1})\| + \kappa\sqrt{\mu_{k-1}} \\
&\quad \vdots \\
&\leq \|\Phi_{\lambda\mu_{k_j}}(\mathbf{x}^{k_j})\| + \kappa\sqrt{\mu_{k_j}} \\
&\leq \|\Phi_\lambda(\mathbf{x}^{k_j})\| + \|\Phi_{\lambda\mu_{k_j}}(\mathbf{x}^{k_j}) - \Phi_\lambda(\mathbf{x}^{k_j})\| + \kappa\sqrt{\mu_{k_j}} \\
&\leq \|\Phi_\lambda(\mathbf{x}^{k_j})\| + \kappa\sqrt{\mu_{k_j}} + \kappa\sqrt{\mu_{k_j}} \\
&= \beta_{k_j} + 2\kappa\sqrt{\mu_{k_j}},
\end{aligned}$$

lo cual, prueba la desigualdad (5.1), para todo $k \in \mathbb{N}$. □

Lema 5.4. *Suponga que el conjunto K definido en (4.9) está ordenado de la siguiente manera $k_0 = 0 < k_1 < k_2 < \dots$. Sean $k \in \mathbb{N}$, arbitrario pero fijo y k_j el entero más grande en K tales que $k_j \leq k$ entonces*

$$\sqrt{\mu_{k_j}} \leq \frac{1}{2^{j+1}} \frac{\alpha}{\kappa} \|\Phi_\lambda(\mathbf{x}^0)\| \quad (5.3)$$

y

$$\beta_{k_j} \leq r^j \|\Phi_\lambda(\mathbf{x}^0)\|, \quad (5.4)$$

donde

$$r = \max \left\{ \frac{1}{2}, \eta \right\}. \quad (5.5)$$

Demostración. Para $j = 0$, tenemos que $k_0 = 0$. Usando la definición de los parámetros iniciales μ_0 y β_0 , siguen las siguientes igualdades,

$$\sqrt{\mu_{k_0}} = \sqrt{\mu_0} = \frac{\alpha}{2\kappa} \|\Phi_\lambda(\mathbf{x}^0)\|$$

$$\beta_{k_0} = \beta_0 = r^0 \|\Phi_\lambda(\mathbf{x}^0)\|,$$

Supongamos que, $j \geq 1$ entonces del paso **P.4.** del algoritmo y por (5.2),

$$\beta_{k_j} \leq \beta_{k_{j-1}} = \eta\beta_{k_{j-1}} \leq r\beta_{k_{j-1}}, \quad (5.6)$$

para todo $k_j \in K_1$.

De (5.6) y nuevamente, por el **Corolario 2.2**,

$$\begin{aligned}
\beta_{k_j} &\leq \frac{1}{\alpha} \|\Phi_\lambda(\mathbf{x}^{k_j}) - \Phi_{\lambda\mu_{k_j-1}}(\mathbf{x}^{k_j})\| \\
&\leq \frac{1}{\alpha} \kappa \sqrt{\mu_{k_j-1}} \leq \frac{1}{\alpha} \kappa \sqrt{\mu_{k_j-1}} \\
&\leq \frac{1}{\alpha} \kappa \sqrt{\left(\frac{\alpha}{2\kappa} \|\Phi_\lambda(\mathbf{x}^{k_j})\|\right)^2} = \frac{1}{2} \|\Phi_\lambda(\mathbf{x}^{k_{j-1}})\| = \frac{1}{2} \beta_{k_{j-1}} \\
&\leq r \beta_{k_{j-1}},
\end{aligned}$$

para todo $k_j \in K_2$. En forma análoga, obtenemos que

$$\mu_{k_j} \leq \mu_{k_{j-1}} \leq \frac{1}{4} \mu_{k_{j-1}}. \quad (5.7)$$

Usando la definición de μ_0 y β_0 tenemos que

$$\sqrt{\mu_{k_j}} \leq \frac{1}{2^j} \sqrt{\mu_0} \leq \frac{1}{2^{j+1}} \frac{\alpha}{\kappa} \|\Phi_\lambda(\mathbf{x}^0)\| \quad (5.8)$$

y

$$\beta_{k_j} \leq r^j \beta_0 = r^j \|\Phi_\lambda(\mathbf{x}^0)\|, \quad (5.9)$$

lo cual, completa la demostración. \square

Usando los resultados anteriores demostraremos, que las iteraciones \mathbf{x}_k permanecen en un mismo conjunto de nivel. Esto es importante porque estamos minimizando diferentes funciones de mérito y el decrecimiento de una, no implica el decrecimiento de la otra. Dicho conjunto puede hacerse arbitrariamente cercano al conjunto de nivel $\Psi_\lambda(\mathbf{x}^0)$.

Proposición 5.1. *La sucesión generada por el **Algoritmo 4.1** permanece en el conjunto de nivel*

$$L_0 = \{\mathbf{x} \in \mathbb{R}^n : \Psi_\lambda(\mathbf{x}) \leq (1 + \alpha)^2 \Psi_\lambda(\mathbf{x}^0)\}. \quad (5.10)$$

Demostración. Supongamos sin pérdida de generalidad, que el conjunto K definido en (4.9) está ordenado de la siguiente forma $k_0 = 0 < k_1 < k_2 < \dots$, lo cual, no indica que el conjunto sea infinito. Consideremos $k \in \mathbb{N}$, arbitrario pero fijo y k_j el entero más grande en K de tal forma que $k_j \leq k$.

Usando los **Lemas 5.3 y 5.4** tenemos que,

$$\|\Phi_\lambda(\mathbf{x}^k)\| \leq \beta_{k_j} + 2\kappa\sqrt{\mu_{k_j}} \leq r^j \|\Phi_\lambda(\mathbf{x}^0)\| + \frac{\alpha}{2^j} \|\Phi_\lambda(\mathbf{x}^0)\|,$$

donde r está definido por (5.5). Es decir que,

$$\|\Phi_\lambda(\mathbf{x}^k)\| \leq r^j(1 + \alpha)\|\Phi_\lambda(\mathbf{x}^0)\| \quad (5.11)$$

luego,

$$\|\Phi_\lambda(\mathbf{x}^k)\| \leq (1 + \alpha)\|\Phi_\lambda(\mathbf{x}^0)\|,$$

Por lo tanto, $\mathbf{x}^k \in L_0$. □

Una consecuencia de la desigualdad (5.11) es la siguiente proposición.

Proposición 5.2. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1** y suponga que el conjunto K es infinito. Entonces todo punto de acumulación de la sucesión $\{\mathbf{x}^k\}$ es una solución del PCNL(F).*

Proposición 5.3. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1** y sea $\{\mathbf{x}^k\}_L$ una subsucesión que converge a un punto $\mathbf{x}^* \in \mathbb{R}^n$. Si $\mathbf{d}^k = -\nabla\Psi_\lambda(\mathbf{x}^k)$ para todo $k \in L$, entonces \mathbf{x}^* es un punto estacionario de Ψ_λ .*

Demostración. Inicialmente consideramos la cantidad de elementos del conjunto de índices K . Si este conjunto es infinito, por la **Proposición 5.2** se concluye que el punto de acumulación \mathbf{x}^* es solución del PCNL(F) y así, este sería un mínimo global y por tanto, un punto estacionario de la función Ψ_λ .

Si K es finito, sin pérdida de generalidad podemos suponer que $K \cap L = \emptyset$ puesto que la sucesión tiene una cantidad infinita de elementos. Por lo anterior, se tiene que regla de actualización (2.9) se satisface para todo $k \in L$, y con ello, la sucesión $\{\mu_k\}$ converge a cero.

Dado que K es finito, existe el elemento más grande en dicho conjunto, que llamaremos \hat{k} . Usando las reglas de actualización definidas en el paso **P4** del **Algoritmo 4.1**, tenemos las siguientes desigualdades,

$$\mu_k \leq \mu_{\hat{k}}, \quad \beta_k = \beta_{\hat{k}} = \|\Phi_\lambda(\mathbf{x}^{\hat{k}})\|, \quad (5.12)$$

$$\|\Phi_\lambda(\mathbf{x}^k)\| > \eta \beta_{k-1} = \eta \|\Phi_\lambda(\mathbf{x}^{\hat{k}})\| > 0. \quad (5.13)$$

$$\alpha \|\Phi_\lambda(\mathbf{x}^k)\| > \|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_{k-1}}(\mathbf{x}^k)\|. \quad (5.14)$$

Ahora de (5.13), tenemos que

$$\Psi_\lambda(\mathbf{x}^k) > \eta^2 \Psi_\lambda(\mathbf{x}^{\hat{k}}) > 0 \quad (5.15)$$

para todo $k > \hat{k}$.

Supongamos, por contradicción que \mathbf{x}^* es un punto estacionario de Ψ_λ . Es decir, $\nabla \Psi_\lambda(\mathbf{x}^*) \neq 0$. En primer lugar, demostraremos que

$$\liminf_{k \in L} t_k = 0.$$

Para lo cual, supongamos que $\liminf_{k \in L} t_k = t^* > 0$. Dado que $d^k = -\nabla \Psi_\lambda(\mathbf{x}^k)$ para todo $k \in L$, obtenemos por la regla de *Armijo* (4.5) que,

$$\Psi_\lambda(\mathbf{x}^{k+1}) - \Psi_\lambda(\mathbf{x}^k) \leq -\sigma t_k \|\nabla \Psi_\lambda(\mathbf{x}^k)\|^2 \leq -\frac{c}{2}, \quad (5.16)$$

para todo $k \in L$ suficientemente grande, donde $c := \sigma t_* \|\nabla \Psi_\lambda(\mathbf{x}^k)\|^2 > 0$.

Dado que $\{\mu_k\}$ converge a cero, el **Corolario 2.2** garantiza que

$$|\Psi_{\lambda\mu_k}(\mathbf{x}^{k+1}) - \Psi_\lambda(\mathbf{x}^{k+1})| \leq \frac{c}{4}$$

y

$$|\Psi_{\lambda\mu_k}(\mathbf{x}^k) - \Psi_\lambda(\mathbf{x}^k)| \leq \frac{c}{4},$$

para todo $k \in \mathbb{N}$, suficientemente grande.

Usando nuevamente, el hecho que la sucesión $\{\mu_k\}$ converge a cero; teniendo en cuenta que la sucesión $\{\|\Phi_\lambda(\mathbf{x}^k)\|\}$ es acotada; recordando que K es finito entonces por la **Proposición 5.1**, tenemos que

$$2\kappa \sqrt{\mu_k} \|\Phi_\lambda(\mathbf{x}^k)\| + 2\kappa^2 \mu_k \leq \frac{c}{4}, \quad (5.17)$$

para todo $k \in \mathbb{N}$, suficientemente grande.

Supongamos ahora, que $L = \{l_0, l_1, \dots\}$, entonces para todo l_j , suficientemente grande, tenemos en forma análoga a la **Proposición 5.1** y usando (5.17) que,

$$\begin{aligned}
\Psi_\lambda(\mathbf{x}^{l_{j+1}}) &= \frac{1}{2} \|\Phi_\lambda(\mathbf{x}^{l_{j+1}})\| \\
&\leq \frac{1}{2} (\|\Phi_\lambda(\mathbf{x}^{l_{j+1}})\| + 2\kappa\sqrt{\mu_{l_{j+1}}})^2 \\
&= \Psi_\lambda(\mathbf{x}^{l_{j+1}}) + 2\kappa\sqrt{\mu_{l_{j+1}}} \|\Phi_\lambda(\mathbf{x}^{l_{j+1}})\| + 2\kappa^2\mu_{l_{j+1}} \\
&\leq \Psi_\lambda(\mathbf{x}^{l_{j+1}}) + \frac{c}{4}.
\end{aligned} \tag{5.18}$$

Por (5.16) y (5.18), obtenemos que

$$\Psi_\lambda(\mathbf{x}^{l_{j+1}}) - \Psi_\lambda(\mathbf{x}^{l_j}) = [\Psi_\lambda(\mathbf{x}^{l_{j+1}}) - \Psi_\lambda(\mathbf{x}^{l_{j+1}})] + [\Psi_\lambda(\mathbf{x}^{l_{j+1}}) - \Psi_\lambda(\mathbf{x}^{l_j})] \leq -\frac{c}{4},$$

para todo l_j , suficientemente grande. Luego, la sucesión $\{\Psi_\lambda(\mathbf{x}^{l_j})\} \rightarrow -\infty$ cuando $j \rightarrow \infty$, lo que contradice el hecho de que $\Psi_\lambda(\mathbf{x}) \geq 0$ para todo $\mathbf{x} \in \mathbb{R}^n$. Así,

$$\liminf_{k \in L} t_k = 0.$$

Si es necesario, a partir de una subsucesión, podemos suponer que $\lim_{k \in L} t_k = 0$.

Debido a que $\lim_{k \in L} t_k = 0$, un tamaño del paso completo nunca es aceptado para todo $k \in L$, suficientemente grande. Así, a partir de la regla de *Armijo* (4.5) obtenemos

$$\Psi_\lambda(\mathbf{x}^k + \lambda^{m_k-1} \mathbf{d}^k) > \Psi_\lambda(\mathbf{x}^k) - \sigma \theta^{m_k-1} \|\mathbf{d}^k\|^2$$

o equivalentemente,

$$\frac{\Psi_\lambda(\mathbf{x}^k + \lambda^{m_k-1} \mathbf{d}^k) - \Psi_\lambda(\mathbf{x}^k)}{\theta^{m_k-1}} > -\sigma \|\mathbf{d}^k\|^2. \tag{5.19}$$

Tomando el límite cuando $k \rightarrow \infty$ sobre L , De (5.19) por la diferenciabilidad de la función Ψ_λ , $\mathbf{d}^k = -\nabla \Psi_\lambda(\mathbf{x}^k)$ para todo $k \in L$ y el hecho que $\theta^{m_k-1} \rightarrow 0$ cuando $k \rightarrow \infty$, en L , obtenemos

$$-\nabla \Psi_\lambda(\mathbf{x}^*)^T \nabla \Psi_\lambda(\mathbf{x}^*) \geq -\sigma \nabla \Psi_\lambda(\mathbf{x}^*)^T \nabla \Psi_\lambda(\mathbf{x}^*).$$

Esto implica que $\sigma \geq 1$, lo cual, es claramente una contradicción con la elección inicial de σ . Por tanto,

$$\nabla \Psi_\lambda(\mathbf{x}^*) = 0. \quad \square$$

Los dos lemas siguientes son resultados útiles para la demostración del teorema principal de convergencia, **Teorema 5.1**, relacionados con la dirección de *Newton*. En ambos resultados se considera el caso particular en el que K es un conjunto finito y una subsucesión de una sucesión generada por el algoritmo propuesto.

Lema 5.5. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1** y sea $\{\mathbf{x}^k\}_L$ una subsucesión que converge a un punto $\mathbf{x}^* \in \mathbb{R}^n$. Si \mathbf{d}^k es la dirección de *Newton* para todo $k \in L$, y el conjunto de índices K es finito entonces la sucesión $\{\|\mathbf{d}^k\|\}_L$ es acotada, es decir, existen constantes m y M tales que*

$$0 < m \leq \|\mathbf{d}^k\| \leq M, \quad \forall k \in L. \quad (5.20)$$

Demostración. Supongamos que \mathbf{d}^k es la dirección de *Newton* para todo $k \in L$, así, para estos índices se verifica que,

$$\|\Phi_\lambda(\mathbf{x}^k)\| = \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\mathbf{d}^k\| \leq \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\|\|\mathbf{d}^k\|. \quad (5.21)$$

Si $\{\|\mathbf{d}^k\|\}_{\hat{L}} \rightarrow 0$ en un subconjunto $\hat{L} \subseteq L$, entonces por (5.21) se tendría que $\{\|\Phi_\lambda(\mathbf{x}^k)\|\}_{\hat{L}} \rightarrow 0$, puesto que la sucesión $\{\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\|\}_{\hat{L}}$ es acotada sobre la sucesión convergente $\{\mathbf{x}^k\}_{\hat{L}}$. La continuidad de Φ_λ implicaría que $\Phi_\lambda(\mathbf{x}^*) = \mathbf{0}$ y por el **Lema 5.1**, K sería infinito, contradiciendo la hipótesis de que K es finito.

Por otra parte, de (4.2) tenemos que para todo $k \in L$,

$$-\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_\lambda(\mathbf{x}^k)\|\|\mathbf{d}^k\| \leq \Phi_\lambda(\mathbf{x}^k)^T \Phi'_{\lambda\mu_k}(\mathbf{x}^k)\mathbf{d}^k \leq -\rho\|\mathbf{d}^k\|^p. \quad (5.22)$$

Dado que la sucesión $\{\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_\lambda(\mathbf{x}^k)\}_L$ es convergente (ya sea por el **Lema 3.2** o porque μ_k es constante, para k suficientemente grande) y por tanto acotada, luego existe una constante $C > 0$ tal que

$$\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_\lambda(\mathbf{x}^k)\| \leq C$$

para todo $k \in L$. De lo anterior y (5.22) tenemos

$$\rho\|\mathbf{d}^k\|^p \leq \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T \Phi_\lambda(\mathbf{x}^k)\|\|\mathbf{d}^k\| \leq C\|\mathbf{d}^k\| \quad (5.23)$$

para todo $k \in L$. Dado que $p > 1$ esto muestra que $\{\|\mathbf{d}^k\|\}_L$ es acotada, lo cual garantiza (5.20). \square

Lema 5.6. Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1** y sea $\{\mathbf{x}^k\}_L$ una subsucesión que converge a un punto $\mathbf{x}^* \in \mathbb{R}^n$. Si \mathbf{d}^k es la dirección de Newton para todo $k \in L$, y el conjunto de índices K es finito entonces

$$\liminf_{k \in L} t_k = 0.$$

Demostración. Supongamos que $\liminf_{k \in L} t_k = t_* > 0$. Entonces de (5.15) y de la búsqueda lineal (4.4), tenemos que para todo $k \in L$ suficientemente grande:

$$\Psi_{\lambda\mu_k}(\mathbf{x}^{k+1}) - \Psi_{\lambda\mu_k}(\mathbf{x}^k) \leq -2\sigma t_k \Psi_\lambda(\mathbf{x}^k) \leq -\sigma t_* \eta^2 \Psi_\lambda(\mathbf{x}^k) < 0. \quad (5.24)$$

Sea $c = t_* \eta^2 \Psi_\lambda(\mathbf{x}^k)$.

Dado que la sucesión $\{\mu_k\}$ es monótonamente decreciente y acotada inferiormente, ella converge a algún $\mu_* \geq 0$. Consideramos dos casos.

Caso 1: si $\{\mu_k\} \rightarrow \mu_* > 0$. En este caso, por la regla de actualización del paso **P.4** del algoritmo, μ_k es constante, para k lo suficientemente grande, $k \in \mathbb{N}$. Así, $\mu_k = \mu_*$.

De la regla de actualización (5.24) para todo $k \in L$ suficientemente grande:

$$\Psi_{\lambda\mu_*}(\mathbf{x}^{k+1}) - \Psi_{\lambda\mu_*}(\mathbf{x}^k) = \Psi_{\lambda\mu_k}(\mathbf{x}^{k+1}) - \Psi_{\lambda\mu_k}(\mathbf{x}^k) \leq -c \quad (5.25)$$

Dado que μ_k es eventualmente constante, la actualización (4.8) excluye la existencia de un paso del gradiente para $k \in \mathbb{N}$ suficientemente grande.

Si $L = \{l_0, l_1, \dots\}$ obtenemos de la primera parte del **Lema 5.2** que para todo, l_j suficientemente grande,

$$\Psi_{\lambda\mu_*}(\mathbf{x}^{l_j+1}) - \Psi_{\lambda\mu_*}(\mathbf{x}^{l_j}) = \Psi_{\lambda\mu_k}(\mathbf{x}^{l_j+1}) - \Psi_{\lambda\mu_k}(\mathbf{x}^{l_j}) \leq -c.$$

lo cual, implica que $\Psi_{\lambda\mu_*}(\mathbf{x}^{l_j}) \rightarrow -\infty$, para $j \rightarrow \infty$. Esto contradice el hecho $\Psi_{\lambda\mu_*}(\mathbf{x}) \geq 0$, para todo $\mathbf{x} \in \mathbb{R}^n$

Caso 2: si $\{\mu_k\} \rightarrow 0$. En este caso, el **Corolario 2.2** garantiza que:

$$|\Psi_{\lambda\mu_k}(\mathbf{x}^{k+1}) - \Psi_{\lambda\mu_{k+1}}(\mathbf{x}^k)| \leq \frac{c}{4} \quad y \quad |\Psi_{\lambda\mu_k}(\mathbf{x}^k) - \Psi_\lambda(\mathbf{x}^k)| \leq \frac{c}{4}, \quad (5.26)$$

para todo k suficientemente grande. Nuevamente, si $L = \{l_0, l_1, \dots\}$ entonces de (5.24)

y (5.26), se cumple que,

$$\begin{aligned}
\Psi_\lambda(\mathbf{x}^{l_j+1}) - \Psi_\lambda(\mathbf{x}^{l_j}) &= -(\Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j+1}) - \Psi_\lambda(\mathbf{x}^{l_j+1})) + (\Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j}) - \Psi_\lambda(\mathbf{x}^{l_j})) \\
&\quad + (\Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j+1}) - \Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j})) \\
&\leq |\Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j+1}) - \Psi_\lambda(\mathbf{x}^{l_j+1})| + |\Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j}) - \Psi_\lambda(\mathbf{x}^{l_j})| \\
&\quad + \Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j+1}) - \Psi_{\lambda\mu_{l_j}}(\mathbf{x}^{l_j}) \\
&\leq -\frac{c}{2}.
\end{aligned} \tag{5.27}$$

La prueba continúa en forma análoga a lo hecho en demostración de la **Proposición 5.3**. En particular para l_j , suficientemente grande, se cumplen las siguientes desigualdades que son esencialmente las mismas de la **Proposición 5.3** (Para esto puede ver (5.18) y recordar que K es finito),

$$\Psi_\lambda(\mathbf{x}^{l_j+1}) \leq \Psi_\lambda(\mathbf{x}^{l_j+1}) + \frac{c}{4}. \tag{5.28}$$

De (5.27) y (5.28) se sigue que

$$\begin{aligned}
\Psi_\lambda(\mathbf{x}^{l_j+1}) - \Psi_\lambda(\mathbf{x}^{l_j}) &= \Psi_\lambda(\mathbf{x}^{l_j+1}) - \Psi_\lambda(\mathbf{x}^{l_j+1}) + \Psi_\lambda(\mathbf{x}^{l_j+1}) - \Psi_\lambda(\mathbf{x}^{l_j}) \\
&\leq \frac{c}{4} - \frac{c}{2} = -\frac{c}{4}.
\end{aligned}$$

Esto implica que $\Psi_\lambda(\mathbf{x}^{l_j}) \rightarrow -\infty$ cuando $j \rightarrow \infty$, lo que contradice el hecho que $\Psi_\lambda(\mathbf{x}) \geq 0$, para todo $\mathbf{x} \in \mathbb{R}^n$. Dado que en los **Casos 1** y **2** se llegó a una contradicción, se sigue que

$$\liminf_{k \in L} t_k = 0. \quad \square$$

El siguiente teorema es el principal resultado de convergencia para el algoritmo propuesto.

Teorema 5.1. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1**. Entonces todo punto de acumulación de la sucesión $\{\mathbf{x}^k\}$ es un punto estacionario de Ψ_λ .*

Demostración. Si el conjunto de índices K es infinito, la proposición (5.2) garantiza la conclusión del teorema.

Si K es un conjunto finito y \hat{k} denota su índice más grande entonces las desigualdades (5.12) a (5.15) se satisfacen, para todo $k \geq \hat{k}$.

Por otra parte, sea \mathbf{x}^* un punto de acumulación de la sucesión $\{\mathbf{x}^k\}$, entonces existe una subsucesión $\{\mathbf{x}^k\}_L$ de $\{\mathbf{x}^k\}$ que converge a \mathbf{x}^* . Si $\mathbf{d}^k = -\nabla\Psi_\lambda(\mathbf{x}^k)$ para una cantidad infinita de índices $k \in L$, la **Proposición 5.3** garantiza que \mathbf{x}^* es un punto estacionario de Ψ_λ .

Por lo anterior, sin pérdida de generalidad supongamos que \mathbf{d}^k es la dirección de *Newton* para todo $k \in L$. Dado que K es finito podemos suponer para todo $k \in L$, se cumple que, $k \notin K$, con lo cual, las reglas de actualización (4.7) y (4.8) no se verifican, para dichos índices.

Procedemos por contradicción y suponemos que \mathbf{x}^* no es un punto estacionario de Ψ_λ . Es decir, $\Psi_\lambda(\mathbf{x}^*) \neq 0$.

Por el **Lema 5.6**

$$\liminf_{k \in L} t_k = 0.$$

Sea L_0 una subsucesión de L tal que $\{t_k\}_{L_0}$ converge a 0. Entonces, $m_k > 0$ para todo $k \in L_0$, suficientemente grande, donde $m_k \in \mathbb{N}$ denota el exponente de la búsqueda lineal en (4.4). Luego,

$$-2\sigma\theta^{m_k-1}\Psi_\lambda(\mathbf{x}^k) < \Psi_{\lambda\mu_k}(\mathbf{x}^k + \theta^{m_k-1}\mathbf{d}^k) - \Psi_{\lambda\mu_k}(\mathbf{x}^k),$$

para todo $k \in L_0$, suficientemente grande. Dividiendo ambas desigualdades por θ^{m_k-1} , obtenemos

$$-2\sigma\Psi_\lambda(\mathbf{x}^k) < \frac{\Psi_{\lambda\mu_k}(\mathbf{x}^k + \theta^{m_k-1}\mathbf{d}^k) - \Psi_{\lambda\mu_k}(\mathbf{x}^k)}{\theta^{m_k-1}}.$$

Sea μ_* el límite de la sucesión $\{\mu_k\}$. Si $\mu_* = 0$ entonces escribimos $\nabla\Psi_{\lambda\mu_*}(\mathbf{x}^*)$ para el gradiente de la función no perturbada Ψ_λ en el punto límite \mathbf{x}^* .

De (5.20) podemos asumir (se puede extraer una subsucesión) que $\{\mathbf{d}^k\}_{L_0} \rightarrow \mathbf{d}^* \neq 0$. Tomando el límite cuando $k \rightarrow \infty$ tenemos que

$$-2\sigma\Psi_\lambda(\mathbf{x}^*) \leq \nabla\Psi_\lambda(\mathbf{x}^*)^T \mathbf{d}^*. \quad (5.29)$$

Para $\mu_* = 0$ lo anterior se verifica por el **Lema 3.3**. Si $\mu_* > 0$, entonces $\mu_k = \mu_*$ para k suficientemente grande, luego, (5.29) se sigue del Teorema del valor medio.

Usando (4.1), (5.14) y el **Corolario 2.2** tenemos que para todo $k \in \mathbb{L}_0$,

$$\begin{aligned}
\nabla \Psi_{\lambda\mu_k}(\mathbf{x}^k)^T \mathbf{d}^k &= -\Phi_\lambda(\mathbf{x}^k)^T \Phi_{\lambda\mu_k}(\mathbf{x}^k) \\
&= -2\Psi_\lambda(\mathbf{x}^k) + \Phi_\lambda(\mathbf{x}^k)^T (\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_k}(\mathbf{x}^k)) \\
&\leq -2\Psi_\lambda(\mathbf{x}^k) + \|\Phi_\lambda(\mathbf{x}^k)\| \|\Phi_\lambda(\mathbf{x}^k) - \Phi_{\lambda\mu_{k-1}}(\mathbf{x}^k)\| \\
&\quad + \|\Phi_\lambda(\mathbf{x}^k)\| \|\Phi_{\lambda\mu_{k-1}}(\mathbf{x}^k) - \Phi_{\lambda\mu_k}(\mathbf{x}^k)\| \\
&\leq -2\Psi_\lambda(\mathbf{x}^k) + 2\alpha\Psi_\lambda(\mathbf{x}^k) + \kappa\|\Phi_\lambda(\mathbf{x}^k)\|(\sqrt{\mu_{k-1}} - \sqrt{\mu_k}) \\
&= -2(1 - \sigma)\Psi_\lambda(\mathbf{x}^k) + \kappa\|\Phi_\lambda(\mathbf{x}^k)\|(\sqrt{\mu_{k-1}} - \sqrt{\mu_k}),
\end{aligned}$$

tomando límite, cuando $k \rightarrow \infty$, $k \in L_0$, obtenemos de (5.29) (y del **Lema 3.2**, si $\mu_* = 0$),

$$-2\sigma\Psi_\lambda(\mathbf{x}^*) \leq \nabla \Psi_{\lambda\mu_*}(\mathbf{x}^*)^T \mathbf{d}^* \leq -2(1 - \alpha)\Psi_\lambda(\mathbf{x}^*), \quad (5.30)$$

Por la **Proposición 5.1**, $\{\Psi_\lambda(\mathbf{x}^k)\}$ es acotada. Además, $(\sqrt{\mu_{k-1}} - \sqrt{\mu_k}) \rightarrow 0$, puesto que $\{\mu_k\}$ converge. Tenemos que $\Psi_\lambda(\mathbf{x}^*) > 0$, ya que en otro caso, se tendría que K sería infinito. Por lo tanto, de (5.30) $\sigma \geq (1 - \alpha)$ lo cual, contradice que $\sigma < (1 - \alpha)$. Esto completa la prueba del teorema. \square

Capítulo 6

Convergencia local

En este capítulo demostraremos, bajo ciertas hipótesis, que el algoritmo propuesto converge local y q -superlineal o q -cuadráticamente. Para esto, presentamos inicialmente, algunos resultados preliminares. En particular, el primero de ellos, está basado en la **Proposición 3.3**, da una condición suficiente para la convergencia de una sucesión generada por el **Algoritmo 4.1**.

Teorema 6.1. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1**. Si uno de sus puntos de acumulación, \mathbf{x}^* , es una solución aislada del PCNL(F) entonces $\{\mathbf{x}^k\}$ converge a \mathbf{x}^* .*

Demostración. Sea \mathbf{x}^* una solución aislada del PCNL(F). Luego, por el **Lema 5.1**, el conjunto de índices K definido en (4.9) es infinito y $\{\mu_k\}$ converge a cero y por tanto, la **Proposición 5.2** garantiza que cada punto de acumulación de la sucesión $\{\mathbf{x}^k\}$ es también una solución del PCNL(F). Así \mathbf{x}^* debe ser un punto de acumulación aislado de $\{\mathbf{x}^k\}$.

Supongamos que $\{\mathbf{x}^k\}_L$ es una subsucesión arbitraria de $\{\mathbf{x}^k\}$ que converge a \mathbf{x}^* . Usando la regla de actualización del paso **P.3** del **Algoritmo 4.1** tenemos,

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\| = \theta^{m_k} \|\mathbf{d}^k\| \leq \|\mathbf{d}^k\|. \quad (6.1)$$

Teniendo en cuenta la desigualdad (6.1), es suficiente probar que $\{\mathbf{d}^k\}_L \rightarrow 0$. Dado que Ψ_λ es continuamente diferenciable y que la solución del PCNL(F), \mathbf{x}^* , es un punto

estacionario de Ψ_λ , tenemos que

$$\{\nabla\Psi_\lambda(\mathbf{x}^k)\}_L \rightarrow \nabla\Psi_\lambda(\mathbf{x}^*) = 0. \quad (6.2)$$

Si la sucesión $\{\mathbf{d}^k\}_L$ tiene solo una cantidad finita de direcciones de *Newton* entonces ella converge a cero. Por ello, supongamos que existe una subsucesión $\{\mathbf{d}^k\}_{L_0}$ de $\{\mathbf{d}^k\}_L$ tal que \mathbf{d}^k es la solución del sistema 4.1, para todo $k \in L_0$.

De (4.3),

$$\rho\|\mathbf{d}^k\|^p \leq -(\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T\Phi_\lambda(\mathbf{x}^k))^T\mathbf{d}^k \leq \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T\Phi_\lambda(\mathbf{x}^k)\|\|\mathbf{d}^k\|,$$

para todo $k \in L_0$, de donde,

$$\|\mathbf{d}^k\| \leq \left(\frac{\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T\Phi_\lambda(\mathbf{x}^k)\|}{\rho} \right)^{\frac{1}{p-1}} \quad (6.3)$$

puesto que $p > 1$. Además, recordando que $\{\mathbf{x}^k\} \rightarrow \mathbf{x}^*$ y $\{\mu_k\} \rightarrow 0$, con $k \in L_0$ el **Lema 3.2** permite concluir que,

$$\lim_{k \rightarrow \infty, k \in L_0} \Phi'_{\lambda\mu_k}(\mathbf{x}^k)^T\Phi_\lambda(\mathbf{x}^k) = \nabla\Psi_\lambda(\mathbf{x}^*) = 0,$$

lo que a su vez implica que, el lado derecho de (6.3) converge a cero, con lo cual, $\{\mathbf{d}^k\}_{L_0}$ también converge a cero. De (5.24), se infiere que $\{\mathbf{d}^k\}_{L \setminus L_0} \rightarrow 0$, si $L \setminus L_0$ es infinito. Así, de (6.1) tenemos que,

$$\{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|\}_L \rightarrow 0$$

y por tanto, por la **Proposición 3.3** se concluye que $\{\mathbf{x}^k\}$ converge a \mathbf{x}^* . \square

Los dos siguientes resultados son lemas técnicos que serán usados en la demostración del **Teorema 6.2**. El primero de ellos garantiza que, para todo $k \in K$, las matrices $\Phi'_{\lambda\mu_k}(\mathbf{x}^k)$ son no singulares y sus inversas están acotadas, mientras que, en el segundo resultado se tiene que, para $k \in K$, con k suficientemente grande, la dirección de *Newton* satisface la condición de descenso (4.2).

Lema 6.1. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1**. Si uno de sus puntos límite, \mathbf{x}^* , es una solución R -regular del PCNL(F) entonces para todo $k \in K$, suficientemente grande las matrices $\Phi'_{\lambda\mu_k}(\mathbf{x}^k)$ son no singulares y sus inversas satisfacen*

$$\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}\| \leq 2c, \quad (6.4)$$

para cierta constante positiva c .

Demostración. Por hipótesis, \mathbf{x}^* es una solución R-regular del PCNL(F) entonces todas las matrices del C -subdiferencial $\partial_C\Phi_\lambda(\mathbf{x}^*)$ son no singulares [19]. Además, por la **Proposición 2.5** en [31] y la **Proposición 3.1**, \mathbf{x}^* es una solución aislada de $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$ y por ello, una solución del PCNL(F). Por tanto, del **Teorema 6.1** tenemos que la sucesión $\{\mathbf{x}^k\}$ converge a \mathbf{x}^* . En este caso, el conjunto K definido en (4.9) es infinito puesto que \mathbf{x}^* es una solución del PCNL(F). En particular, $\{\mathbf{x}^k\}_K \rightarrow \mathbf{x}^*$.

Por la convergencia $\{\mathbf{x}^k\} \rightarrow \mathbf{x}^*$, la R-regularidad y semicontinuidad del C -subdiferencial tenemos que, para todo $k \in \mathbb{N}$ suficientemente grande las matrices $V_k \in \partial_C\Phi_\lambda(\mathbf{x}^k)$ son no singulares con $\|V_k^{-1}\| \leq c$, para alguna constante positiva c . Veamos que lo anterior también se cumple para las matrices $\Phi'_{\lambda\mu_k}(\mathbf{x}^k)$.

De [19] se tiene que el conjunto $\partial_C\Phi_\lambda(\mathbf{x}^k)$ es no vacío y compacto, por ello podemos considerar $H_k \in \partial_C\Phi_\lambda(\mathbf{x}^k)$, tal que

$$\text{dist}_F(\Phi'_{\lambda\mu_k}(\mathbf{x}^k), \partial_C\Phi_\lambda(\mathbf{x}^k)) = \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k) - H_k\|_F$$

de lo anterior y por (4.11), tenemos que,

$$\|H_k - \Phi'_{\lambda\mu_k}(\mathbf{x}^k)\| \leq \|H_k - \Phi'_{\lambda\mu_k}(\mathbf{x}^k)\|_F \leq \gamma\beta_k, \quad (6.5)$$

para todo $k \in K$. Así

$$\|I - H_k^{-1}\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\| \leq \|H_k^{-1}\| \|H_k - \Phi'_{\lambda\mu_k}(\mathbf{x}^k)\| \leq \gamma\beta_k c. \quad (6.6)$$

Dado que K es infinito y la regla de actualización en el paso **P.4** del **Algoritmo 4.1** tenemos que $\beta_k \rightarrow 0$. Por tanto, para $k \in K$ suficientemente grande tal que $\beta_k \leq (2\gamma c)^{-1}$,

$$\|I - H_k^{-1}\Phi'_{\lambda\mu_k}(\mathbf{x}^k)\| \leq \frac{1}{2}.$$

Del **Teorema 3.1.4** en [13] tenemos que $\Phi'_{\lambda\mu_k}(\mathbf{x}^k)$ es no singular para todo $k \in K$ suficientemente grande, además

$$\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}\| \leq 2\|H_k^{-1}\| \leq 2c. \quad (6.7)$$

Entonces el sistema (4.1) tiene solución para todo $k \in K$, suficientemente grande. \square

Lema 6.2. *Bajo las hipótesis del Lema 6.1, sea $k \in K$ suficientemente grande, la solución del sistema lineal (4.1) satisface la condición descenso (4.2).*

Demostración. Del lema anterior tenemos que el sistema (4.1) tiene una única solución \mathbf{d}_k , para $k \in K$ suficientemente grande y además, las matrices $\Phi'_{\lambda\mu_k}(\mathbf{x}^k)$ son no singulares y sus inversas están acotadas superiormente por la constante positiva $2c$. Así, a partir de (4.1) tenemos la siguiente desigualdad:

$$\|\mathbf{d}^k\| \leq \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}\| \|\Phi_\lambda(\mathbf{x}^k)\| \leq 2c \|\Phi_\lambda(\mathbf{x}^k)\|,$$

de donde

$$-\|\Phi_\lambda(\mathbf{x}^k)\|^2 \leq -\rho_1 \|\mathbf{d}^k\|^2 \tag{6.8}$$

con, $\rho_1 = (4c^2)^{-1}$.

Por otra parte, multiplicando (4.1) por $\Phi_\lambda(\mathbf{x})^T$ y usando (6.8), tenemos que

$$\Phi_\lambda(\mathbf{x})^T \Phi'_{\lambda\mu_k}(\mathbf{x}) \mathbf{d}^k = -\|\Phi_\lambda(\mathbf{x}^k)\|^2 \leq -\rho_1 \|\mathbf{d}^k\|^2, \tag{6.9}$$

para todo $k \in K$ suficientemente grande. Con lo cual se demuestra que \mathbf{d}^k satisface la condición descenso (4.2), con $\rho_1 = (4c^2)^{-1}$.

En la demostración del lema anterior, se garantiza que $\{\|\mathbf{d}^k\|\} \rightarrow 0$, con lo cual se puede ver que, (4.2) se sigue eventualmente a partir de (6.9), para cualquier $\rho > 0$ y $p > 2$. Por tanto, para todo $k \in K$ suficientemente grande, la dirección \mathbf{d}^k siempre está dada por (4.1). \square

El siguiente lema garantiza que los índices de las iteraciones \mathbf{x}^k permanecerán en el conjunto K . La repetición de este argumento garantizará que $k \in K$ y $t_k = 1$ para todo, $k \in \mathbb{N}$ suficientemente grande.

Lema 6.3. *Asuma las hipótesis del Lema 6.1. Existe un índice $\hat{k} \in K$, tal que para todo $k \geq \hat{k}$, el índice $k + 1$ también pertenece a K y $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{d}^k$.*

Demostración. Por el Lema 6.1 tenemos que, existe una constante positiva c tal que $\|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}\| \leq 2c$, para todo $k \in K$ suficientemente grande. Además, para estos k ,

por el **Algoritmo 4.1** y por (6.5) se tiene que,

$$\begin{aligned}
\|\mathbf{x}^k + \mathbf{d}^k - \mathbf{x}^*\| &= \|\mathbf{x}^k - \mathbf{x}^* - \Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}\Phi_\lambda(\mathbf{x}^k)\| \\
&= \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}(\Phi'_{\lambda\mu_k}(\mathbf{x}^k)(\mathbf{x}^k - \mathbf{x}^*) - \Phi_\lambda(\mathbf{x}^k) + \Phi_\lambda(\mathbf{x}^*))\| \\
&\leq \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1}\|(\|(\Phi'_{\lambda\mu_k}(\mathbf{x}^k)^{-1} - H_k)(\mathbf{x}^k - \mathbf{x}^*)\| \\
&\quad + \|H_k(\mathbf{x}^k - \mathbf{x}^*) - \Phi_\lambda(\mathbf{x}^k) + \Phi_\lambda(\mathbf{x}^*)\|) \\
&\leq 2c(\gamma\beta_k\|\mathbf{x}^k - \mathbf{x}^*\| + \|H_k(\mathbf{x}^k - \mathbf{x}^*) - \Phi_\lambda(\mathbf{x}^k) + \Phi_\lambda(\mathbf{x}^*)\|),
\end{aligned} \tag{6.10}$$

aquí, $H_k \in \partial_C\Phi_\lambda(\mathbf{x}^k)$ es tal que (al igual que en **Lema 6.1**)

$$\text{dist}_F(\Phi'_{\lambda\mu_k}(\mathbf{x}^k), \partial_C\Phi_\lambda(\mathbf{x}^k)) = \|\Phi'_{\lambda\mu_k}(\mathbf{x}^k) - H_k\|_F.$$

Usando la **Proposición 3.1** y teniendo en cuenta que $\beta_k \rightarrow 0$, se sigue que

$$\|\mathbf{x}^k + \mathbf{d}^k - \mathbf{x}^*\| = o(\|\mathbf{x}^k - \mathbf{x}^*\|), \tag{6.11}$$

para $k \in K$ con $k \rightarrow \infty$. Lo anterior y la **Proposición 3.4** muestran que

$$\|\Phi_\lambda(\mathbf{x}^k + \mathbf{d}^k)\| = o(\|\Phi_\lambda(\mathbf{x}^k)\|), \tag{6.12}$$

para $k \in K$ con $k \rightarrow \infty$.

Sea

$$\omega := \max\left\{\frac{1}{2} - \frac{(1 - \alpha - 2\sigma)^2}{2(2 + \alpha)^2}, \frac{(1 - \eta)^2}{2}\right\},$$

donde α , η y σ son las constantes definidas en el **Algoritmo 4.1**. A partir de (6.12), existe un índice $\hat{k} \in K$ tal que

$$\Psi_{\lambda\mu_k}(\mathbf{x}^k + \mathbf{d}^k) \leq \Psi_{\lambda\mu_k}(\mathbf{x}^k) - 2\omega\Psi_\lambda(\mathbf{x}^k), \tag{6.13}$$

para todo $k \in K$ con $k \geq \hat{k}$. Luego, por el **Lema 3.4** concluimos que

$$\Psi_{\lambda\mu_k}(\mathbf{x}^k + \mathbf{d}^k) \leq \Psi_{\lambda\mu_k}(\mathbf{x}^k) - 2\sigma\Psi_\lambda(\mathbf{x}^k), \tag{6.14}$$

para todo $k \in K$ con $k \geq \hat{k}$. Lo que indica que, el paso completo será aceptado para todo $k \geq \hat{k}$, $k \in K$. En particular, $\mathbf{x}^{\hat{k}+1} = \mathbf{x}^{\hat{k}} + \mathbf{d}^{\hat{k}}$. Además, de (6.13) y usando la definición de ω , se obtiene que

$$\|\Phi_\lambda(\mathbf{x}^{\hat{k}+1})\| \leq \sqrt{1 - 2\omega}\|\Phi_\lambda(\mathbf{x}^{\hat{k}})\| \leq \eta\|\Phi_\lambda(\mathbf{x}^{\hat{k}})\| = \eta\beta_{\hat{k}},$$

lo cual, implica que $\hat{k} + 1 \in K$. Repitiendo el proceso anterior, se demuestra que para todo $\hat{k} \geq k$ se cumple que $k \in K$ y $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{d}^k$. \square

En siguiente teorema da condiciones suficientes para la tasa de convergencia de una sucesión generada por el **Algoritmo 4.1**.

Teorema 6.2. *Sea $\{\mathbf{x}^k\}$ una sucesión generada por el **Algoritmo 4.1**. Si uno de sus puntos límite, \mathbf{x}^* , es una solución R -regular del PCNL(F) entonces $\{\mathbf{x}^k\}$ converge a \mathbf{x}^* por lo menos q -superlinealmente a \mathbf{x}^* . Si $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es continuamente diferenciable con matriz jacobiana localmente Lipschitz entonces la convergencia es q -cuadrática.*

Demostración. El **Lema 6.3** garantiza que, para $k \in \mathbb{N}$ suficientemente grande, $k \in K$ y $t_k = 1$. Luego la convergencia q -superlineal se obtiene de (6.11). Por otro lado, si $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es continuamente diferenciable con matriz jacobiana localmente Lipschitz entonces la **Proposición 3.1** garantiza que

$$\|H_k(\mathbf{x}^k + \mathbf{d}^k) - \Phi_\lambda(\mathbf{x}^k) + \Phi_\lambda(\mathbf{x}^*)\| = O(\|\mathbf{x}^k - \mathbf{x}^*\|^2).$$

Dado que Φ_λ es localmente Lipschitz, se tiene que

$$\beta_k = \|\Phi_\lambda(\mathbf{x}^k)\| = \|\Phi_\lambda(\mathbf{x}^k) - \Phi_\lambda(\mathbf{x}^*)\| = O(\|\mathbf{x}^k - \mathbf{x}^*\|).$$

Usando estas dos igualdades en (6.10), se tiene que,

$$\|\mathbf{x}^k + \mathbf{d}^k - \mathbf{x}^*\| = O(\|\mathbf{x}^k - \mathbf{x}^*\|^2),$$

Por lo tanto, la sucesión $\{\mathbf{x}^k\}$ converge q -cuadráticamente a \mathbf{x}^* .

\square

Capítulo 7

Pruebas numéricas

En este capítulo analizamos numéricamente el desempeño del **Algoritmo 4.1**, propuesto en el **Capítulo 4**. Realizamos diversas pruebas numéricas variando el parámetro λ en el intervalo $(0, 4)$ en forma dinámica [4] y en forma aleatoria, respectivamente. También consideramos el caso $\lambda = 2$ que reduce nuestra algoritmo al propuesto en [20]. Con lo anterior tenemos tres versiones del **Algoritmo 4.1**.

Otro punto clave en las comparaciones es el de los puntos iniciales ya que el carácter global de un algoritmo está asociado con su convergencia desde puntos iniciales arbitrarios. En este sentido, consideramos en primer lugar, puntos tomados de la literatura sobre problemas de prueba para algoritmos de complementariedad no lineal y posteriormente, generamos puntos iniciales aleatorios.

Comparamos el desempeño numérico de nuestra propuesta algorítmica con tres métodos globales que resuelven problemas de complementariedad lineal, propuestos en [4], [20] y recientemente en [36]. El primero de estos de tipo *cuasi-Newton*, el segundo, de tipo *Newton* con Jacobiano suavizado y el tercero, tipo *Levenberg- Marquardt* [13].

Finalizamos el capítulo con algunas pruebas numéricas que ilustran la ventaja de tener un algoritmo global frente a uno local. Para esto, implementamos una versión local de nuestra propuesta algorítmica.

7.1. Problemas de prueba, puntos iniciales y valores de los parámetros

Para las pruebas numéricas consideramos diez problemas de complementariedad no lineal asociados a las siguientes funciones. En cada caso, reportamos el nombre con el que haremos referencia del problema en la tablas de resultados y, la solución(es) del problema.

Problema 7.1. Kojima Shindo (*Koj-Shi*) [28]. $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, definido por,

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

Soluciones: $\mathbf{x}_1^* = (\frac{1}{2}\sqrt{6} \ 0 \ 0 \ \frac{1}{2})^T$ y $\mathbf{x}_2^* = (1 \ 0 \ 3 \ 0)^T$. La solución \mathbf{x}_1^* es una *degenerada*, mientras que \mathbf{x}_2^* es *no degenerada*.

Problema 7.2. Kojima Josephy (*Koj-Jo*) [17]. $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, definido por,

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

Solución única: $\mathbf{x}^* = (\frac{1}{2}\sqrt{6} \ 0 \ 0 \ \frac{1}{2})^T$.

Problema 7.3. Mathiesen Modificado (*Math mod*) [25]. $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definido por,

$$F(\mathbf{x}) = \begin{pmatrix} -x_2 + x_3 + x_4 \\ x_1 - 0.9 \left(\frac{5x_3 + 3x_4}{x_2 + 1} \right) \\ 5 - x_1 - 0.1 \left(\frac{5x_3 + 3x_4}{x_2 + 1} \right) \\ 3 - x_1 \end{pmatrix}.$$

Infinitas soluciones de la forma $\mathbf{x}^* = (a \ 0 \ 0 \ 0)^T$, con $a \in [0, 3]$.

Problema 7.4. Mathiesen (Mathiesen) [28]. $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definido por,

$$F(\mathbf{x}) = \begin{pmatrix} -x_2 + x_3 + x_4 \\ x_1 - \alpha \left(\frac{b_1 x_3 + b_2 x_4}{x_2} \right) \\ b_1 - x_1 - (1 - \alpha) \left(\frac{b_1 x_3 + b_2 x_4}{x_3} \right) \\ b_2 - x_1 \end{pmatrix}.$$

Donde las constantes $b_1, b_2 > 0$ y $\alpha \in (0, 1)$. Para el problema asociado a esta función, se desea encontrar una solución con $x_1 > 0$, $x_2 > 0$ y $x_2 + x_3 + x_4 = 1$. La solución al problema de complementariedad es:

i) Si $\alpha b_1 > b_2$, para cualquier $\tau > 0$,

$$\mathbf{x}^* = \left(b_2 \quad \tau \frac{R}{S} \quad \tau \quad \tau \left(\frac{R}{S} - 1 \right) \right)^T,$$

$$\text{donde } R = \frac{b_1 - b_2}{b_2}, \quad S = \frac{1 - \alpha}{\alpha}.$$

ii) Si $\alpha b_1 < b_2$ entonces para cualquier $\tau > 0$,

$$\mathbf{x}^* = \left(\alpha b_1 \quad \frac{\tau}{2} \quad \frac{\tau}{2} \quad 0 \right)^T,$$

Problema 7.5. Billups (Billups) [5]. $F: \mathbb{R} \rightarrow \mathbb{R}$ definido por,

$$F(x) = (x - 1)^2 - 1.1.$$

Solución: $\mathbf{x}^* = 2.048808849324785$.

Problema 7.6. Nash Cournot (Nash Co). [17] $F: \mathbb{R}_+^n \rightarrow \mathbb{R}^n$ definido por,

$$F_i(\mathbf{x}) = g'_i(x_i) - h \left(\sum_{j=1}^n x_j \right) - x_i h' \left(\sum_{j=1}^n x_j \right), \quad i = 1, \dots, n,$$

donde

$$g(x_i) = c_i x_i + \frac{b_i}{1+b_i} l_i \frac{1}{b_i} x_i^{1+\frac{1}{b_i}}, \quad h(Q) = 5000 \gamma Q^{-\frac{1}{\gamma}}.$$

Los valores c_i, l_i, b_i son positivos y $\gamma > 0$. O en forma más compacta

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

donde $\mathbf{c} = (c_1 \cdots c_n)^T$, $\mathbf{b} = (b_1 \cdots b_n)^T$, $\mathbf{L} = (l_1 \cdots l_n)^T$, $\mathbf{e} = (1 \cdots 1)^T$.

Haremos dos pruebas numéricas de este problema, con $n = 5$ y $n = 10$. Los parámetros son tomados de [17], para mayor claridad en la lectura del documento, los incluimos a continuación.

Para $n = 5$, llamaremos el problema *Nash Co (5)*,

$$\begin{aligned} \mathbf{c} &= (10 \ 8 \ 6 \ 4 \ 2)^T \\ \mathbf{b} &= (1.2 \ 1.1 \ 1 \ 0.9 \ 0.8)^T \\ \mathbf{L} &= (5 \ 5 \ 5 \ 5 \ 5)^T \\ \mathbf{e} &= (1 \ 1 \ 1 \ 1 \ 1)^T \end{aligned}$$

y $\gamma = 1.1$. La solución es $\mathbf{x}^* = (15.429 \ 12.499 \ 9.6635 \ 7.1651 \ 5.1326)^T$.

Para $n = 10$, llamaremos el problema *Nash Co (10)*,

$$\begin{aligned} \mathbf{c} &= (5 \ 3 \ 8 \ 4 \ 5 \ 1 \ 3 \ 7 \ 4 \ 6 \ 3)^T \\ \mathbf{b} &= (1.2 \ 1 \ 0.9 \ 0.6 \ 1.5 \ 1 \ 0.7 \ 1.1 \ 0.95 \ 0.75)^T \\ \mathbf{L} &= (10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10)^T \\ \mathbf{e} &= (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1)^T \end{aligned}$$

y $\gamma = 1.2$. Solución:

$\mathbf{x}^* = (7.4415 \ 4.0978 \ 2.5906 \ 0.9354 \ 17.9490 \ 4.0978 \ 1.3047 \ 5.5901 \ 3.2222 \ 1.6771)^T$.

Problema 7.7. Hock Schittkowski (HH 66) [36]. $F: \mathbb{R}^8 \rightarrow \mathbb{R}^8$ definido por,

$$F(\mathbf{x}) = \begin{pmatrix} -0.8 + x_4 e^{x_1} + x_6 \\ -x_4 + x_5 e^{x_2} + x_7 \\ 0.2 - x_5 + x_8 \\ x_2 - e^{x_1} \\ x_3 - e^{x_2} \\ 100 - x_1 \\ 100 - x_2 \\ 10 - x_3 \end{pmatrix}.$$

Solución: $\mathbf{x}^* = (0.1841 \ 1.2022 \ 3.3273 \ 0.6655 \ 0.2 \ 0 \ 0 \ 0)^T$.

Problema 7.8. Geiger Kanzow (Geiger-Kanzow) [16] $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por

$$F(\mathbf{x}) = M\mathbf{x} + \mathbf{q},$$

donde

$$M = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 4 \end{pmatrix}, \quad \mathbf{q} = (-1 \ -1 \ \cdots \ -1)^T.$$

Con $n = 500$, la solución es:

$$\mathbf{x}^* = (0.3660 \ 0.4641 \ 0.4904 \ 0.4974 \ 0.4993 \ 0.4998 \ 0.5 \ \cdots \\ 0.5 \ 0.4998 \ 0.4993 \ 0.4974 \ 0.4904 \ 0.4641 \ 0.3660)^T.$$

Problema 7.9. Ahn (Ahn) [8]. $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por,

$$F(\mathbf{x}) = M\mathbf{x} + \mathbf{q},$$

donde la matriz M y el vector \mathbf{q} están dados, respectivamente, por,

$$M = \begin{pmatrix} 4 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 4 & -2 & \cdots & 0 & 0 \\ 0 & 1 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix}, \quad q = (-1 \quad -1 \quad \cdots \quad -1)^T.$$

Con $n = 3000$, la solución es:

$$\mathbf{x}^* = (0.4082 \quad 0.3165 \quad 0.3371 \quad 0.3325 \quad 0.3335 \quad 0.3333 \quad \cdots \\ 0.3333 \quad 0.3332 \quad 0.3331 \quad 0.3328 \quad 0.3321 \quad 0.3306 \quad 0.3272 \quad 0.3197 \quad 0.3031 \quad 0.2660 \quad 0.1835)^T.$$

Para escribir el código del **Algoritmo 4.1** y de las funciones de prueba usamos el *software* MATLAB[®] y realizamos las pruebas en un computador con procesador Intel (R) Celeron (R) de 1.5 GHz.

En cuanto a los vectores iniciales, consideramos algunos fijos, tomados de [4] [36], y otros aleatorios, según el experimento numérico a realizar. Con el objetivo de hacer referencia rápida, posterior de los puntos iniciales fijos, los enumeramos a continuación.

$$\begin{aligned} \mathbf{x}_1 &= (1 \quad 1 \quad 1 \quad 1)^T & \mathbf{x}_{12} &= (1 \quad 0 \quad 0 \quad 0)^T \\ \mathbf{x}_2 &= 1 & \mathbf{x}_{13} &= (1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0)^T \\ \mathbf{x}_3 &= (10 \quad 10 \quad 10 \quad 10 \quad 10)^T & \mathbf{x}_{14} &= 0 \\ \mathbf{x}_4 &= (10 \quad 10 \quad \cdots \quad 10)^T & \mathbf{x}_{15} &= (1 \quad 1 \quad 1 \quad 1 \quad 1)^T \\ \mathbf{x}_5 &= 10 \mathbf{x}_1 & \mathbf{x}_{16} &= (100 \quad \cdots \quad 100)^T \\ \mathbf{x}_6 &= (-1 \quad \cdots \quad -1)^T & \mathbf{x}_{17} &= -2 \mathbf{x}_1 \\ \mathbf{x}_7 &= 6 \mathbf{x}_1 & \mathbf{x}_{18} &= (1 \quad 4 \quad 4 \quad 1)^T \\ \mathbf{x}_8 &= (1 \quad 2 \quad 3 \quad 4)^T & \mathbf{x}_{19} &= 3 \mathbf{x}_1 \\ \mathbf{x}_9 &= (2 \quad -3 \quad -3 \quad 2)^T & \mathbf{x}_{20} &= (0 \quad \cdots \quad 0)^T \\ \mathbf{x}_{10} &= 100 \mathbf{x}_1 & \mathbf{x}_{21} &= (-1 \quad \cdots \quad -1 \quad 1 \quad \cdots \quad 1)^T \\ \mathbf{x}_{11} &= (1 \quad 0 \quad 1 \quad 0)^T \end{aligned}$$

Los parámetros requeridos fueron tomados de [20], a saber:

$$\rho = 10^{-18}, \quad p = 2.1 \quad \theta = 0.5 \quad \sigma = 10^{-4}, \quad \gamma = 30, \quad \alpha = 0.95 \quad \eta = 0.9. \\ \epsilon_1 = 10^{-12}, \quad \epsilon_2 = 10^{-6}, \quad k_{\text{máx}} = 300, \quad t_{\text{mín}} = 10^{-16},$$

donde ϵ_1, ϵ_2 son las tolerancias tomadas para $\Phi_\lambda(\mathbf{x}_k)$ y $\|\nabla\Psi_\lambda(\mathbf{x}_k)\|$, respectivamente, y t_{\min} indica el tamaño mínimo del paso en la búsqueda lineal.

7.2. Pruebas numéricas

Teniendo en cuenta que el parámetro λ del **Algoritmo 4.1** puede tomar valores en el intervalo $(0, 4)$, iniciamos las pruebas numéricas variando este valor en dicho intervalo y considerando, en primer lugar puntos iniciales fijos (tomados [4] [36]), y posteriormente, otros generados aleatoriamente. Dicha variación de λ la haremos de tres formas (con las cuales tendremos tres versiones de nuestro algoritmo) que describimos a continuación.

Método 1. Usamos $\lambda = 2$. En este caso, la familia de funciones de complementariedad se reduce a la función de *Fisher* y nuestro algoritmo coincide con el propuesto en [20].

Método 2. Usamos la escogencia dinámica para el parámetro λ utilizada en [4], es decir,

1. $\lambda = 2$
2. **Si** $\Psi_\lambda(\mathbf{x}_k) \leq \gamma_1$ **entonces** $\lambda = \Psi_\lambda(\mathbf{x}_k)$
 Si no $\lambda = \min\{c_1\Psi(\mathbf{x}_k), \lambda\}$
3. **Si** $\Psi(\mathbf{x}_k) \leq \gamma_2$ **entonces** $\lambda = \min\{c_2, \lambda\}$,

donde $\gamma_1 = 10^{-2}$, $\gamma_2 = 10^{-4}$, $c_1 = 10$, $c_2 = 10^{-8}$.

Esta estrategia combina la eficiencia de la función de *Fisher* lejos de la solución, con la de la función mínimo, cerca de la misma.

Método 3. Variamos aleatoriamente el parámetro λ de forma uniforme en $(0, 4)$.

Los resultados obtenidos en cuanto a número de iteraciones, criterio de parada y número de pasos de *Newton*, es decir, cuando $t = 1$, partir de puntos iniciales fijos con las tres variantes del parámetro λ descritas anteriormente son presentados en la **Tabla 7.1**, cuyas columnas contienen la siguiente información: nombre del problema de prueba (**Problema**), dimensión del mismo (**Dim**), número de iteraciones (k), punto inicial (\mathbf{x}_0), solución encontrada (\mathbf{x}^*), valor de la función Ψ_λ en \mathbf{x}^* ($\Psi_\lambda(\mathbf{x}^*)$), valor de la norma

del gradiente de Ψ_λ en \mathbf{x}^* ($\|\nabla\Psi_\lambda(\mathbf{x}^*)\|$) y número de pasos completos en la búsqueda monótona (**np**).

Problema	\mathbf{x}_0	Método	Dim	k	$\Psi_\lambda(\mathbf{x}^*)$	$\ \nabla\Psi_\lambda(\mathbf{x}^k)\ $	np
<i>Koj-Shi</i>	\mathbf{x}_1	1	4	7	9.2e-17	9.1e-08	3
		2	4	7	3.6e-17	1.3e-07	3
		3	4	7	1.9e-14	2.9e-07	3
<i>Koj-Jo</i>	\mathbf{x}_1	1	4	6	3.7e-13	7.4e-06	3
		2	4	6	1.9e-21	9.2e-10	3
		3	4	6	1.7e-13	2.5e-06	3
<i>Math mod</i>	\mathbf{x}_1	1	4	4	6.8e-18	3.7e-09	4
		2	4	4	5.6e-32	6.3e-16	4
		3	4	4	1.3e-27	9.4e-14	4
<i>Billups</i>	\mathbf{x}_2	1	1	4	2.6e-15	1.5e-07	3
		2	1	5	1.2e-17	2.0e-08	4
		3	1	4	1.8e-18	7.4e-10	4
<i>Nash Co (5)</i>	\mathbf{x}_3	1	5	6	3.1e-16	4.0e-07	5
		2	5	6	5.1e-27	3.3e-12	5
		3	5	7	2.3e-15	5.4e-07	6
<i>Nash Co (10)</i>	\mathbf{x}_4	1	10	11	2.1e-22	7.2e-10	4
		2	10	10	1.8e-19	4.3e-08	3
		3	10	11	5.4e-23	3.9e-10	4
<i>Mathiesen</i>	\mathbf{x}_5	1	4	7	8.3e-14	5.8e-07	6
		2	4	6	4.9e-15	3.0e-07	5
		3	4	8	5.1e-21	1.6e-10	8
<i>HH 66</i>	\mathbf{x}_6	1	8	9	1.6e-15	1.4e-07	8
		2	8	8	1.7e-16	1.1e-07	7
		3	8	9	4.0e-23	3.5e-11	8
<i>Geiger-Kanzow</i>	\mathbf{x}_6	1	500	5	1.7e-17	2.4e-08	5
		2	500	4	7.0e-17	9.5e-08	4
		3	500	5	9.6e-20	2.5e-09	5
<i>Ahn</i>	\mathbf{x}_1	1	3000	6	9.9e-22	1.4e-10	5
		2	3000	5	1.4e-16	9.9e-08	4
		3	3000	6	9.1e-22	1.3e-10	5

Tabla 7.1: Algoritmo 4.1 variando λ con puntos iniciales fijos.

Cabe mencionar que en todos los casos, el Algoritmo 4.1 usó la dirección de *Newton*,

lo cual es por supuesto deseable pues hace su convergencia más rápida. Cabe mencionar que nuestro algoritmo encontró la solución no degenerada del problema *Koj-Shi* y se encontraron soluciones distintas en el caso del problema *Math mod* con las tres variantes de λ .

Por otra parte, la **Tabla 7.1** permite observar que el criterio de parada, $\Psi_\lambda(\mathbf{x}_k) < 10^{-12}$, impuesto al algoritmo fue alcanzado en todos los casos, es decir hubo convergencia en dichos casos. Se observa además, que el criterio impuesto sobre $\|\nabla\Psi_\lambda(\mathbf{x}^k)\|$ también se alcanzó en prácticamente todos los casos.

Observemos que, en el 77% de las pruebas numéricas, el número de pasos de *Newton* completos (**np**) en la búsqueda monótona, supera o es igual a la mitad de iteraciones ($k/2$), lo que explica la rápida convergencia que tiene nuestro algoritmo, para este conjunto inicial de experimentos. Además, refleja lo mostrado teóricamente en el **Capítulo 6**, donde se demostró que cerca de la solución el paso de *Newton* es completo ($t_k = 1$).

En cuanto a las diferentes formas de elegir λ (**Métodos 1, 2 y 3**), se observa que el número de iteraciones no varía mucho. Por otro lado, en las pruebas para los problemas *Math mod*, *Nash Co (5)*, *Mathisen*, *Geiger-Kanzow* y *Ahn*, con alguno de los métodos, el número de pasos de *Newton* es igual al número de iteraciones, es decir, no se usa la búsqueda lineal o en otros términos, el algoritmo se comporta como un algoritmo local. Esta aparente ventaja nos motiva a hacer pruebas numéricas variando el punto inicial.

Continuando con las pruebas numéricas y con el propósito de analizar la convergencia global del **Algoritmo 4.1**, generamos 100 vectores aleatorios iniciales para cada uno de los problemas de complementariedad no lineal con cada método de elección de λ .

Los resultados de estos experimentos los presentamos en la **Tabla 7.2** que además del problema y métodos utilizados, contiene el tiempo promedio de ejecución medido en segundos (t); número promedio de iteraciones (k) y porcentaje de éxito obtenido en el experimento numérico (**Éxito (%)**), es decir, número de veces en las que el algoritmo encuentra la solución del problema.

De la columna **Éxito (%)**, observamos que el **Algoritmo 4.1** encontró una solución para la mayoría de los experimentos con una tasa de éxito del 100% a diferencia del problema *Math mod* que no convergió para aproximadamente el 15% de las pruebas. Es decir, nuestro algoritmo tiene un buen comportamiento global para este conjunto de experimentos.

Como se mencionó anteriormente, la **Tabla 7.1** no refleja gran diferencia en el número de

iteraciones entre los métodos, esto podría hacer pensar que la elección del parámetro λ no juega un gran papel en la convergencia del nuestro algoritmo, sin embargo, los resultados obtenidos en la **Tabla 7.2** dejan ver que no es así, pues el número de iteraciones con el **Método 3** es muy superior a los demás, en la mayoría de experimentos.

Problema	Método	k	t (seg)	Éxito (%)
<i>Koj-Shi</i>	1	15.79	0.0137	99
	2	14.87	0.0141	99
	3	12.85	0.0055	97
<i>Koj-Jo</i>	1	15.95	0.0161	100
	2	14.66	0.0170	100
	3	24.06	0.0101	100
<i>Math mod</i>	1	10.3165	0.0104	79
	2	9.5135	0.0123	74
	3	8.8028	0.0066	71
<i>Billups</i>	1	15.85	0.0048	100
	2	17.65	0.0093	100
	3	38.65	0.0040	61
<i>Nash Co (5)</i>	1	8.22	0.0095	100
	2	8.09	0.0148	100
	3	8.36	0.0096	100
<i>Nash Co (10)</i>	1	12.54	0.0252	100
	2	12.02	0.0325	100
	3	12.71	0.0398	100
<i>Mathiesen</i>	1	9.29	0.0052	100
	2	10.91	0.0129	99
	3	10.91	0.0056	100
<i>HH 66</i>	1	11.82	0.0107	100
	2	10.54	0.0128	100
	3	12.1	0.0096	100
<i>Geiger-Kanzow</i>	1	7	1.25	100
	2	6	1.342	100
	3	6.61	1.179	100
<i>Ahn</i>	1	6	128.804	100
	2	—	—	—
	3	—	—	—

Tabla 7.2: Algoritmo 4.1 variando λ con puntos iniciales aleatorios.

Se observa además, que para el problema *Ahn*, los **Métodos 2 y 3**,

Con el objetivo de analizar el desempeño global del **Algoritmo 4.1**, lo comparamos con tres métodos globales que resuelven problemas de complementariedad lineal. El primero de estos algoritmos, que llamaremos **Algoritmo 5.2**, propuesto en [4], es de tipo cuasi *Newton* no diferenciable.

El segundo, propuesto en [20], es de tipo jacobiano suavizado, el cual a diferencia de nuestra propuesta usa una suavización de la función de *Fischer* (un caso particular de la familia de funciones de complementariedad considerada en nuestro trabajo) para reformular el problema de complementariedad no lineal. Llamaremos a este algoritmo, **Algoritmo 5.3**.

El tercero, que llamaremos **Algoritmo 5.4**, de tipo *Newton* suavizado fue propuesto recientemente en [36], usa la familia de funciones de complementariedad (1.4) y la suavización presentada en [35], para la reformulación del problema de complementariedad no lineal. Su propuesta algorítmica difiere considerablemente de la nuestra: ellos aproximan la parte que contiene la información de segundo orden en la matriz hessiana de la función de mérito, por un múltiplo de la matriz identidad (tipo algoritmos *Levenberg-Marquardt*, para mínimos cuadrados no lineales [13]), lo que hace que puedan garantizar la no singularidad de la matriz del sistema lineal a resolver en cada iteración, evitando así el uso de la dirección del gradiente de la función de mérito.

Previo a la comparación de los algoritmos, procedemos de la siguiente forma: consideramos el problema de complementariedad no lineal asociado a cada una de las diez funciones descritas al inicio del capítulo, hacemos pruebas numéricas del **Algoritmo 4.1** con vectores iniciales fijos usados en [4] y [36] y variando λ con los **Métodos 1,2 y 3**. Medimos el promedio de iteraciones, tiempo promedio de ejecución del algoritmo y el porcentaje de éxito en el experimento numérico. Los resultados obtenidos los reportamos en la tabla **Tabla 7.3**, a continuación.

Problema	Método	\mathbf{x}_0	\bar{t} (seg)	\bar{k}	Éxito (%)
<i>Koj-Shi</i>	1	\mathbf{x}_7	0.0090	14.00	100
<i>Koj-Shi</i>	2	\mathbf{x}_7	0.0064	11.00	100
<i>Koj-Shi</i>	3	\mathbf{x}_7	0.0062	11.82	95
<i>Koj-Shi</i>	1	\mathbf{x}_8	0.0062	11.00	100
<i>Koj-Shi</i>	2	\mathbf{x}_8	0.0069	12.00	100
<i>Koj-Shi</i>	3	\mathbf{x}_8	0.0081	13.92	97

Continúa en la siguiente página

Continuación de la página anterior

Problema	Método	\mathbf{x}_0	\bar{t} (seg)	\bar{k}	Éxito (%)
<i>Koj-Shi</i>	1	\mathbf{x}_9	0.0057	10.00	100
<i>Koj-Shi</i>	2	\mathbf{x}_9	0.0062	10.00	100
<i>Koj-Shi</i>	3	\mathbf{x}_9	0.0412	30.56	93
<i>Koj-Jo</i>	1	\mathbf{x}_{10}	0.0354	31.00	100
<i>Koj-Jo</i>	2	\mathbf{x}_{10}	0.0360	31.00	100
<i>Koj-Jo</i>	3	\mathbf{x}_{10}	0.0381	32.58	99
<i>Koj-Jo</i>	1	\mathbf{x}_{11}	0.0025	6.00	100
<i>Koj-Jo</i>	2	\mathbf{x}_{11}	0.0027	6.00	100
<i>Koj-Jo</i>	3	\mathbf{x}_{11}	0.0026	6.08	100
<i>Koj-Jo</i>	1	\mathbf{x}_{12}	0.0041	10.00	100
<i>Koj-Jo</i>	2	\mathbf{x}_{12}	0.0046	10.00	100
<i>Koj-Jo</i>	3	\mathbf{x}_{12}	0.0044	10.31	100
<i>Math mod (4)</i>	1	\mathbf{x}_{10}	0.0129	9.00	100
<i>Math mod (4)</i>	2	\mathbf{x}_{10}	0.0130	9.00	100
<i>Math mod (4)</i>	3	\mathbf{x}_{10}	0.0094	10.68	96
<i>Math mod (4)</i>	1	\mathbf{x}_1	0.0017	4.00	100
<i>Math mod (4)</i>	2	\mathbf{x}_1	0.0018	4.00	100
<i>Math mod (4)</i>	3	\mathbf{x}_1	0.0016	3.84	100
<i>Math mod (4)</i>	1	\mathbf{x}_{11}	0.0017	4.00	100
<i>Math mod (4)</i>	2	\mathbf{x}_{11}	0.0015	3.00	100
<i>Math mod (4)</i>	3	\mathbf{x}_{11}	0.0015	3.64	100
<i>Billups</i>	1	\mathbf{x}_{14}	0.0092	20.00	100
<i>Billups</i>	2	\mathbf{x}_{14}	0.0097	19.00	100
<i>Billups</i>	3	\mathbf{x}_{14}	0.0065	16.62	100
<i>Billups</i>	1	\mathbf{x}_2	0.0005	4.00	100
<i>Billups</i>	2	\mathbf{x}_2	0.0008	5.00	100
<i>Billups</i>	3	\mathbf{x}_2	0.0005	3.99	100
<i>Nash Co (5)</i>	1	\mathbf{x}_{15}	0.0097	8.00	100
<i>Nash Co (5)</i>	2	\mathbf{x}_{15}	0.0103	8.00	100
<i>Nash Co (5)</i>	3	\mathbf{x}_{15}	0.0095	8.50	100
<i>Nash Co (5)</i>	1	\mathbf{x}_3	0.0069	6.00	100
<i>Nash Co (5)</i>	2	\mathbf{x}_3	0.0081	6.00	100
<i>Nash Co (5)</i>	3	\mathbf{x}_3	0.0070	6.07	100
<i>Nash Co (5)</i>	1	\mathbf{x}_{16}	0.0117	9.00	100
<i>Nash Co (5)</i>	2	\mathbf{x}_{16}	0.0131	9.00	100
<i>Nash Co (5)</i>	3	\mathbf{x}_{16}	0.0118	9.24	100

Continúa en la siguiente página

Continuación de la página anterior

Problema	Método	\mathbf{x}_0	\bar{t} (seg)	\bar{k}	Éxito (%)
<i>Nash Co (10)</i>	1	\mathbf{x}_{17}	0.0120	8.00	100
<i>Nash Co (10)</i>	2	\mathbf{x}_{17}	0.0117	7.00	100
<i>Nash Co (10)</i>	3	\mathbf{x}_{17}	0.0109	7.82	100
<i>Nash Co (10)</i>	1	\mathbf{x}_{18}	0.0181	11.00	100
<i>Nash Co (10)</i>	2	\mathbf{x}_{18}	0.0189	10.00	100
<i>Nash Co (10)</i>	3	\mathbf{x}_{18}	0.0194	11.31	100
<i>Nash Co (10)</i>	1	\mathbf{x}_{19}	0.0981	7.00	100
<i>Nash Co (10)</i>	2	\mathbf{x}_{19}	0.0907	7.00	100
<i>Nash Co (10)</i>	3	\mathbf{x}_{19}	0.1100	8.90	99
<i>Mathiesen</i>	1	\mathbf{x}_6	0.0034	8.00	100
<i>Mathiesen</i>	2	\mathbf{x}_6	0.0036	8.00	100
<i>Mathiesen</i>	3	\mathbf{x}_6	0.0188	19.50	100
<i>Mathiesen</i>	1	\mathbf{x}_{20}	0.0021	5.00	100
<i>Mathiesen</i>	2	\mathbf{x}_{20}	0.0020	4.00	100
<i>Mathiesen</i>	3	\mathbf{x}_{20}	0.0021	4.67	100
<i>Mathiesen</i>	1	\mathbf{x}_{21}	0.0115	11.00	100
<i>Mathiesen</i>	2	\mathbf{x}_{21}	—	—	—
<i>Mathiesen</i>	3	\mathbf{x}_{21}	0.0355	21.18	96
<i>HH 66</i>	1	\mathbf{x}_6	0.0077	9.00	100
<i>HH 66</i>	2	\mathbf{x}_6	0.0063	8.00	100
<i>HH 66</i>	3	\mathbf{x}_6	0.0062	9.53	100
<i>HH 66</i>	1	\mathbf{x}_4	0.0053	8.00	100
<i>HH 66</i>	2	\mathbf{x}_4	0.0049	7.00	100
<i>HH 66</i>	3	\mathbf{x}_4	0.0056	8.89	100
<i>HH 66</i>	1	\mathbf{x}_{16}	0.0045	8.00	100
<i>HH 66</i>	2	\mathbf{x}_{16}	0.0042	7.00	100
<i>HH 66</i>	3	\mathbf{x}_{16}	0.0049	8.80	100
<i>Geiger-Kanzow</i>	1	\mathbf{x}_6	1.2476	5.00	100
<i>Geiger-Kanzow</i>	2	\mathbf{x}_6	1.0896	4.00	100
<i>Geiger-Kanzow</i>	3	\mathbf{x}_6	2.8282	13.97	99
<i>Geiger-Kanzow</i>	1	\mathbf{x}_4	1.1148	5.00	100
<i>Geiger-Kanzow</i>	2	\mathbf{x}_4	0.8795	4.00	100
<i>Geiger-Kanzow</i>	3	\mathbf{x}_4	4.2205	15.03	88

Tabla 7.3: Algoritmo 4.1 con los Métodos 1,2 y 3.

Para la comparación con el **Algoritmo 5.2**, tenemos en cuenta los resultados de la **Tabla 7.3** asociados a los problemas *Koj-Jo*, *Koj-Shi*, *Math mod* y *Billups* que también fueron utilizados en [4] para analizar el desempeño de su algoritmo. Además, hubo convergencia para algunos puntos iniciales para los cuales, el **Algoritmo 5.2** no convergió. Por ejemplo, el problema *Koj-Jo* con el punto inicial \mathbf{x}_{10} y *Billups* con valor inicial \mathbf{x}_2 .

Para la comparación con el **Algoritmo 5.3**, el cual recordemos es un caso particular de nuestro algoritmo ($\lambda = 2$) tenemos en cuenta los resultados de la **Tabla 7.3** con los **Métodos 2 y 3** versus los resultados de la misma tabla, con el **Método 1**.

Observamos que el número de iteraciones y tiempo de ejecución con los **Métodos 1 y 2** es muy similar, prácticamente no hay una diferencia significativa. En cuanto al algoritmo con el **Método 1** frente al mismo con el **Método 3** observamos diferencias, en algunos casos, en el número de iteraciones, mostrando un mejor desempeño cuando se considera el **Método 1**, tal es el caso de los problemas *Mathiesen* (8 iteraciones en promedio con el **Método 1** y 19.50 con el **Método 3**); *Geiger-Kanzow* (5 iteraciones en promedio con el **Método 1** y 13.97 con el **Método 3**). Por otra parte, el tiempo de ejecución, con estos dos métodos es muy similar. Vale la pena mencionar que el porcentaje de éxito cuando se usa el **Método 3**, en general, es menor que 100, a diferencia de cuando se usan los **Métodos 1 y 2** que el porcentaje es 100 %.

Teniendo en cuenta los resultados para los problemas *Koj-Shi*, *Mathiesen*, *HH 66*, *Geiger-Kanzow* y *Ahn*, reportados en la **Tabla 7.3** para el **Algoritmo 4.1** y en [36], para el **Algoritmo 5.4**, encontramos que en todos los casos nuestro algoritmo con los **Métodos 1 y 2** encontró una solución en menos iteraciones. No obstante, con el **Método 3** superó, en general, el número de iteraciones reportado en [36]. Quizá esto se deba al carácter aleatorio del parámetro λ .

Finalmente, para apreciar la ventaja de disponer de un algoritmo global frente a uno local, comparamos el **Algoritmo 4.1** con una versión local del mismo. Esta última, que llamaremos **Algoritmo Local**, la obtenemos suprimiendo el paso **P3** de nuestro algoritmo que corresponde a la búsqueda lineal no monótona, la estrategia de globalización usada.

Los resultados obtenidos por el **Algoritmo local** en cuanto a convergencia y número de iteraciones son presentados en la **Tabla 7.4**, en la cual el símbolo – indica que hubo divergencia ya que el algoritmo superó el número de iteraciones permitido.

Problema	\mathbf{x}_0	Método	Dim	k	$\Psi_\lambda(\mathbf{x}^*)$	$\ \nabla\Psi_\lambda(\mathbf{x}^k)\ $
<i>Koj-Shi</i>	\mathbf{x}_1	1	4	34	1.1e-17	4.1e-08
		2	4	32	1.3e-25	7.7e-12
		3	4	-	-	-
<i>Koj-Jo</i>	\mathbf{x}_1	1	4	9	2.8e-18	6.7e-09
		2	4	9	1.1e-20	2.2e-09
		3	4	-	-	-
<i>Math mod</i>	\mathbf{x}_1	1	4	4	6.8e-18	3.7e-09
		2	4	4	5.6e-32	6.3e-16
		3	4	4	7.2e-22	5.9e-11
<i>Billups</i>	\mathbf{x}_2	1	1	6	6.1e-22	7.3e-11
		2	1	7	2.3e-23	2.9e-11
		3	1	-	-	-
<i>Nash Co (5)</i>	\mathbf{x}_3	1	5	7	2.3e-16	3.5e-07
		2	5	7	2.8e-27	2.4e-12
		3	5	7	1.8e-20	4.3e-09
<i>Nash Co (10)</i>	\mathbf{x}_4	1	10	-	-	-
		2	10	-	-	-
		3	10	-	-	-
<i>Mathiesen</i>	\mathbf{x}_5	1	4	8	8.9e-18	6.0e-09
		2	4	7	8.2e-21	3.6e-10
		3	4	7	7.6e-13	2.0e-06
<i>HH 66</i>	\mathbf{x}_6	1	8	11	8.4e-13	4.5e-06
		2	8	10	2.5e-16	1.5e-07
		3	8	12	1.3e-15	2.9e-07
<i>Geiger-Kanzow</i>	\mathbf{x}_6	1	500	5	1.7e-17	2.4e-08
		2	500	4	7.0e-17	9.5e-08
		3	500	55	1.4e-12	7.8e-07
<i>Ahn</i>	\mathbf{x}_1	1	3000	6	1.0e-15	2.0e-07
		2	3000	6	7.4e-29	7.4e-14
		3	3000	-	-	-

Tabla 7.4: Desempeño del Algoritmo local.

En la **Tabla 7.4** observamos que el **Algoritmo local** no convergió para el 23.3% de los problemas, en contraste a los resultados reportados en **Tabla 7.1** para el **Algoritmo global 4.1** donde se observa que hubo convergencia del 100% de los problemas analizados.

Respecto al número de iteraciones, observamos un claro aumento de estas, en el **Algoritmo local** con las diferentes variantes de λ . Por ejemplo, el problema *Koj-Shi*, el número de iteraciones se cuadruplicó, respecto al mismo problema con el **Algoritmo global 4.1**. Además, los datos reportados en la **Tabla 7.4** muestran que el **Método 3**, no convergió para el 50% de los experimentos numéricos, lo anterior evidencia el papel importante que juega la estrategia de globalización, en este caso la búsqueda lineal monótona.

En el **capítulo 6** se demostró, bajo ciertas hipótesis, que el **Algoritmo 4.1** converge superlineal y, hasta cuadráticamente, lo que es muy deseable para cualquier método iterativo. Esto nos motivó a realizar pruebas numéricas que evidencien estos tipos de convergencia; con este fin elegimos las funciones *Koj-Shi*, *Math Mod*, *Billups* y *Ahn* y los problemas asociados a cada una de estas funciones.

Las Figuras 7.1 a 7.4 ilustran la convergencia superlineal del **Algoritmo 4.1** para cada uno de los cuatro problemas considerados. En cada caso, la gráfica presenta el número de iteraciones versus el cociente R_k , definido por

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

Para ilustrar la convergencia cuadrática del **Algoritmo 4.1**, con la escogencia de λ dada por el **Método 2**, presentamos en las **Tablas 7.5 a 7.8**, las diez primeras iteraciones del algoritmo y registramos el valor de la función de mérito Ψ_λ en \mathbf{x}_k . Cabe mencionar que para la mayoría de problemas el algoritmo converge en menos de diez iteraciones, pero esto se hace con el objetivo de evidenciar en forma clara dicho tipo de convergencia.

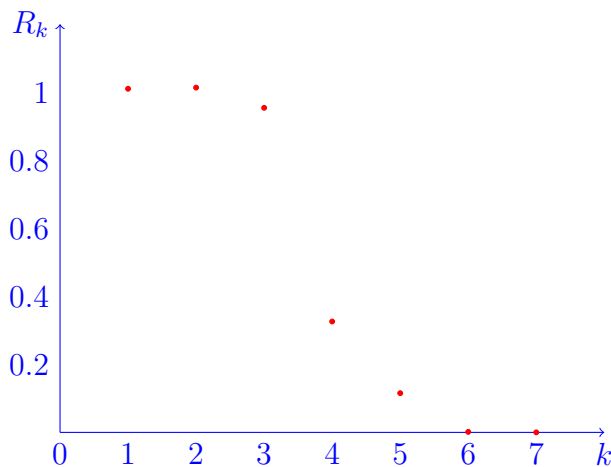


Figura 7.1: Koj-Shi

k	$\Psi_\lambda(\mathbf{x}^k)$
1	1.7e+00
2	9.4e-01
3	4.0e-01
4	2.2e-01
5	2.4e-01
6	9.6e-03
7	1.5e-07
8	4.4e-17
9	1.6e-30
10	0.0e+00

Tabla 7.5: Koj-Shi

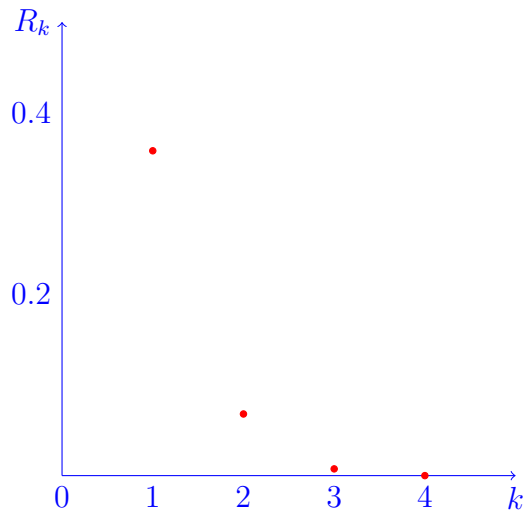


Figura 7.2: Math Mod

k	$\Psi_\lambda(\mathbf{x}^k)$
1	1.0e+01
2	2.0e-01
3	1.8e-03
4	3.9e-12
5	5.6e-32
6	5.6e-32
7	5.6e-32
8	5.6e-32
9	5.6e-32
10	5.6e-32

Tabla 7.6: Math Mod

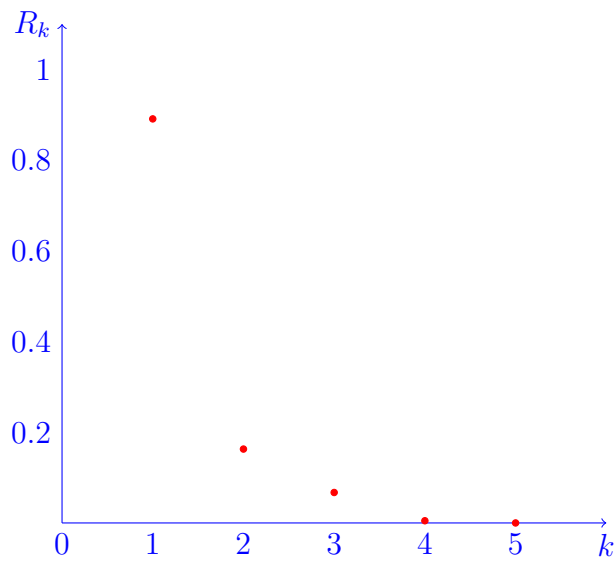


Figura 7.3: Billups

k	$\Psi_\lambda(\mathbf{x}^k)$
1	1.3e+00
2	1.4e+00
3	1.4e-01
4	9.3e-04
5	2.1e-08
6	1.2e-17
7	0.0e+00
8	0.0e+00
9	0.0e+00
10	0.0e+00

Tabla 7.7: Billups

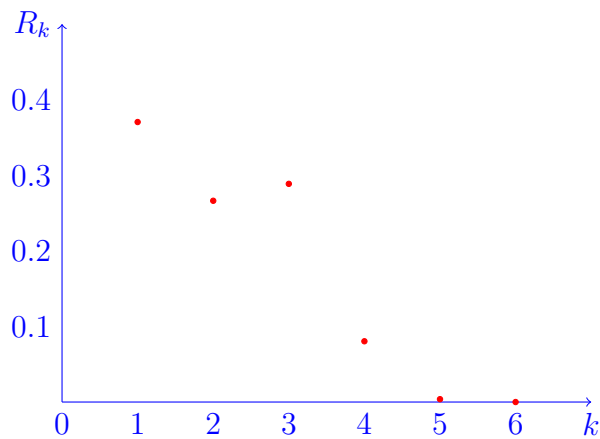


Figura 7.4: Ahn

k	$\Psi_\lambda(\mathbf{x}^k)$
1	8.8e+02
2	2.2e+02
3	3.4e+01
4	5.9e+00
5	1.0e-01
6	4.4e-06
7	3.8e-23
8	2.9e-28
9	7.4e-29
10	7.3e-29

Tabla 7.8: Ahn

De las Figuras 7.1 a 7.4 observamos que, en todos los casos, el cociente R_k tiende a cero cuando k se hace muy grande, es decir que en los problemas *Koj-Shi*, *Math Mod*, *Billups* y *Ahn*, nuestro algoritmo converge superlinealmente¹

Por otra parte, De las **Tablas 7.5 a 7.8** observamos que el número de decimales exactos de la función de mérito $\Psi(\mathbf{x}_k)$, en cada uno de los problemas, se dobla a partir de una cierta iteración, esto indica que el algoritmo converge cuadráticamente.

¹Una sucesión $\{\mathbf{x}_k\}$ converge superlinealmente a \mathbf{x}^* , si $\lim_{k \rightarrow \infty} R_k = 0$. Dicha sucesión converge cuadráticamente, si para algún $c > 0$ se satisface que $\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq c \|\mathbf{x}_k - \mathbf{x}^*\|^2$.

Comentarios finales

Resolver el PCNL(F) trae consigo un inconveniente, ya que tratar de resolverlo directamente es poco práctico, es por ello, que se buscan diferentes estrategias para resolverlo; la más popular consiste en reformularlo, mediante el uso de funciones de complementariedad, como un sistema de ecuaciones no lineales que resulta no diferenciable. Esto último impide la aplicación de algoritmos clásicos tipo *Newton* para resolver el problema, pero ha motivado la creación de los llamados métodos tipo *Newton* generalizados para su solución [32] [19] [4].

Otra forma de lidiar con la no diferenciabilidad de la reformulación del problema de complementariedad es usar la llamada estrategia del jacobiano suavizado [10] [20], que permite aproximar la reformulación no diferenciable del PCNL(F) por una sucesión de sistemas no lineales diferenciables que dependen de un parámetro positivo μ . La idea es hacer que este parámetro tienda a cero. En la versión local del método se resuelve, en cada iteración, una ecuación de *Newton* híbrida, cuya matriz de coeficientes es la matriz jacobiana de la aproximación suavizada.

En el presente trabajo de investigación propusimos un algoritmo global con jacobiano suavizado que generaliza el algoritmo del mismo tipo, propuesto en [20] a todos los miembros de la familia φ_λ definida en (1.3). Desarrollamos la teoría de convergencia del método y realizamos pruebas numéricas para analizar el desempeño del método. Nuestra propuesta resuelve el PCNL(F) y ofrece ventajas en cuanto a la robustez del método, pues el uso de puntos iniciales aleatorios frente a otros métodos [36] [4], evidenció en las pruebas numéricas presentadas en el **Capítulo 7** un comportamiento similar al propuesto

en [20], además se amplía el margen de vectores que se pueden tomar como punto inicial, lo cual está acuerdo con el carácter global del algoritmo.

Finalmente, queremos mencionar que las variaciones realizadas a nuestro algoritmo nos permitieron formular conclusiones a priori sobre el desempeño del mismo, por lo tanto, creemos que es necesario estudiar una relación entre el parámetro λ y el problema de complementariedad no lineal, que permita mejorar la efectividad en cuanto a convergencia del algoritmo. También sería interesante aplicar la estrategia del jacobiano suavizado usando la familia (1.4).

Bibliografía

- [1] Jozsef Abaffy, Charles Broyden, and Emilio Spedicato. A class of direct methods for linear systems. *Numerische Mathematik*, 45(3):361–376, 1984.
- [2] Mihai Anitescu, James F Cremer, and Florian A Potra. On the existence of solutions to complementarity formulations of contact problems with friction. *Complementary and Variational Problems: State of the Art*, pages 12–21, 1997.
- [3] Favián E Arenas, Héctor J Martínez, and Rosana Pérez. Un método de Newton con jacobiano suavizado para problemas de complementariedad no lineal. Reporte técnico, Universidad del Cauca, 2018.
- [4] Carlos A Arias. Un algoritmo cuasi Newton global para problemas de complementariedad no lineal. Tesis de maestría, Universidad del Cauca, 2014.
- [5] Stephen C Billups. Algorithms for complementarity problems and generalized equations. Technical report, 1995.
- [6] Charles George Broyden, JE Dennis Jr, and Jorge J Moré. On the local and superlinear convergence of quasi-Newton methods. *IMA Journal of Applied Mathematics*, 12(3):223–245, 1973.
- [7] Sandra Buhmiller and Nataša Krejić. A new smoothing quasi-Newton method for nonlinear complementarity problems. *Journal of Computational and Applied Mathematics*, 211(2):141–155, 2008.
- [8] AHN Byong-Hun. Iterative methods for linear complementarity problems with upperbounds on primary variables. *Mathematical Programming*, 26(3):295–315, 1983.

-
- [9] Anthony 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(7):3020–3031, 2010.
- [10] Xiaojun Chen, Liqun Qi, and Defeng Sun. Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities. *Mathematics of Computation of the American Mathematical Society*, 67(222):519–540, 1998.
- [11] Frank H Clarke. Generalized gradients and applications. *Transactions of the American Mathematical Society*, 205:247–262, 1975.
- [12] Frank H Clarke. *Optimization and nonsmooth analysis*, volume 5. Siam, 1990.
- [13] John E Dennis Jr and Robert B Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16. Siam, 1996.
- [14] Michael C Ferris and Jong-Shi Pang. Engineering and economic applications of complementarity problems. *Siam Review*, 39(4):669–713, 1997.
- [15] Andreas Fischer and Christian Kanzow. On finite termination of an iterative method for linear complementarity problems. *Mathematical Programming*, 74(3):279–292, 1996.
- [16] Carl Geiger and Christian Kanzow. On the resolution of monotone complementarity problems. *Computational Optimization and Applications*, 5(2):155–173, 1996.
- [17] Patrick T Harker. Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities. *Mathematical Programming*, 41(1-3):29–59, 1988.
- [18] Sheng-Long Hu, Zheng-Hai Huang, and Jein-Shan Chen. Properties of a family of generalized ncp-functions and a derivative free algorithm for complementarity problems. *Journal of Computational and Applied Mathematics*, 230(1):69–82, 2009.
- [19] Christian Kanzow and Helmut Kleinmichel. A new class of semismooth Newton-type methods for nonlinear complementarity problems. *Computational Optimization and Applications*, 11(3):227–251, 1998.
- [20] Christian Kanzow and Heiko Pieper. Jacobian smoothing methods for nonlinear complementarity problems. *SIAM Journal on Optimization*, 9(2):342–373, 1999.
- [21] Michael M Kostreva. Elasto-hydrodynamic lubrication: A non-linear complementarity problem. *International Journal for numerical methods in fluids*, 4(4):377–397, 1984.

-
- [22] 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(1-4):71–97, 2001.
- [23] Vera Lúcia Rocha Lopes, José Mario Martínez, and Rosana Pérez. On the local convergence of quasi-Newton methods for nonlinear complementarity problems. *Applied Numerical Mathematics*, 30(1):3–22, 1999.
- [24] Changfeng Ma. A new smoothing quasi-Newton method for nonlinear complementarity problems. *Applied mathematics and computation*, 171(2):807–823, 2005.
- [25] Lars Mathiesen. An algorithm based on a sequence of linear complementarity problems applied to a walrasian equilibrium model: An example. *Mathematical programming*, 37(1):1–18, 1987.
- [26] Jorge J Moré and Danny C Sorensen. Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing*, 4(3):553–572, 1983.
- [27] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.
- [28] Jong-Shi Pang and Liqun Qi. Nonsmooth equations: motivation and algorithms. *SIAM Journal on optimization*, 3(3):443–465, 1993.
- [29] Rosana Pérez and Vera Lucia Rocha Lopes. Recent applications and numerical implementation of quasi-Newton methods for solving nonlinear systems of equations. *Numerical Algorithms*, 35(2-4):261–285, 2004.
- [30] L Qi. C-differentiability, C-differential operators and generalized Newton methods. Technical report, Applied Mathematics Report AMR96/5, University of New South Wales, Sydney, Australia, 1996.
- [31] Liqun Qi. Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of operations research*, 18(1):227–244, 1993.
- [32] Liqun Qi and Jie Sun. A nonsmooth version of Newton’s method. *Mathematical programming*, 58(1-3):353–367, 1993.
- [33] Nicol N Schraudolph, Jin Yu, and Simon Günter. A stochastic quasi-Newton method for online convex optimization. In *Artificial Intelligence and Statistics*, pages 436–443, 2007.
- [34] Andrew H Sherman. On Newton-iterative methods for the solution of systems of nonlinear equations. *SIAM Journal on Numerical Analysis*, 15(4):755–771, 1978.

-
- [35] Jianguang Zhu and Binbin Hao. A new class of smoothing functions and a smoothing Newton method for complementarity problems. *Optimization Letters*, 7(3):481–497, 2013.
- [36] Jianguang Zhu and Binbin Hao. A new smoothing method for solving nonlinear complementarity problems. *Open Mathematics*, 17(1):104–119, 2019.