

Un algoritmo local cuasi-Newton suavizado para resolver el problema de complementariedad no lineal

A local smoothing quasi-Newton algorithm for solving nonlinear complementarity problem

Camila Quirá ¹, Rosana Pérez ^{2a}, Favián Arenas ^{2b}, Diego Correa ^{2c}

¹ Departamento de Matemáticas, Universidad Autónoma Indígena Intercultural, Popayán, Colombia. Orcid: 0000-0001-7016-7816. Correo electrónico: c.quirá@uaiinpebi-cric.edu.co

² Grupo de Optimización, Departamento de Matemáticas, Universidad del Cauca, Popayán, Colombia. Orcid 0000-0003-0279-8522 ^a, 0000-0002-7781-7559 ^b, 0000-0002-1766-7851 ^c. Correos electrónicos: rosana@unicauca.edu.co ^a, farenas@unicauca.edu.co ^b, dicorrea@unicauca.edu.co ^c

Recibido: 2 febrero, 2023. Aceptado: 16 junio, 2023. Versión final: 20 noviembre, 2023.

Resumen

Debido a la importancia y efectividad del método con jacobiano suavizado o cuasi-Newton suavizado para resolver indirectamente el problema de complementariedad no lineal, y que la función de complementariedad Mínimo no ha sido usada en conexión con dichos métodos, en el presente trabajo se propone un algoritmo de ese tipo, con dicha función y una suavización de la misma. Se demuestra que bajo ciertas hipótesis el algoritmo propuesto converge local y q-cuadráticamente. Además, se presentan pruebas numéricas que muestran un buen desempeño del algoritmo.

Palabras clave: complementariedad no lineal; función de complementariedad; jacobiano suavizado; cuasi-Newton suavizado; convergencia cuadrática; problemas de complementariedad no lineal; programación no lineal; métodos cuasi Newton; sistemas de ecuaciones no lineales; jacobiano generalizado.

Abstract

Due to the importance and effectiveness of a Jacobian smoothing method or smoothing quasi-Newton method for solving the nonlinear complementarity problem and that the Minimum complementarity function has not been used in connection with such methods, in the present work, we propose an algorithm of that type, which uses the minimum function and a smoothing of it. We show that under certain hypotheses the proposed algorithm converges locally and q-quadratically. In addition, we present numerical tests that show a good performance of the algorithm.

Keywords: nonlinear complementarity; minimum complementarity function; Jacobian smoothing method; smoothing quasi-Newton method; quadratic convergence; nonlinear complementarity problems; nonlinear programming; Quasi-Newton methods; systems of nonlinear equations; generalized jacobian.

1. Introducción

El Problema de Complementariedad No Lineal, asociado a $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))^T$, no lineal y

continuamente diferenciable, (PCNL(F)), consiste en encontrar un vector $\mathbf{x} \in \mathbb{R}^n$ tal que

$$x_i \geq 0, F_i(x) \geq 0, x_i F_i(x) = 0, \forall i = 1, \dots, n. \quad (1)$$

La tercera condici3n en (1) exige que \mathbf{x} y $F(\mathbf{x})$ sean ortogonales, y es la que le da el calificativo de ‘‘complementariedad’’ al problema. Son numerosas sus aplicaciones en ingeniera, economa y fısica [1], [2], [3], [4]. Su importancia se debe a que el concepto de complementariedad es sin3nimo de sistema en equilibrio.

La estrategia mas popular para resolver el PCNL(F) consiste en reformularlo como un sistema de ecuaciones no lineales mediante un operador $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$, definido por $\Phi(\mathbf{x}) = (\varphi(x_1, F_1(\mathbf{x})), \dots, \varphi(x_n, F_n(\mathbf{x})))^T$, con $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$, llamada *funci3n de complementariedad*, que satisface la equivalencia: $\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0$, la cual permite demostrar que resolver el PCNL(F) es equivalente a resolver el sistema de ecuaciones no lineales [5]:

$$\Phi(x) = (\varphi(x_1, F_1(x)), \dots, \varphi(x_n, F_n(x)))^T = 0 \quad (2)$$

que resulta ser no diferenciable debido a la no suavidad de la funci3n de complementariedad φ . El sistema (2), se conoce como reformulaci3n del PCNL(F) como un sistema de ecuaciones no lineales.

Existen numerosas funciones de complementariedad [6], [7], [8] pero, quiza, las mas populares debido a su amplio uso en problemas de complementariedad son las funciones Mınimo [9] y Fischer [10] definidas, respectivamente, por: $\varphi(a, b) = \min\{a, b\}$ y $\varphi(a, b) = \sqrt{a^2 + b^2} - a - b$.

Para resolver el sistema (2) e indirectamente el PCNL(F), se han propuesto algoritmos tipo Newton generalizados [11], cuasi-Newton generalizados [12], [13] y recientemente, algoritmos suavizados y con jacobiano suavizado [14], [15]. Estos ultimos como alternativa a la no diferenciable de la reformulaci3n del PCNL(F).

La idea basica de los metodos suavizados es aproximar Φ por $\Phi_\mu: \mathbb{R}^n \rightarrow \mathbb{R}^n, \mu > 0$, definido por $\Phi_\mu(\mathbf{x}) = (\varphi_\mu(x_1, F_1(\mathbf{x})), \dots, \varphi_\mu(x_n, F_n(\mathbf{x})))^T$, donde φ_μ es una suavizaci3n de la funci3n de complementariedad φ , y ası, aproximar el sistema de ecuaciones no lineales (2) por una sucesi3n de sistemas suavizados $\Phi_\mu(\mathbf{x}) = 0$ y hacer tender el parametro μ a cero. Estos metodos han resultado muy eficientes, en particular los llamados en [14] de *jacobiano suavizado* que se caracterizan por resolver, en cada iteraci3n, un sistema de ecuaciones lineales suavizado de la forma:

$$\Phi'_\mu(\mathbf{x}_k)\mathbf{s}_k = -\Phi(\mathbf{x}_k), \quad (3)$$

En lugar del sistema Newtoniano,

$$H_k\mathbf{s}_k = -\Phi(\mathbf{x}_k), \quad (4)$$

Con $H_k \in \partial\Phi(\mathbf{x}_k)$ o $H_k \in \partial_C\Phi(\mathbf{x}_k)$, donde $\partial\Phi(\mathbf{x})$ es el jacobiano generalizado de Φ en \mathbf{x} , definido para Φ Lipschitz continua como el siguiente conjunto [16], $\partial\Phi(\mathbf{x}) = \text{conv}\{H = \lim_{k \rightarrow \infty} \Phi'(\mathbf{x}_k) \in \mathbb{R}^{n \times n} : \mathbf{x}_k \rightarrow \mathbf{x}, \mathbf{x}_k \in D_\Phi\}$, donde D_Φ es el conjunto de puntos de \mathbb{R}^n en los que Φ es diferenciable y $\text{conv}\{A\}$ es la envolvente convexa del conjunto A . En general, $\partial\Phi(\mathbf{x})$ es difıcil de calcular. Una alternativa basada en el hecho de que $\partial\Phi(\mathbf{x})^T \subseteq \partial\Phi_1(\mathbf{x}) \times \dots \times \partial\Phi_n(\mathbf{x})$ [16], es usar el C-subdiferencial [9] de Φ en \mathbf{x} definido por, $\partial_C\Phi(\mathbf{x})^T = \partial\Phi_1(\mathbf{x}) \times \dots \times \partial\Phi_n(\mathbf{x})$. Es decir, el conjunto de matrices en $\mathbb{R}^{n \times n}$ cuya i -esima columna es el gradiente generalizado de la i -esima funci3n componente Φ_i .

A partir de (3) y (4), se infiere que los metodos con jacobiano suavizado son cuasi-Newton suavizados. Por ello, en este documento se dira metodos con jacobiano suavizado o cuasi-Newton suavizados.

Debido a la importancia y efectividad de los metodos cuasi-Newton suavizados para resolver indirectamente el PCNL(F), y que la funci3n Mınimo no ha sido usada en conexi3n con dichos metodos, en el presente trabajo se propone un algoritmo de este tipo, en el cual se usa la funci3n Mınimo y una suavizaci3n de la misma propuesta en [17]. Se demuestra, bajo ciertas hip3tesis, que el algoritmo propuesto converge local y q -cuadraticamente. Ademas, se presentan pruebas numericas que muestran el buen desempeno del algoritmo propuesto.

Este documento se organiza de la siguiente forma. En la Secci3n 2 se presentan algunos resultados te3ricos importantes acerca de la funci3n Mınimo y su suavizaci3n; tambien, se deducen cotas ıtiles para el desarrollo te3rico presentado en la Secci3n 4. En la Secci3n 3 se reformula el PCNL(F) como un sistema de ecuaciones no lineales usando la funci3n Mınimo y se caracteriza la matriz jacobiana de Φ_μ en un punto \mathbf{x} . En la primera parte de la Secci3n 4, se propone un algoritmo cuasi-Newton suavizado para resolver el PCNL(F) y luego, bajo ciertas hip3tesis se desarrolla su teora de convergencia local. En la Secci3n 5 se analiza numericamente el desempeno local del algoritmo propuesto. Finalmente, en la Secci3n 6, se presentan algunos comentarios finales y propuestas de trabajos futuros.

2. La funci3n Mınimo y su suavizaci3n

En esta secci3n se analizan algunas propiedades de la funci3n Mınimo y de una suavizaci3n de la misma propuesta en [17]. Ademas, se deducen algunas cotas

útiles en el desarrollo teórico posterior. Para mayor claridad en la lectura del capítulo se incluyen a continuación, a manera de definición, la función Mínimo y el concepto de suavización de una función de complementariedad.

Definición 1. La función $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por

$$\varphi(a, b) = \min\{a, b\}, \quad (5)$$

se conoce como función *Mínimo*.

La función *Mínimo* se puede escribir, equivalentemente, como:

$$\begin{aligned} \varphi(a, b) &= \frac{1}{2}((a + b) - |a - b|) \\ &= \frac{1}{2}\left((a + b) - \sqrt{(a - b)^2}\right), \end{aligned} \quad (6)$$

que permite observar su no diferenciabilidad en los puntos de la forma (a, a) como lo ilustra la [Figura 1](#).

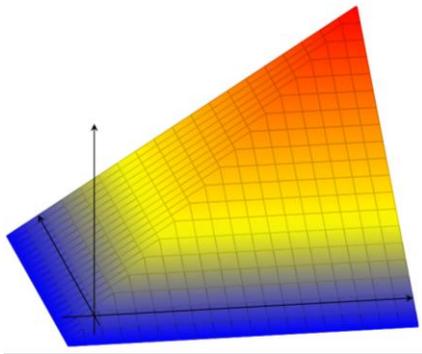


Figura 1. La superficie $\varphi(a, b) = \min(a, b)$

Definición 2. Una suavización de una función de complementariedad φ , es una función $\varphi_\mu: \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

$$\varphi_\mu(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = \mu. \quad (7)$$

A continuación se presenta la suavización φ_μ propuesta en [17] y conocida como *función suavizada CHKS* (Chen-Harker-Kanzow-Smale).

Definición 3. La función $\varphi_\mu: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por

$$\varphi_\mu(a, b) = \frac{1}{2}\left(a + b - \sqrt{(a - b)^2 + 4\mu}\right), \mu > 0, \quad (8)$$

es una suavización de la función *Mínimo*.

Un caso particular de esta suavización, se ilustra en la [Figura 2](#), para el parámetro $\mu = 0.5$, en ella se observa la suavidad de dicha superficie, en contraste con la no suavidad de la [Figura 1](#).

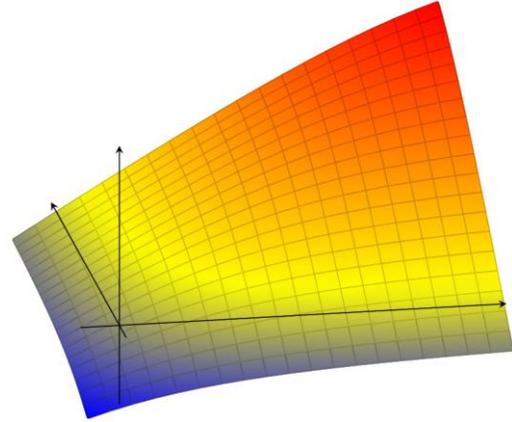


Figura 2. Función mínima suavizada con $\mu = 0.5$

El siguiente lema garantiza que φ_μ , es en efecto una suavización de la función *mínimo*.

Lema 1. La función φ_μ es una suavización de la función *Mínimo*.

Demostración. Se debe verificar que la suavización φ_μ satisface la equivalencia (7).

(i) Supongamos que $\varphi_\mu(a, b) = 0$. Esto es,

$$\frac{1}{2}\left[a + b - \sqrt{(a - b)^2 + 4\mu}\right] = 0,$$

de donde

$$a + b = \sqrt{(a - b)^2 + 4\mu} > 0. \quad (9)$$

Elevando al cuadrado ambos lados de (9) y organizando términos, se tiene que

$$a^2 + 2ab + b^2 = a^2 - 2ab + b^2 + 4\mu,$$

con lo cual $ab = \mu$, y como $\mu > 0$, a y b deben tener el mismo signo. De la desigualdad (9) se sigue que $a \geq 0, b \geq 0$.

(ii) Suponga que $a \geq 0, b \geq 0, ab = \mu$.

$$\begin{aligned} \varphi_\mu(a, b) &= \frac{1}{2}\left[a + b - \sqrt{(a - b)^2 + 4\mu}\right] \\ &= \frac{1}{2}\left[a + b - \sqrt{(a - b)^2 + 4ab}\right] \\ &= \frac{1}{2}\left[a + b - |a + b|\right] = \frac{1}{2}(0) = 0, \end{aligned}$$

con lo cual, queda demostrado el lema.

Con el fin de simplificar notacion en el analisis de resultados posteriores se definen dos nuevas funciones G y G_μ , que hacen parte de las funciones φ y φ_μ , respectivamente.

Definicion 4. La funcion $G: \mathbb{R}^2 \rightarrow \mathbb{R}$ esta definida por,

$$G(a, b) = \sqrt{(a - b)^2}, \quad (10)$$

para todo $(a, b) \in \mathbb{R}^2$.

Definicion 5. La funcion $G_\mu: \mathbb{R}^2 \rightarrow \mathbb{R}$ esta definida por

$$G_\mu(a, b) = \sqrt{(a - b)^2 + 4\mu}, \quad (11)$$

Observacion 1. Note que $(a - b)^2 \leq (a - b)^2 + 4\mu$, para todo $(a, b) \in \mathbb{R}^2$, de ahı que

$$G(a, b) \leq G_\mu(a, b).$$

El siguiente lema proporciona una cota superior, en terminos de μ , para la distancia entre la funcion φ y su suavizacion φ_μ .

Lema 2. Las funciones φ y φ_μ satisfacen la siguiente desigualdad

$$|\varphi_\mu(a, b) - \varphi(a, b)| \leq \sqrt{\mu},$$

para todo $(a, b) \in \mathbb{R}^2$,

Usando las Definiciones 4 y 5, se tiene que

$$\begin{aligned} & |\varphi_\mu(a, b) - \varphi(a, b)| \\ &= \frac{1}{2} |[a + b - G_\mu(a, b)] \\ &\quad - [a + b - G(a, b)]| \\ &= \frac{1}{2} |G_\mu(a, b) - G(a, b)| \\ &= \frac{1}{2} (G_\mu(a, b) - G(a, b)) \\ &\leq \frac{1}{2} [G(a, b) + \sqrt{4\mu} - G(a, b)] = \sqrt{\mu} \end{aligned}$$

la tercera igualdad sigue de la Observacion 1.

Por lo tanto, $|\varphi_\mu(a, b) - \varphi(a, b)| \leq \sqrt{\mu}$.

Corolario 1. Las funciones G y G_μ satisfacen la siguiente desigualdad para todo $(a, b) \in \mathbb{R}^2$.

$$|G_\mu(a, b) - G(a, b)| \leq 2\sqrt{\mu}.$$

Corolario 2. Cuando μ tiende a 0, φ_μ tiende a φ .

Como se menciono anteriormente, el proposito de las suavizaciones es lidiar con la no diferenciabilidad de las funciones de complementariedad. Recuerde que, en el caso particular de la funcion *Mınimo* su no diferenciabilidad se presenta en los puntos de la forma (a, a) , lo cual se soluciona con la suavizacion φ_μ definida en (8). Esto se formaliza en el siguiente lema.

Lema 3. La funcion φ_μ definida por (8), para todo $(a, b) \in \mathbb{R}^2$, es continuamente diferenciable en los puntos de la forma (a, a) .

Demostracion. Calculos sencillos permiten verificar las derivadas parciales de φ_μ en (a, b) existen y son continuas lo que implica el resultado del lema.

Dado que, la funcion φ_μ es una funcion continuamente diferenciable en \mathbb{R}^2 , se tiene que para todo $(a, b) \in \mathbb{R}^2$, su vector gradiente esta definido por:

$$\begin{aligned} \nabla \varphi_\mu(a, b) = \frac{1}{2} & \left(1 - \frac{a - b}{\sqrt{(a - b)^2 + 4\mu}}, 1 \right. \\ & \left. + \frac{a - b}{\sqrt{(a - b)^2 + 4\mu}} \right)^T \end{aligned} \quad (12)$$

Notemos que para G_μ definida en (11),

$$\nabla G_\mu(a, b) = \frac{1}{\sqrt{(a - b)^2 + 4\mu}} \begin{pmatrix} a - b \\ b - a \end{pmatrix}. \quad (13)$$

Observacion 2. Para todo $(a, b) \in \mathbb{R}^2$, la norma euclidiana del gradiente de G_μ en (a, b) satisface la siguiente desigualdad

$$\|\nabla G_\mu(a, b)\|_2 < \sqrt{2}. \quad (14)$$

Lo anterior se tiene usando la Observacion 1 en la siguiente igualdad

$$\|\nabla G_\mu(a, b)\|_2^2 = \frac{2(a - b)^2}{(a - b)^2 + 4\mu}.$$

De (12) y (13)

$$\nabla \varphi_\mu(a, b) = \frac{1}{2} \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \nabla G_\mu(a, b) \right] \quad (15)$$

A partir de (15), usando la desigualdad triangular y (14), se puede calcular una cota superior para la norma

euclidiana del gradiente de φ_μ en (a, b) , lo cual se garantiza en la siguiente observación.

Observación 3. Para todo $(a, b) \in \mathbb{R}^2$ la norma euclidiana del gradiente de φ_μ en (a, b) satisface la siguiente desigualdad

$$\|\nabla\varphi_\mu(a, b)\|_2 < \sqrt{2}. \quad (16)$$

Para su uso posterior, las derivadas parciales de G y de G_μ , se denotarán por:

$$\begin{aligned} \chi(a, b) &= \frac{\partial G}{\partial a}(a, b) = \frac{a-b}{G(a, b)} = \begin{cases} 1 & \text{si } a > b, \\ -1 & \text{si } a < b. \end{cases} \\ \psi(a, b) &= \frac{\partial G}{\partial b}(a, b) = \frac{b-a}{G(a, b)} = \begin{cases} -1 & \text{si } a > b, \\ 1 & \text{si } a < b. \end{cases} \end{aligned} \quad (17)$$

$$\begin{aligned} \chi_\mu(a, b) &= \frac{\partial G_\mu}{\partial a}(a, b) = \frac{a-b}{G_\mu(a, b)} \\ \psi_\mu(a, b) &= \frac{\partial G_\mu}{\partial b}(a, b) = \frac{b-a}{G_\mu(a, b)} \end{aligned} \quad (18)$$

El siguiente lema da una cota superior para $G_\mu(a, b)$ la cual será muy útil en la demostración del Lema 6.

Lema 4. Para todo $(a, b) \in \mathbb{R}^2$, se satisface que $G_\mu(a, b) \geq 2\sqrt{\mu}$.

Demostración. Para todo $(a, b) \in \mathbb{R}^2$, se tiene que

$$G_\mu(a, b)^2 = (a-b)^2 + 4\mu \geq 4\mu,$$

por lo tanto $G_\mu(a, b) \geq 2\sqrt{\mu}$.

Por el Lema 4 se deduce que, para todo $(a, b) \in \mathbb{R}^2$

$$\frac{1}{G_\mu(a, b)} \leq \frac{1}{2\sqrt{\mu}}.$$

Por otra parte,

$$(a-b)^2 = (a \ b) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (a \ b)K \begin{pmatrix} a \\ b \end{pmatrix},$$

donde

$$K = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (19)$$

es una matriz simétrica y semidefinida positiva con valores propios $\lambda_1 = 0$ y $\lambda_2 = 2$. Luego,

$$\|K\|_2 = 2. \quad (20)$$

Lema 5. La función φ_μ es Lipschitz continua con constante $\sqrt{2}$. Es decir, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, se satisface la desigualdad

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

Demostración. Sean $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Por el Teorema del Valor Medio existe un vector \mathbf{z} en el segmento $[\mathbf{x}, \mathbf{y}]$ tal que

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

Usando la desigualdad de Cauchy-Schwartz y la cota dada en (16), se tiene que

$$\begin{aligned} |\varphi_\mu(\mathbf{x}) - \varphi_\mu(\mathbf{y})| &\leq \|\nabla\varphi_\mu(\mathbf{z})\|_2 \|\mathbf{x} - \mathbf{y}\|_2 \\ &< \sqrt{2} \|\mathbf{x} - \mathbf{y}\|_2, \end{aligned}$$

con lo cual queda demostrado el lema.

Corolario 3. La función G_μ es Lipschitz continua con constante $\sqrt{2}$. Es decir, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, se satisface

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

Demostración. La prueba es análoga a la del lema anterior.

Usando (19), para (a, b) con $a \neq b$, se tiene que

$$\begin{aligned} \nabla G(a, b) &= \frac{1}{G(a, b)} K \begin{pmatrix} b \\ a \end{pmatrix} \\ \nabla G_\mu(a, b) &= \frac{1}{G_\mu(a, b)} K \begin{pmatrix} b \\ a \end{pmatrix} \end{aligned} \quad \text{y} \quad (21)$$

Lema 6. Sea φ_μ la función definida en (7), $\mathbf{w} \in \mathbb{R}^2$, con $0 < \|\mathbf{w}\|_2 \leq r$, y $\mathcal{B}(\mathbf{w}; r)$ una bola con centro en \mathbf{w} y radio r . Entonces existe $\kappa > 0$ tal que para todo $\mathbf{u}, \mathbf{v} \in \mathcal{B}(\mathbf{w}; r)$, se satisface que

$$\|\nabla\varphi_\mu(\mathbf{u}) - \nabla\varphi_\mu(\mathbf{v})\|_2 \leq \kappa \|\mathbf{u} - \mathbf{v}\|_2.$$

Demostración. Sean \mathbf{u} y $\mathbf{v} \in \mathbb{R}^2$. De (15), (20) y (21), se tiene

$$\begin{aligned} \|\nabla\varphi_\mu(\mathbf{u}) - \nabla\varphi_\mu(\mathbf{v})\|_2 &= \frac{1}{2} \|\nabla G_\mu(\mathbf{v}) - \nabla G_\mu(\mathbf{u})\|_2 \\ &= \frac{1}{2} \left\| \frac{1}{G_\mu(\mathbf{v})} K\mathbf{v} - \frac{1}{G_\mu(\mathbf{u})} K\mathbf{u} \right\|_2 \\ &\leq \frac{1}{2} \|K\|_2 \left\| \frac{1}{G_\mu(\mathbf{v})} \mathbf{v} - \frac{1}{G_\mu(\mathbf{u})} \mathbf{u} \right\|_2 \\ &= \left\| \frac{1}{G_\mu(\mathbf{v})} \mathbf{v} - \frac{1}{G_\mu(\mathbf{u})} \mathbf{u} \right\|_2. \end{aligned}$$

Sumando y restando la expresion $\frac{1}{G_\mu(\mathbf{u})}\mathbf{v}$, se deduce que

$$\|\nabla\varphi_\mu(\mathbf{u}) - \nabla\varphi_\mu(\mathbf{v})\|_2 \leq \left| \frac{1}{G_\mu(\mathbf{u})} - \frac{1}{G_\mu(\mathbf{v})} \right| \|\mathbf{v}\|_2 + \frac{1}{G_\mu(\mathbf{u})} \|\mathbf{u} - \mathbf{v}\|_2.$$

Para el primer sumando de la anterior desigualdad, y puesto que, por Corolario 3, G_μ es Lipschitz continua con constante $\sqrt{2}$, se tiene que

$$\left| \frac{1}{G_\mu(\mathbf{u})} - \frac{1}{G_\mu(\mathbf{v})} \right| = \left| \frac{G_\mu(\mathbf{v}) - G_\mu(\mathbf{u})}{G_\mu(\mathbf{u})G_\mu(\mathbf{v})} \right| \leq \frac{\sqrt{2}}{G_\mu(\mathbf{u})G_\mu(\mathbf{v})} \|\mathbf{u} - \mathbf{v}\|_2.$$

Por lo anterior y el Lema 4,

$$\begin{aligned} \|\nabla\varphi_\mu(\mathbf{u}) - \nabla\varphi_\mu(\mathbf{v})\|_2 &\leq \frac{1}{G_\mu(\mathbf{u})} \left(\frac{\sqrt{2}\|\mathbf{v}\|_2}{G_\mu(\mathbf{v})} + 1 \right) \|\mathbf{u} - \mathbf{v}\|_2 \\ &\leq \frac{1}{4\mu} (\sqrt{2}\|\mathbf{v}\|_2 + 2\sqrt{\mu}) \|\mathbf{u} - \mathbf{v}\|_2. \end{aligned}$$

Por otro lado, $\mathbf{v} \in \mathcal{B}(\mathbf{w}; r)$, $\|\mathbf{v} - \mathbf{w}\|_2 < \varepsilon$, de donde $\|\mathbf{v}\|_2 < r + \|\mathbf{w}\|_2$. Dado que $\|\mathbf{w}\|_2 < r$, se tiene que $\|\mathbf{v}\|_2 < 2r$. Finalmente, para $\mathbf{u}, \mathbf{v} \in \mathcal{B}(\mathbf{w}; r)$, se concluye que

$$\|\nabla\varphi_\mu(\mathbf{u}) - \nabla\varphi_\mu(\mathbf{v})\|_2 \leq \frac{1}{2\mu} (\sqrt{2}r + \sqrt{\mu}) \|\mathbf{u} - \mathbf{v}\|_2,$$

luego $\kappa = \frac{1}{2\mu} (\sqrt{2}r + \sqrt{\mu})$.

3. Reformulacion usando la funcion Mınimo

Si en (2), φ es la funcion Mınimo (5), se tiene la siguiente reformulacion del PCNL(F),

$$\Phi(x) = (\min(x_1, F_1(x)), \dots, \min(x_n, F_n(x)))^T = \mathbf{0}, \quad (22)$$

la cual se usa en el desarrollo de este trabajo. Por ello, a partir de ahora, cada vez que se haga referencia a la reformulacion $\Phi = \mathbf{0}$, se estara pensando en (22).

Observe que, si existe $i \in \{1, \dots, n\}$ tal que $x_i = F_i(x)$ la funcion es no diferenciable, por lo cual el sistema de ecuaciones no lineales (22) sera no diferenciable. El conjunto de estos ındices se denotara β . Es decir, $\beta = \{i: x_i = F_i(x)\}$.

Dado que por el Lema 5 la funcion Mınimo es Lipschitz continua, Φ tambien lo es, y por tanto, su jacobiano generalizado existe. Formalizando los calculos presentados en (17), se tiene el siguiente Lema.

Lema 7. Para todo $(a, b) \in \mathbb{R}^2$ con $a \neq b$,

$$\nabla G(a, b) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ o } \nabla G(a, b) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \quad (23)$$

Demostracion. El resultado sigue inmediatamente de (17).

Corolario 4. Sea G la funcion definida por (10). Para todo $(a, b) \in \mathbb{R}^2$ con $a \neq b$, se satisface que

$$\|\nabla G(a, b)\|_2 = \sqrt{2}.$$

Demostracion. Sigue directamente de (23) y de la definicion de norma 2.

El siguiente lema permite caracterizar las matrices en $\partial_C \Phi(x)$,

Lema 8. Dado $x \in \mathbb{R}^n$, $H \in \partial_C \Phi(x)$, es de la forma

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

donde la fila i -esima esta dada por

$$[H]_i = \begin{cases} e_i^T & \text{si } x_i < F_i(x) \\ \nabla F_i(x)^T & \text{si } x_i > F_i(x) \\ C_1 e_i^T + C_2 \nabla F_i(x)^T & \text{si } x_i = F_i(x) \end{cases} \quad (24)$$

donde c_1 y c_2 son constantes no negativas tales que $c_1 + c_2 = 1$.

Demostracion. A partir de la Proposicion 2.6.2 (e) en [16], tenemos que

$$\partial \Phi(x)^T \subseteq \partial \Phi_1(x) \times \dots \times \partial \Phi_n(x) = \partial_C \Phi(x),$$

donde Φ_i denota la i -esima funcion componente de Φ .

Si $(x_i, F_i(x))$ es tal que $x_i \neq F_i(x)$, Φ continuamente diferenciable, por lo tanto,

$$\begin{aligned} \partial \Phi_i(x) &= \{\nabla \Phi_i(x)\} \\ &= \left\{ \frac{1}{2} \left[1 - \frac{x_i - F_i(x)}{\sqrt{(x_i - F_i(x))^2}} \right] e_i \right. \\ &\quad \left. + \frac{1}{2} \left[1 + \frac{x_i - F_i(x)}{\sqrt{(x_i - F_i(x))^2}} \right] \nabla F_i(x) \right\}. \end{aligned}$$

donde e_i es el i -esimo vector canonico de \mathbb{R}^n .

Para $x_i \neq F_i(x)$, tenemos que

- Si $x_i > F_i(x)$ entonces $\nabla \Phi_i(x) = \nabla F_i(x)$.
- Si, $x_i < F_i(x)$ entonces $\nabla \Phi_i(x) = e_i$

Por otra parte, si $(x_i, F_i(x))$ es tal que $x_i = F_i(x)$, cada $H_i \in \partial \Phi_i(x)$ es de la forma

$$\partial \Phi_i(x) = \text{conv}[H]_i = \lim_{y_k \rightarrow x} \nabla \Phi_i(y_k), \text{ cuando } k \rightarrow \infty.$$

Ahora, denotando la componente i -ésima del vector y_k , por y_i^k , se tiene que

$$\begin{aligned} [H]_i &= \lim_{y_k \rightarrow x} \nabla \Phi_i(y_k) \\ &= \lim_{k \rightarrow \infty} \left\{ \frac{1}{2} \left[1 - \frac{y_i^k - F_i(y_k)}{\sqrt{(y_i^k - F_i(y_k))^2}} \right] e_i \right. \\ &\quad \left. + \frac{1}{2} \left[1 + \frac{y_i^k - F_i(y_k)}{\sqrt{(y_i^k - F_i(y_k))^2}} \right] \nabla F_i(y_k) \right\} \\ &= \frac{1}{2} [(1 - \xi_i) e_i + (1 - \tau_i) \nabla F_i(x)]. \end{aligned}$$

Por el Lema 7, $(\xi_i, \tau_i) = (1, -1)$ o $(\xi_i, \tau_i) = (-1, 1)$. En el primer caso, $[H]_i = \nabla F_i(x)$ y para el segundo, $[H]_i = e_i$. Luego, para $(x_i, F_i(x))$, con $x_i = F_i(x)$, se tiene $\partial \Phi_i(x)$, *conv* $\{e_i, \nabla F_i(x)\}$ definido por:

$$\{c_1 e_i + c_2 \nabla F_i(x) : c_1, c_2 \in \mathbb{R}^+ \cup \{0\}, c_1 + c_2 = 1\}.$$

Por lo tanto, una matriz $H \in \partial_C \Phi(x)$ es de la forma:

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

donde

$$[H]_i = \begin{cases} e_i^T & \text{si } x_i < F_i(x) \\ \nabla F_i(x)^T & \text{si } x_i > F_i(x) \\ c_1 e_i^T + c_2 \nabla F_i(x)^T & \text{si } x_i = F_i(x), \end{cases} \quad (25)$$

donde c_1 y c_2 son constantes no negativas tales que $c_1 + c_2 = 1$.

Sin perder generalidad, en lo que sigue, se considera

$$[H]_i = e_i^T \text{ o } [H]_i = \nabla F_i(x)^T, \text{ cuando } x_i = F_i(x).$$

Es decir, se escoge $c_1 = 1$ y $c_2 = 0$ o $c_1 = 0$ y $c_2 = 1$. Esta elección es natural ya que en el caso $x_i = F_i(x)$, se tiene que $\min\{x_i, F_i(x)\} = x_i = F_i(x)$. Por ello, es razonable escoger $[H]_i = e_i^T$ o $[H]_i = \nabla F_i(x)^T$.

Por otra parte, asociada a una solución x^* definimos $H_* \in \partial_C \Phi(x^*)$ por:

$$H_* = \begin{pmatrix} [H_*]_1 \\ \vdots \\ [H_*]_n \end{pmatrix}, \quad (26)$$

donde

$$[H_*]_i = \begin{cases} e_i^T & \text{si } x_i^* < F_i(x^*) \\ \nabla F_i(x^*)^T & \text{si } x_i^* > F_i(x^*) \\ e_i^T \text{ o } \nabla F_i(x^*)^T & \text{si } x_i^* = F_i(x^*). \end{cases} \quad (26)$$

Para finalizar la sección, se deduce una forma explícita de la matriz jacobiana de la suavización Φ_μ en x . A partir de (8), se define $\Phi_\mu: \mathbb{R}^n \rightarrow \mathbb{R}^n$ de la siguiente forma

$$\Phi_\mu(x) = \begin{pmatrix} \varphi_\mu(x_1, F_1(x)) \\ \vdots \\ \varphi_\mu(x_n, F_n(x)) \end{pmatrix}, \quad (27)$$

y con ello, el sistema $\Phi_\mu(x) = \mathbf{0}$ es una suavización de (22). En general, la matriz jacobiana de Φ_μ está dada por,

$$\Phi'_\mu(x) = \begin{pmatrix} \nabla \varphi_\mu(x_1, F_1(x))^T \\ \vdots \\ \nabla \varphi_\mu(x_n, F_n(x))^T \end{pmatrix} = \begin{pmatrix} [\Phi'_\mu(x)]_1 \\ \vdots \\ [\Phi'_\mu(x)]_n \end{pmatrix}. \quad (28)$$

Para cualquier $(a, b) \in \mathbb{R}^2$ de (12) y (18) se tiene que

$$\nabla \varphi_\mu(a, b) = \frac{1}{2} \begin{pmatrix} 1 - \chi_\mu(a, b) \\ 1 - \psi_\mu(a, b) \end{pmatrix},$$

con $\chi_\mu(a, b)$ y $\psi_\mu(a, b)$ dadas por (18).

Para obtener una forma explícita de la matriz (28), se definen las variables auxiliares u y v por $u = x_1$ y $v = F_1(x)$. Usando la regla de la cadena, se tiene que

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

De ahí que,

$$\begin{aligned} &\nabla \varphi_\mu(u, v) \\ &= \frac{1}{2} \begin{pmatrix} (1 - \chi_\mu(u, v))1 + (1 - \psi_\mu(u, v)) \frac{\partial F_1(x)}{\partial x_1} \\ \vdots \\ (1 - \chi_\mu(u, v))0 + (1 - \psi_\mu(u, v)) \frac{\partial F_1(x)}{\partial x_n} \end{pmatrix} \\ &= \frac{1}{2} [(1 - \chi_\mu(u, v)) e_1 + (1 - \psi_\mu(u, v)) \nabla F_1(x)]. \end{aligned}$$

Luego, la primera fila de matriz $\Phi'_\mu(x)$ está dada por:

$$[\Phi'_\mu(x)]_1 = \frac{1}{2} [(1 - \chi_\mu(x_1, F_1(x))) e_1^T + (1 - \psi_\mu(x_1, F_1(x))) \nabla F_1(x)^T],$$

y de forma analoga, la fila i -esima de la matriz $\Phi'_\mu(\mathbf{x})$ esta dada por:

$$[\Phi'_\mu(\mathbf{x})]_i = \frac{1}{2} [(1 - \chi_\mu(x_i, F_i(\mathbf{x}))) \mathbf{e}_i^T + (1 - \psi_\mu(x_i, F_i(\mathbf{x}))) \nabla F_i(\mathbf{x})^T]. \quad (29)$$

Por lo tanto, una expresion explıcita de la matriz jacobiana $\Phi'_\mu(\mathbf{x})$ es la siguiente

$$\Phi'_\mu(\mathbf{x}) = \frac{1}{2} \begin{pmatrix} (1 - \chi_\mu(x_1, F_1(\mathbf{x}))) \mathbf{e}_1^T + (1 - \psi_\mu(x_1, F_1(\mathbf{x}))) \nabla F_1(\mathbf{x})^T \\ \vdots \\ (1 - \chi_\mu(x_n, F_n(\mathbf{x}))) \mathbf{e}_n^T + (1 - \psi_\mu(x_n, F_n(\mathbf{x}))) \nabla F_n(\mathbf{x})^T \end{pmatrix}, \quad (30)$$

donde χ_μ , ψ_μ y G_μ estan definidas en (18) y (11), respectivamente.

3.1. Algoritmo y resultados de convergencia

En esta seccion, se presenta un nuevo algoritmo tipo cuasi-Newton suavizado para resolver el problema de complementariedad no lineal, indirectamente a traves de su reformulacion como un sistema de ecuaciones no lineales que usa la funcion de complementariedad Mınimo, el cual, a su vez, se aproxima por una sucesion de sistemas suavizados que usan una suavizacion de dicha funcion. Siguiendo la filosofıa de los metodos con jacobiano suavizado [14], una iteracion basica del metodo que se propone es de la forma:

$$\begin{aligned} \Phi'(x_k) s_k &= -\Phi(x_k) \\ x_{k+1} &= s_k + x_k \end{aligned} \quad (31)$$

donde de $\Phi'_\mu(\mathbf{x}_k)$ es la matriz jacobiana de Φ_μ en \mathbf{x}_k , dada en (30).

A diferencia del metodo de Newton generalizado, se usa la matriz $\Phi'_\mu(\mathbf{x}_k)$ en lugar de una matriz $H_k \in \partial_C \Phi(\mathbf{x}_k)$. Desde este punto de vista, la iteracion (31), puede verse como una iteracion cuasi-Newton, donde la matriz $\Phi'_\mu(\mathbf{x}_k)$ es una aproximacion de $H_k \in \partial_C \Phi(\mathbf{x}_k)$, o $H_k \in \partial \Phi(\mathbf{x}_k)$.

El siguiente lema garantiza que cuando el parametro μ se acerca a cero, la distancia entre la matriz $\Phi'_\mu(\mathbf{x}_k)$ y el conjunto $\partial_C \Phi(\mathbf{x}_k)$, tambien tiende a cero, lo cual justifica reemplazar la iteracion del metodo de Newton generalizado por (31).

Lema 9. Sean $\mathbf{x} \in \mathbb{R}^n$ arbitrario pero fijo y $\mu > 0$ entonces

$$\lim_{\mu \rightarrow 0} \text{dist}(\Phi'_\mu(\mathbf{x}), \partial_C \Phi(\mathbf{x})) = 0 \quad (32)$$

Demostracion. Puesto que

$$\begin{aligned} \lim_{\mu \rightarrow 0} \text{dist}(\Phi'_\mu(\mathbf{x}), \partial_C \Phi(\mathbf{x})) &= \lim_{\mu \rightarrow 0} \inf_{H \in \partial_C \Phi(\mathbf{x}_k)} \|\Phi'_\mu(\mathbf{x}) - H\|_F \\ &= \inf_{H \in \partial_C \Phi(\mathbf{x}_k)} \lim_{\mu \rightarrow 0} \|\Phi'_\mu(\mathbf{x}) - H\|_F, \end{aligned}$$

para que (32) se satisfaga, basta demostrar que existe $H \in \partial_C \Phi(\mathbf{x})$, la cual es el lımite de $\Phi'_\mu(\mathbf{x})$ cuando $\mu \rightarrow 0$. Para ello, se considera el conjunto de ındices $\beta = \beta(\mathbf{x}) = \{i : x_i = F_i(\mathbf{x}) = 0\}$, y usando (29), se tiene que $\lim_{\mu \rightarrow 0} [\Phi'_\mu(\mathbf{x})]_i$ esta definido por:

$$\begin{cases} \frac{1}{2} [(1 - \chi(x_i, F_i(\mathbf{x}))) \mathbf{e}_i^T + (1 - \psi(x_i, F_i(\mathbf{x}))) \nabla F_i(\mathbf{x})^T], & i \notin \beta \\ \frac{1}{2} \mathbf{e}_i^T + \frac{1}{2} \nabla F_i(\mathbf{x})^T, & i \in \beta, \end{cases}$$

equivalentemente,

$$\begin{cases} \nabla F_i(\mathbf{x})^T, & \text{si } x_i > F_i(\mathbf{x}) \\ \mathbf{e}_i^T, & \text{si } x_i < F_i(\mathbf{x}) \\ \frac{1}{2} \mathbf{e}_i^T + \frac{1}{2} \nabla F_i(\mathbf{x})^T, & \text{si } x_i = F_i(\mathbf{x}), \end{cases} = [H]_i.$$

Note que la matriz H tiene la estructura que describe el Lema 8, donde los escalares de la combinacion lineal son iguales a $\frac{1}{2}$ para $i \in \beta(\mathbf{x})$. Por lo tanto, existe $H \in \partial_C \Phi(\mathbf{x})$, la cual es el lımite de $\Phi'_\mu(\mathbf{x})$ cuando $\mu \rightarrow 0$, de manera que el ınfimo buscado es cero y satisface (32).

Las hipotesis para la teorıa de convergencia del algoritmo propuesto son las siguientes.

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

H2. La matriz jacobiana de F es *Lipschitz continua* en una vecindad de $\mathbf{x}^* \in \mathbb{R}^n$. Es decir, existen constantes $\gamma > 0$ y $\delta > 0$ tales que $\|F'(\mathbf{x}) - F'(\mathbf{x}^*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}^*\|$, para todo $\mathbf{x} \in \mathcal{B}(\mathbf{x}^*; \delta)$ donde $\|\cdot\|$ denota una norma en \mathbb{R}^n y tambien su respectiva norma matricial inducida.

H3. Las matrices de $\partial_C \Phi(\mathbf{x}^*)$ son no singulares.

Por H3 y la compacidad de $\partial \Phi(\mathbf{x}^*)$, existe una constante positiva σ tal que, para toda $H^* \in \partial_C \Phi(\mathbf{x}^*)$ se satisface que $\|H^{*-1}\| \leq \sigma$.

Observacion 4. Las matrices en $\partial_C \Phi(\mathbf{x}^*)$ forman un conjunto finito de 2^r matrices.

En efecto, suponga que \mathbf{x}^* es una solucion degenerada del PCNL(F), la cual tiene r degeneraciones, que sin

pérdida de generalidad, se presentan en las primeras r componentes de \mathbf{x}^* es decir,

$$\begin{aligned} x_i^* &= F_i(\mathbf{x}^*) \text{ para } i = 1, \dots, r \\ x_i^* &\neq F_i(\mathbf{x}^*) \text{ para } i = r + 1, \dots, n. \end{aligned}$$

Note que por cada degeneración en \mathbf{x}^* se obtienen dos matrices en $\partial_C \Phi(\mathbf{x}^*)$ entonces existirán 2^r matrices en dicho conjunto.

El siguiente lema da una cota superior para la distancia, en norma infinito, entre las matrices H y H_* .

Lema 10. Sean $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función continuamente diferenciable cuya matriz jacobiana satisface H2, $\epsilon > 0$ dado, $H \in \partial \Phi(\mathbf{x})$ definida por (24) y $H_* \in \partial_C \Phi(\mathbf{x}^*)$, definida por (26). Entonces para cada $\mathbf{x} \in B(\mathbf{x}^*; \epsilon)$, existe una constante α tal que

$$\|H_* - H\|_\infty \leq \alpha.$$

Por definición de norma infinito:

$$\begin{aligned} \|H_* - H\|_\infty &= \max_{1 \leq i \leq n} \{ |[H_*]_i - [H]_i| \}_1 \\ &= \|[H_*]_j - [H]_j\|_1, \end{aligned}$$

para algún $j \in \{1, \dots, n\}$. Por (24) y ([26]), se tiene que

- Si $x_i > F_i(\mathbf{x})$, $\|H_* - H\|_\infty = \|\nabla F_i(\mathbf{x}^*)^T - \nabla F_i(\mathbf{x})^T\|_1 \leq \gamma \|\mathbf{x} - \mathbf{x}^*\|_1 \leq \gamma \epsilon$
- Si $x_i < F_i(\mathbf{x})$, $\|H_* - H\|_\infty = \|\mathbf{e}_i^T - \mathbf{e}_i^T\|_1 = 0$
- Si $x_i < F_i(\mathbf{x})$, $\|H_* - H\|_\infty = \|\mathbf{e}_i^T - \mathbf{e}_i^T\|_1 = 0$.
Luego, $\|H_* - H\|_\infty = \|\mathbf{e}_i^T - \mathbf{e}_i^T\|_1 = 0$.
Si $[H]_i = [H_*]_i = \nabla F_i(\mathbf{x})^T$,
 $\|H_* - H\|_\infty = \|\nabla F_i(\mathbf{x}^*)^T - \nabla F_i(\mathbf{x})^T\|_1 \leq \gamma \|\mathbf{x} - \mathbf{x}^*\|_1 \leq \gamma \epsilon$.

Por tanto, existe $\alpha = \gamma \epsilon$, tal que $\|H_* - H\|_\infty < \alpha$.

El siguiente es un lema técnico que será útil en la demostración del Teorema 1.

Lema 11. Sea $T(\mathbf{x}) = \Phi(\mathbf{x}) - \Phi(\mathbf{x}^*) - H(\mathbf{x} - \mathbf{x}^*)$, Existe una constante positiva ρ tal que, si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \epsilon$ entonces $T(\mathbf{x})$ satisface

$$\|T(\mathbf{x})\|_\infty \leq \rho \|\mathbf{x} - \mathbf{x}^*\|_\infty.$$

Además, si F' es Lipschitz continua en una vecindad de \mathbf{x}^* existe $\gamma > 0$ tal que

$$\|T(\mathbf{x})\|_\infty \leq \gamma \|\mathbf{x} - \mathbf{x}^*\|_\infty^2$$

Demostración. Por la definición de Φ , se tiene que

$$\Phi(\mathbf{x}) - \Phi(\mathbf{x}^*) = \begin{pmatrix} \phi_1(\mathbf{x}, F_1(\mathbf{x})) - \phi_1(\mathbf{x}^*, F_1(\mathbf{x}^*)) \\ \vdots \\ \phi_n(\mathbf{x}, F_n(\mathbf{x})) - \phi_n(\mathbf{x}^*, F_n(\mathbf{x}^*)) \end{pmatrix},$$

luego, la fila i -ésima de este vector está dada por:

$$[\Phi(\mathbf{x}) - \Phi(\mathbf{x}^*)]_i = \begin{cases} x_i - x_i^* & \text{si } x_i^* < F_i(\mathbf{x}^*) \\ F_i(\mathbf{x}) - F_i(\mathbf{x}^*) & \text{si } x_i^* > F_i(\mathbf{x}^*) \end{cases} \quad (33)$$

Por otra parte, la fila i -ésima de $H(\mathbf{x} - \mathbf{x}^*)$ es dada por:

$$[H(\mathbf{x} - \mathbf{x}^*)]_i = \begin{cases} x_i - x_i^* & \text{si } x_i^* < F_i(\mathbf{x}^*) \\ \nabla F_i(\mathbf{x})^T (\mathbf{x} - \mathbf{x}^*) & \text{si } x_i^* > F_i(\mathbf{x}^*) \end{cases} \quad (34)$$

Además, cuando $x_i^* = F_i(\mathbf{x}^*)$, se puede escoger \mathbf{e}_i^T o $\nabla F_i(\mathbf{x})^T$ (o una combinación lineal de estos). Ahora, por la definición de la función T , (33) y (34), se tiene que

$$T_i(\mathbf{x}) = \begin{cases} 0 & \text{si } x_i^* < F_i(\mathbf{x}^*) \\ F_i(\mathbf{x}) - F_i(\mathbf{x}^*) - \nabla F_i(\mathbf{x})^T (\mathbf{x} - \mathbf{x}^*) & \text{si } x_i^* > F_i(\mathbf{x}^*) \end{cases}$$

luego,

$$\|T(\mathbf{x})\|_\infty = \max_{1 \leq i \leq n} |F_i(\mathbf{x}) - F_i(\mathbf{x}^*) - \nabla F_i(\mathbf{x})^T (\mathbf{x} - \mathbf{x}^*)| = |F_k(\mathbf{x}) - F_k(\mathbf{x}^*) - \nabla F_k(\mathbf{x})^T (\mathbf{x} - \mathbf{x}^*)|,$$

para algún $k \in \{1, \dots, n\}$. Dado que F es continuamente diferenciable, para cualquier $\rho > 0$ existe $\epsilon > 0$ tal que si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \epsilon$ entonces $\|T(\mathbf{x})\|_\infty \leq \rho \|\mathbf{x} - \mathbf{x}^*\|_\infty$.

Por otra parte, como F' es Lipschitz continua, para cualquier ρ_1 existe $\hat{\epsilon} > 0$ tal que, si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \hat{\epsilon}$ entonces $\|T(\mathbf{x})\|_\infty \leq \rho_1 \|\mathbf{x} - \mathbf{x}^*\|_\infty^2$.

Teorema 1. Existe una constante positiva ϵ tal que, si $\|\mathbf{x} - \mathbf{x}^*\|_\infty < \epsilon$ la función Q definida por:

$$Q(\mathbf{x}) = \mathbf{x} - \Phi'_\mu(\mathbf{x})^{-1} \Phi(\mathbf{x}) \quad (35)$$

está bien definida, y dado $r \in (0, 1)$

$$\|Q(\mathbf{x}) - \mathbf{x}^*\|_\infty \leq r \|\mathbf{x} - \mathbf{x}^*\|_\infty. \quad (36)$$

Además, si la matriz jacobiana de F es Lipschitz continua en una vecindad de $\mathbf{x}^* \in \mathbb{R}^n$,

$$\|Q(\mathbf{x}) - \mathbf{x}^*\|_\infty \leq c \|\mathbf{x} - \mathbf{x}^*\|_\infty^2. \quad (37)$$

Demostracion. En primer lugar, se demuestra que (35) esta bien definida, es decir, existe $\Phi'_\mu(x)^{-1}$. Por la definicion del C-subdiferencial, se tiene que la matriz $H_* \in \partial_C \Phi(x)$ es de la forma

$$H_* = \begin{pmatrix} [H_*]_1 \\ \vdots \\ [H_*]_n \end{pmatrix} = \lim_{k \rightarrow \infty} \Phi'(y^k) = \begin{pmatrix} \lim_{k \rightarrow \infty} \nabla \varphi(y_1^k, F_1(y^k))^T \\ \vdots \\ \lim_{k \rightarrow \infty} \nabla \varphi(y_n^k, F_n(y^k))^T \end{pmatrix},$$

donde $\{y^k\}$ es una sucesion de vectores de \mathbb{R}^n , que converge a x^* y tal que $\Phi'(y^k)$ existe. Se analiza a continuacion $\|\Phi'_\mu(x) - H\|_\infty$.

Para algun $j \in \{1, \dots, n\}$, se tiene que

$$\begin{aligned} \|\Phi'_\mu(x) - H_*\|_\infty &= \|\Phi'_\mu(x)_j - [H_*]_j\|_1 \\ &\leq n \|\nabla \varphi_\mu(x_j, F_j(x))^T - \lim_{k \rightarrow \infty} \nabla \varphi(y_j^k, F_j(y^k))^T\|_\infty \\ &\leq n \lim_{k \rightarrow \infty} \|\nabla \varphi_\mu(x_j, F_j(x))^T - \nabla \varphi(y_j^k, F_j(y^k))^T\|_\infty \end{aligned}$$

Por el Lema 6,

$$\begin{aligned} &\|\Phi'_\mu(x) - H_*\|_\infty \\ &\leq nk \lim_{k \rightarrow \infty} \left\| \begin{pmatrix} x_j - y_j^k \\ F_j(x) - F_j(y^k) \end{pmatrix} \right\|_2 \\ &\leq \sqrt{2nk} \left\| \begin{pmatrix} x_j - y_j^k \\ F_j(x) - F_j(x^*) \end{pmatrix} \right\|_\infty \\ &= \sqrt{2nk} \max\{|x_j - x_j^*|, |F_j(x) - F_j(x^*)|\}. \end{aligned} \quad (38)$$

Dado que F es continua, para todo $\hat{\varepsilon}$ existe $\hat{\delta} > 0$ tal que, si $\|x - x^*\|_\infty < \hat{\delta}$ entonces $|F_j(x) - F_j(x^*)| < \hat{\varepsilon}$.

Sea $\varepsilon' = \min\{\hat{\varepsilon}, \hat{\delta}\}$. Se consideran las dos posibilidades para el maximo en (38).

i Si $\max\{|x_j - x_j^*|, |F_j(x) - F_j(x^*)|\} = |x_j - x_j^*| \leq \|x - x^*\|_\infty \leq \varepsilon' < \hat{\varepsilon}$.

ii Si $\max\{|x_j - x_j^*|, |F_j(x) - F_j(x^*)|\} = |F_j(x) - F_j(x^*)| < \hat{\varepsilon}$.

Por las consideraciones anteriores y por (38) para cualquier $\hat{\varepsilon}$, se tiene que

$$\|\Phi'_\mu(x) - H_*\|_\infty < \sqrt{2} nk \hat{\varepsilon}.$$

Sea $\varepsilon < \frac{1}{2\sqrt{2}nk\hat{\varepsilon}\sigma}$ entonces $\|\Phi'_\mu(x) - H_*\|_\infty < \frac{1}{2\sigma}$.

Ahora,

$$\|H_*^{-1}\Phi'_\mu(x) - I\|_\infty \leq \|H_*^{-1}\|_\infty \|\Phi'_\mu(x) - H_*\|_\infty \leq \sigma \frac{1}{2\sigma} = \frac{1}{2}.$$

De donde $\|H_*^{-1}\Phi'_\mu(x) - I\|_\infty < 1$. Por el Lema de Banach [18], existe $\Phi'_\mu(x)^{-1}$ y por lo tanto, la funcion \mathcal{Q} esta bien definida. Ademas,

$$\|\Phi'_\mu(x)^{-1}\|_\infty \leq \frac{\|H_*^{-1}\|_\infty}{1 - \|H_*^{-1}\Phi'_\mu(x) - I\|_\infty} \leq \frac{\sigma}{1 - \frac{1}{2}} = 2\sigma.$$

Para la segunda parte de la prueba, se demuestra (36) usando (35). Mediante algunas manipulaciones algebraicas, se tiene que

$$\begin{aligned} \|\mathcal{Q}(x) - x^*\|_\infty &= \|x - x^* - \Phi'_\mu(x)^{-1}\Phi(x)\|_\infty \\ &= \left\| \Phi'_\mu(x)^{-1} \left[\Phi'_\mu(x)(x - x^*) - \Phi(x) + \Phi(x^*) \right] \right\|_\infty \\ &\leq \left\| \Phi'_\mu(x)^{-1} \right\|_\infty \|\Phi'_\mu(x)(x - x^*) - \Phi(x) + \Phi(x^*)\|_\infty \\ &\leq 2\sigma \left\| (\Phi'_\mu(x)(x - x^*) - H(x - x^*) + H(x - x^*) - \Phi(x) + \Phi(x^*)) \right\|_\infty \end{aligned} \quad (39)$$

$$\leq 2\sigma \left[\|\Phi'_\mu(x) - H\|_\infty \|x - x^*\|_\infty + \|H(x - x^*) - \Phi(x) + \Phi(x^*)\|_\infty \right] \quad (40)$$

Por el Lema 9, para todo $\delta > 0$, existe $\varepsilon_1 > 0$ tal que si $\|x - x^*\|_\infty < \varepsilon_1$ entonces $\|\Phi'_\mu(x^*) - H\| < \delta$, para $H \in \partial_C \Phi(x)$. En particular, para $\delta = \frac{r}{4\sigma}$ existe ε_r tal que, si $\|x - x^*\|_\infty < \varepsilon_r$,

$$\|\Phi'_\mu(x) - H\| < \frac{r}{4\sigma}. \quad (41)$$

Ademas, por el Lema 11, para todo $\rho > 0$, existe $\varepsilon_2 > 0$ tal que si $\|x - x^*\|_\infty < \varepsilon_1$ se tiene que

$$\|H(x - x^*) - \Phi(x) + \Phi(x^*)\|_\infty < \rho \|x - x^*\|_\infty.$$

En particular, para $\rho < \frac{r}{4\sigma}$ existe $\bar{\varepsilon}_r > 0$ tal que si $\|x - x^*\|_\infty < \bar{\varepsilon}_r$ se tiene que

$$\|H(x - x^*) - \Phi(x) + \Phi(x^*)\|_\infty < \frac{r}{4\sigma} \|x - x^*\|_\infty \quad (42)$$

Luego para $\varepsilon = \min\{\varepsilon', \varepsilon_r, \bar{\varepsilon}_r\}$, por (40), (41) y (42) se tiene,

$$\|Q(x) - x^*\|_\infty \leq r \|x - x^*\|_\infty.$$

La tercera parte de la prueba consiste en demostrar (37) bajo la hipótesis de que la matriz jacobiana de F es *Lipschitz* continua en $x^* \in \mathbb{R}^n$. Por el Lema 11 existe una constante positiva γ tal que,

$$\|H(x - x^*) - \Phi(x) + \Phi(x^*)\|_\infty \leq \gamma \|x - x^*\|_\infty^2. \quad (43)$$

Por otra parte, sumando y restando la matriz H_* en $\|\Phi'_\mu(x^*) - H\|_\infty$ y aplicando la desigualdad triangular, se tiene que

$$\|\Phi'_\mu(x^*) - H\|_\infty \leq \|\Phi'_\mu(x^*) - H_*\|_\infty + \|H_* - H\|_\infty. \quad (44)$$

Por (38), el primer término del lado derecho de (44) está acotado superiormente por $\sqrt{2}n\kappa M$, donde $M = \max\{|x_j - x_j^*|, |F_j(x) - F_j(x^*)|\}$ y por el Lema 10, el segundo término está acotado por $\gamma n \|x - x^*\|_\infty$. Observe que,

- i. Si $M = |x_j - x_j^*|$ entonces $M \leq \|x - x^*\|_\infty$.
- ii. Si $M = |F_j(x) - F_j(x^*)|$ entonces $M \leq \|F(x) - F(x^*)\|_\infty \leq \zeta \|x - x^*\|_\infty$.

La última desigualdad se tiene porque F es continuamente diferenciable.

Sea $\hat{\zeta} = \max\{1, \zeta\}$,

$$\begin{aligned} \|\Phi'_\mu(x^*) - H\|_\infty &\leq \sqrt{2}n\kappa \hat{\zeta} \|x - x^*\|_\infty \\ &\quad + \gamma n \|x - x^*\|_\infty \left(\sqrt{2}n\kappa \hat{\zeta} \right. \\ &\quad \left. + \gamma \right) \|x - x^*\|_\infty \\ &= \theta \|x - x^*\|_\infty \end{aligned} \quad (45)$$

con $\theta = n(\sqrt{2}\kappa \hat{\zeta} + \gamma)$, luego de (40), (43) y (45),

$$\begin{aligned} \|Q(x) - x^*\|_\infty &\leq 2\sigma(\theta \|x - x^*\|_\infty^2 + \gamma n \|x - x^*\|_\infty^2) \\ &= 2\sigma(\theta + \gamma) \|x - x^*\|_\infty^2. \end{aligned}$$

Por lo tanto, $\|Q(x) - x^*\|_\infty \leq c \|x - x^*\|_\infty^2$, donde $c = 2\sigma(\theta + \gamma)$.

El siguiente teorema garantiza que el algoritmo propuesto está bien definido, converge linealmente y además da una condición suficiente para la convergencia cuadrática.

Teorema 2. Existe una constante $\varepsilon_0 > 0$ tal que si $\|x_0 - x^*\|_\infty \leq \varepsilon_0$, entonces la sucesión $\{x_k\}$, generada por

$$x_{k+1} = x_k - \Phi'_\mu(x_k)^{-1} \Phi(x_k), \quad (46)$$

está bien definida, converge a x^* y dado $r \in (0, 1)$, satisface

$$\|x_{k+1} - x^*\|_\infty \leq r \|x_k - x^*\|_\infty, \quad (47)$$

Además, si la matriz jacobiana de F es *Lipschitz* continua en una vecindad de x^* entonces

$$\|x_{k+1} - x^*\|_\infty \leq c \|x_k - x^*\|_\infty^2, \quad (48)$$

donde c es la constante del Teorema 1.

Demostración. Sea Q definida en (35). Para $k = 0, 1, \dots$

$$x_{k+1} = Q(x_k) = x_k - \Phi'_\mu(x_k)^{-1} \Phi(x_k).$$

Sean $r \in (0, 1)$ y $\varepsilon_0 \in (0, \varepsilon)$, donde ε es la constante del Teorema 1. La demostración es inducción sobre k .

- Para $k = 0$. Si $\|x_0 - x^*\|_\infty \leq \varepsilon_0 < \varepsilon$, por el Lema 11, $x_1 = Q(x_0)$ está bien definida y satisface $\|x_1 - x^*\|_\infty \leq r \|x_0 - x^*\|_\infty$. Además, si la matriz jacobiana de F es *Lipschitz* continua en una vecindad de x^* , $\|x_1 - x^*\|_\infty \leq c \|x_0 - x^*\|_\infty^2$.
- **Hipótesis inductiva.** Supongamos que $\|x_k - x^*\|_\infty < \varepsilon_0$, para todo $0 \leq k \leq m - 1$. Entonces por el Lema 11, $x_m = Q(x_{m-1})$ está bien definida y satisface

$$\begin{aligned} \|x_m - x^*\|_\infty &= \|Q(x_{m-1}) - x^*\|_\infty \leq r \\ &\quad \|x_{m-1} - x^*\|_\infty. \end{aligned} \quad (49)$$

Además, $\|x_m - x^*\|_\infty \leq c \|x_{m-1} - x^*\|_\infty^2$, si la matriz jacobiana de F es *Lipschitz* continua en una vecindad de x^* .

A partir de (49), se tiene

$$\begin{aligned} \|x_m - x^*\|_\infty &\leq r \|x_{m-1} - x^*\|_\infty \leq r^m \|x_0 - x^*\|_\infty \\ &\leq r^m \varepsilon_0 < \varepsilon, \end{aligned}$$

así, el Lema 11 garantiza que x_{m+1} , está bien definido y satisface que

$$\|x_{m-1} - x^*\|_\infty \leq r \|x_m - x^*\|_\infty,$$

y si la matriz jacobiana de F es *Lipschitz* continua en una vecindad de x^* , entonces x_{m+1} , satisface que

$$\|x_{m-1} - x^*\|_\infty \leq c \|x_m - x^*\|_\infty^2.$$

Por lo tanto, se concluye que (47) y (48) se cumplen para todo $k = 0, 1, \dots$

3.2. Experimentación numérica

En esta sección se analiza el desempeño numérico del algoritmo cuasi-Newton propuesto en la sección anterior que se llamará Algoritmo CNS. Para ello, se compara con dos algoritmos, el primero usa una suavización de la función de Fischer-Burmeister en su reformulación [15], el cual se llamará Algoritmo SF, y el segundo es un algoritmo tipo Newton generalizado que usa la función Mínimo, Algoritmo NM.

Se resolvieron diez problemas de complementariedad asociados con las funciones *Billups* [19], *Nash-Cornout* (*Nash-Co*) [20], *Kojima-Shindo* (*Koj-Sh*), *Kojima-Josephy* (*Koj-Jo*), *Mathiesen modificado* (*Math mod*) [9], *Hock-Schittkowski* (*HH-66*) [21], *Geiger-Kanzow* (*Gei-Kan*) [22], *Ahn* [23], la función del *Ejemplo A* en [24] y finalmente, un problema que se llamó *Ejemplo 1*, asociado con la función $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definida por $F(x) = ((x_1 - 2)^2 + x_2^2 - 4, (x_1 - 3)^2 + 4x_2^2 - 4)^T$ el cual se construyó para hacer ilustraciones geométricas.

Los algoritmos y las funciones de prueba fueron codificados utilizando el software MATLAB R2022b, y las pruebas numéricas fueron realizadas en un computador con procesador Intel(R) Core(TM) i5-1035G1 CPU @ 1.00GHz 1.19 GHz. Se dirá que el algoritmo converge si $\|\Phi(x_k)\| < 10^{-6}$ y diverge, si el número de iteraciones es mayor que 200, lo cual se indica con el símbolo (-). Para la variación de μ_k en los Algoritmos ACNS y ASF se usaron las sucesiones $\{2^{-k}\}$ y $\{100^{-k}\}$.

3.2.1. Experimento 1

En primer experimento numérico se analiza el comportamiento local de los Algoritmos CNS, SF y NM, en términos de número de iteraciones usando los 10 problemas mencionados anteriormente y los siguientes puntos iniciales.

$$\begin{array}{ll} x_0^1 = 1 & x_0^{10} = (1,1,1,1)^T \\ x_0^2 = 1.5 & x_0^{11} = (1,0,1,1)^T \\ x_0^3 = 2 & x_0^{12} = (1,1,0,1)^T \\ x_0^4 = (4.5,0.8)^T & x_0^{13} = (1,-1,2,-2,5)^T \\ x_0^5 = (3,1)^T & x_0^{14} = (1,1,1,1,1)^T \\ x_0^6 = (10,10,10)^T & x_0^{15} = (1,1,1,1,1,1,1,1)^T \\ x_0^7 = (10,10,8)^T & x_0^{16} = (1,0,1,0,1,0,1,0)^T \\ x_0^8 = (1,1,1)^T & x_0^{17} = (1,1, \dots, 1)^T \\ x_0^9 = (1,0,0,1)^T & x_0^{18} = (-1,-1, \dots, -1)^T. \end{array}$$

Los resultados obtenidos se presentan en la *Tabla 1* que contiene la siguiente información: nombre del problema (Problema), dimensión del problema (Dim), punto inicial

(x_0), número de iteraciones con el Algoritmo CNS (k_{CNS}), número de iteraciones con el Algoritmo SF (k_{SF}), número de iteraciones con el Algoritmo NM (k_{NM}). La sucesión $\{\mu_k\}$ usada se indica en la parte superior de los resultados respectivos.

En la *Tabla 1*, se observa que el Algoritmo CNS que se propone en este trabajo, tiene un buen comportamiento respecto al número de iteraciones, comparado con los otros dos. En general, dicho número es menor o igual que el del Algoritmo FS. No obstante, este último converge en casi todos los casos. En contraste el Algoritmo NM, cuando converge lo hace en menos iteraciones que los suavizados, lo cual no es sorprendente puesto que dicho algoritmo es tipo Newton, y los otros son cuasi-Newton. También se observa que el algoritmo propuesto converge en un número mayor de casos que el Algoritmo NM.

Un aspecto interesante que ilustra la *Tabla 1* es que la elección: $\mu_k = \{100^{-k}\}$ influye en el número de iteraciones de los algoritmos suavizados, y el tiempo de ejecución dado en segundos, en ambos casos este número es menor o igual que los reportados con la sucesión $\mu_k = \{2^{-k}\}$.

Para complementar este experimento, a partir de los datos reportados en la *Tabla 1*, se calcularon los *índices de robustez* (porcentaje de éxito del método), *eficiencia* (velocidad del método en sentido de número de iteraciones) y *eficiencia combinada* (combinación de los dos previos) [25] de los tres algoritmos. Para todos los índices, entre más cercano a uno esté su valor, el método será mejor.

Los resultados se ilustran en la *Figura 3*, en la cual se observa que el Algoritmo CNS tiene el mayor índice de eficiencia y eficiencia combinada mientras que, el Algoritmo SF es el más robusto.

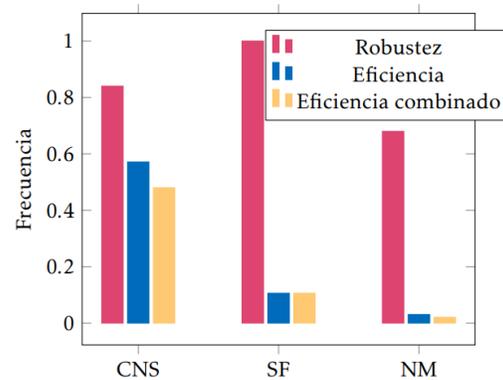


Figura 3. Índices de robustez, eficiencia y eficiencia combinada de los algoritmos.

3.2.2. Experimento 2

Un aspecto importante en el trabajo con algoritmos locales es determinar su radio de convergencia que corresponde al radio de una vecindad de la solución, en la cual cualquiera de sus puntos puede tomarse como inicial y el algoritmo converge. El propósito de este segundo experimento es determinar y comparar los radios de convergencia de los tres algoritmos usando los diez problemas mencionados anteriormente.

Se considera la solución de cada problema x^* indicada en la **Tabla 2** y se procede de la siguiente manera. Se generan 1000 puntos iniciales arbitrarios en una vecindad de esa solución, $B_r(x^*)$, con un radio inicial $r = 0.1$, y se evalúa cada uno de los algoritmos; si para la totalidad de los puntos iniciales se logra convergencia, el radio se incrementa y se repite el proceso.

En caso contrario, se detiene y se registra el valor del radio antes del incremento. Para la métrica de las vecindades se usa la norma infinita y cada uno de los 1000 puntos aleatorios x se genera teniendo en cuenta

que $x = (x_1, x_2, x_3, \dots, x_n)^T$ pertenece a $B_r(x^*)$ si x_i pertenece al intervalo $[x_i^* - r, x_i^* + r]$, para todo $i = 1, \dots, n$. Estas componentes se calcularon mediante la parametrización, $x_i = x_i^* - r + 2t_i$, con $t_i \in [0, 1]$. Es de mencionar que en este experimento, se utilizó la sucesión $\{\mu_k\} = \{100^{-k}\}$.

Los resultados de este experimento se presentan en la **Tabla 3**, donde r_{CNS} , r_{SF} y r_{NM} indican el valor del radio en el que hubo convergencia para todos los puntos iniciales, el símbolo 0.01 – indica que el radio de convergencia es menor que 0.01 y 3+ indica que el radio de convergencia es mayor que 3.

La **Tabla 3** permite observar que en general, los radios de convergencia de los tres algoritmos son pequeños, lo cual está de acuerdo con el carácter local de dichos algoritmos. Además, para la mitad de los problemas, el radio de convergencia del Algoritmo SF supera el de los otros dos, y se presenta un radio de convergencia grande en los dos problemas asociados a funciones lineales.

Tabla 1. Comparación de los Algoritmos CNS, FS y NM

Problema	Dim	x_0	$\{\mu_k\} = \{2^{-k}\}$				$\{\mu_k\} = \{100^{-k}\}$				k_{NM}	t_{NM} (seg.)
			k_{CNS}	t_{CNS} (seg.)	k_{SF}	t_{SF} (seg.)	k_{CNS}	t_{CNS} (seg.)	k_{SF}	t_{SF} (seg.)		
Billups	1	x_0^1	—	—	8	0.78137	—	—	6	0.88772	—	—
Billups	1	x_0^2	—	—	5	0.80843	—	—	3	0.79331	4	0.056681
Billups	1	x_0^3	4	0.054801	4	0.79741	3	0.027482	2	0.77131	3	0.032134
Ejemplo 1	2	x_0^4	7	0.389780	5	0.73847	4	0.12911	4	0.37972	—	—
Ejemplo 1	2	x_0^5	7	0.126870	5	0.36402	5	0.12065	5	0.36268	—	—
Nash-Co	3	x_0^6	4	0.719580	5	0.44706	4	0.40676	5	0.47816	4	0.13309
Nash-Co	3	x_0^7	4	0.421950	5	0.47344	4	0.42574	5	0.47525	4	0.10743
Nash-Co	3	x_0^8	—	—	9	0.53133	—	—	9	0.50928	8	0.18300
Koj-Sh	4	x_0^9	6	0.451030	33	0.79821	6	0.18392	7	0.43828	4	0.33183
Koj-Sh	4	x_0^{10}	—	—	37	0.82768	—	—	10	0.41490	—	—
Koj-Jo	4	x_0^9	7	0.525300	7	0.89756	—	—	19	0.87702	—	—
Koj-Jo	4	x_0^{10}	7	0.299450	6	0.11853	5	0.11039	5	0.70142	1	0.04345
Koj-Jo	4	x_0^{11}	7	0.095853	6	0.94979	5	0.088716	5	0.84041	—	—
Math mod	4	x_0^{10}	7	0.423735	6	0.62618	—	—	—	—	2	0.235610
Math mod	4	x_0^9	12	0.184210	6	0.32572	5	0.08134	4	0.37082	1	0.072393
Math mod	4	x_0^{10}	7	0.074039	6	0.41756	—	—	—	—	1	0.046164
Math mod	4	x_0^{11}	7	0.143370	6	0.43367	3	0.064714	—	—	1	0.050788
Ejemplo A	5	x_0^{13}	5	0.557180	32	0.78761	4	0.22266	7	0.48678	2	0.295510
Ejemplo A	5	x_0^{14}	4	0.115810	32	0.79268	4	0.080743	7	0.49015	2	0.071479
HH-66	8	x_0^{15}	11	1.05220	10	0.59224	—	—	7	0.55213	—	—
HH-66	8	x_0^{16}	10	0.59206	10	0.60842	—	—	7	0.57736	—	—
Gei-Kan	500	x_0^{17}	8	104.680	7	97.680	4	63.672	5	70.414	2	14.904
Gei-Kan	500	x_0^{18}	8	105.290	8	114.480	3	47.440	5	70.982	1	13.506
Ahn	1000	x_0^{17}	10	701.660	7	401.580	4	300.65	6	344.77	2	54.331
Ahn	1000	x_0^{18}	10	720.000	8	446.720	4	301.09	5	288.62	1	44.583

Tabla 2. Soluciones utilizadas en cada problema

Problema	x^*
<i>Billups</i>	2.0488
<i>Ejemplo 1</i>	$(3.7748, 0.9219)^T$
<i>Nash-Co</i>	$(16.6034, 13.7308, 10.7756)^T$
<i>Koj-Sh</i>	$(\sqrt{6}/2, 0, 0, 1/2)^T$
<i>Koj-Jo</i>	$(1, 0, 3, 0)^T$
<i>Math mod</i>	$(1.5, 0, 0, 0)^T$
<i>Ejemplo A</i>	$(0, 3, 1, 0.5, 0)^T$
<i>HH-66</i>	$(0.1841, 1.2022, 3.3273, 0.6655, 0.2, 0, 0, 0)^T$
<i>Gei-Kan</i>	$(0.3660, 0.4641, 0.4904, 0.4974, \dots, 0.4641, 0.3660)^T$
<i>Ahn</i>	$(0.4082, 0.3165, 0.3371, \dots, 0.2660, 0.1835)^T$

Tabla 3. Radios de convergencia en cada problema

Problema	Dim	r_{CNS}	r_{SF}	r_{NM}
<i>Billups</i>	1	0.78	1.07	0.75
<i>Ejemplo 1</i>	2	0.01-	0.01-	0.11
<i>Nash-Co</i>	3	0.11	0.55	0.10
<i>Koj-Sh</i>	4	0.01-	0.61	0.87
<i>Koj-Jo</i>	4	2.20	2.11	3+
<i>Math mod</i>	4	0.31	0.38	0.14
<i>Ejemplo A</i>	5	0.15	1.18	0.11
<i>HH-66</i>	8	3+	3	3+
<i>Gei-Kan</i>	4	3+	3+	3+
<i>Ahn</i>	4	3+	3+	3+

3.2.3. Experimento 3

Para ilustrar la regi3n de convergencia de los tres algoritmos se considera el problema llamado *Ejemplo 1*, descrito al inicio de esta Secci3n, el cual tiene dos soluciones: $x_1^* = (3.7748, 0.9219)^T$ y $x_2^* = (0, 0)^T$. Se construye una malla rectangular de puntos iniciales igualmente espaciados a una distancia de 0.01 unidades, contenida en $[1.77, 5.77] \times [-1.07, 2.92]$ para un total de 160.801 puntos.

La Figuras 4, 5 y 6 ilustran la regi3n de convergencia los Algoritmo CNS, FS y NM, respectivamente. En color verde lim3n aparecen los puntos iniciales para los cuales el algoritmo converge a x_1^* ; en aguamarina los que convergen a x_2^* y en azul los que divergen. Ademas, las curvas $(x_1 - 2)^2 + x_2^2 = 4$ y $(x_1 - 3)^2 + 4x_2^2 = 4$ aparecen dibujadas en color negro.

La Figura 4 ilustra la regi3n de convergencia del algoritmo propuesto. En este caso, hay convergencia para el 99 % de los puntos iniciales generados y divergencia para el 1 % restante. Ademas, los resulta-

dos muestran que, en caso de convergencia, solo con el 5.2 % de los puntos iniciales el algoritmo converge a la soluci3n x_1^* .

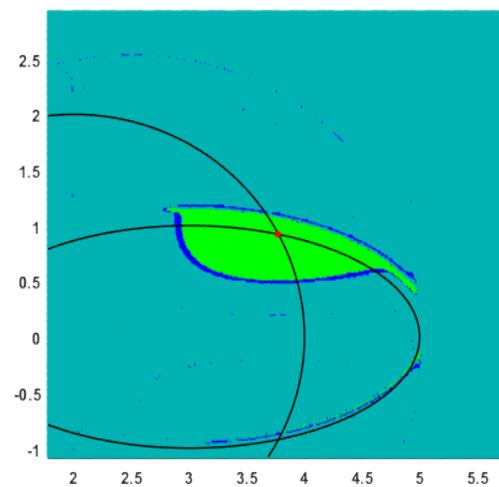


Figura 4. Regi3n de convergencia del Algoritmo CNS.

En la **Figura 5**ab, correspondiente al Algoritmo FS, se observa una amplia región de convergencia formada por el 99.95 % de los puntos iniciales generados, a partir de los cuales con 67.8 % el algoritmo converge a x_1^* . Por otra parte, el algoritmo diverge para el 0.05 % de dichos puntos.

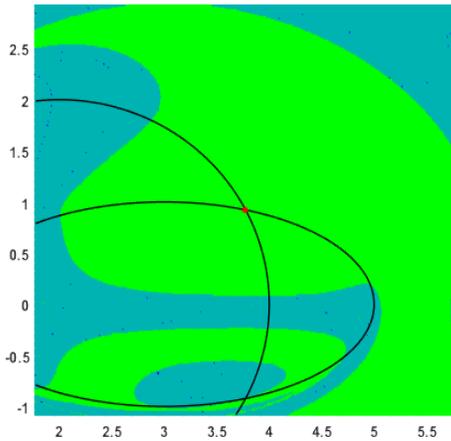


Figura 5. Región de convergencia del Algoritmo FS

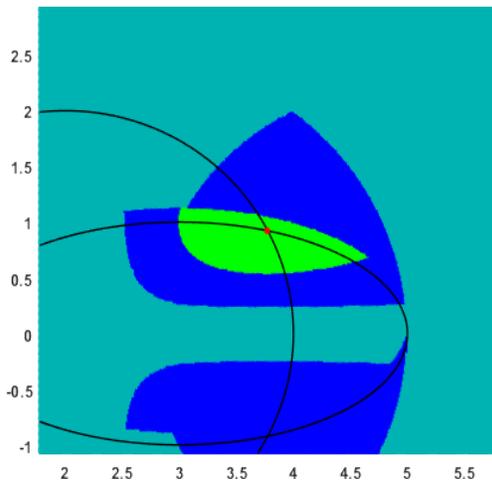


Figura 6. Región de convergencia del Algoritmo NM

La **Figura 6** muestra que el Algoritmo NM tiene una región de convergencia más pequeña que los otros dos algoritmos, y una región de divergencia mayor que los mismos. En efecto, converge para el 75 % de los puntos iniciales (con el 5.48% de estos puntos converge a x_1^*), y diverge para el 25% restante. Este experimento permite observar dos aspectos interesantes: la región de convergencia de los algoritmos no necesariamente es un conjunto convexo y su tamaño puede aumentar usando una suavización, lo que hace de esta estrategia una buena alternativa para tener en cuenta en las propuestas algorítmicas que resuelven el problema de complementariedad.

3.2.4. Experimento 4

Finalmente, con el fin de analizar numéricamente la rapidez de convergencia del algoritmo propuesto, se calcula para cada problema usando la sucesión $\mu_0 2^{-k}$, el cociente R_k definido por

$$R_k = \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2}$$

Los resultados para cuatro problemas (para los otros, ver [26]), se presentan en la **Tabla 4**, donde se reporta el número de iteraciones (k) y el valor de R_k . A pesar de que el algoritmo propuesto converge en un número menor de iteraciones que las reportadas en las siguientes tablas, se ha decidido mostrar más iteraciones para tener una mejor idea de la posible acotación del cociente R_k .

Los resultados de la **Tabla 4** evidencian que el cociente R_k , para los cuatro problemas considerados, parece estar acotado, lo cual significa que el algoritmo propuesto puede ser cuadráticamente convergente. En efecto, dicho algoritmo converge cuadráticamente, bajo ciertas hipótesis, como se demostró en el Sección 4.

Tabla 4. Cociente R_k del Algoritmo CNS

(a) <i>Billups</i>		(b) <i>Koj-Shi</i>	
k	R_k	k	R_k
0	0.508438	0	0.486029
1	0.423711	1	0.244749
2	0.567450	2	0.312568
3	0.458637	3	0.315030
4	0.500002	4	0.315037
5	0.465112	5	0.315037
6	0.500002	6	0.315037
7	0.465112	7	0.315037
8	0.500002	8	0.315037
9	0.465112	9	0.315037
10	0.500002	10	0.315037
11	0.465112	11	0.315037

(c) <i>HH-66</i>		(d) <i>Ejem A</i>	
k	R_k	k	R_k
0	0.151527	0	0.020098
1	0.181261	1	0.147162
2	9.445162	2	0.095277
3	0.000005	3	0.095238
4	0.112197	4	0.095238
5	0.045331	5	0.095238
6	0.047551	6	0.095238
7	0.047530	7	0.095238
8	0.047530	8	0.095238
9	0.047530	9	0.095238
10	0.047530	10	0.095238
11	0.047530	11	0.095238

4. Conclusiones

Recientemente mtodos con jacobiano suavizado han resultado muy eficientes para resolver el problema de complementariedad no lineal, va su reformulacin como un sistema de ecuaciones no lineales, el cual resulta ser no diferenciable.

En este trabajo se propone un nuevo algoritmo local cuasi-Newton suavizado para complementariedad no lineal usando la funcin Mnimo y una la suavizacin de la misma. Se demuestran algunas propiedades de estas dos funciones junto con la estrategia del jacobiano suavizado para obtener resultados de convergencia.

Pruebas numricas preliminares muestran un buen desempeo local del algoritmo propuesto; adems, se determina su regin de convergencia, y se analiza numricamente su rapidez de convergencia, la cual parece ser cuadrtica, coincidiendo con lo demostrado tericamente. En las pruebas numricas se evidencia, que en la mayora de casos, el mtodo con jacobiano suavizado que utiliza una suavizacin de la funcin de complementariedad de Fischer es ms exitoso que el propuesto, pero en cuanto a los promedios de iteraciones y tiempo, el algoritmo que proponemos lo supera.

Finalmente, sera interesante usar una estrategia hbrida, quiz usando la funcin de Fischer, para globalizar el algoritmo y analizar su desempeo numrico. Esto permitira usar las ventajas ya conocidas mencionadas en la introduccin de cada una de esas funciones.

Financiacin

No aplica.

Contribuciones de los autores

C. Quir: Anlisis Formal, Escritura – Borrador original. R. Prez: Conceptualizacin, Metodologa, Investigacin, Redaccin, Revisin y Edicin. F. Arenas: Conceptualizacin, Investigacin, Escritura Revisin y Edicin. D. Correa: Investigacin y Pruebas Numricas.

Todos los autores han ledo y aceptado la versin publicada del manuscrito.

Conflictos de inters

Los autores declaran no tener conflicto de intereses.

Declaracin de la Junta de Revisin Institucional

No aplica.

Declaracin de consentimiento informado

No aplica

Referencias

- [1] M. M. Kostreva, “Elasto-hydrodynamic lubrication: A non-linear complementarity problem,” *International Journal for Numerical Methods in Fluids*, vol. 4, no. 4, pp. 377–397, 1984, doi: <https://doi.org/10.1002/flid.1650040407>
- [2] M. Anitescu, J. F. Cremer, F. A. Potra, “On the existence of solutions to complementarity formulations of contact problems with friction,” *Complementarity and Variational Problems: State of the art*. SIAM Publications, 1997, pp. 12–21.
- [3] M. C. Ferris and J. S. Pang, “Engineering and economic applications of complementarity problems,” *SIAM Review*, vol. 39, pp. 669–713, 1997, doi: <https://doi.org/10.1137/S0036144595285963>
- [4] A. Chen, J.-S. Oh, D. Park, W. Recker, “Solving the bicriteria traffic equilibrium problem with variable demand and nonlinear path costs,” *Applied Mathematics and Computation*, vol. 217, no. 7, pp. 3020–3031, 2010, doi: <https://doi.org/10.1016/j.amc.2010.08.035>
- [5] R. Prez, F. Arenas, H. J. Martnez, C. Arias, *El problema de complementariedad no lineal: Teora, aplicaciones y nuevos algoritmos para su solucin*. Programa Editorial de la Universidad del Valle, 2019.
- [6] C. Kanzow and H. Kleinmichel, “A new class of semismooth newton-type methods for nonlinear complementarity problems,” *Computational Optimization and Applications*, vol. 11, no. 3, pp. 227–251, 1998, doi: <https://doi.org/10.1023/A:1026424918464>
- [7] S. L. Hu, Z. H. Huang, and J. S. Chen, “Properties of a family of generalized ncp functions and a derivative free algorithm for complementarity problems,” *Journal of Computational and Applied Mathematics*, vol. 230, no. 1, pp. 69–82, 2009, doi: <https://doi.org/10.1016/j.cam.2008.10.056>
- [8] L. Yong, “Nonlinear complementarity problem and solution methods,” in *Artificial Intelligence and Computational Intelligence*, F. L. Wang, H. Deng, Y. Gao, and J. Lei, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 461–469. doi: https://doi.org/10.1007/978-3-642-16530-6_55

- [9] J. S. Pang and L. Qi, “Nonsmooth equations: Motivation and algorithms,” *SIAM Journal on Optimization*, vol. 3, no. 3, pp. 443–465, 1993, doi: <https://doi.org/10.1137/0803021>
- [10] A. Fischer and C. Kanzow, “On finite termination of an iterative method for linear complementarity problems,” *Math. Program.*, vol. 74, no. 3, pp. 279–292, 1996, doi: <https://doi.org/10.1007/BF02592200>
- [11] A. Sherman, “On newton-iterative methods for the solution of systems of nonlinear equations,” *SIAM Journal on Numerical Analysis*, vol. 15, no. 4, pp. 755–711, 1978, doi: <https://doi.org/10.1137/0715050>
- [12] C. G. Broyden, J. E. Dennis, and J. J. Moré, “On the Local and Superlinear Convergence of Quasi-Newton Methods,” *IMA J. Appl. Math.*, vol. 12, no. 3, pp. 223–245, 1973, doi: <https://doi.org/10.1093/imat/12.3.223>
- [13] V. L. R. Lopes, J. M. Martínez, and R. Pérez, “On the local convergence of quasi-newton methods for nonlinear complementary problems,” *Applied Numerical Mathematics*, vol. 30, pp. 3–22, 1999, doi: [https://doi.org/10.1016/S0168-9274\(98\)00080-4](https://doi.org/10.1016/S0168-9274(98)00080-4)
- [14] C. Kanzow and H. Pieper, “Jacobian smoothing methods for nonlinear complementarity problems,” *SIAM Journal on Optimization*, vol. 9, no. 2, pp. 342–373, 1999, doi: <https://doi.org/10.1137/S1052623497328781>
- [15] F. E. Arenas, H. J. Martínez, and R. Pérez, “A local jacobian smoothing method for solving nonlinear complementarity problems,” *Universitas Scientiarum*, vol. 25, no. 1, p. 149–174, May 2020, doi: <https://doi.org/10.11144/JAVERIANA.SC25-1.ALJS>
- [16] F. H. Clarke, *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, 1990.
- [17] X. Chen, “Smoothing methods for complementarity problems and their applications: a survey,” *J. Oper. Res. Soc. Japan*, vol. 43, no. 1, pp. 32–47, 2000.
- [18] J. E. Dennis and R. B. Schnabel, *Numerical methods for unconstrained optimization and non-linear equations*. Society for Industrial and Applied Mathematics, 1996.
- [19] S. C. Billups, “Algorithms for complementarity problems and Generalized Equations,” Ph.D. dissertation, University of Wisconsin, 1995.
- [20] P. T. Harker, “Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities,” *Mathematical Programming*, vol. 41, pp. 29–59, 1988, doi: <https://doi.org/10.1007/BF01580752>
- [21] J. Zhu and B. Hao, “A new smoothing method for solving nonlinear complementarity problems,” *Open Mathematics*, vol. 17, no. 1, pp. 104–119, 2019, doi: <https://doi.org/10.1515/math-2019-0011>
- [22] C. Geiger and C. Kanzow, “On the resolution of monotone complementarity problems,” *Computational Optimization and Applications*, vol. 5, no. 2, pp. 155–173, 1996, doi: <https://doi.org/10.1515/math-2019-0011>
- [23] A. Byong-Hun, “Iterative methods for linear complementarity problems with upperbounds on primary variables,” *Mathematical Programming*, vol. 26, no. 3, pp. 295–315, oct 1983.
- [24] Y. Xia, H. Leung, and J. Wang, “A projection neural network and its application to constrained optimization problems,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 4, pp. 447–458, 2002, doi: <https://doi.org/10.1109/81.995659>
- [25] S. Buhmiller and N. Krejić, “A new smoothing quasi-newton method for nonlinear complementarity problems,” *Journal of Computational and Applied Mathematics*, vol. 211, no. 2, pp. 141–155, feb 2008, doi: <https://doi.org/10.1016/j.cam.2006.11.007>
- [26] C. Quirá, “Un algoritmo local cuasi-Newton suavizado para resolver el problema de complementariedad no lineal,” Ph.D. dissertation, Universidad del Cauca, Popayán, Colombia, Agosto 2021.