

Un algoritmo híbrido de región de confianza y
búsqueda lineal para el problema de
complementariedad no lineal usando el método de
Levenberg-Marquardt



Carlos Andrés Leiton Piamba

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Maestría en Ciencias Matemáticas
Popayán
2022.

**Un algoritmo híbrido de región de confianza y
búsqueda lineal para el problema de
complementariedad no lineal usando el método de
Levenberg-Marquardt**

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

Linea de investigación-Optimización

Carlos Andrés Leiton Piamba

Director:

Mg. Favián Enrique Arenas Aparicio

Profesor de la Universidad del Cauca

Universidad del Cauca

Facultad de Ciencias Naturales, Exactas y de la Educación

Maestría en Ciencias Matemáticas

Popayán

2022.

Nota de aceptación

Director _____

Mg. Favián Enrique Arenas Aparicio

Jurado _____

Dr. Carlos Andrés Arias Torres

Jurado _____

Dr. Héctor Jairo Martínez Romero

Lugar y fecha de sustentación: Popayán, 13 de diciembre de 2022.

Agradecimientos

*Puedes hacer todos los planes que quieras, pero el propósito de Dios prevalecerá. (**Proverbios 19:21**)*

A mi familia por todo su apoyo incondicional. A mi padre por ser siempre un ejemplo y modelo a seguir. A mis hermanos Felipe, Wilmer, Eider quienes con su apoyo y consejos me han guiado en las etapas más importantes de mi vida. A mi hermana quien me inspiro y motivo a continuar con cada uno de mis sueños, a ella le dedico este logro profesional.

A mi director de tesis, Mg. Favián Enrique Arenas Aparicio, por su acompañamiento y poner a disposición toda su experiencia y conocimiento para que lográramos alcanzar nuestros objetivos en este trabajo, pero sobre todo, por su valiosa amistad.

Al Dr. Héctor Jairo Martínez por sus aportes, los cuales fueron de suma importancia para dar precisiones que se necesitaban el documento.

Al Dr. Carlos Andrés Arias Torres, por sus acertadas sugerencias y por brindarme un mejor panorama sobre la teoría del problema que se trabajó en este documento.

A mis compañeros y docentes por todos sus aportes en pro de mi formación académica.

Carlos Andrés Leiton Piamba

Índice general

Índice general	3
1. Introducción	5
2. Preliminares	11
2.1. Existencia de soluciones	11
2.2. Reformulación del PCNL	19
2.3. Propiedades del operador φ_λ	26
2.4. Estrategias de globalización	36
2.4.1. Región de confianza para un problema de minimi- zación sin restricciones	37
2.4.2. Región de Confianza para el PCNL(F)	44

2.4.3.	Búsqueda lineal para un problema de minimización sin restricciones	48
2.4.4.	Búsqueda lineal para el $PCNL(F)$	48
2.4.5.	Combinación de las técnicas de globalización región de confianza y búsqueda lineal para $PCNL(F)$	49
2.5.	Teoremas para el análisis de la convergencia del algoritmo	51
3.	Algoritmo y Teoría de convergencia	55
3.1.	Algoritmo	55
3.2.	Convergencia global	57
3.3.	Convergencia q-superlineal y q-cuadrática	71
4.	Resultados numéricos	79
4.1.	Resumen	109
4.2.	Comparación numérica	110
4.3.	Comentarios finales	116
	Bibliografía	118

Introducción

Dada $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x}))^T$ 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, un vector $\mathbf{z} \in \mathbb{R}^n$ satisface que $\mathbf{z} \geq 0$ si y solo si $z_i \geq 0$ para cada $i = 1, \dots, n$. La tercera condición en el planteamiento del PCNL(F), exige que los vectores \mathbf{x} y $F(\mathbf{x})$ sean ortogonales, por ello, es llamada condición de complementariedad, la cual trae consigo implícitamente la búsqueda de un equilibrio entre la variable \mathbf{x} y $F(\mathbf{x})$. Así, el concepto de complementariedad es equivalente al concepto de sistema en equilibrio.

La motivación para el desarrollo investigativo del PCNL no solo desde el punto de vista teórico si no algorítmico, radica en las numerosas aplicaciones que surgen en problemas de ingeniería, como son los equilibrios de tráfico [10], en física, en problemas de contacto físico de estructuras mecánicas [24], en problemas de lubricación elasto-hidrodinámicas [2], así como en problemas relacionados con modelos de equilibrio económicos[25].

En la literatura sobre este tema, existen dos formas de afrontar este problema, estos son:

1. Como un sistema de ecuaciones no lineales, mediante el uso de ciertas funciones de complementariedad [40].
2. Como un problema de minimización usando las funciones de mérito [37, 28].

Estas reformulaciones se pueden resolver mediante diferentes métodos, como por ejemplo, redes neuronales [5], tipo Newton o tipo cuasi-Newton no diferenciables [57, 59, 8, 42, 43, 4], tipo Levenberg-Marquardt [20, 17].

Los métodos mencionados anteriormente tienen una desventaja, el punto inicial debe estar cerca de la solución, es decir, son métodos locales. Por eso, una de las características deseables en los métodos numéricos iterativos es la convergencia global, es decir, sin importar donde se tome el punto inicial el método converge. A razón de esto se aborda el PCNL como un problema de minimización y para este, las dos estrategias más usadas para globalizar un algoritmo local que resuelve (1.1) son búsqueda lineal y región de confianza.

La mayoría de algoritmos que existen la literatura del PCNL han aplicado la estrategia de búsqueda lineal, que se describe a continuación:

1. Búsqueda direccional, el algoritmo elige una dirección de descenso d_k según el caso (*Gradiente*, *Newton* y *cuasi-Newton*).
2. Partiendo desde el iterante actual, buscamos a lo largo de esa dirección un nuevo punto en el cual la función objetivo decrezca de manera suficiente, para ello se puede utilizar la condición de *Armijo* [31], la condición de *Goldstein*[18, 31] o condiciones libres de derivadas [41].
3. Finalmente, se actualizan los valores y se repite el proceso.

La segunda estrategia de globalización es la ya mencionada región de confianza, donde destacamos los siguientes procesos del método:

1. A partir de la función objetivo G , se construye un modelo cuadrático q_k por medio de una aproximación de *Taylor*.
2. Como el modelo q_k puede no ser una buena aproximación de G cuando \mathbf{x} está lejos de \mathbf{x}_k , restringimos la búsqueda del siguiente punto en un región de confianza, donde esperamos que la aproximación del modelo sea aceptable.

Es decir, encontramos el candidato \mathbf{d}_k resolviendo de manera aproximada el siguiente subproblema:

$$\begin{aligned} \mathbf{d}_k &:= \text{mín } q_k(\mathbf{d}) \\ &\|\mathbf{d}\| \leq \Delta_k \end{aligned}$$

3. Si la solución aproximada no produce un decrecimiento suficiente en G , se concluye que la región de confianza es demasiado grande y debe reducirse, para después resolver el nuevo subproblema asociado.

Desde la introducción del Problema de Complementariedad No Lineal en 1964 por *Cottle* [14] hasta la actualidad, se han escrito una gran cantidad de artículos y libros sobre el tema. Se puede ordenar en una línea de tiempo, las herramientas que han permitido resolver este problema de la siguiente manera:

En 1973 *Clark* presentó en su tesis doctoral el concepto de Jacobiano generalizado y algunas de sus propiedades, este concepto de análisis es fundamental para la construcción de los métodos siguientes [11].

Posteriormente en 1976, *Mangasarian* demuestra que el PCNL es equivalente a resolver un sistema de ecuaciones no lineales, no diferenciable [46].

Es importante mencionar que dentro la literatura del problema, las funciones de Complementariedad más usadas por los investigadores se han ido introduciendo desde 1991, donde *Pang* reformula el PCNL como un sistema de ecuaciones no lineales usando la *función de complementariedad mínimo*, además proponen un método *tipo Newton generalizado* donde, demuestra convergencia superlineal y convergencia global [51].

Después, en 1992 *Fischer* presenta de manera implícita la función de complementariedad ahora conocida como la *función de Fischer-Burmeister*, es de aclarar que los autores desde 1985 ya habían trabajado con esta función [26].

En 1994, *Gabriel* y *Pang* desarrollan y analizan la convergencia de un método de región de confianza para el problema de complementariedad no lineal [28] y en este mismo año *Kanzow* establece seis funciones de complementariedad para la reformulación del PCNL como un sistema de ecuaciones no lineales [35].

En 1996, *De Luca*, *Facchinei* y *Kanzow* propone una forma de elegir elementos del Jacobiano generalizado, además usan la *función de Fischer-Burmeister* para la reformulación del PCNL y presentan un algoritmo *tipo Newton*, donde demuestran su convergencia local y global [15].

En 1997, *Pérez* y *Martínez* introducen una familia de métodos secantes para el PCNL mediante la reformulación como un sistema de ecuaciones no lineales usando tanto la *función de mínimo* como la *función de Fischer-Burmeister*, además demuestran convergencia superlineal y comparan las dos reformulaciones teórica y numéricamente [43].

En el año de 1998, *Kanzow* y *Kleinmichel* presentan varios resultados importantes, entre ellos, introducen la familia uniparamétrica de funciones de complementariedad, y presentan una investigación detallada de las propiedades del operador de la ecuación no lineal, así como de la función de

mérito correspondiente, otro de los resultados de suma importancia es que caracteriza el C-subdiferencial de una función localmente *Lipschitz* continua. Además proponen un algoritmo *tipo Newton*, y por último muestran los resultados numéricos obtenidos [38].

En el mismo año, *Houyuan y Fukushima* propone un algoritmo no monótono de región de confianza usando la *función de Fischer-Burmeister* para la reformulación del problema generalizado de complementariedad, demuestra convergencia global, superlineal y cuadrática [34], con base en esto, *Kanzow* propone un algoritmo inexacto de región de confianza para el PCNL usando la familia de funciones uniparamétricas de complementariedad para la reformulación del problema, demuestra convergencia global, superlineal y cuadrática, además realiza una comparación numérica entre su método y el método propuesto por *Fukushima* [39].

En 2010, *Shou-Qiang y Gao* muestran dos algoritmos *tipo Levenberg-Marquardt* no monótonos usando la técnica de región de confianza, utilizan la *función de Fischer-Burmeister* para la reformulación del problema como un sistema de ecuaciones no lineales y además demuestran que son globalmente convergentes [17].

En el año de 2015, *Arenas, Martínez y Pérez* presentan un método *cuasi Newton* no diferenciable para resolver el PCNL usando la familia uniparamétrica de funciones de complementariedad para la reformulación como un sistema no lineal, desarrollan su teoría de convergencia local y presentan ciertas propiedades dicha función [4].

En 2016, con base al artículo presentado en [4] *Arias, Martínez y Pérez* proponen un algoritmo *cuasi Newton global* para resolver el PCNL, mediante su reformulación como un problema de minimización de la función de mérito mediante la técnica de búsqueda lineal y desarrollan la teoría de convergencia global del algoritmo [6].

Recientemente, en [20], los autores proponen un algoritmo tipo *Levenberg-Marquardt* no monótono combinando las técnicas de región de confianza y búsqueda lineal para resolver el Problema de Complementariedad No Lineal usando la función de complementariedad de *Fischer-Bermeinster* para la reformulación del sistema no lineal y obtienen convergencia global, superlineal y cuadrática, además realizan una comparación numérica con el algoritmo propuesto en [45].

Motivados por los resultados obtenidos en [20], buscamos proponer un método híbrido que permita resolver Problemas de Complementariedad no Lineal que sea eficiente y eficaz. Es por esto que; en este trabajo de investigación, proponemos un algoritmo híbrido que utiliza región de confianza y búsqueda lineal, utilizando en cada paso el método de *Levenberg-Marquardt* y para la reformulación del sistema no lineal usamos la familia uniparamétrica presentada por [38].

Este documento se ha organizado de la siguiente forma: En el **Capítulo 2** presentamos algunos resultados teóricos que son útiles en el desarrollo de la teoría de convergencia del algoritmo que proponemos. En el **Capítulo 3**, presentamos un algoritmo híbrido para el PCNL, el algoritmo combina las técnicas de región de confianza y búsqueda lineal. Además, demostraremos la convergencia global del método propuesto y bajo ciertas hipótesis, la convergencia q-superlineal y q-cuadrática. En el **Capítulo 4**, presentamos los resultados numéricos, al comparar nuestro algoritmo con otros métodos que resuelven el PCNL(F) (1.1), en la segunda parte de este capítulo realizamos una comparación con el algoritmo propuestos por los autores en [20], para ello determinamos un $\lambda \in (0, 4)$ óptimo para el algoritmo 1 y así comparar estos dos métodos. Por ultimo en el **Capítulo 5**, presentamos comentarios finales sobre los resultados numéricos y algunas propuestas que motivan futuros trabajos en el tema.

Capítulo 2

Preliminares

En la literatura relacionada con problemas de complementariedad no lineal, una de las estrategias más utilizadas consiste en reformular el PCNL como un sistema de ecuaciones no lineales, no diferenciales. Luego, mediante métodos que resuelven este tipo de sistemas se encuentran las soluciones al problema, si estas existen.

2.1. Existencia de soluciones

Partiendo de la importancia del PCNL, queremos saber ¿cuando el problema tiene solución?, ¿para qué casos existe única solución? y ¿en qué casos tiene múltiples soluciones?.

Para el caso del Problema de Complementariedad No Lineal, la existencia y unicidad están ligadas a algunas condiciones suficientes que debe cumplir la función que define el problema (1.1). Los autores en [53], realizan el estudio analítico del problema de existencia y unicidad de soluciones para el PCNL(F).

Nosotros, con la intención de ilustrar la existencia de la solución en estos problemas, vamos a analizar geoméricamente el problema (1.1) para el caso en que, $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ donde

$$F(x, y) = \begin{pmatrix} F_1(x, y) \\ F_2(x, y) \end{pmatrix}.$$

Para ello, vemos necesario presentar la siguiente definición en forma general.

Definición 2.1. *Un vector $\mathbf{x} \in \mathbb{R}^n$ es factible para el PCNL(F), si $\mathbf{x} \geq 0$ y $F(\mathbf{x}) \geq 0$, y estrictamente factible, si $\mathbf{x} > 0$ y $F(\mathbf{x}) > 0$. Además, al conjunto de puntos factibles para el PCNL(F) se le llamara región factible y lo denotaremos por Γ .*

Observemos que si restringimos el problema a 2 dimensiones, las tres condiciones del problema nos permiten inferir lo siguiente:

- Si existe la solución, esta debe ser un punto en el primer cuadrante, por la primera condición.
- Las restricciones $F_1(x, y) \geq 0$ y $F_2(x, y) \geq 0$ definen regiones en el plano cartesiano.
- Si la intersección entre estas regiones es vacío, el problema no tiene solución; en caso contrario, los puntos de dicha intersección que están en el primer cuadrante son los candidatos a ser soluciones del problema.
- La restricción de complementariedad permite determinar cual o cuales de esos puntos es o son soluciones.

A continuación presentamos algunos ejemplos que describen lo anterior.

Ejemplo 2.1. Dada la función $F(x, y)$ y su gráfica como en la Figura (2.1), se tiene que los posibles casos para la solución de $PCNL(F)$, pueden ser los cortes con los ejes coordenados y el origen, es decir A , B , C y O .

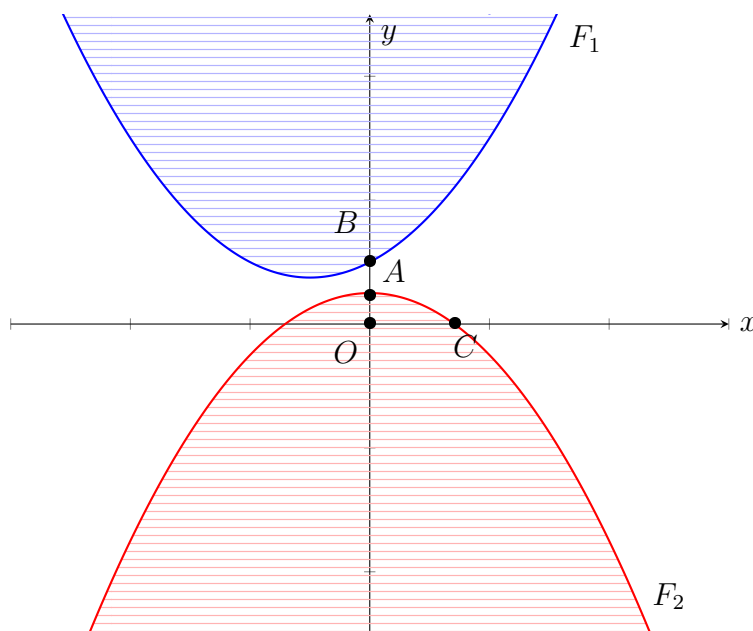


Figura 2.1: El $PCNL(F)$ no tiene solución.

Observemos que al evaluar el punto $A \geq 0$, se tiene que la función componente $F_1(A) < 0$, es decir, A no cumple la segunda condición, por tanto A no es una solución para el $PCNL(F)$, de igual manera se puede verificar que para los puntos $B \geq 0$, $C \geq 0$ y $O \geq 0$ al evaluarlos se tiene $F_2(B) < 0$, $F_1(C) < 0$ y $F_1(O) < 0$. De ahí que, B , C y O no cumplen la segunda condición, por tanto el $PCNL(F)$ no tiene solución, esto ocurre cuando la intersección entre las dos regiones es vacía.

Ahora, para la función

$$F(x, y) = \begin{pmatrix} F_1(x, y) \\ F_2(x, y) \end{pmatrix},$$

como lo muestran las Figuras 2.2, 2.3, 2.4 y 2.5; se ha ilustrado distintas

intersecciones con el fin analizar las posibles soluciones para el PCNL, estas pueden ser cortes con los ejes coordenados, el punto de intersección de las dos gráficas y el origen, es decir $A = (0, y_1)$, $B = (a, b)$, $C = (0, y_2)$ y $O = (0, 0)$.

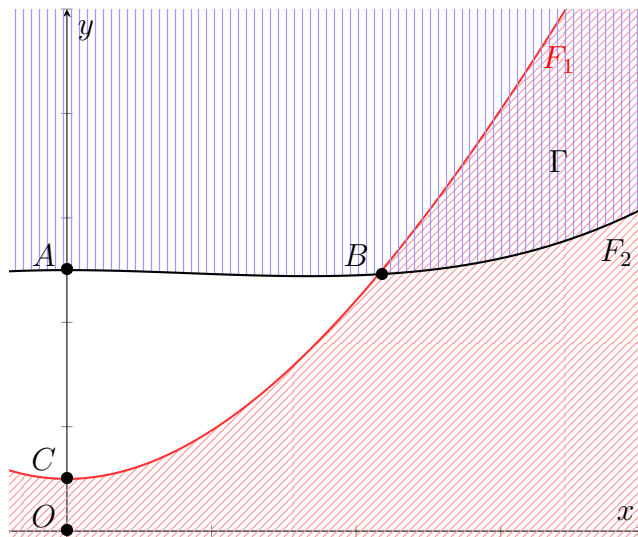


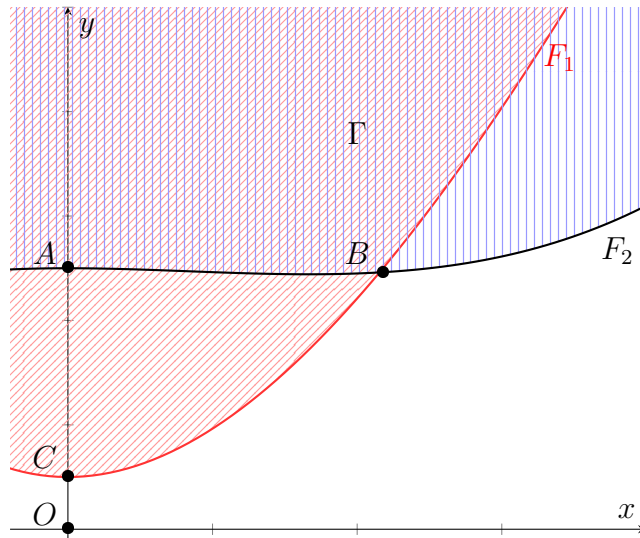
Figura 2.2: Solución única B

Ejemplo 2.2. La región factible que definen el PCNL(F) está dada como lo muestra la Figura 2.2; se tiene que la condición primera del problema se cumple para los puntos A , B , C y O por estar en el primer cuadrante. Luego, para la segunda condición, al evaluar los puntos en las funciones componentes, se tiene que:

$F_1(A) < 0$, $F_2(A) = 0$, $F_1(C) = 0$, $F_2(C) < 0$, $F_1(O) > 0$, $F_2(A) < 0$, es decir A , C y O no cumplen la segunda condición, por tanto A , C , O no son soluciones para el PCNL(F). Por otro lado, se tiene que B es tal que:

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

cumple la segunda y tercera condición del problema, por tanto B es la única solución para el PCNL(F).

Figura 2.3: Dos soluciones B y A

Ejemplo 2.3. La región factible que define el $\text{PCNL}(F)$ se muestra en la región doblemente sombreada de la Figura 2.3; el mismo argumento del ejemplo anterior, nos implica que los puntos A , B , C y O cumplen la primera condición del problema. Para el punto A se tiene que $F_2(A) = 0$, y existe un $r > 0$, tal que $F_1(A) = r$, además

$$F(A)^T A = \begin{pmatrix} r & 0 \end{pmatrix} \begin{pmatrix} 0 \\ y_1 \end{pmatrix} = 0,$$

con lo cual, A cumple también la segunda y tercera condición del problema; con respecto al punto B , tenemos que $F(B)^T = (0, 0) \geq 0$, y $F(B)^T B = 0$, así que B también cumple las tres condiciones del problema, por tanto, A y B son soluciones para el $\text{PCNL}(F)$.

Por otra parte, se puede observar que $F_1(O) < 0$, $F_2(O) < 0$, $F_1(C) = 0$, $F_2(C) < 0$, es decir O y C no cumplen la condición segunda, de ahí que O y C no son soluciones para el $\text{PCNL}(F)$.

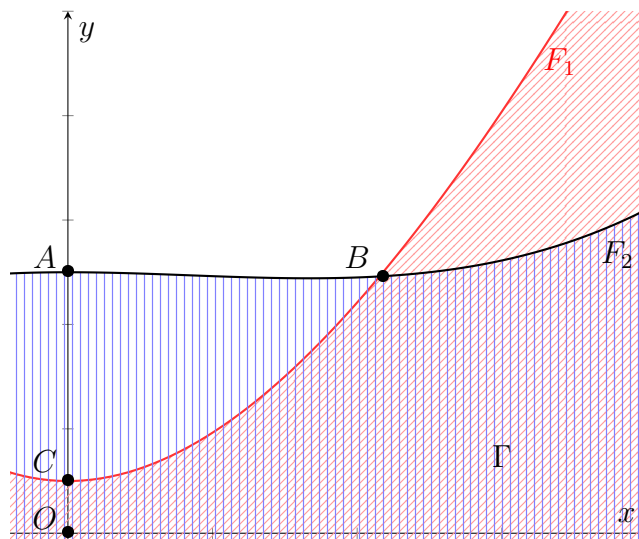


Figura 2.4: Dos o tres soluciones para el PCNL(F)

Ejemplo 2.4. La región factible que definen el PCNL(F) está dado como lo muestra la Figura 2.4. Como ya sabemos, los puntos A , B , C y O cumplen la condición primera del problema. Luego, al verificar la segunda condición para los puntos A y C se tiene que $F_1(A) < 0$, $F_2(A) = 0$, $F_1(C) = 0$, $F_2(C) > 0$, es decir A no cumple la segunda condición, de ahí que A no es solución para el PCNL(F), además $F(C) \geq 0$, pero para $r > 0$, con $F_2(C) = r$ se tiene

$$F(C)^T C = \begin{pmatrix} 0 & r \end{pmatrix} \begin{pmatrix} 0 \\ y_2 \end{pmatrix} = ry_2 \neq 0,$$

es decir, C no cumple la tercera condición del problema, por tanto C no es solución.

Podemos observar que si la posición de las funciones componentes de F se invierten, se tiene una función diferente; pero la región factible es la misma, la nueva función es de la forma:

$$\tilde{F}(x, y) = \begin{pmatrix} F_2(x, y) \\ F_1(x, y) \end{pmatrix},$$

entonces, C sería solución para el PCNL(\tilde{F}), por tanto dar un interpretación geométrica sobre las soluciones sin conocer explícitamente las funciones componentes que definen a F puede llevar a cometer errores.

Por último, se tiene que $F_1(O) > 0$, $F_2(O) > 0$, y $F(B)^T = (0, 0) \geq 0$, es decir son puntos factibles, cumplen la condición primera y segunda del problema. Además, $F(B)^T B = 0 = F(O)^T O$ por tanto, las soluciones para el PCNL(F) son O y B .

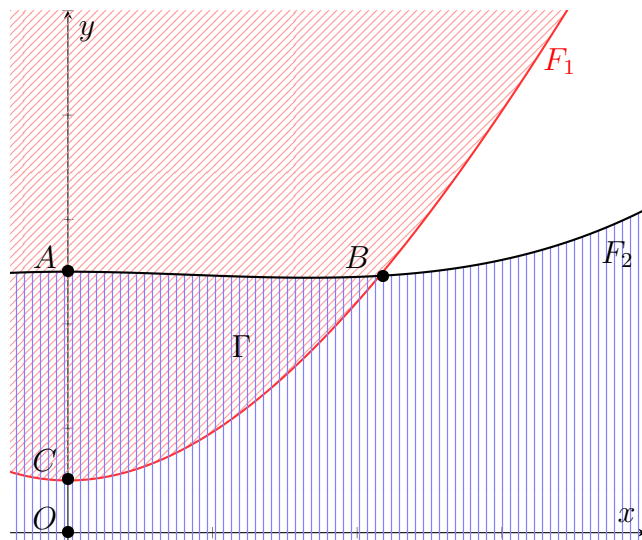


Figura 2.5: Dos o tres soluciones para el PCNL(F)

Ejemplo 2.5. La región factible que definen el PCNL(F) está dado como lo muestra la Figura 2.5.

Como antes, solo basta con verificar la segunda y tercera condición del problema, esto es $F_1(A) > 0$, $F_2(A) = 0$, $F_1(B) = 0$, $F_2(B) = 0$, $F_1(C) = 0$, $F_2(C) > 0$, $F_1(O) < 0$, $F_2(O) > 0$. De ahí que O no es solución, ahora verifiquemos si A , B , C cumplen la tercera condición, $F(A)^T A = 0$, $F(B)^T B = 0$ y $F(C)^T C \neq 0$, por tanto A y B son las únicas soluciones para el PCNL(F).

De acuerdo a lo anterior, se puede caracterizar las soluciones para PCNL(F)

cuando $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, estas soluciones se encuentran en la frontera del cuadrante positivo, o son las raíces no negativas de F .

A continuación presentaremos un ejemplo para una función en particular, en él se determinará los posibles candidatos a ser solución del problema, y entre ellos verificaremos cuales cumplen las tres condiciones del problema (1.1), además se ilustra en la Figura (2.6) la región factible.

Ejemplo 2.6. Sea $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ una función continuamente diferenciable, la cual definimos de la siguiente manera,

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

En la Figura 2.6, los posibles puntos en los cuales se puede encontrar

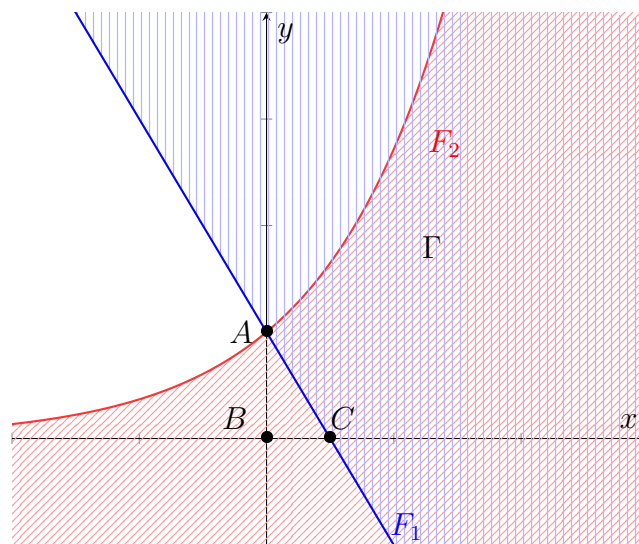


Figura 2.6: Dos soluciones A y C

la solución al PCNL(F) son; $A = (0, 1/2)$, $B = (0, 0)$, $C = (1/2, 0)$, los cuales provienen de los cortes con los ejes coordenados, el origen, o el punto de intersección de las dos gráficas.

Al evaluar el punto $B \geq 0$, se tiene que $F_1(B) < 0$ es decir, no cumple la

condición de no negatividad de su imagen. Además, se tiene que el punto A que es la intersección entre las curvas y C que es la intersección con la frontera del cuadrante positivo, son soluciones del PCNL(F).

Puesto que A y C son mayores o iguales a cero, donde al evaluarlos se cumple $F(A)^T = (0, 0)$, $F(C)^T = (0, e^{\frac{1}{2}})$, además, también satisfacen la condición de complementariedad,

$$C^T F(C) = \left(\frac{1}{2} \ 0\right) \begin{pmatrix} 0 \\ e^{\frac{1}{2}} \end{pmatrix} = 0 = A^T F(A) = \left(0 \ \frac{1}{2}\right) \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

entonces, C y A son soluciones para el problema.

2.2. Reformulación del PCNL

Para la reformulación del problema (1.1) como un sistema de ecuaciones no lineales, no diferenciables, se considera ciertas funciones especiales, llamadas funciones de complementariedad [65].

Definición 2.2. Sea $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, φ es una función de complementariedad si cumple:

$$\varphi(a, b) = 0 \Leftrightarrow a \geq 0, b \geq 0, ab = 0. \quad (2.1)$$

En la literatura sobre el problema de complementariedad no lineal, como ya se había mencionado anteriormente existen numerosas funciones de complementariedad, entre las más usadas se encuentran:

La función mínimo [52],

$$\varphi_{\text{mín}}(a, b) = \text{mín}(a, b) = \frac{(a + b) - |a - b|}{2},$$

la cual no es diferenciable en los puntos de la forma (a, a) .

La *función de Fischer-Bermeinster* [26], definida como

$$\varphi_{FB}(a, b) = \sqrt{a^2 + b^2} - a - b,$$

donde la no diferenciability se presenta en $(0, 0)$.

Por último en 1998, *Kanzow y Kleinmichel* [37], presentaron la familia uniparamétrica de funciones de complementariedad φ_λ , definida por

$$\varphi_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} - a - b, \quad \lambda \in (0, 4) \quad (2.2)$$

donde la no diferenciability se presenta en $(0, 0)$, además es fácil verificar que para $\lambda = 2$, φ_λ se reduce a la *función de Fischer-Bermeinster* y cuando λ converge a cero, φ_λ tiende a un múltiplo de la *función mínimo*.

Luego, para reformular el problema de complementariedad no lineal (1.1), como un sistema de ecuaciones no lineales, consideremos una función de complementariedad φ , definamos $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ como,

$$\Phi(\mathbf{x}) = \begin{pmatrix} \varphi(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi(x_n, F_n(\mathbf{x})) \end{pmatrix} \quad (2.3)$$

y el sistema de ecuaciones no lineales,

$$\Phi(\mathbf{x}) = \mathbf{0}, \quad (2.4)$$

que debido a la falta de suavidad de φ , se tiene que (2.4) es un sistema de ecuaciones no lineal, no diferenciable, aspecto clave para tener en cuenta a la hora de resolverlo.

Nuestro interés se centra en utilizar la familia uniparamétrica de funciones de complementariedad (2.2) en la función (2.3) a la cual llamaremos

$\Phi_\lambda(\mathbf{x})$, y es de la forma,

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

Ahora, para el Ejemplo 2.6, si reformulamos el PCNL(F) como el sistema de ecuaciones no lineales (2.4) usando la familia uniparamétrica de funciones de complementariedad se tiene lo siguiente,

Ejemplo 2.7. *Dada la función*

$$F(x, y) = \begin{pmatrix} x + y - \frac{1}{2} \\ e^x - 2y \end{pmatrix},$$

se tiene que la reformulación está dada como sigue,

$$\begin{aligned} \Phi_\lambda(\mathbf{x}) &= \begin{pmatrix} \varphi_\lambda(x_1, F_1(\mathbf{x})) \\ \varphi_\lambda(x_2, F_2(\mathbf{x})) \end{pmatrix} \\ &= \begin{pmatrix} \varphi_\lambda(x, x + y - \frac{1}{2}) \\ \varphi_\lambda(y, e^x - 2y) \end{pmatrix}. \end{aligned}$$

Luego,

$$\Phi_\lambda(\mathbf{x}) = \begin{pmatrix} \sqrt{(\frac{1}{2} - y)^2 + \lambda x(x + y - \frac{1}{2})} - 2x - y + \frac{1}{2} \\ \sqrt{(3y - e^x)^2 + \lambda y(e^x - 2y)} + y - e^x \end{pmatrix}.$$

Es decir el sistema $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$, es de la siguiente forma:

$$\begin{pmatrix} \sqrt{(\frac{1}{2} - y)^2 + \lambda x(x + y - \frac{1}{2})} - 2x - y + \frac{1}{2} \\ \sqrt{(3y - e^x)^2 + \lambda y(e^x - 2y)} + y - e^x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (2.6)$$

Las curvas de la Figura 2.7 representan las ecuaciones (2.6) trazadas con $\lambda = 2$.

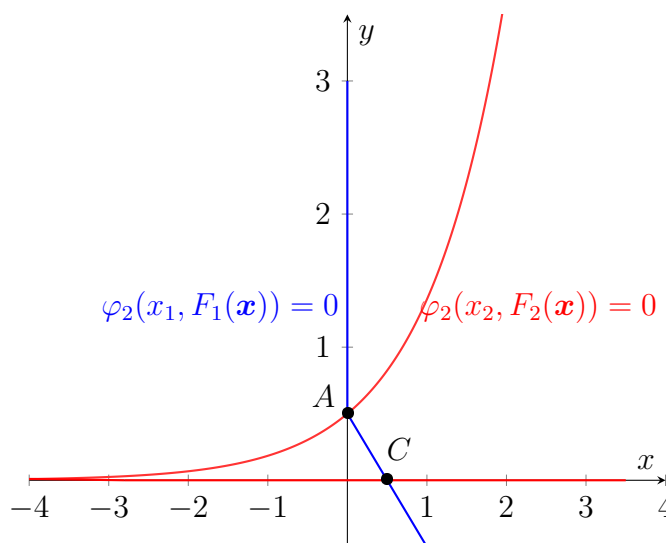


Figura 2.7: Funciones componentes de $\Phi_2(\mathbf{x})$

En la Figura 2.7, se puede observar la gráfica de las funciones componentes de $\Phi_2(\mathbf{x}) = (\varphi_2(x_1, F_1(\mathbf{x})), \varphi_2(x_2, F_2(\mathbf{x})))$, la función Φ_2 es no diferenciable; esto no cambia al variar λ , aunque las curvas si cambian levemente. Además, los puntos A y C que corresponden a las intersecciones entre las dos gráficas, son las soluciones para el $PCNL(F)$.

Las gráficas 2.6 y 2.7 representan problemas diferentes que tiene la misma solución, es más, estos dos problemas son equivalentes como se demuestra en la siguiente proposición.

Proposición 2.1. *Un vector $\mathbf{x}^* \in \mathbb{R}^n$ resuelve el sistema (2.4), si y sólo*

si, \mathbf{x}^* resuelve el PCNL(F)(1.1).

En efecto, supongamos que \mathbf{x}^* resuelve (2.4), esto es que para todo $i = 1, 2, \dots, n$, se cumple la siguiente equivalencia,

$$\begin{aligned} \Phi(\mathbf{x}^*) = \mathbf{0} &\iff \varphi(x_i^*, F_i(\mathbf{x}^*)) = 0, \\ &\iff x_i^* \geq 0, \quad F_i(\mathbf{x}^*) \geq 0, \quad x_i^* F_i(\mathbf{x}^*) = 0, \\ &\iff \mathbf{x}^* \geq 0, \quad F(\mathbf{x}^*) \geq 0, \quad \mathbf{x}^* F(\mathbf{x}^*) = 0, \\ &\iff \mathbf{x}^* \text{ resuelve el PCNL}(F). \end{aligned}$$

Para resolver sistemas de ecuaciones no lineales, y diferenciales de la forma $G(\mathbf{x}) = \mathbf{0}$, el método más utilizado es el *método de Newton*, el cual en cada iteración se resuelve el sistema lineal de la forma,

$$J(\mathbf{x}_k) \mathbf{s}_k = -G(\mathbf{x}_k) \tag{2.7}$$

donde J es la matriz Jacobiana de G en el punto \mathbf{x}_k .

Cuando una función G no es diferenciable, como es el caso de la función Φ , hablar de “la matriz Jacobiana” no tiene sentido, en su lugar se utiliza el concepto de *Jacobiano generalizado* presentado por *Clark* [12]. Para ello, es necesario recordar la siguiente definición.

Definición 2.3. Sea Ω un subconjunto abierto de \mathbb{R}^n , una función G , es una función Lipschitz continua en Ω , si para todo $\mathbf{x}, \mathbf{y} \in \Omega$, existe un $K > 0$ tal que:

$$\|G(\mathbf{x}) - G(\mathbf{y})\| \leq K \|\mathbf{x} - \mathbf{y}\|.$$

Si la función G es Lipschitz en una vecindad $B(x, \epsilon)$ con $\epsilon > 0$, se dice que G es Lipschitz en \mathbf{x} . Ahora, la función G es locamente Lipschitz continua en una vecindad $V \subseteq \Omega$ de \mathbf{x}_0 , si para todo $\mathbf{x}, \mathbf{y} \in V$, existe un $K > 0$ tal

que:

$$\|G(\mathbf{x}) - G(\mathbf{y})\| \leq K\|\mathbf{x} - \mathbf{y}\|.$$

Se dice que G es *Lipschitz* continua (localmente *Lipschitz* continua) con constante *Lipschitz* K .

Definición 2.4. Sea $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función localmente *Lipschitz* continua, se define *B-Jacobiano* generalizado de G en \mathbf{x} como:

$$\partial_B G(\mathbf{x}) = \left\{ \lim_{k \rightarrow \infty} G'(\mathbf{x}_k) \in \mathbb{R}^{n \times n} : \lim_{k \rightarrow \infty} \mathbf{x}_k \rightarrow \mathbf{x}, \mathbf{x}_k \in D_G \right\}, \quad (2.8)$$

donde D_G es el conjunto de todos los puntos de \mathbb{R}^n donde G es diferenciable, y definimos el *Jacobiano* generalizado de G en \mathbf{x} como el siguiente conjunto

$$\partial G(\mathbf{x}) = \text{conv} \{ \partial_B G(\mathbf{x}) \}, \quad (2.9)$$

donde $\text{conv} \{A\}$ representa la envolvente convexa de A , se ha demostrado que este conjunto es no vacío, convexo y compacto, los detalles de la demostración se encuentra en [12]. Es de notar, para el caso en que G es diferenciable en el punto \mathbf{x} , el conjunto $\partial G(\mathbf{x})$ tiene un único elemento, la matriz Jacobiana de G en \mathbf{x} , esto es $G'(\mathbf{x})$.

El siguiente ejemplo ilustra cual es el *Jacobiano* generalizado para una función que no es diferenciable en un punto,

Ejemplo 2.8. Consideremos la función $F(x) = |x|$ la cual es *Lipschitz* continua, dado que:

$$|f(x) - f(y)| = ||x| - |y|| \leq |x - y|, \quad \text{para todo } x, y \in \mathbb{R}$$

de ahí que, podemos determinar $\partial F(0)$ de la siguiente forma.

Al calcular la derivada de F , esto es $F'(x) = \frac{x}{|x|}$ se puede observar que no está definida para $x = 0$, por tanto el *Jacobiano* generalizado de F en

$x = 0$, se puede determinar como sigue

$$\partial F(0) = \text{conv} \left\{ \lim_{k \rightarrow \infty} F'(x_k) \in \mathbb{R} : \lim_{k \rightarrow \infty} x_k \rightarrow 0, x_k \in D_F \right\}$$

donde

$$\lim_{k \rightarrow \infty} F'(x_k) = \begin{cases} -1 & \text{si } x_k < 0 \\ 1 & \text{si } x_k > 0 \end{cases}$$

De ahí que,

$$\begin{aligned} \partial F(0) &= \text{conv} \left\{ \lim_{k \rightarrow \infty} F'(x_k) : \lim_{k \rightarrow \infty} x_k \rightarrow 0, x_k \in D_F \right\} \\ &= \text{conv} \{-1, 1\} \\ &= \{y \in \mathbb{R} : y = t - 1(1 - t), t \in [0, 1]\} \\ &= \{y \in \mathbb{R} : y = 2t - 1, t \in [0, 1]\} \\ &= [-1, 1] \end{aligned}$$

Por lo general, el conjunto $\partial G(\mathbf{x})$ es difícil de calcular. La Proposición 2.6.2 (e) *Clarke* [13] proporciona una subestimación de ∂G dada por,

$$\partial G(\mathbf{x})^T \subseteq \partial G_1(\mathbf{x}) \times \partial G_2(\mathbf{x}) \times \dots \times \partial G_n(\mathbf{x}).$$

Motivado por lo anterior Qi [57], define el conjunto

$$\partial_C G(\mathbf{x})^T := \partial G_1(\mathbf{x}) \times \partial G_2(\mathbf{x}) \times \dots \times \partial G_n(\mathbf{x}).$$

Llamado el C-subdiferencial de G en \mathbf{x} ; el cual es un conjunto de matrices de $\mathbb{R}^{n \times n}$ con la i -ésima columna dada por el gradiente generalizado de la i -ésima componente de la función G .

Arenas, en [4] presenta algunas propiedades importantes de φ_λ y del operador Φ_λ , así como también los autores en [53]. Presentamos a continuación las más relevantes para este documento.

2.3. Propiedades del operador φ_λ

Proposición 2.2. *La función φ_λ definida en (2.2) está bien definida [53].*

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

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

Es una norma en \mathbb{R}^2 [3].

Observemos que la familia uniparamétrica de funciones de complementariedad puede ser representada por:

$$\varphi_\lambda(a, b) = G_\lambda(a, b) - (a + b). \quad (2.10)$$

donde para cualquier vector no nulo en \mathbb{R}^2 , el vector gradiente de φ_λ está definido por,

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

en forma compacta,

$$\nabla \varphi_\lambda(a, b) = \nabla G_\lambda(a, b) - \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (2.11)$$

Además, se denotarán las derivadas parciales de G_λ en (a, b) por,

$$\begin{aligned} \chi(a, b) &= \frac{\partial G_\lambda}{\partial a} = \frac{2(a - b) + \lambda b}{2G_\lambda(a, b)}, \\ \psi(a, b) &= \frac{\partial G_\lambda}{\partial b} = \frac{-2(a - b) + \lambda a}{2G_\lambda(a, b)}. \end{aligned} \quad (2.12)$$

El siguiente lema nos proporciona la una constante *Lipschitz* para la familia uniparamétrica de funciones de complementariedad

Lema 2.2. [53] *La familia uniparamétrica de funciones de complementariedad es Lipschitz continua con constante $4\sqrt{2}$. Es decir, para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$*

$$|\varphi_\lambda(\mathbf{x}) - \varphi_\lambda(\mathbf{y})| \leq 4\sqrt{2}|\mathbf{x} - \mathbf{y}|.$$

Además, se han determinado cotas superiores para las derivadas parciales.

Lema 2.3. [53] *Sea $\lambda \in (0, 4)$. Las derivadas parciales de G_λ definidas en (2.12) satisfacen las siguientes desigualdades*

$$|\chi(a, b)| \leq 1 \quad y \quad |\psi(a, b)| \leq 1, \quad (2.13)$$

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

La proposición siguiente es importante, pues esta nos garantiza la existencia del Jacobiano generalizado para el operador Φ_λ .

Proposición 2.3. *La función Φ_λ es una función localmente Lipschitz continua.*

Demostración. Sea $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, demostremos que existe un $K > 0$ tal que:

$$\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\|_\infty \leq K\|\mathbf{x} - \mathbf{y}\|_\infty.$$

Sabemos por el **Lema 2.2**, que φ_λ es *Lipschitz* continua, por tanto para algún $i \in \{1, 2, \dots, n\}$ se tiene que:

$$\begin{aligned} \|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\|_\infty &= |\varphi_\lambda(\mathbf{x}_i, F_i(\mathbf{x})) - \varphi_\lambda(\mathbf{y}_i, F_i(\mathbf{y}))| \\ &\leq K_1 \left\| \begin{pmatrix} x_i - y_i \\ F_i(\mathbf{x}) - F_i(\mathbf{y}) \end{pmatrix} \right\|_\infty \\ &= K_1 \text{máx} \{|x_i - y_i|, |F_i(\mathbf{x}) - F_i(\mathbf{y})|\} \end{aligned}$$

Luego, tenemos los dos casos:

1. Si $\max\{|x_i - y_i|, |F_i(\mathbf{x}) - F_i(\mathbf{y})|\} = |x_i - y_i|$, por otro lado se tiene que: $|x_i - y_i| \leq \|\mathbf{x} - \mathbf{y}\|_\infty$ entonces se cumple que para $K = K_1$,

$$\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\|_\infty \leq K\|\mathbf{x} - \mathbf{y}\|_\infty.$$

2. Si $\max\{|x_i - y_i|, |F_i(\mathbf{x}) - F_i(\mathbf{y})|\} = |F_i(\mathbf{x}) - F_i(\mathbf{y})|$ y dado que F es una función continuamente diferenciable, entonces F es localmente *Lipschitz* [58].

Si denotemos a K_2 como la constante *Lipschitz* de F , luego se tiene que,

$$|F_i(\mathbf{x}) - F_i(\mathbf{y})| \leq \|F(\mathbf{x}) - F(\mathbf{y})\|_\infty \leq K_2\|\mathbf{x} - \mathbf{y}\|_\infty,$$

basta tomar $K = \max\{K_1, K_1K_2\}$ para que se cumpla,

$$\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\|_\infty \leq K\|\mathbf{x} - \mathbf{y}\|_\infty.$$

De (1) y (2) se concluye que $\Phi_\lambda(\mathbf{x})$ es localmente *Lipschitz*. \square

Por tanto el jacobiano generalizado $\partial\Phi_\lambda(\mathbf{x})$ existe. Esto es que,

$$\partial\Phi_\lambda(\mathbf{x}) = \text{conv} \left\{ \lim_{k \rightarrow \infty} \Phi'_\lambda(\mathbf{x}_k) \in \mathbb{R}^{n \times n} : \lim_{k \rightarrow \infty} \mathbf{x}_k \rightarrow \mathbf{x}, \mathbf{x}_k \in D_{\Phi_\lambda} \right\}, \quad (2.14)$$

existe.

Ejemplo 2.9. Consideremos la función $F(x) = |x|$ y determinemos $\partial\Phi_\lambda(0)$.

Se tiene que,

$$\begin{aligned} \Phi_\lambda(x) &= \varphi_\lambda(x_1, F_1(x)) \\ &= \varphi_\lambda(x, |x|) \\ &= \sqrt{(x - |x|)^2 + \lambda x|x|} - x - |x| \end{aligned}$$

observemos los dos casos siguientes:

1. Si $x > 0$ entonces,

$$\begin{aligned}\Phi_\lambda(x) &= \sqrt{\lambda x^2} - 2x \\ &= (\sqrt{\lambda} - 2)x\end{aligned}$$

Luego,

$$\Phi'_\lambda(x) = \sqrt{\lambda} - 2$$

de ahí que, $H_1 = \lim_{k \rightarrow \infty} \Phi'_\lambda(x_k) = \sqrt{\lambda} - 2$, para cualquier sucesión $\{\mathbf{x}_k\}$.

2. Si $x < 0$ entonces,

$$\begin{aligned}\Phi_\lambda(x) &= \sqrt{4x^2 - \lambda x^2} \\ &= -(\sqrt{4 - \lambda})x\end{aligned}$$

Luego,

$$\Phi'_\lambda(x) = -\sqrt{4 - \lambda}$$

de ahí que, $H_2 = \lim_{k \rightarrow \infty} \Phi'_\lambda(x_k) = -\sqrt{4 - \lambda}$, para cualquier sucesión $\{\mathbf{x}_k\}$.

De otro lado, sabemos por (2.14) que

$$\partial\Phi_\lambda(x) = \text{conv} \left\{ \lim_{k \rightarrow \infty} \Phi'_\lambda(x_k) \in \mathbb{R} : \lim_{k \rightarrow \infty} x_k \rightarrow x, x_k \in D_{\Phi_\lambda} \right\}$$

por tanto,

$$\begin{aligned}\partial\Phi_\lambda(0) &= \text{conv} \left\{ \lim_{k \rightarrow \infty} \Phi'_\lambda(x_k) \in \mathbb{R} : \lim_{k \rightarrow \infty} x_k \rightarrow 0, x_k \in D_G \right\} \\ &= \text{conv} \{H_1, H_2\} \\ &= \left\{ y \in \mathbb{R} : y = t(\sqrt{\lambda} - 2) - (1 - t)\sqrt{4 - \lambda}, t \in [0, 1] \right\} \\ &= [-\sqrt{4 - \lambda}, \sqrt{\lambda} - 2].\end{aligned}$$

Observemos que cuando $\lambda \rightarrow 0$, se tiene que $\partial\Phi_\lambda(0) \rightarrow -2$ y cuando $\lambda \rightarrow 4$, $\partial\Phi_\lambda(0) \rightarrow 0$; para el caso en que $\lambda = 2$, el Jacobiano generalizado de Φ_λ en $x = 0$ es el segmento de recta $[-\sqrt{2}, \sqrt{2} - 2]$. En la Figura (2.8) podemos ver el comportamiento del Jacobiano generalizado para distintos valores de λ .

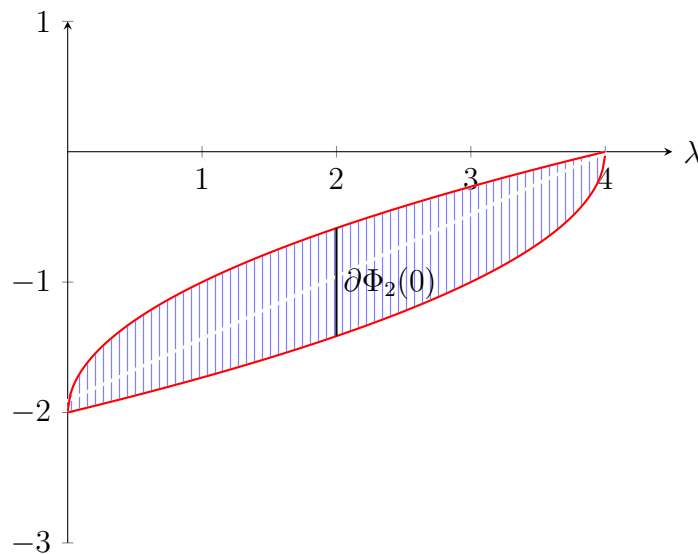


Figura 2.8: Jacobiano generalizado ($\partial\Phi_\lambda(0)$) al variar λ .

Ahora, para la construcción de las matrices H_k , usaremos la estrategia presentada en [53].

Supongamos que \mathbf{x}^* es una solución para el problema (1.1), se tiene el siguiente conjunto de índices asociados a la solución del PCNL(F).

$$\alpha := \{i: \mathbf{x}_i^* > 0, F_i(\mathbf{x}^*) = 0\}, \quad (2.15)$$

$$\beta := \{i: \mathbf{x}_i^* = 0, F_i(\mathbf{x}^*) = 0\}, \quad (2.16)$$

$$\gamma := \{i: \mathbf{x}_i^* = 0, F_i(\mathbf{x}^*) > 0\}. \quad (2.17)$$

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

Con el fin de construir matrices en $\partial_B \Phi_\lambda$, se considera una sucesión de vectores $\{\mathbf{y}_k\}$ en \mathbb{R}^n que convergen a \mathbf{x} , tal que $\Phi'_\lambda(\mathbf{y}_k)$ existe, para ello, se considera la siguiente sucesión

$$\mathbf{y}_k = \mathbf{x} + \epsilon_k \mathbf{z}$$

donde, $\{\epsilon_k\}$ es una sucesión de números positivos tales que $\epsilon_k \rightarrow 0$ cuando $k \rightarrow \infty$ y el vector \mathbf{z} se elige de tal forma que $\mathbf{z}_i \neq 0$ cuando $i \in \beta$.

Este novedoso método fue propuesto por *De Luca* [15], además es claro que $\mathbf{y}_k \rightarrow \mathbf{x}$ cuando $k \rightarrow \infty$ de ahí que la componente H_i de la matriz $H \in \partial_B \Phi_\lambda$ es de la forma,

$$H_i = \begin{cases} (\chi(\mathbf{z}_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) - 1) \mathbf{e}_i^T + (\psi(\mathbf{z}_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) - 1) \nabla F_i(\mathbf{x})^T & \text{si } i \in \beta \\ (\chi(\mathbf{x}_i, F_i(\mathbf{x})) - 1) \mathbf{e}_i^T + (\psi(\mathbf{x}_i, F_i(\mathbf{x})) - 1) \nabla F_i(\mathbf{x})^T & \text{si } i \notin \beta \end{cases} \quad (2.18)$$

donde,

$$\chi(\mathbf{z}_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) = \frac{2(\mathbf{z}_i - \nabla F_i(\mathbf{x})^T \mathbf{z}) + \lambda \nabla F_i(\mathbf{x})^T \mathbf{z}}{2\sqrt{(\mathbf{z}_i - \nabla F_i(\mathbf{x})^T \mathbf{z})^2 + \nabla \mathbf{z}_i F_i(\mathbf{x})^T \mathbf{z}}}$$

$$\psi(\mathbf{z}_i, \nabla F_i(\mathbf{x})^T \mathbf{z}) = \frac{-2(\mathbf{z}_i - \nabla F_i(\mathbf{x})^T \mathbf{z}) + \lambda \mathbf{z}_i}{2\sqrt{(\mathbf{z}_i - \nabla F_i(\mathbf{x})^T \mathbf{z})^2 + \nabla \mathbf{z}_i F_i(\mathbf{x})^T \mathbf{z}}}$$

$$\chi(\mathbf{x}_i, F_i(\mathbf{x})) = \frac{2(\mathbf{x}_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})}{2\sqrt{(\mathbf{x}_i - F_i(\mathbf{x}))^2 + \lambda \mathbf{x}_i F_i(\mathbf{x})}}$$

$$\psi(\mathbf{x}_i, F_i(\mathbf{x})) = \frac{-2(\mathbf{x}_i - F_i(\mathbf{x})) + \lambda \mathbf{x}_i}{2\sqrt{(\mathbf{x}_i - F_i(\mathbf{x}))^2 + \lambda \mathbf{x}_i F_i(\mathbf{x})}}$$

Para resolver el sistema de ecuaciones no lineales, no diferenciables (2.4) el cual es de la forma $G(\mathbf{x}) = \mathbf{0}$, usando matrices del conjunto (2.8). Se

presentan cuatro métodos importantes:

Qi-Liqun [56] propone el llamado método de *Newton generalizado* el cual, en cada iteración, resuelve un sistema lineal de la forma,

$$H_k \mathbf{s}_k = -G(\mathbf{x}_k) \quad (2.19)$$

donde $H_k \in \partial G(\mathbf{x}_k)$.

De igual manera, para problemas de complementariedad no lineal, también se han propuesto métodos *cuasi-Newton*¹ no diferenciable [4, 6]. donde se debe resolver un sistema lineal de la forma,

$$B_k \mathbf{s}_k = -G(\mathbf{x}_k) \quad (2.20)$$

donde $B_k \approx H_k \in \partial G(\mathbf{x}_k)$.

Como un aporte de la teoría de mínimos cuadrados no lineales, consideremos el método *Gauss-Newton* [5], donde se debe resolver en cada iteración el sistema lineal de la forma,

$$(H_k^T H_k) \mathbf{s}_k = -H^T G(\mathbf{x}_k) \quad (2.21)$$

con $H_k \in \partial G(\mathbf{x}_k)$.

Este método generalmente es usado para matrices no cuadradas de rango completo, observemos que si H_k es de rango completo, al resolver $H_k^T H_k$ sería una matriz invertible.

Para el caso particular en el que estamos trabajando, matrices cuadradas de orden n , si la matriz H_k no es invertible, entonces $H_k^T H_k$ no es inverti-

¹Para evitar el cálculo del elemento H_k en el Jacobiano generalizado, se pueden utilizar alternativas que reduzcan ese cálculo, entre estas alternativas se conoce los *métodos cuasi-Newton*, porque aproximan la matriz H_k por medio de otra, para ello usan varias estrategias, como por ejemplo: métodos secantes (donde las actualizaciones más exitosas de estos métodos se encuentra, actualización Broyden buena, y Broyden mala, entre otras).

ble, puesto que $\text{Rank}(H_k) = \text{Rank}(H_k^T H_k)$, pero se trae a mención, dado que el método que se utilizará en este documento es una modificación del método *Gauss-Newton*.

El método de *Levenberg-Marquardt* [20, 17], publicado por primera vez en el año de 1944 por *Kenneth Levenberg*² y posteriormente, en 1963, *Donald Marquardt*³ realizó una modificación del algoritmo, este método es conocido actualmente como el *método Levenberg-Marquardt*, presenta una ligera modificación del método de *Gauss-Newton*. En cada iteración, se debe resolver el siguiente sistema lineal,

$$(H_k^T H_k + \mu_k I_n) s_k = -H_k^T G(\mathbf{x}_k) \quad (2.22)$$

con $H_k \in \partial G(\mathbf{x}_k)$ y μ_k es un parámetro que se actualiza en cada iteración, en la **Sección 2.4.2** hablaremos de una propuesta para selección del parámetro μ_k .

En general, los métodos vistos anteriormente para resolver sistemas de ecuaciones no lineales (2.4) garantizan, bajo ciertas condiciones, una *tasa*

²**Kenneth Levenberg:** Nació el 6 de agosto de 1919, fue un estadístico estadounidense y autor original del algoritmo de ajuste de mínimos cuadrados no lineales ampliamente utilizado, mejorado más tarde por *Donald Marquardt*, conocido actualmente como el algoritmo de Levenberg-Marquardt. Levenberg publicó por primera vez el algoritmo en 1944 mientras trabajaba en el Frankford Arsenal. Más tarde trabajó para Boeing, donde desarrolló modelos matemáticos utilizados para diseñar el Boeing 737. Terminó su carrera en el departamento de matemáticas de la Universidad de Hawaii, Hilo y murió en Hawaii en 1973. uno de los reconocimientos más grandes de su carrera fue en el año de 1970 cuando fue incluido en American Men of Science.

³**Donald W. Marquardt:** Nació el 13 de marzo de 1929, Ciudad de Nueva York, fue un estadístico estadounidense, el redescubridor del algoritmo de ajuste por mínimos cuadrados no lineales, conocido como el algoritmo de Levenberg-Marquardt. Marquardt se educó en la Universidad de Columbia con una licenciatura en 1950 en física y matemáticas, luego en 1956 en la Universidad de Delaware realizó una maestría en matemáticas y estadística. En 1953 se incorporó a DuPont y trabajó allí durante 39 años. También fundó y dirigió el Centro de Tecnología y Gestión de Calidad de DuPont. En 1963 publicó su famoso artículo “algoritmo para la estimación por mínimos cuadrados de problemas no lineales” en la revista SIAM. Marquardt desarrolló su algoritmo para resolver el ajuste de modelos químicos no lineales a datos de laboratorio. Dentro de los reconocimientos mas destacados se encuentra que en año de 1975 fue elegido miembro de la Asociación Estadounidense de Estadística, Ganó la medalla Shewhart en 1986, y murió de un infarto a la edad de 68 años.

de convergencia q -cuadrática siempre y cuando el punto inicial del algoritmo esté suficientemente cerca de la solución del sistema, esto quiere decir que son *métodos locales* [16].

Una de las propiedades deseables en muchos algoritmos (iterativos) diseñados para resolver el PCNL(F) dado en (1.1) es la convergencia global. Esta propiedad garantiza que el método propuesto converge a la solución del problema independientemente del punto inicial que se escoja. Así pues, la técnica de globalización del algoritmo local respectivo, se basa el uso de las *funciones de mérito* [49]. Para el caso del sistema de ecuaciones no lineales $G(\mathbf{x}) = \mathbf{0}$, la función de mérito más utilizada es $R : \mathbb{R}^n \rightarrow \mathbb{R}$, definida por $R(\mathbf{x}) = \|G(\mathbf{x})\|_2^2$. Teniendo en cuenta como base esta función, una nueva iteración en un algoritmo global será aceptada siempre y cuando el valor de la función R decrezca con respecto a la iteración actual del mismo; este requerimiento conduce a resolver el problema de minimización [6, 49].

$$\begin{aligned} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{Minimizar}} \quad g(\mathbf{x}) &= \frac{1}{2} \|G(\mathbf{x})\|_2^2. \end{aligned} \quad (2.23)$$

con cero el valor óptimo, y el factor $\frac{1}{2}$ es introducido por conveniencia algebraica; además, en caso de que la función $g(\mathbf{x}) > 0$, se tiene que los minimizadores locales de (2.23) no necesariamente son raíces de G .

Si consideremos la reformulación del PCNL(F), $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$, se tiene que una función de mérito está dada por,

$$\Psi_\lambda(x) := \frac{1}{2} \Phi_\lambda(x)^T \Phi_\lambda(x) = \frac{1}{2} \|\Phi_\lambda(x)\|_2^2. \quad (2.24)$$

Con esta nueva función, se puede reformular nuevamente el sistema de ecuaciones no lineales, no diferenciables (2.5) como el *problema de mini-*

mización siguiente,

$$\begin{aligned} \text{Minimizar } & \Psi_\lambda(\mathbf{x}). \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned} \quad (2.25)$$

Entonces, cualquier solución de (1.1) es un minimizador global de Ψ_λ y cualquier minimizador global de Ψ_λ resuelve (1.1), esto si (1.1) es solucio- nable. Un resultado importante dado por *Kanzow y Kleinmichel*, en 1998 [38], ellos demuestran que Ψ_λ es continuamente diferenciable, además su gradiente está dado por,

$$\nabla \Psi_\lambda(\mathbf{x}) = H^T \Phi_\lambda(\mathbf{x}), \quad (2.26)$$

Otro resultado, que se considera importante para el estudio del tipo de convergencia para nuestro algoritmo, radica en determinar bajo que con- diciones, la norma de las matrices H_k (matrices del Jacobiano generalizado de Φ_λ), es acotada.

Lema 2.4. *Sea $M \in \partial_B \Phi_\lambda(\mathbf{x})$, entonces existe una constante $L = L(x)$ tal que:*

$$\|M\|_2 \leq L.$$

Demostración. Consideremos la matriz M definida como en 2.18, se tiene que

$$\|M\|_\infty = \max \{\|M_i\|_1\}, \quad (2.27)$$

y supongamos que el máximo en (2.27) se alcanza en la fila j , entonces se tienen los dos casos siguientes:

- Si $j \notin \beta$, entonces $\mathbf{x}_j \neq 0$ o $F(\mathbf{x}_j) \neq 0$, así se tiene por **Lema 2.3** que:

$$|(\chi(\mathbf{x}_j, F_j(\mathbf{x})))| \leq 1 \quad \text{y} \quad |(\psi(\mathbf{x}_j, F_j(\mathbf{x})))| \leq 1,$$

de otro lado,

$$\begin{aligned}\|M\|_\infty &= \|(\chi(\mathbf{x}_j, F_j(\mathbf{x})) - 1) \mathbf{e}_j^T + (\psi(\mathbf{x}_j, F_j(\mathbf{x})) - 1) \nabla F_j(\mathbf{x})^T\|_\infty \\ &\leq \|(\chi(\mathbf{x}_j, F_j(\mathbf{x})) - 1)\|_\infty + \|(\psi(\mathbf{x}_j, F_j(\mathbf{x})) - 1)\|_\infty \|\nabla F_j(\mathbf{x})^T\|_\infty \\ &\leq 2 + 2\|\nabla F_j(\mathbf{x})^T\|_\infty.\end{aligned}$$

- Para el caso en que $j \in \beta$, el proceso es análogo, y se concluye que

$$\|M\|_\infty \leq 2 + 2\|\nabla F_i(\mathbf{x})^T\|_\infty,$$

de ahí que, si tomamos $\eta = 2 + 2\|\nabla F_i(\mathbf{x})^T\|_\infty$, entonces

$$\|M\|_\infty \leq \eta,$$

Pero por propiedades de las normas matriciales [30],

$$\|M\|_2 \leq \sqrt{n}\|M\|_\infty$$

de esta manera, tomando $L = \eta\sqrt{n}$ se concluye que

$$\|M\|_2 \leq L.$$

□

2.4. Estrategias de globalización

En general, el problema

$$\begin{aligned}\min G(\mathbf{x}), \\ \mathbf{x} \in \mathbb{R}^n\end{aligned}\tag{2.28}$$

donde, $G : \mathbb{R}^n \rightarrow \mathbb{R}$ y $G \in C^2$, se puede resolver utilizando métodos como: el método del gradiente, tipo Newton y cuasi-Newton, entre otros; estos dos últimos métodos coinciden en que una iteración básica requiere

resolver un sistema de ecuaciones lineales para posteriormente, generar la iteración siguiente; además, tienen excelentes propiedades de convergencia; sin embargo, todos estos métodos son locales, es decir, solo se puede garantizar la convergencia si el punto inicial está cerca de la solución; es por ello que se necesita una estrategia de globalización que permita lograr, a partir de un punto inicial no necesariamente cercano a la solución, la convergencia del algoritmo.

Las estrategias de globalización más usadas son región de confianza y búsqueda lineal, de ellas hablaremos en las **secciones 2.4.1 y 2.4.3** respectivamente. Particularmente, para la solución del problema de optimización (2.25), es decir, la técnica de reformulación de (1.1) como un problema de minimización (2.25), en las **secciones 2.4.2 y 2.4.4** introducimos estas estrategias de globalización.

2.4.1. Región de confianza para un problema de minimización sin restricciones

La estrategia de globalización, conocida como región de confianza, la describimos de la siguiente manera; a partir de un punto inicial \mathbf{x}_k se fija un tamaño de paso inicial. Luego, resolvemos un subproblema de minimización restringido a la vecindad del punto inicial \mathbf{x}_k , donde la función a minimizar es un modelo que está relacionado con el método a implementar y aproxima la función objetivo G , en particular para el método de Newton el modelo es:

$$q_k(\mathbf{d}) = G(\mathbf{x}_k) + \nabla G(\mathbf{x}_k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T B_k \mathbf{d}, \quad (2.29)$$

donde $B_k = \nabla^2 G(\mathbf{x}_k)$ es la matriz Hessiana de G en \mathbf{x}_k , para los métodos cuasi-Newton en el modelo (2.29) se debe cambiar B_k por $B_k \approx \nabla^2 G(\mathbf{x}_k)$, y bajo un criterio se acepta o rechaza la actualización del iterante. si rechaza, entonces se reduce el tamaño de paso y se repite el proceso; si

se acepta, entonces el tamaño del paso se puede conservar o aumentar, y repetimos el proceso.

Para reforzar esta noción, describimos los pasos a continuación:

1. A partir de la función objetivo G , donde $G: \mathbb{R}^n \rightarrow \mathbb{R}$, se construye una función modelo $q_k(\mathbf{d})$ que aproxime la función objetivo $G(\mathbf{x}_k + \mathbf{d})$ cerca del punto de iteración actual \mathbf{x}_k .
2. Como el modelo q_k puede no ser una buena aproximación de G cuando $\|\mathbf{d}\|$ sea muy grande, restringimos la búsqueda del siguiente punto a una bola cerrada que llamaremos *región de confianza*, donde esperamos que la aproximación del modelo sea aceptable, esto es, resolver el subproblema.

$$\begin{aligned} & \text{mín } q_k(\mathbf{d}) \\ & \|\mathbf{d}\| \leq \Delta_k \end{aligned} \quad (2.30)$$

donde Δ_k se le denomina radio de confianza.

3. Si la solución del subproblema (2.30) no es “aceptada”, se concluye que la región de confianza es demasiado grande y debe reducirse Δ_k , para después resolver el subproblema (2.30) nuevamente.
4. Si la solución del subproblema (2.30) es “aceptada”, se calcula el nuevo iterante y se decide conservar o aumentar el radio de confianza Δ_k y se repite el proceso [16, 49].

El siguiente lema (para el modelo cuadrático) nos garantiza la existencia de una solución del subproblema (2.30).

Lema 2.5. [49] *Sea $G: \mathbb{R}^n \rightarrow \mathbb{R}$ dos veces diferenciable, $B_k \in \mathbb{R}^{n \times n}$, simétrica y definida positiva. La solución del problema (2.30) es*

$$s(\alpha) = -(B_k + \alpha I)^{-1} \nabla G(\mathbf{x}_k) \quad (2.31)$$

para un único $\alpha \geq 0$ tal que $\|s(\alpha)\| = \Delta_k$, a menos que $\|s(0)\| \leq \Delta_k$, en tal caso $s(0)$ es la solución. Para cualquier $\alpha \geq 0$, $s(\alpha)$ define una dirección de descenso para G desde \mathbf{x}_k .

El lema anterior no nos da una forma explícita para calcular el α , sin embargo nos da una idea de como aproximar la solución del problema (2.30); de esta forma se construyen los métodos numéricos para el subproblema de región de confianza conocidos como *Newton*, *Levenberg-Marquardt*, el punto de *Cauchy*, *Dogleg*, *doble Dogleg*, *Hook* entre otros [16, 18, 39, 44, 45]; a continuación explicamos para qué parámetros se obtienen los métodos más conocidos en región de confianza.

- **Paso de Newton de región de confianza**

Se puede observar en (2.31) que si $\|s(0)\| \leq \Delta_k$, entonces

$$s(0) = -(B_k)^{-1} \nabla G(\mathbf{x}_k), \quad (2.32)$$

es la solución del subproblema (2.30). Si B_k es la matriz Hessiana de G en \mathbf{x}_k y es definida positiva, $s(0)$ es el paso de *Newton*. Si B_k es una aproximación a la matriz Hessiana de G y es definida positiva, $s(0)$ es el paso *Cuasi-Newton*.

- **Paso del gradiente para región de confianza**

Además en (2.31), si α toma un valor muy grande, el paso tiende a un múltiplo escalar de $-\nabla G(\mathbf{x}_k)$, es decir

$$s(\alpha) \approx -\frac{1}{\alpha} \nabla G(\mathbf{x}_k), \quad (2.33)$$

es la solución para el subproblema (2.30).

- **Paso de Levenberg-Marquardt para región de confianza** A partir de (2.31), si $\|s(0)\| > \Delta_k$, entonces $\|s(\alpha_k)\| = \Delta_k$ es la solución, dado que determinar el valor de α de manera exacta es difícil (costoso computacionalmente), se aproxima α_k por un valor tal que $B_k + \alpha_k I$

sea suficientemente definida positiva, de esta forma una aproximación a la solución de (2.30) es

$$s_k^{LM} = - (B_k + \alpha_k I)^{-1} \nabla G(\mathbf{x}_k) \quad (2.34)$$

donde $\|s_k^{LM}\| = \Delta_k$ y B_k es la matriz Hessiana de G en \mathbf{x}_k .

Una vez se obtiene una aproximación de la solución del subproblema (2.30), mediante los algunos de los métodos ya mencionadas anteriormente, se determinará si el paso es aceptado o será rechazado usando el siguiente criterio [49, 16]: sean $\rho_1 \in (0, 1)$, $\rho_2 \in (\rho_1, 1)$ y

$$\tilde{\rho}_k =: \frac{Ared_k}{Pred_k},$$

donde,

$$Ared_k = G(\mathbf{x}) - G(\mathbf{x} + \mathbf{d}_k),$$

$$Pred_k = q_k(\mathbf{0}) - q_k(\mathbf{d}_k) = -\nabla G(\mathbf{x}_k)^T \mathbf{d}_k - \frac{1}{2} \mathbf{d}_k^T B_k \mathbf{d}_k.$$

Si $\tilde{\rho}_k \geq \rho_1$ el paso es aceptado, de lo contrario es rechazado.

Luego de clasificar si el punto es aceptado o rechazado, procedemos a decidir si aumentamos el radio de la región de la confianza, lo disminuimos o lo dejamos igual, para ello en [49, 16] usan el siguiente criterio:

$$\Delta_{k+1} = \begin{cases} \frac{1}{2} \Delta_k & \text{si } \tilde{\rho}_k \leq \rho_1 \\ \Delta_k & \text{si } \tilde{\rho}_k \in (\rho_1, \rho_2) \\ \text{mín} \{2\Delta_k, \Delta_{máx}\} & \text{si } \tilde{\rho}_k \geq \rho_2 \end{cases} \quad (2.35)$$

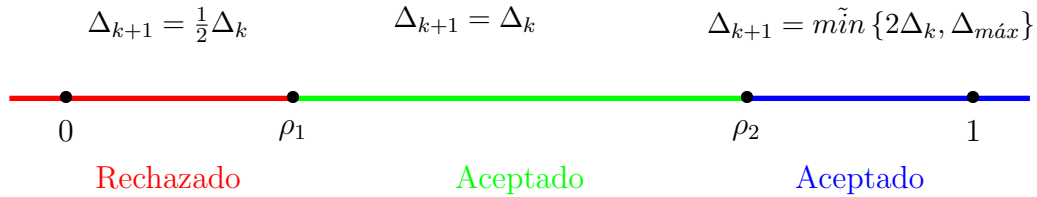


Figura 2.9: Franja de aceptación y rechazo del paso para $\tilde{\rho}_k$

La estructura anterior es conocida como la región de confianza tradicional o convencional, en la mayoría de los problemas funciona de manera eficaz; el algoritmo tiene convergencia global fuerte, es decir todo punto de acumulación es un punto estacionario, sin embargo, para algunos casos puede tener convergencia débil [66, 22], es decir, existe al menos un punto de acumulación que es un punto estacionario.

Con el fin de superar esta dificultad, la doctora ***Jin-Yan Fan*** en [22], propuso un nuevo algoritmo tipo ***Levenberg-Maquardt*** basado en la técnica de región de confianza, este método está diseñado para resolver sistemas de ecuaciones no lineales $R(\mathbf{x}) = \mathbf{0}$, donde $J_k = R'(\mathbf{x}_k)$ y la relación $\tilde{\rho}_k$ entre la reducción real y la reducción prevista juega un papel clave en la decisión de aceptar el paso de prueba \mathbf{d}_k y cómo actualizar el parámetro de ***Levenberg-Maquardt***.

Primero, se resuelve el sistema de ecuaciones lineales

$$(J_k^T J_k + \mu_k I_n) \mathbf{d} = -J_k^T R(\mathbf{x}_k), \quad (2.36)$$

para obtener \mathbf{d}_k , después se determina

$$\tilde{\rho}_k = \frac{A_{red}}{P_{red}},$$

y se determina si el paso es acepto o rechazado, luego se calcula la actualización del parámetro de L.M, para facilitar su lectura lo describimos a continuación.

sean $\rho_1 \in (0, 1)$, $\rho_2 \in (\rho_1, 1)$, $0 < \tau_1 < 1 < \tau_2$ y

$$\beta_{k+1} = \begin{cases} \tau_2 \beta_k & \text{si } \tilde{\rho}_k \leq \rho_1 \\ \beta_k & \text{si } \tilde{\rho}_k \in (\rho_1, \rho_2) \\ \text{máx} \{ \tau_1 \beta_k, \Delta_{\min} \} & \text{si } \tilde{\rho}_k \geq \rho_2 \end{cases} \quad (2.37)$$

se define

$$\mu_{k+1} = \beta_{k+1} \|R(\mathbf{x}_k)\|,$$

en virtud al lema (2.31), se puede decir que \mathbf{d}_k es una aproximación a la solución del subproblema.

$$\begin{aligned} & \text{mín } \frac{1}{2} \|R(\mathbf{x}) + J_k \mathbf{d}\|_2^2 \\ & \text{s.a } \|\mathbf{d}\|_2 \leq \Delta_k =: \| (J_k^T J_k + \mu_k I_n)^{-1} J_k^T R(\mathbf{x}_k) \|_2. \end{aligned} \quad (2.38)$$

Observemos de la ecuación (2.46) que cuando el paso es rechazado, el valor de β_k se aumenta y de esa forma también aumenta μ_k , recíprocamente cuando el paso es aceptado μ_k se reduce, así que es natural hacerse la pregunta. ¿Este método tipo *Levenberg-Maquardt* usa realmente la estrategia de región de confianza?

La respuesta a esta pregunta es si, en el año 1978 *Moré* [48] mostró la relación entre el radio de confianza y el parámetro de *Levenberg-Maquardt*, a continuación realizamos un esquema del argumento que lo llevo a dicha relación, para facilitar la lectura denotamos $J_k = J$, $\mu_k = \mu$ y $R(\mathbf{x}_k) = R$ y de (2.38), tenemos que:

$$\Delta = \| (J^T J + \mu I_n)^{-1} J^T R \|_2,$$

y sea $J = U \Sigma V^T$ la descomposición en valores singulares de J , con U y

V matrices ortogonales, entonces

$$\begin{aligned}
\Delta &= \left\| [J^T J + \mu I_n]^{-1} J^T R \right\|_2 \\
&= \left\| [(U \Sigma V^T)^T (U \Sigma V^T) + \mu I_r]^{-1} (U \Sigma V^T)^T R \right\|_2 \\
&= \left\| [V \Sigma^2 V^T + \mu V V^T]^{-1} (V \Sigma U^T) R \right\|_2 \\
&= \left\| [V (\Sigma^2 + \mu I_r) V^T]^{-1} (V \Sigma U^T) R \right\|_2 \\
&= \left\| V (\Sigma^2 + \mu I_r)^{-1} V^T (V \Sigma U^T) R \right\|_2 \\
&= \left\| V [(\Sigma^2 + \mu I_r)^{-1} \Sigma U^T R] \right\|_2,
\end{aligned}$$

por la invarianza de la norma 2 bajo transformaciones ortogonales, se tiene que

$$\Delta = \left\| (\Sigma^2 + \mu I_r)^{-1} \Sigma U^T R \right\|_2$$

Por otro lado, definamos $\mathbf{z} = U^T R$, $\text{Rank}(J) = r \leq n$ y $\sigma_1, \dots, \sigma_r$ los valores singulares de la matriz J , de ahí que

$$\begin{aligned}
\Delta &= \left\| \text{diag} \left(\frac{1}{\sigma_1^2 + \mu}, \dots, \frac{1}{\sigma_r^2 + \mu} \right) \text{diag} (\sigma_1, \dots, \sigma_r) \mathbf{z} \right\|_2 \\
&= \left\| \text{diag} \left(\frac{\sigma_1}{\sigma_1^2 + \mu}, \dots, \frac{\sigma_r}{\sigma_r^2 + \mu} \right) \mathbf{z} \right\|_2 \\
&= \left[\sum_{i=1}^r \left(\frac{\sigma_i \mathbf{z}_i}{\sigma_i^2 + \mu} \right)^2 \right]^{\frac{1}{2}}
\end{aligned}$$

por tanto,

$$\Delta^2 = \sum_{i=1}^r \left(\frac{\sigma_i \mathbf{z}_i}{\sigma_i^2 + \mu} \right)^2, \quad (2.39)$$

la anterior relación, puede ilustrarse con el gráfico de la figura (2.10).

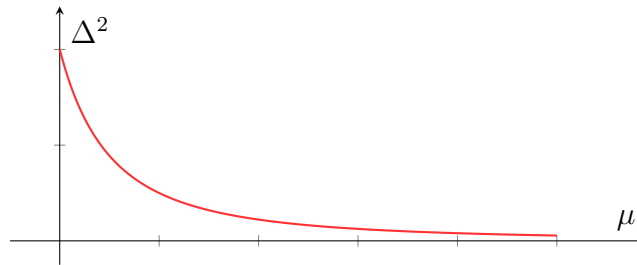


Figura 2.10: Relación entre parámetro L.M y radio de confianza

Es decir, si aumentamos el parámetro de μ , se está reduciendo el radio de confianza y viceversa si reducimos el tamaño del parámetro μ entonces estamos aumentando el radio de confianza.

En [21], usan $\tau_1 = \frac{1}{4}$ y $\tau_2 = 4$, el resto de parámetros son los mismos escogidos en el método de región de confianza tradicional, donde $\Delta_{min} > 0$ se escoge de tal forma que μ_k no sea demasiado pequeño; motivados por los resultados de esta nueva propuesta, los autores *Bin Fan* et al, proponen una nueva actualización [20], usan $\tau_1 = \frac{1}{10}$, $\tau_2 = 10$ y definen,

$$\mu_{k+1} = \beta_{k+1} \|R(\mathbf{x}_k)\|^2. \quad (2.40)$$

la cual usaremos también, en este documento.

El método (2.36), del cual existen varias versiones que varían según la estrategia para escoger μ_k , fue sugerido por *Levenberg* en (1944) y luego por *Marquardt* en (1963), en [64] usan $\mu_k = \|R(\mathbf{x}_k)\|^2$, en [22] usan $\mu_k = \min(\|R(\mathbf{x}_k)\|, \|J_k^T R(\mathbf{x}_k)\|)$.

2.4.2. Región de Confianza para el PCNL(F)

Sean Φ_λ la función dada por (2.5) la cual está asociada al PCNL(F) (1.1) [46], y Ψ_λ la función de merito asociada a Φ_λ (2.14) [19]. Como ya mencionamos anteriormente, se debe resolver el problema de minimización

sin restricciones (2.25), para ello, usaremos un método tipo Newton, como en la sección 2.3, recordemos que estos métodos son locales, por ello, para globalizar dicho método se puede usar la estrategia de región de confianza. Como se explicó al inicio de esta sección; primero redefinamos el modelo cuadrático (2.29) el cual está dado de la siguiente forma.

$$q_k(\mathbf{x}_k + \mathbf{d}) =: \frac{1}{2} \|\Phi_\lambda(\mathbf{x}_k) + H_k \mathbf{d}\|_2^2 \quad (2.41)$$

donde $H_k \in \partial\Phi_\lambda(\mathbf{x}_k)$, observe que q_k en (2.41) se puede escribir como una cuadrática, esto es:

$$q_k(\mathbf{x}_k + \mathbf{d}) = \frac{1}{2} \|\Phi_\lambda(\mathbf{x}_k)\|_2^2 + [\Phi_\lambda(\mathbf{x}_k)^T H_k] \mathbf{d} + \frac{1}{2} \mathbf{d}^T [H_k^T H_k] \mathbf{d}, \quad (2.42)$$

lo que es equivalente,

$$q_k(\mathbf{x}_k + \mathbf{d}) = \Psi_\lambda(\mathbf{x}_k) + \nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d} + \frac{1}{2} \mathbf{d}^T [H_k^T H_k] \mathbf{d}, \quad (2.43)$$

Dado el radio $\Delta_k = \| (H_k^T H_k + \mu_k I_n)^{-1} H_k^T \Phi_\lambda(\mathbf{x}_k) \|_2$, se resuelve el subproblema,

$$\begin{aligned} \text{mín } q_k(\mathbf{x}_k + \mathbf{d}) \\ \|\mathbf{d}\|_2 = \Delta_k \end{aligned} \quad (2.44)$$

Los autores en [23, 20, 62] proponen un método tipo *Levenberg-Marquardt* (2.34) para resolver el subproblema (2.44) de manera aproximada. Es decir, resuelven el sistema lineal:

$$(H_k^T H_k + \mu_k I_n) \mathbf{d}_k = -H_k^T \Phi_\lambda(\mathbf{x}_k), \quad (2.45)$$

donde $H_k \in \partial_B \Phi_\lambda(\mathbf{x}_k)$. En [20] toman $\mu_k = \beta_k \|\Phi_{FB}(\mathbf{x}_k)\|^2$, en nuestro algoritmo usaremos $\mu_k = \beta_k \|\Phi_\lambda(\mathbf{x}_k)\|^2$.

Si $\tilde{\rho}_k \leq \rho_1$, el paso es rechazado, de lo contrario es aceptado, en cualquier caso se actualiza el parámetro μ_k según el criterio escogido, en el algoritmo

propuesto usaremos el siguiente.

$$\beta_{k+1} = \begin{cases} \tau_2 \beta_k & \text{si } \tilde{\rho}_k \leq \rho_1 \\ \beta_k & \text{si } \tilde{\rho}_k \in (\rho_1, \rho_2) \\ \max\{\tau_1 \beta_k, \Delta_{min}\} & \text{si } \tilde{\rho}_k \geq \rho_2 \end{cases} \quad (2.46)$$

y definamos,

$$\mu_{k+1} = \beta_{k+1} \|\Phi_\lambda(\mathbf{x}_k)\|^2.$$

Por último, se actualiza el iterante $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$ y se repite el proceso.

El siguiente diagrama nos ilustra, la estructura de un algoritmo tipo *Levenberg-Marquardt* que utiliza solo región de confianza.

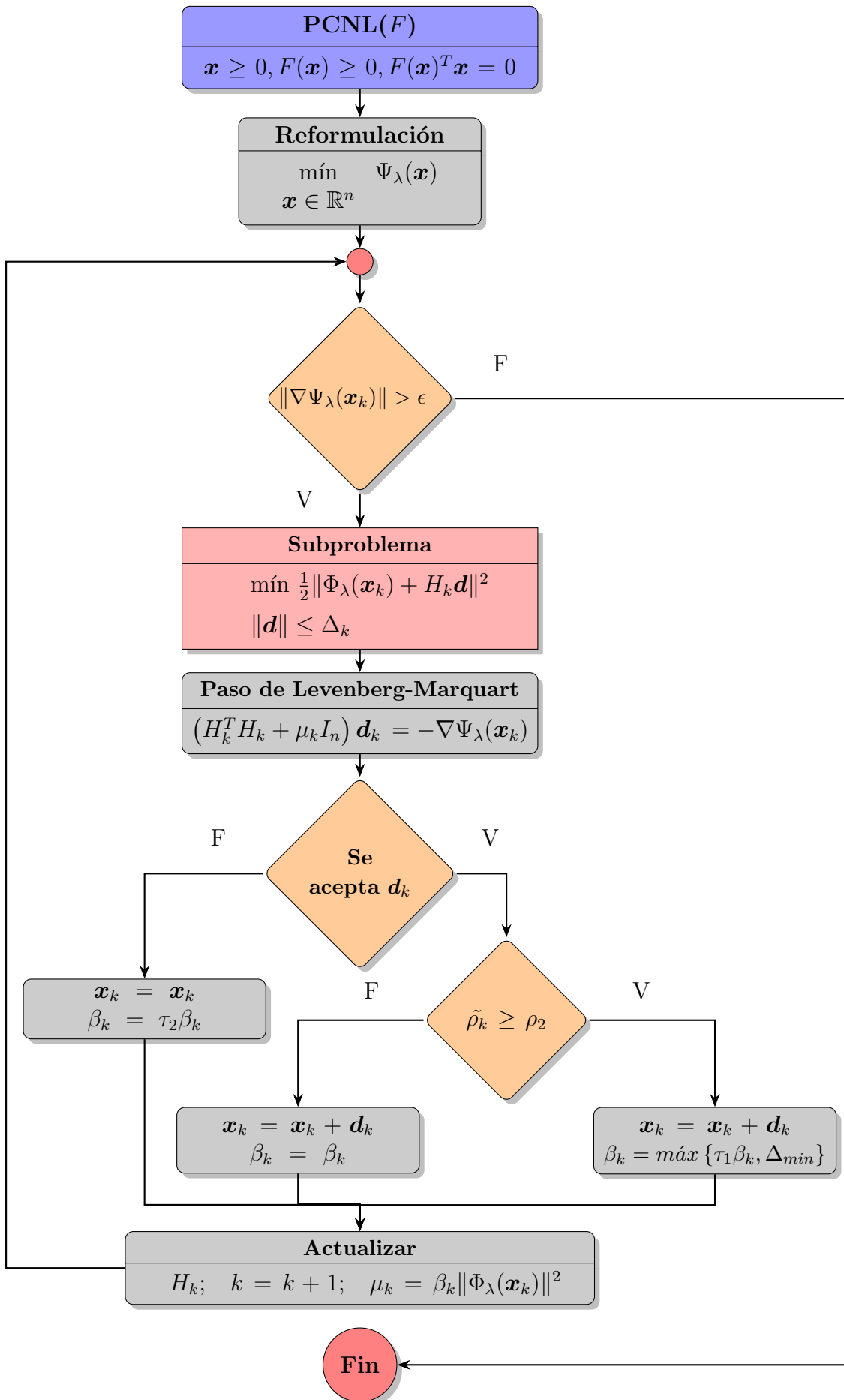


Figura 2.11: Diagrama de flujo del algoritmo de región de confianza L.M.

2.4.3. Búsqueda lineal para un problema de minimización sin restricciones

Como ya se había mencionando anteriormente, muchos algoritmos globales usan la estrategia de búsqueda lineal para resolver el problema (2.28), es por ello que destacamos los siguientes procesos que consideramos relevantes:

1. Búsqueda direccional, el algoritmo elige una dirección de descenso \mathbf{d}_k según el caso (*Gradiente*, *Newton* o *cuasi-Newton*).
2. Partiendo desde el iterante actual \mathbf{x}_k , buscamos a lo largo de esa dirección un nuevo punto en el cual la función que define el problema (2.28) decrezca de manera suficiente, para ello utilizamos la condición de *Armijo* [31] o la condición de *Goldstein*[18, 31]. Finalmente, se actualiza el vector \mathbf{x}_{k+1} y se repite el proceso.

2.4.4. Búsqueda lineal para el $PCNL(F)$

Esta es una de las estrategias más usadas para la implementación de los algoritmos globales que resuelven el $PCNL(F)$, donde el problema a resolver está dado como en (2.25).

1. A partir de la dirección dada en (2.45).
2. Para un $\delta \in (0, 1)$ fijo, se busca un α_k tal que

$$\Psi_\lambda(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq \Psi_\lambda(\mathbf{x}_k) + \delta \alpha_k \nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k$$

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

Se repite el proceso.

2.4.5. Combinación de las técnicas de globalización región de confianza y búsqueda lineal para $PCNL(F)$

Las técnicas de globalización mencionadas en las anteriores secciones, poseen ventajas y desventajas; por ejemplo, con los métodos *Cuasi Newton* de región de confianza; estos se catalogan como robustos, estables y tienen excelentes propiedades de convergencia. Sin embargo, resolver el subproblema de región de confianza es computacionalmente costoso, especialmente para problemas grandes.

Por otro lado, los métodos tipo *Newton*, *Cuasi Newton*, método del gradiente de búsqueda lineal requieren pocos cálculos para determinar un nuevo punto iterativo, pero dichos métodos tiene algunas desventajas. Por ejemplo, cuando la búsqueda direccional, en el algoritmo de búsqueda lineal es casi ortogonal a la dirección máximo descenso se tiene que el tamaño de paso es muy pequeño; en algunos casos, los errores de redondeo pueden hacer que falle la búsqueda lineal [61]. Observando las ventajas de estos dos métodos, han surgido nuevos de algoritmos que combinan las dos estrategias. En el años de 1998 *Nocedal y Yuan* proponen nuevo algoritmo el cual conserva la buenas propiedades de convergencia del método de región de confianza pero que es más económico(costo computacional) cuando el tamaño de las variables sean muy grandes [50]. De ahí que, motivados por los resultados obtenidos en [50, 61, 20], construimos un algoritmo híbrido en el cual; si se rechaza un paso de prueba del método de región de confianza, se pueda usar un método de búsqueda de línea para obtener el siguiente punto iterativo, además que comparta las ventajas de estos dos métodos, es decir, que sea eficiente y eficaz. El nuevo algoritmo hídrido que propondremos está diseñado especialmente para resolver problemas de complementariedad no lineal. Ahora, con el fin de darnos una idea de como funciona el algoritmo que proponemos, realizamos el siguiente diagrama flujo; en él, se ve de forma explicita como se combinan las técnicas región de confianza y búsqueda lineal.

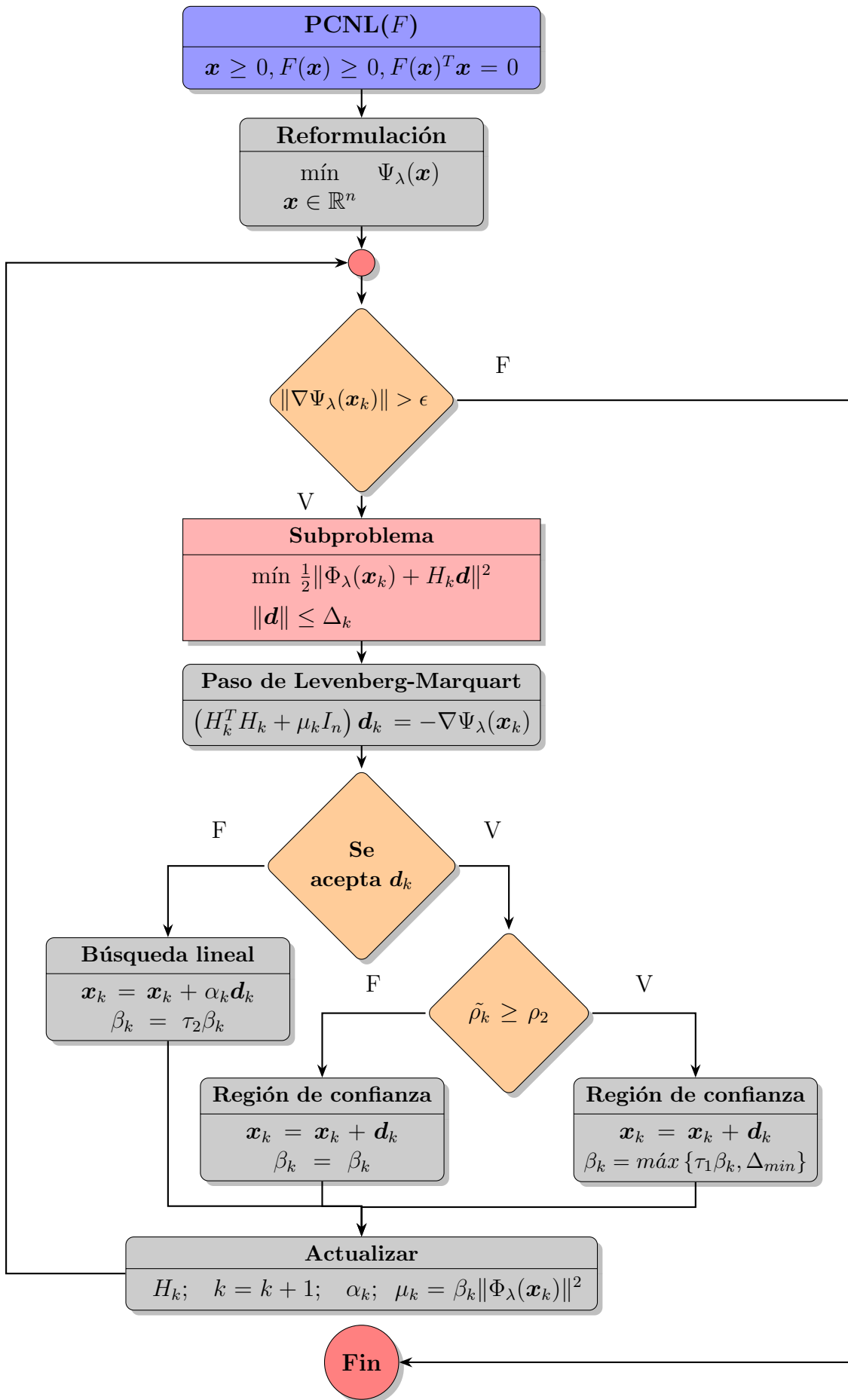


Figura 2.12: Diagrama de flujo del algoritmo híbrido de región de confianza y búsqueda lineal (LMHIB)

2.5. Teoremas para el análisis de la convergencia del algoritmo

En esta sección, presentamos algunas definiciones como también teoremas que se asocian con el problema (1.1), estos son de suma importancia para análisis sobre la convergencia global, superlineal y cuadrática de nuestro algoritmo.

Es importante establecer condiciones suficientes para la no singularidad de las matrices del Jacobiano generalizado en una solución del problema (1.1). Por ello, para desarrollo de la teoría de convergencia de los algoritmos que resuelvan el PCNL, se introduce el concepto de regularidad.

Definición 2.5. Sea \mathbf{x}^* una solución del PCNL(F).

1. Si todas las matrices $H \in \partial_B \Phi_\lambda(\mathbf{x}^*)$ son no singulares, \mathbf{x}^* es llamada una solución *BD-regular*.
2. Si la submatriz $F'(\mathbf{x}^*)_{\alpha\alpha}$ ⁴ es no singular y el complemento de Shurt,

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

es una *P-matriz*⁵, \mathbf{x}^* es llamada una solución *R-regular*, donde α, β, γ están dados por 2.15, 2.16, 2.17.

Una condición suficiente para garantizar la no singularidad de los elementos del Jacobiano generalizado de Φ_λ en una solución del PCNL, fue presentada por *Kanzow y Kleinmichel* [38] dada en el siguiente teorema.

Teorema 2.1. [36] Si $\mathbf{x}^* \in \mathbb{R}^n$ es una solución *R-regular* del PCNL,

⁴Dada una matriz $A = (a_{ij})$ de tamaño $m \times n$ y conjuntos de índices η y τ , la matriz $A_{\eta\tau}$ es aquella con componentes a_{ij} tal que $i \in \eta$ y $j \in \tau$.

⁵Una matriz $M \in \mathbb{R}^{n \times n}$ es una *P-matriz*, si para cualquier vector no nulo $\mathbf{y} \in \mathbb{R}^n$, existe un índice $i_0 = i_0(\mathbf{y}) \in \{1, \dots, n\}$, tal que $y_{i_0}[M\mathbf{y}]_{i_0} > 0$.

entonces las matrices en el Jacobiano generalizado $\partial\Phi_\lambda(\mathbf{x}^*)$ son no singulares.

Por tanto, una consecuencia inmediata del anterior teorema es que toda solución R-regular del PCNL, es una solución BD-regular.

Ahora, queremos ver bajo qué hipótesis se garantiza la semisuavidad y fuerte semisuavidad de la función Φ_λ , para ello, recordemos las siguientes definiciones.

Definición 2.6. Sea $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función localmente Lipschitz continua, se dice que G es semisuave en $\mathbf{x} \in \mathbb{R}^n$, si G en \mathbf{x} es direccionalmente diferenciable y,

$$G(\mathbf{x} + \mathbf{h}) - G(\mathbf{x}) - H\mathbf{h} = o(\|\mathbf{h}\|)$$

para todo $H \in \partial_C G(\mathbf{x} + \mathbf{h})$ cuando $\mathbf{h} \rightarrow \mathbf{0}$.

Definición 2.7. Sea $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función localmente Lipschitz continua, se dice que G es fuertemente semisuave en $\mathbf{x} \in \mathbb{R}^n$, si G en \mathbf{x} es direccionalmente diferenciable y,

$$G(\mathbf{x} + \mathbf{h}) - G(\mathbf{x}) - H\mathbf{h} = O(\|\mathbf{h}\|^2)$$

para todo $H \in \partial G_C(\mathbf{x} + \mathbf{h})$ cuando $\mathbf{h} \rightarrow \mathbf{0}$.

La demostración del siguiente teorema se puede verificar en [38, 57]. Este resultado se usa para comprobar la convergencia superlineal y cuadrática de los algoritmos que resuelven PCNL.

Teorema 2.2. Si $\{\mathbf{x}_k\} \subset \mathbb{R}^n$ es una sucesión convergente con punto límite $\mathbf{x}^* \in \mathbb{R}^n$, entonces se cumplen las siguientes afirmaciones:

1. Si la función Φ_λ es semisuave, implica que para cualquier $H_k \in$

$$\partial_C \Phi_\lambda(\mathbf{x}_k),$$

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

2. Si $F'(\mathbf{x})$ es Lipschitz continua, entonces la función Φ_λ es fuertemente semisuave, lo que implica que para cualquier $H_k \in \partial_C \Phi(\mathbf{x}_k)$,

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

La demostración de la **Proposición 2.4** es análoga a la demostración de la **proposición 3.1** en [57].

Proposición 2.4. Si para todo $H \in \partial G(\mathbf{x})$ es no singular, entonces existe una vecindad $N(\mathbf{x})$ de \mathbf{x} y una constante C tal que para cualquier $\mathbf{y} \in N(\mathbf{x})$ y para todo $H \in \partial G(\mathbf{y})$, H es no singular y $\|H^{-1}\| \leq C$.

Ahora, el siguiente lema nos garantiza una cota para la norma de la matriz inversa de la matriz H .

Lema 2.6. Supongamos que \mathbf{x}^* es la solución del sistema no lineal $\Phi_\lambda(\mathbf{x}) = 0$, y cualquier $H^* \in \partial_B \Phi_\lambda(\mathbf{x}^*)$ no singular. Entonces existe una vecindad $B(\mathbf{x}^*, r)$ y una constante $\nu > 0$ talque

$$\|H^{-1}\| \leq \nu \quad \forall H \in \partial_B \Phi_\lambda(\mathbf{x}), \forall \mathbf{x} \in B(\mathbf{x}^*, r)$$

Demostración. En efecto, dado que si \mathbf{x}^* es la solución $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$ y se tiene que todas las matrices $H^* \in \partial_B \Phi_\lambda(\mathbf{x}^*)$ son no singulares, se hace uso de la Proposición 2.4, y con ello se garantiza la demostración. \square

Definición 2.8. Sea $G: \mathbb{R}^n \rightarrow \mathbb{R}$ se dice que es una función SC^1 , si G es continuamente diferenciable y su gradiente es semisuave.

El siguiente resultado presenta una característica particular de la función de mérito, la cual se quiere minimizar, su demostración recoge el hecho

de que la función de mérito, la podemos ver como una composición de funciones de clase SC^1 [39].

Teorema 2.3. *Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, una función continuamente diferenciable, sea Φ_λ la función dada en (2.5) asociada a el problema (1.1) y Ψ_λ la función de mérito asociada a Φ_λ . Si para todo $i = 1, \dots, n$, la función F_i es de clase SC^1 , entonces Ψ_λ es también de clase SC^1 .*

Demostración. En efecto, se tiene que las funciones componentes de Φ_λ están dadas como sigue

$$\Phi_{\lambda_i}(\mathbf{x}^*) = (\varphi_{\lambda_i}(x_i^*, F_i(\mathbf{x}^*))) \quad \text{para todo } i = 1, 2, \dots, n.$$

Dado que las funciones componentes de F son de clase SC^1 , entonces φ_{λ_i} es de clase SC^1 , esto para todo $i = 1, 2, \dots, n$.

Por otro lado se tiene que,

$$\Psi_\lambda(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^n (\varphi_{\lambda_i}(x_i, F_i(\mathbf{x})))^2,$$

además dado que composición y suma de funciones SC^1 sigue siendo de clase SC^1 [27], se puede concluir que Ψ_λ es de clase SC^1 . \square

Capítulo 3

Algoritmo y Teoría de convergencia

En este capítulo, proponemos un algoritmo híbrido (1) que utiliza región de confianza y búsqueda lineal, para resolver el PCNL(F) (1.1), el cual ha sido reformulado como un problema de minimización mediante la familia uniparamétrica presentada por [38], utilizando en cada paso el método de *Levenberg-Marquardt*[45], además desarrollamos la teoría de convergencia global del algoritmo.

3.1. Algoritmo

El algoritmo en forma de diagrama de flujo presentado en el anterior capítulo (Figura 2.12) nos da una idea general de como es el proceso de selección y aceptación del paso, así como actualizaciones del radio de confianza, parámetro de *Levenberg-Marquardt* y paso. Ahora, incluiremos todos los parámetros que intervienen en el algoritmo, es por ello que a continuación lo escribimos en el ambiente de algoritmo.

Algoritmo 1 Algoritmo Híbrido de Levenberg-M.

Entrada: Sea $\mathbf{x}_0 \in \mathbb{R}^n$, $0 < \rho_1 < \rho_2 < 1$, $0 < \tau_1 < 1 < \tau_2$, $\beta_0 = \mu_0 > \Delta_{min} > 0$,
 $\delta \in (0, 1)$, $t \in (0, 1)$, $\lambda \in (0, 4)$ $k := 0$.

1: $\nabla\Psi_\lambda(\mathbf{x}_k) \leftarrow H_k^T \Phi_\lambda(\mathbf{x}_k)$, donde $H_k \in \partial\Phi_\lambda(\mathbf{x}_k)$.

2: $\mathbf{d}_k \leftarrow - (H_k^T H_k + \mu_k I_n)^{-1} \nabla\Psi_\lambda(\mathbf{x}_k)$.

3: **mientras** $\|\nabla\Psi_\lambda(\mathbf{x}_k)\| < 10^{-6}$ y $k < 200$ **hacer**

4: $\tilde{\rho}_k \leftarrow Ared_k / Pred_k$.

donde, $Pred_k \leftarrow -\nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k - \frac{1}{2} \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k$

$Ared_k \leftarrow \Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_k + \mathbf{d}_k)$

5: **si** $\tilde{\rho}_k \leq \rho_1$ **entonces**

6: Calcular $\alpha_k := \max \{t^l | l = 0, 1, \dots\}$ tal que

$$\Psi_\lambda(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq \Psi_\lambda(\mathbf{x}_k) + \delta \alpha_k \nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k$$

7: $\mathbf{x}_k \leftarrow \mathbf{x}_k + \alpha_k \mathbf{d}_k$; $\beta_k \leftarrow \tau_2 \beta_k$

8: **si no**

9: **si** $\tilde{\rho}_k \geq \rho_2$ **entonces**

10: $\mathbf{x}_k \leftarrow \mathbf{x}_k + \mathbf{d}_k$; $\beta_k \leftarrow \max \{\tau_1 \beta_k, \Delta_{min}\}$

11: **si no**

12: $\mathbf{x}_k \leftarrow \mathbf{x}_k + \mathbf{d}_k$; $\beta_k \leftarrow \beta_k$

13: **fin si**

14: **fin si**

15: actualizar,

16: $\mu_k \leftarrow \beta_k \|\Phi_\lambda(\mathbf{x}_k)\|^2$,

17: $\nabla\Psi_\lambda(\mathbf{x}_k) \leftarrow H_k \Phi_\lambda(\mathbf{x}_k)$,

18: $\mathbf{d}_k \leftarrow - (H_k^T H_k + \mu_k I_n)^{-1} \nabla\Psi_\lambda(\mathbf{x}_k)$.

19: $k \leftarrow k + 1$.

20: **fin mientras**

21: **Salida:** \mathbf{x}_{k+1} .

Después de calculado $\tilde{\rho}_k$, se decide como actualizar \mathbf{x}_k , si la disminución de la función Ψ es por lo menos una fracción (ρ_1) de la disminución de la función modelo q_k , simplemente sumamos a \mathbf{x}_k el vector \mathbf{d}_k y aumentamos β_k ; en caso contrario en vez de rechazar el punto, se utiliza búsqueda lineal, como lo muestra el paso 6, además, en el paso 10, se puede observar que $\Delta_{min} > 0$ es una constante de valor pequeño, esto con el fin de evitar dar

pasos muy grandes cuando estamos cerca de la solución del problema, y por ende evita que el parámetro μ_k sea demasiado pequeño, es decir Δ_{min} es una cota inferior, cuando k es lo suficientemente grande.

En la siguiente sección, nos ocuparemos de demostrar que el **Algoritmo 1** propuesto anteriormente está bien definido, así como también la convergencia global del algoritmo.

3.2. Convergencia global

Para comenzar, debemos garantizar que la sucesión generada por el **Algoritmo 1** está bien definida y una parte importante de esto es demostrar que el paso de *Levenberg-Marquardt* (L.M) está bien definido, para ello basta demostrar la siguiente proposición:

Proposición 3.1. *Sean $H_k \in \partial\Phi_\lambda(\mathbf{x}_k)$ y $\mu_k > 0$ (parámetro Levenberg-Marquardt). Si $B = H_k^T H_k + \mu_k I$ entonces B es una matriz simétrica definida positiva.*

Demostración. Primero demostraremos la simetría de la matriz B , es decir verifiquemos que $B^T = B$, observemos que:

$$\begin{aligned} B^T &= (H_k^T H_k + \mu_k I)^T \\ &= (H_k^T H_k)^T + (\mu_k I)^T \\ &= H_k^T H_k + \mu_k I \\ &= B, \end{aligned}$$

por último, demostremos que la matriz B es definida positiva, es decir verifiquemos que $\mathbf{x}^T B \mathbf{x} > 0$, para todo $\mathbf{x} \in \mathbb{R}^n$ no nulo. En efecto,

$$\begin{aligned}
\mathbf{x}^T B \mathbf{x} &= \mathbf{x}^T (H_k^T H_k + \mu_k I) \mathbf{x} \\
&= \mathbf{x}^T (H_k^T H_k) \mathbf{x} + \mathbf{x}^T (\mu_k I) \mathbf{x} \\
&= (H_k \mathbf{x})^T H_k \mathbf{x} + \mu_k \mathbf{x}^T \mathbf{x} \\
&= \|H_k \mathbf{x}\|_2^2 + \mu_k \|\mathbf{x}\|_2^2 > 0
\end{aligned}$$

□

En seguida, consideremos las siguientes observaciones, así como también el **Lema 3.1** que serán de gran ayuda para demostrar que el paso 4 en el **Algoritmo 1** está bien definido.

Observación 1:

- a) De ahora en adelante, la norma Euclidiana, o la norma matricial $\|\cdot\|_2$ la denotaremos de la siguiente forma $\|\cdot\|$, en caso que utilicemos otra norma la vamos a especificar.
- b) Dado que $H_k^T H_k + \mu_k I$ es una matriz definida positiva, en el **Algoritmo 1**, el paso 2 tiene solución para todo k .
- c) Si suponemos que el algoritmo no termina en un número finito de iteraciones, entonces se tiene que

$$\|\nabla \Psi_\lambda(\mathbf{x}_k)\| \neq 0, \quad \text{para todo } k \geq 0 \quad (3.1)$$

- d) Dado que el algoritmo combina el método de región de confianza y búsqueda lineal, se define los siguientes conjuntos de índices.

$$I = \{k : \rho_k \geq \rho_1\},$$

es decir, es el conjunto de índices donde se aceptó el paso usando el

método de región de confianza y

$$J = \{k : \rho_k < \rho_1\},$$

es el conjunto de índices donde se utiliza búsqueda lineal.

El siguiente lema es un resultado famoso de *Powell* [55], el cual nos indica que la función modelo tiene un decrecimiento, es decir \mathbf{d}_k es un minimizador en la región de confianza.

Lema 3.1. [20, 39] Sean $H_k \in \partial\Phi_\lambda(\mathbf{x}_k)$, \mathbf{d}_k como en el paso 2 del **algoritmo 1**, para todo $k \geq 0$ se cumple

$$Pred_k \geq \frac{1}{2} \|\nabla\Psi_\lambda(\mathbf{x}_k)\| \min \left\{ \|\mathbf{d}_k\|, \frac{\|\nabla\Psi_\lambda(\mathbf{x}_k)\|}{\|H_k^T H_k\|} \right\}.$$

Ahora, vamos a verificar que efectivamente la función objetivo está decreciendo en cada paso, para ello basta probar el **Lema 3.2**; su demostración es análoga a la presentada en [20].

Lema 3.2. Supongamos que la sucesión $\{\mathbf{x}_k\}$ es generada por el **Algoritmo 1**, entonces

$$\Psi_\lambda(\mathbf{x}_{k+1}) < \Psi_\lambda(\mathbf{x}_k), \quad \text{para todo } k > 0. \quad (3.2)$$

Demostración. Para esta demostración, consideramos dos casos, los cuales dependen si $k \in I$ o $k \in J$:

1. Si $k \in I$ entonces $\tilde{\rho}_k \geq \rho_1$,

$$\Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1}) \geq \rho_1 Pred_k,$$

luego por el **Lema 3.1** se tiene que

$$\Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1}) \geq \frac{1}{2} \rho_1 \|\nabla\Psi_\lambda(\mathbf{x}_k)\| \min \left\{ \|\mathbf{d}_k\|, \frac{\|\nabla\Psi_\lambda(\mathbf{x}_k)\|}{\|H_k^T H_k\|} \right\} \geq 0.$$

Por tanto,

$$\Psi_\lambda(\mathbf{x}_{k+1}) \leq \Psi_\lambda(\mathbf{x}_k)$$

2. Si $K \in J$, es decir $\tilde{\rho}_k < \rho_1$, usamos el método de búsqueda lineal dado en el paso 8 del **Algoritmo 1**; esto es, se busca α_k tal que:

$$\Psi(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq \Psi(\mathbf{x}_k) + \delta \alpha_k \nabla \Psi(\mathbf{x}_k)^T \mathbf{d}_k$$

Pero se puede observar que por el paso de L.M.

$$[H_k^T H_k + \mu_k I] \mathbf{d}_k = -\nabla \Psi_\lambda(\mathbf{x}_k)$$

$$\mathbf{d}_k^T [H_k^T H_k + \mu_k I] \mathbf{d}_k = -\mathbf{d}_k^T \nabla \Psi_\lambda(\mathbf{x}_k) \quad (3.3)$$

luego, por la proposición 3.1, inferimos que

$$\mathbf{d}_k^T \nabla \Psi_\lambda(\mathbf{x}_k) < 0 \quad (3.4)$$

en consecuencia, existe α_k tal que:

$$\Psi(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq \Psi(\mathbf{x}_k) + \delta \alpha_k \nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k$$

o lo que es equivalente,

$$\Psi(\mathbf{x}_k + \alpha_k \mathbf{d}_k) - \delta \alpha_k \nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k \leq \Psi(\mathbf{x}_k),$$

además, dado que $\delta \alpha_k > 0$ y $\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k < 0$ para todo k , entonces se tiene que $-\delta \alpha_k \mathbf{d}_k^T B \mathbf{d}_k > 0$, de ahí que:

$$\Psi_\lambda(\mathbf{x}_{k+1}) < \Psi_\lambda(\mathbf{x}_k).$$

Por tanto, por 1 y 2, se garantiza la desigualdad (3.2). \square

Por último, nos falta garantizar que el paso 8 en el **Algoritmo 1** se termina en un número finito de iteraciones. Todo esto se resume en la

demostración del siguiente lema.

Lema 3.3. *Supongamos que el **Algoritmo 1** genera una sucesión infinita $\{\mathbf{x}_k\}$, entonces el algoritmo está bien definido, esto es: el paso 2 está bien definido, $Pred_k > 0$ y el proceso de búsqueda del tamaño del paso con búsqueda lineal se logra en un número finito de pasos (no se quedará en un ciclo infinito).*

Demostración. Primero vamos a demostrar que $Pred_k > 0$.

- Partiendo de la ecuación (3.3), se infiere que:

$$\mathbf{d}_k^T H_k^T H_k \mathbf{d}_k + \mathbf{d}_k^T \mu_k \mathbf{d}_k = -\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k, \quad (3.5)$$

en consecuencia

$$\begin{aligned} \mu_k \|\mathbf{d}_k\|^2 &= -\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k - \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k \\ &\leq -\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k - \frac{1}{2} \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k \\ &= Pred_k \end{aligned}$$

es decir,

$$\mu_k \|\mathbf{d}_k\|^2 \leq Pred_k \quad (3.6)$$

Pero dado que, $\mu_k \|\mathbf{d}_k\|^2 > 0$, entonces concluimos que $Pred_k > 0$.

- Ahora, resta garantizar que para cada k existe un entero l , tal que

$$\Psi(\mathbf{x}_k + t^l \mathbf{d}_k) \leq \Psi(\mathbf{x}_k) + \delta t^l \nabla \Psi(\mathbf{x}_k)^T \mathbf{d}_k$$

Por la desigualdad (3.4), \mathbf{d}_k es una dirección de descenso, entonces por **Proposición 5.2** [54], se garantiza que existe un entero l_k para cada k tal que se verifica

$$\Psi(\mathbf{x}_k + t^{l_k} \mathbf{d}_k) \leq \Psi(\mathbf{x}_k) + \delta t^{l_k} \nabla \Psi(\mathbf{x}_k)^T \mathbf{d}_k$$

es decir la búsqueda lineal presentada en el **Algoritmo 1** termina en un número finito de pasos.

Lo anterior, demuestra que el algoritmo está bien definido. \square

Ahora, para demostrar la convergencia global del **Algoritmo 1** vamos a definir el conjunto de nivel inferior.

Definición 3.1. Sea $\mathbf{x}_0 \in \mathbb{R}^n$, se define el conjunto de nivel inferior de Ψ_λ en \mathbf{x}_0 como:

$$\zeta(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^n : \Psi_\lambda(\mathbf{x}) \leq \Psi_\lambda(\mathbf{x}_0)\}. \quad (3.7)$$

Además, en este contexto necesitamos suponer algunas hipótesis que usan para el estudio del análisis de convergencia de nuestro algoritmo.

- **H.1** El conjunto de nivel $\zeta(\mathbf{x}_0)$ es un conjunto cerrado y acotado, además $\nabla \Psi_\lambda(\mathbf{x}_k)$ es semisuave en $\zeta(\mathbf{x}_0)$.
- **H.2** $\Phi_\lambda(\mathbf{x})$ es *Lipschitz continua* en $\zeta(\mathbf{x}_0)$, es decir existe un $K > 0$ tal que:

$$\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\| \leq K \|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in \zeta(\mathbf{x}_0)$$

- **H.3** Existe una constante $\nu_1 > 0$ tal que

$$\mathbf{d}_k^T [H_k^T H_k] \mathbf{d}_k \geq \nu_1 \|\mathbf{d}\|^2 \quad \text{para todo } k \geq 0.$$

Observación 2: A modo de resumen, tenemos las siguientes afirmaciones:

a) Si F es una P-función uniforme ¹ o $F(\mathbf{x}) = M\mathbf{x} + \mathbf{q}$, donde $\mathbf{q} \in \mathbb{R}^n$

¹Sea $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$; se dice que es una P-función uniforme en un conjunto S si existe un $\kappa > 0$ tal

y $M \in \mathbb{R}^{n \times n}$ es una P_0 -Matriz², entonces el conjunto de nivel $\zeta(\mathbf{x}_0)$ definido (3.7) es acotado [20].

b) Si Φ_λ es una función SC^1 , entonces $\nabla \Psi_\lambda$ es una función semisuave[39].

c) [20, 21] Para todo $\mathbf{x} \in \zeta(\mathbf{x}_0)$, se cumple en virtud al lema 2.4.

$$\|H\| \leq L, \quad \forall H \in \partial_B \Phi_\lambda(\mathbf{x}). \quad (3.8)$$

Los **Lemas 3.4** y **3.5** serán de utilidad para la demostración de la convergencia global; el primero garantiza que la sucesión generada por el **Algoritmo 1** converge en el conjunto de nivel inferior (3.7), y con ello garantiza que $\{\Psi_\lambda(\mathbf{x}_k)\}$ converge; el segundo garantiza que la sucesión generada por $\{\mu_k\}$ (parámetro L.M) es acotada; para la demostración de los **Lemas 3.4** y **3.5** y el **Teorema 3.1** nos basamos en las ideas propuestas por [20].

Lema 3.4. *Sea $\{\mathbf{x}_k\}$ la sucesión generada por el **Algoritmo 1**, entonces $\{\mathbf{x}_k\}$ permanece en $\zeta(\mathbf{x}_0)$, además la sucesión $\{\Psi_\lambda(\mathbf{x}_k)\}$ es una sucesión convergente.*

Demostración. En primer lugar, demostremos que todos los puntos de la sucesión $\{\mathbf{x}_k\}$ se encuentran en el conjunto de nivel $\zeta(\mathbf{x}_0)$, para ello, usaremos el principio de inducción matemática sobre k .

Si $k = 0$, entonces es claro que $\mathbf{x}_0 \in \zeta(\mathbf{x}_0)$, puesto que $\Psi_\lambda(\mathbf{x}_0) \leq \Psi_\lambda(\mathbf{x}_0)$.

Ahora, asumamos que $\mathbf{x}_m \in \zeta(\mathbf{x}_0)$ para todo $m = 1, 2, \dots, k$ y demostremos que $\mathbf{x}_{k+1} \in \zeta(\mathbf{x}_0)$, en efecto: dado que,

$$\mathbf{x}_m \in \zeta(\mathbf{x}_0), \quad m = 1, 2, \dots, k.$$

que para todo $\mathbf{x}, \mathbf{y} \in S$,

$$\max_{1 \leq i \leq n} (\mathbf{x}_i - \mathbf{y}_i)(G_i(\mathbf{x}) - G_i(\mathbf{y})) \leq \kappa \|\mathbf{x} - \mathbf{y}\|^2$$

²Sea $M \in \mathbb{R}^{n \times n}$; se dice que M es una P_0 -Matriz si todos sus menores principales son positivos.

se tiene que:

$$\Psi_\lambda(\mathbf{x}_k) \leq \Psi_\lambda(\mathbf{x}_0),$$

pero por (3.2) sabemos que,

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

por tanto,

$$\Psi_\lambda(\mathbf{x}_{k+1}) \leq \Psi_\lambda(\mathbf{x}_0),$$

es decir,

$$\mathbf{x}_{k+1} \in \zeta(\mathbf{x}_0).$$

Ahora, probemos que $\{\Psi_\lambda(\mathbf{x}_k)\}$ es una sucesión convergente, dado que Ψ_λ es una función continua, basta probar que la sucesión $\{\mathbf{x}_k\}$ converge en el conjunto de nivel, en efecto:

Dado que $\{\mathbf{x}_k\}$ es una sucesión de puntos de $\zeta(\mathbf{x}_0) \subset \mathbb{R}^n$ el cual es un conjunto cerrado y acotado, es decir es compacto, por el teorema *Bolzano-Weierstrass*, se tiene que existe una subsucesión $\{\mathbf{x}_{k_j}\}$ convergente, es decir $\mathbf{x}_{k_j} \rightarrow \mathbf{x}$ cuando $j \rightarrow \infty$, de ahí que $\Psi_\lambda(\mathbf{x}_{k_j}) \rightarrow \Psi_\lambda(\mathbf{x})$ cuando $j \rightarrow \infty$. Por tanto, $\{\Psi_\lambda(\mathbf{x}_k)\}$ es una sucesión convergente. \square

Lema 3.5. *Supongamos que se cumplen las hipótesis **H.1**, **H.2**, **H.3** y que el **Algoritmo 1** genera una sucesión infinita $\{\mathbf{x}_k\}$, entonces la secuencia $\{\mu_k\}$ es acotada.*

Demostración. Supongamos que se $\{\mu_k\}$ no es acotada. Luego, observemos que por (3.6) se infiere:

$$Pred_k \geq \|\mathbf{d}_k\|^2$$

Además, dado que se cumplen las hipótesis **H.1**, **H.2**, **H.3** se tiene que:

$$\begin{aligned}
|Ared_k - Pred_k| &= \left| \Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1}) + \nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k + \frac{1}{2} \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k \right| \\
&\leq \left| \Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1}) + \nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k \right| + \frac{1}{2} \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k \\
&\leq O(\|\mathbf{d}_k\|^2) + \|\mathbf{d}_k\|^2 \|H_k\|^2.
\end{aligned}$$

luego,

$$\begin{aligned}
|\tilde{\rho}_k - 1| &= \frac{|Ared_k - Pred_k|}{Pred_k} \\
&\leq \frac{O(\|\mathbf{d}_k\|^2) + \|\mathbf{d}_k\|^2 \|H_k\|^2}{\mu_k \|\mathbf{d}_k\|^2} \\
&= \frac{O(\|\mathbf{d}_k\|^2)}{\mu_k \|\mathbf{d}_k\|^2} + \frac{\|H_k\|^2}{\mu_k},
\end{aligned}$$

luego, cuando $k \rightarrow \infty$, se tiene que

$$\frac{O(\|\mathbf{d}_k\|^2)}{\mu_k \|\mathbf{d}_k\|^2} \rightarrow 0, \quad \frac{\|H_k\|^2}{\mu_k} \rightarrow 0,$$

de ahí que, $\tilde{\rho}_k \rightarrow 1$ cuando $k \rightarrow \infty$. Esto quiere decir que a partir de un $z \in \mathbb{Z}^+$, se cumple $\tilde{\rho}_k > \rho_2$, para todo $k \geq z$ luego por el paso 11 del **Algoritmo 1** se tiene:

$$\beta_k = \max \{ \beta_{k-1} \tau_2, m \}, \quad \text{para todo } k \geq z,$$

si definimos $C = \max \{ \beta_{k-1} \tau_2 \}$ para todo $s = 1, 2, \dots, k-1$, entonces existe $M > \max \{ C, m \}$ tal que:

$$\beta_k < M \quad \text{para todo } k. \tag{3.9}$$

Por otro lado, por el paso 12 del **Algoritmo 1** y desigualdad (3.9).

$$\begin{aligned}\mu_k &= \beta_k \|\Phi_\lambda(\mathbf{x}_k)\|^2 \\ &= \beta_k 2\Psi_\lambda(\mathbf{x}_k) \\ &\leq 2\Psi_\lambda(\mathbf{x}_0)M\end{aligned}$$

Luego, cuando $k \rightarrow \infty$ se tiene que $\mu_k \leq \tilde{M}$, donde $\tilde{M} = 2\Psi_\lambda(\mathbf{x}_0)M$, lo cual es una contradicción. Por tanto la sucesión $\{\mu_k\}$ es acotada. \square

La convergencia global de **Algoritmo 1**, el cual soluciona el problema (1.1), básicamente se resume en demostrar que todo punto de acumulación de la sucesión generada por dicho algoritmo, es un punto estacionario de la función de mérito Ψ_λ asociada a la reformulación Φ_λ en (2.5).

Teorema 3.1. *Supongamos que se cumple las hipótesis **H.1**, **H.2**, **H.3** y que el **Algoritmo 1** genera una sucesión infinita $\{\mathbf{x}_k\}$, entonces*

$$\lim_{k \rightarrow \infty} \|\nabla \Psi_\lambda(\mathbf{x}_k)\| = 0. \quad (3.10)$$

Demostración. Actuando por contradicción, supongamos que

$$\lim_{k \rightarrow \infty} \|\nabla \Psi_\lambda(\mathbf{x}_k)\| \neq 0,$$

por el **Lema 3.1**, se cumple la siguiente desigualdad:

$$\sum_{k=0}^n [\Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1})] \geq \sum_{k=0}^n \left[\frac{1}{2} \rho_1 \|\nabla \Psi_\lambda(\mathbf{x}_k)\| \min \left\{ \|\mathbf{d}_k\|, \frac{\|\nabla \Psi_\lambda(\mathbf{x}_k)\|}{\|H_k^T H_k\|} \right\} \right]$$

luego, cuando $n \rightarrow \infty$ tenemos que,

$$\sum_{k=0}^{\infty} [\Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1})] \geq \sum_{k=0}^{\infty} \left[\frac{1}{2} \rho_1 \|\nabla \Psi_\lambda(\mathbf{x}_k)\| \min \left\{ \|\mathbf{d}_k\|, \frac{\|\nabla \Psi_\lambda(\mathbf{x}_k)\|}{\|H_k^T H_k\|} \right\} \right]$$

Sea $S_k = \Psi_\lambda(\mathbf{x}_0) - \Psi_\lambda(\mathbf{x}_{k+1})$ entonces,

$$\begin{aligned} \sum_{j=0}^{\infty} [\Psi_\lambda(\mathbf{x}_j) - \Psi_\lambda(\mathbf{x}_{j+1})] &= \lim_{k \rightarrow \infty} \sum_{j=0}^k [\Psi_\lambda(\mathbf{x}_j) - \Psi_\lambda(\mathbf{x}_{k+1})] \\ &= \lim_{k \rightarrow \infty} S_k \\ &= \Psi_\lambda(\mathbf{x}_0) - \Psi_\lambda(\mathbf{x}) \end{aligned}$$

La última igualdad se cumple por el **Lema 3.4**, de ahí que la serie,

$$\sum_{k=0}^{\infty} [\Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_{k+1})],$$

es convergente, por tanto la serie

$$\sum_{k=0}^{\infty} \left[\frac{1}{2} \rho_1 \|\nabla \Psi_\lambda(\mathbf{x}_k)\| \min \left\{ \|\mathbf{d}_k\|, \frac{\|\nabla \Psi_\lambda(\mathbf{x}_k)\|}{\|H_k^T H_k\|} \right\} \right],$$

también es convergente, esto implica que

$$\lim_{k \rightarrow \infty} \left[\frac{1}{2} \rho_1 \|\nabla \Psi_\lambda(\mathbf{x}_k)\| \min \left\{ \|\mathbf{d}_k\|, \frac{\|\nabla \Psi_\lambda(\mathbf{x}_k)\|}{\|H_k^T H_k\|} \right\} \right] = 0,$$

esto ocurre si y solo si

$$\lim_{k \rightarrow \infty} \|\mathbf{d}_k\| = 0. \quad (3.11)$$

Por otro lado se tiene que,

$$[H_k^T H_k + \mu_k I] \mathbf{d}_k = -\nabla \Psi_\lambda(\mathbf{x}_k),$$

en consecuencia,

$$\begin{aligned}
\|\nabla\Psi_\lambda(\mathbf{x}_k)\| &= \|[H_k^T H_k + \mu_k I] \mathbf{d}_k\| \\
&\leq \|[H_k^T H_k + \mu_k I]\| \|\mathbf{d}_k\| \\
&\leq [\|H_k^T H_k\| + \mu_k] \|\mathbf{d}_k\| \\
&\leq (\nu_2 + \tilde{M}) \|\mathbf{d}_k\|.
\end{aligned}$$

Luego, cuando $k \rightarrow \infty$, se tiene que $\|\nabla\Psi_\lambda(\mathbf{x}_k)\| \rightarrow 0$, por (3.11), lo cual es una contradicción.

Por tanto se tiene que $\lim_{k \rightarrow \infty} \|\nabla\Psi_\lambda(\mathbf{x}_k)\| = 0$. \square

Corolario 3.1. *Todo punto de acumulación de la sucesión $\{\mathbf{x}_k\}$ generada por el **Algoritmo 1** es un punto estacionario de Ψ_λ .*

Demostración. Supongamos que \mathbf{x}^* es un punto de acumulación de la sucesión $\{\mathbf{x}_k\}$, además dado que $\zeta(\mathbf{x}_0)$ es compacto y $\mathbf{x}_k \subset \zeta(\mathbf{x}_0)$, existe una subsucesión $\{\mathbf{x}_{k_j}\}$ que converge a \mathbf{x}^* .

De otro lado, por el **Teorema (3.1)** y dado que $\nabla\Psi_\lambda$ es semisuave, se tiene que:

$$\begin{aligned}
\lim_{j \rightarrow \infty} \|\nabla\Psi_\lambda(\mathbf{x}_{k_j})\| = 0 &\iff \lim_{j \rightarrow \infty} \nabla\Psi_\lambda(\mathbf{x}_{k_j}) = 0 \\
&\iff \nabla\Psi_\lambda(\lim_{j \rightarrow \infty} \mathbf{x}_{k_j}) = 0 \\
&\iff \nabla\Psi_\lambda(\mathbf{x}^*) = 0,
\end{aligned}$$

es decir, \mathbf{x}^* es un punto estacionario de Ψ_λ . \square

El siguiente lema es útil para la demostración de la convergencia global del **Algoritmo 1** el cual resuelve el PCNL.

Lema 3.6. *Supongamos que se cumplen las hipótesis **H.1**, **H.2**, **H.3** y que el **Algoritmo 1** genera una sucesión infinita $\{\mathbf{d}_k\}$, entonces se cumple que:*

$$\lim_{k \rightarrow \infty} \|\mathbf{d}_k\| = 0. \quad (3.12)$$

Demostración. Partiendo de la ecuación (3.5) se tiene

$$|\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k| = \mathbf{d}_k H_k H_k \mathbf{d}_k + \mathbf{d}_k \mu_k \mathbf{d}_k$$

Luego, por **H.2** existe un $\nu > 0$ tal que:

$$\begin{aligned} |\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k| &\geq \nu \|\mathbf{d}_k\|^2 + \mu_k \|\mathbf{d}_k\|^2 \\ &= (\nu + \mu_k) \|\mathbf{d}_k\|^2 \end{aligned}$$

en consecuencia,

$$\begin{aligned} (\nu + \mu_k) \|\mathbf{d}_k\|^2 &\leq |\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k| \\ &\leq \|\nabla \Psi_\lambda(\mathbf{x}_k)\| \|\mathbf{d}_k\|, \end{aligned}$$

es decir,

$$\|\mathbf{d}_k\| \leq \frac{1}{(\nu + \mu_k)} \|\nabla \Psi_\lambda(\mathbf{x}_k)\|, \quad (3.13)$$

por último, usando el resultado del **Teorema 3.1** se concluye que

$$\lim_{k \rightarrow \infty} \|\mathbf{d}_k\| = 0.$$

□

La siguiente definición se incluye con el fin de dar claridad a un término aparentemente contradictorio en el **Teorema 3.2**.

Definición 3.2. *Sea Ω el conjunto de puntos de acumulación de la sucesión $\{\mathbf{x}_k\}$. Decimos que $\mathbf{x}^* \in \Omega$ es un punto de acumulación aislado, si*

\mathbf{x}^* es un punto aislado en Ω ; es decir, si existe $\delta > 0$ tal que

$$B(\mathbf{x}^*, \delta) \cap \Omega = \{\mathbf{x}^*\}.$$

Para finalizar la sección de la convergencia global del algoritmo propuesto en este documento, el siguiente teorema nos dice cuando un punto de acumulación es solución para el PCNL(F).

Teorema 3.2. *Si \mathbf{x}^* es un punto de acumulación aislado de la sucesión generada por el **Algoritmo 1**, entonces la sucesión converge a \mathbf{x}^* .*

Usando el resultado del **Lema 3.6**, la demostración del **Teorema 3.2** es de manera análoga a la presentada en el **Teorema 3.2** en [6].

3.3. Convergencia q-superlineal y q-cuadrática

En esta sección, estudiaremos el tipo de convergencia del **Algoritmo 1**, para ello agregamos las siguientes hipótesis :

- **H.4** Sea $\mathbf{x}^* \in \mathbb{R}^n$ una solución del sistema no lineal, $\Phi_\lambda(\mathbf{x}) = \mathbf{0}$.
- **H.5** Supongamos que \mathbf{x}^* es una solución BD-regular, entonces cualquier $H_k \in \partial_B \Phi_\lambda(\mathbf{x}^*)$ es no singular.

El lema siguiente nos indica que el **Algoritmo 1** a partir de un cierto entero, solo utiliza el método de región de confianza.

Lema 3.7. *Supongamos que se cumplen las hipótesis **H.1**, **H.2**, **H.3** y **H.4** , entonces existe un entero positivo K_1 , talque $\tilde{\rho}_k \geq \rho_2$ para todo $k \geq K_1$, donde, $\rho_k = Ared_k/Pred_k$ y el parámetro $\rho_2 \in (0, 1)$ esta dado en el **Algoritmo 1**.*

Demostración. Sea $\mathbf{d}_k \in \mathbb{R}^n$ tal que satisface (2.45) y definamos

$$Q_k(\mathbf{d}) = \frac{1}{2} \|\Phi_\lambda(\mathbf{x}_k) + H_k \mathbf{d}\|^2 + \frac{1}{2} \mu_k \|\mathbf{d}\|^2 \quad (3.14)$$

entonces, \mathbf{d}_k también es solución del problema

$$\begin{aligned} & \text{mín } Q_k(\mathbf{d}) \\ & \mathbf{d} \in \mathbb{R}^n \end{aligned}$$

Por el **Lema 3.5**, se tiene que $\{\mu_k\}$ es acotada, además sabemos que $\partial_B \Phi_\lambda(\mathbf{x}_k) \subset \partial_C \Phi_\lambda(\mathbf{x}_k)$.

De otro lado,

$$Q_k(\mathbf{x}^* - \mathbf{x}_k) = \frac{1}{2} \|\Phi_\lambda(\mathbf{x}^* - \mathbf{d}_k) + H_k[\mathbf{x}^* - \mathbf{x}_k]\|^2 + \frac{1}{2} \mu_k \|\mathbf{d}_k\|^2$$

en consecuencia,

$$\begin{aligned} Q_k(\mathbf{x}^* - \mathbf{x}_k) &\geq \frac{1}{2}\mu_k \|\mathbf{d}_k\|^2 \\ \frac{2Q_k(\mathbf{x}^* - \mathbf{x}_k)}{\mu_k} &\geq \|\mathbf{d}_k\|^2, \end{aligned}$$

por (3.14) se tiene que la anterior desigualdad es equivalente a

$$\begin{aligned} \|\mathbf{d}_k\|^2 &\leq \frac{2}{\mu_k} Q_k(\mathbf{x}^* - \mathbf{x}_k) \\ &= \frac{2}{\mu_k} \left[\frac{1}{2} \|\Phi_\lambda(\mathbf{x}_k) + H_k(\mathbf{x}^* - \mathbf{x}_k)\|^2 + \frac{1}{2} \mu_k \|\mathbf{x}^* - \mathbf{x}_k\|^2 \right] \\ &= \frac{1}{\mu_k} \|\Phi_\lambda(\mathbf{x}_k) + H_k(\mathbf{x}^* - \mathbf{x}_k)\|^2 + \|\mathbf{x}^* - \mathbf{x}_k\|^2 \end{aligned}$$

Luego, dado que Φ_λ es semisuave, por **Teorema 2.2** y $\Phi_\lambda(\mathbf{x}^*) = \mathbf{0}$ se cumple;

$$\begin{aligned} \|\mathbf{d}_k\|^2 &\leq \frac{1}{\mu_k} \|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}^*) - H_k(\mathbf{x}_k - \mathbf{x}^*)\|^2 + \|\mathbf{x}_k - \mathbf{x}^*\|^2 \\ &= \frac{o(\|\mathbf{x}_k - \mathbf{x}^*\|^2)}{\mu_k} + \|\mathbf{x}_k - \mathbf{x}^*\|^2 \end{aligned}$$

Por consiguiente, cuando $k \rightarrow \infty$, se tiene que:

$$\frac{o(\|\mathbf{x}_k - \mathbf{x}^*\|^2)}{\mu_k} \rightarrow 0 \quad \text{y} \quad \|\mathbf{x}_k - \mathbf{x}^*\|^2 \rightarrow 0$$

es decir,

$$\|\mathbf{d}_k\|^2 \rightarrow 0, \tag{3.15}$$

de otro lado,

$$Ared_k - \rho_2 Pred_k = \Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_k + \mathbf{d}) - \rho_2 Pred_k,$$

Usando el polinomio de Taylor de segundo orden, de la función Ψ_λ al

rededor del punto \mathbf{x}_k y si definimos $Ared_k - \rho_2 Pred_k = R$, se tiene que:

$$\begin{aligned}
R &= \Psi_\lambda(\mathbf{x}_k) - \Psi_\lambda(\mathbf{x}_k + \mathbf{d}_k) - \rho_2 Pred_k \\
&= -\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k - \frac{1}{2} \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k + o(\|\mathbf{d}_k\|^2) - \rho_2 Pred_k \\
&= Pred_k - \rho_2 Pred_k + o(\|\mathbf{d}_k\|^2) \\
&= (1 - \rho_2) Pred_k + o(\|\mathbf{d}_k\|^2) \\
&= \frac{(1 - \rho_2)}{2} 2 Pred_k + o(\|\mathbf{d}_k\|^2)
\end{aligned}$$

luego,

$$\begin{aligned}
R &= \frac{(1 - \rho_2)}{2} [-2\nabla \Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k - \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k] + o(\|\mathbf{d}_k\|^2) \\
&= \frac{(1 - \rho_2)}{2} [2\mathbf{d}_k^T H_k^T H_k \mathbf{d}_k + 2\mu_k \|\mathbf{d}_k\|^2 - \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k] + o(\|\mathbf{d}_k\|^2) \\
&= \frac{(1 - \rho_2)}{2} [\mathbf{d}_k^T H_k^T H_k \mathbf{d}_k + 2\mu_k \|\mathbf{d}_k\|^2] + o(\|\mathbf{d}_k\|^2) \\
&\geq \frac{(1 - \rho_2)}{2} [\nu_1 \|\mathbf{d}_k\|^2 + 2\mu_k \|\mathbf{d}_k\|^2] + o(\|\mathbf{d}_k\|^2) \\
&= \|\mathbf{d}_k\|^2 \left[\frac{(1 - \rho_2)\nu}{2} + (1 - \rho_2)\mu_k + \frac{o(\|\mathbf{d}_k\|^2)}{\|\mathbf{d}_k\|^2} \right]
\end{aligned}$$

en consecuencia,

$$Ared_k - \rho_2 Pred_k \geq \|\mathbf{d}_k\|^2 \left[\frac{(1 - \rho_2)\nu}{2} + (1 - \rho_2)\mu_k + \frac{o(\|\mathbf{d}_k\|^2)}{\|\mathbf{d}_k\|^2} \right]$$

Luego, para $K_1 \in \mathbb{Z}^+$ lo suficientemente grande, se tiene que

$$\|\mathbf{d}_k\|^2 \left[\frac{(1 - \rho_2)\nu}{2} + (1 - \rho_2)\mu_k + \frac{o(\|\mathbf{d}_k\|^2)}{\|\mathbf{d}_k\|^2} \right] \longrightarrow 0,$$

es decir, existe un entero positivo K_1 , tal que para todo $k \geq K_1$ se cumple

que:

$$Ared_k \geq \rho_2 Pred_k,$$

lo que es equivalente,

$$\frac{Ared_k}{Pred_k} \geq \rho_2.$$

Lo que indica que, $\rho_k \geq \rho_2$ para todo $k \geq K_1$. \square

A continuación, se presenta el **Lema de Banach**, el cual ya esta demostrado en [16].

Lema 3.8. Sean $A, E \in \mathbb{R}^{n \times n}$ y A un matriz no singular tal que $\|A^{-1}E\| < \sigma < 1$, entonces $A + E$ es no singular y satisface la siguiente desigualdad

$$\|(A + E)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}E\|} \quad (3.16)$$

Para cualquier \mathbf{x}_k , definimos

$$B_k = H_k^T H_k + \mu_k I, \quad (3.17)$$

el siguiente lema, nos indica que las matrices definidas B_k^{-1} están acotadas a partir de un cierto número entero.

Lema 3.9. Supongamos que se cumplen las hipótesis **H.1**, **H.2**, **H.3**, **H.4** y **H.5**, entonces existe un entero positivo K y una constante $\nu > 0$ tal que

$$\|B_k^{-1}\| \leq 2\nu^2, \quad \text{para todo } k \geq K. \quad (3.18)$$

Demostración. Por el **Lema 2.6**, entonces existe una vecindad $B(\mathbf{x}^*, r)$ y una constante $\nu > 0$ tal que

$$\|H^{-1}\| \leq \nu \quad \forall H \in \partial_B \Phi_\lambda(\mathbf{x}), \quad \forall \mathbf{x} \in B(\mathbf{x}^*, r).$$

Por otro lado, por el **Lema 3.7** y el paso 11 del **Algoritmo 1**, sabemos

que existe un entero positivo K_1 y una constante $M > m > 0$ tal que

$$\begin{aligned}\mu_k &= \beta_k \|\Phi_\lambda(\mathbf{x}_k)\|^2 \\ &\leq M \|\Phi_\lambda(\mathbf{x}_k)\|^2 \longrightarrow 0 \quad \text{para todo } k \geq K_1\end{aligned}$$

Esto implica que, para todo $\epsilon > 0$, existe $K_2 \in \mathbb{Z}^+$ tal que:

$$\mu_k \leq M \|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}^*)\|^2 < \epsilon, \quad \text{para todo } k > K_2$$

en particular para $\epsilon = \frac{1}{2M\nu^2}$,

$$\mu_k \leq \frac{1}{2\nu^2}, \quad \text{para todo } k > K_2,$$

dado que $\{\mathbf{x}_k\}$ converge a \mathbf{x}^* , entonces existe un entero positivo K_3 tal que:

$$\mathbf{x}_k \in B(\mathbf{x}^*, r), \quad \text{para todo } k > K_3$$

luego, si definimos $K = \max\{K_1, K_2, K_3\}$ entonces tenemos que

$$\begin{aligned}\left\| \mu_k [H_k^T H_k]^{-1} \right\| &= \mu_k \left\| [H_k^T H_k]^{-1} \right\| \\ &\leq \frac{1}{2\nu^2} (\nu^2) \\ &= \frac{1}{2} \quad \text{para todo } k \geq K,\end{aligned}$$

luego, por el **Lema3.8**, se tiene que:

$$\begin{aligned}
\|B_k^{-1}\| &= \left\| (H_k^T H_k + \mu_k I)^{-1} \right\| \\
&\leq \frac{\left\| (H_k^T H_k)^{-1} \right\|}{1 - \left\| [H_k^T H_k]^{-1} \mu_k I \right\|} \\
&\leq \frac{\nu^2}{1 - \frac{1}{2}} \\
&= 2\nu^2
\end{aligned}$$

Por tanto,

$$\|B_k^{-1}\| \leq 2\nu^2 \quad \text{para todo } k > K.$$

□

Para la demostración de la convergencia q-superlineal y q-cuadrática vamos a suponer la siguiente hipótesis:

- **H.6** $F'(\mathbf{x})$ Lipschitz continua en \mathbb{R}^n .

Teorema 3.3. *Supongamos que se cumplen las hipótesis **H.1**, **H.2**, **H.3**, **H.4** y **H.5**, sea $\{\mathbf{x}_k\}$ una sucesión infinita generada por el **Algoritmo 1**, tal que $\{\mathbf{x}_k\}$ converge a \mathbf{x}^* , entonces:*

a) $\{\mathbf{x}_k\}$ converge a \mathbf{x}^* superlinealmente, esto es:

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

b) $\{\mathbf{x}_k\}$ converge a \mathbf{x}^* cuadráticamente si se cumple **H.6**, esto es:

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| = o(\|\mathbf{x}_k - \mathbf{x}^*\|^2)$$

Demostración. Primero vamos a demostrar el inciso:

a) Si se asume como verdaderas las hipótesis **H.1**, **H.2**, **H.3**, **H.4** y **H.5** se tiene que por el **Lema 3.9** existe un entero positivo K y una constante $\nu > 0$ tal que:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k, \quad \text{y} \quad \|B_k^{-1}\| \leq 2\nu^2, \quad \text{para todo } k > K.$$

Por otro lado, por la hipótesis **H.2** y (3.9), existen constantes $C > m > 0$ y L_1 constante *Lipschitz* tal que:

$$\begin{aligned} \mu_k &= \beta_k \|\Phi_\lambda(\mathbf{x}_k)\|^2 \\ &\leq C \|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}^*)\|^2 \\ &\leq CL_1^2 \|\mathbf{x}_k - \mathbf{x}^*\|^2 \quad \text{para todo } k > K. \end{aligned}$$

Luego, por la ecuación (2.45) y (3.8) se deduce que para todo $k > K$.

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

Luego, por (3.17) tenemos

$$\begin{aligned} \|\mathbf{x}_{k+1} - \mathbf{x}^*\| &\leq \|B_k^{-1}\| \|H_k^T H_k (\mathbf{x}_k - \mathbf{x}^*) + \mu_k (\mathbf{x}_k - \mathbf{x}^*) - H_k^T \Phi_\lambda(\mathbf{x}_k)\| \\ &= \|B_k^{-1}\| \|H_k^T [H_k (\mathbf{x}_k - \mathbf{x}^*) - \Phi_\lambda(\mathbf{x}_k)] + \mu_k (\mathbf{x}_k - \mathbf{x}^*)\| \\ &\leq \|B_k^{-1}\| \{ \|H_k^T\| \|H_k (\mathbf{x}_k - \mathbf{x}^*) - \Phi_\lambda(\mathbf{x}_k)\| + \mu_k \|\mathbf{x}_k - \mathbf{x}^*\| \}, \end{aligned}$$

además por (3.8) y la desigualdad triangular se cumple,

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq 2\nu^2 \{ L \|H_k (\mathbf{x}_k - \mathbf{x}^*) - \Phi_\lambda(\mathbf{x}_k) + \Phi_\lambda(\mathbf{x}^*)\| + 2\nu^2 \mu_k \|\mathbf{x}_k - \mathbf{x}^*\| \}$$

en consecuencia,

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq 2\nu^2 L \|\Phi_\lambda(\mathbf{x}_k) - \Phi_\lambda(\mathbf{x}^*) - H_k (\mathbf{x}_k - \mathbf{x}^*)\| + 2C\nu^2 L_1^2 \|\mathbf{x}_k - \mathbf{x}^*\|^3.$$

Como, $\partial_B \Phi_\lambda(\mathbf{x}^*) \subset \partial_C \Phi_\lambda(\mathbf{x}^*)$, por la **Teorema 2.2**, se tiene:

$$\begin{aligned}\|\mathbf{x}_{k+1} - \mathbf{x}^*\| &= o(\|\mathbf{x}_k - \mathbf{x}^*\|) + O(\|\mathbf{x}_k - \mathbf{x}^*\|^3) \\ &= o(\|\mathbf{x}_k - \mathbf{x}^*\|),\end{aligned}$$

esto indica que $\{\mathbf{x}_k\}$ converge a la solución de (2.5) superlinealmente.

b) Para el caso en que $F'(\mathbf{x}_k)$ *Lipschitz continua* en \mathbb{R}^n , por la **Teorema 2.2** se tiene que:

$$\begin{aligned}\|\mathbf{x}_{k+1} - \mathbf{x}^*\| &= O(\|\mathbf{x}_k - \mathbf{x}^*\|^2) + O(\|\mathbf{x}_k - \mathbf{x}^*\|^3) \\ &= O(\|\mathbf{x}_k - \mathbf{x}^*\|^2),\end{aligned}$$

esto indica que $\{\mathbf{x}_k\}$ converge a la solución de (2.5) cuadráticamente.

□

Resultados numéricos

En este capítulo, iniciamos analizando numéricamente el **Algoritmo 1** propuesto en el capítulo anterior, para la selección de λ se utiliza dos técnicas; una es dejar $\lambda = 2$ fijo, y la otra es dejar λ dinámico como en [38, 6], está última la describimos en el **Algoritmo 2**. Para tal fin, vamos a comparar su desempeño numérico con otros cuatro algoritmos los cuales los llamaremos de la siguiente manera:

LM: Algoritmo de región de confianza usando el método de *Levenberg-Marquardt* como en la Figura 2.11.

LMHIB: Algoritmo híbrido de *Levenberg-Marquardt*, que combina región de confianza y búsqueda lineal (**Algoritmo 1**).

LMHDIN: Algoritmo híbrido de *Levenberg-Marquardt*, que combina región de confianza y búsqueda lineal con λ dinámico.

NG: Algoritmo Newton Global [33].

NGDIN: Algoritmo Newton Global con λ dinámico [38].

Algoritmo 2 Parámetro λ dinámico

Entrada: Sean $\lambda_0 = 2$, $\gamma_1 = 10^{-2}$, $\gamma_2 = 10^{-4}$, $c_1 = 10$, $c_2 = 10^{-8}$.

si $\Psi_\lambda(\mathbf{x}_k) \leq \gamma_1$ **entonces**

$\lambda \leftarrow \Psi_\lambda(\mathbf{x}_k)$

si no

$\lambda \leftarrow \min \{c_1 \Psi(\mathbf{x}_k), \lambda\}$

fin si

si $\Psi_\lambda(\mathbf{x}_k) \leq \gamma_2$ **entonces**

$\lambda \leftarrow \min \{c_2, \lambda\}$

fin si

Salida: λ

Los códigos para los algoritmos y funciones de prueba fueron desarrollados en el *software* MATLAB[®]. Los experimentos numéricos se realizaron en un computador con procesador Intel(R) Core(TM) de 2.50 GHz. Los parámetros utilizados en los algoritmos son los siguientes:

Región de confianza	$\rho_1 = 0.25$, $\rho_2 = 0.75$, $\tau_1 = 0.1$, $\tau_2 = 10$, $\Delta_{min} = 10^{-5}$, $\beta_0 = \mu_0 = 10^{-4}$, $\lambda = 2$
Búsqueda lineal	$\lambda = 2$, $t = 0.5$, $\delta = 0.25$
Para λ dinámico	$\lambda_0 = 2$

Se resolvieron 100 experimentos con distintos puntos iniciales, donde dichos puntos son vectores aleatorios tomados de la $N - esfera$ ¹, de radio r y centro en la solución del problema \mathbf{x}^* esto es, $N(\mathbf{x}^*; r)$.

Para las pruebas numéricas, consideramos 9 problemas de complementariedad no lineal y que son considerados problemas “difíciles”, en lo que

¹Definimos $N - esfera$ centrada en $\mathbf{z} \in \mathbb{R}^N$ y radio $r > 0$ como

$$N(\mathbf{z}, r) =: \{\mathbf{x} : \|\mathbf{z} - \mathbf{x}\| = r\},$$

el cual es el conjunto de todos los puntos $\mathbf{x} \in \mathbb{R}^N$ tales que están a una distancia r de \mathbf{z}

respecta a convergencia. Estos problemas están asociados respectivamente a las siguientes funciones:

Problema 4.1. *Kojima Shindo* (*Koj-Shi*) [52] *dimensión* $n = 4$. Sea $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definido por,

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

Este problema tiene una solución degenerada $\mathbf{x}_* = \left(\frac{\sqrt{6}}{2} \ 0 \ 0 \ \frac{1}{2}\right)^T$ y una solución no degenerada $\mathbf{x}^* = (1 \ 0 \ 3 \ 0)^T$.

Problema 4.2. *Kojima Josephy* (*Koj-Jo*)[32] *dimensión* $n = 4$.

Sea $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, definido por,

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

Este problema tiene una única solución $\mathbf{x}^* = \left(\frac{\sqrt{6}}{2} \ 0 \ 0 \ \frac{1}{2}\right)^T$.

Problema 4.3. *Mathiesen-Modificado* (*Math mod*)[47] *dimensión* $n = 4$.

Sea $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definido por,

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

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

Problema 4.4. Mathiesen (Mathiesen)[52] *dimensión* $n = 4$.

Sea $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definido por

$$F(\mathbf{x}) = \begin{pmatrix} -x_2 + x_3 + x_4 \\ x_1 - \alpha \left(\frac{b_1x_3 + b_2x_4}{x_2} \right) \\ b_1 - x_1 - (1 - \alpha) \left(\frac{b_1x_3 + b_2x_4}{x_3} \right) \\ b_2 - x_1 \end{pmatrix}.$$

Para encontrar una solución con $x_1 > 0$, $x_2 > 0$ y $x_2 + x_3 + x_4 = 1$, donde $b_1, b_2 > 0$ y $\alpha \in (0, 1)$ son constantes asociadas a esta función, se tiene que la solución para el problema de complementariedad no lineal es:

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

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

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

b) 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 4.5. Billups (*Billups*) [7] *dimensión* $n = 1$.

Sea $F: \mathbb{R} \rightarrow \mathbb{R}$ definido por,

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

Solución: $\mathbf{x}^* = 1 + \sqrt{1.1}$

Problema 4.6. Nash-Cournot (*Nash Co*) [32] *dimensión* $n = 5$ y $n = 10$.

Sea $F: \mathbb{R}_+^n \rightarrow \mathbb{R}^n$ 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

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

Los valores c_i, l_i, b_i son positivos y $\gamma > 0$. lo anterior se puede escribir

en forma compacta como sigue:

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

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

Para este problema, se harán pruebas numéricas, con $n = 5$ y $n = 10$. Los parámetros son tomados de [32, 20], es decir:

Para $n = 5$, el problema *Nash-Cournot*,

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

la solución a este problemas es $\mathbf{x}^* = (15.429 \ 12.499 \ 9.6635 \ 7.1651 \ 5.1326)^T$.

Para $n = 10$, el problema *Nash-Cournot*,

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

y $\gamma = 1.2$, la solución asociado a este problema es:

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

Problema 4.7. Geiger y Kanzow (*Geiger-Kanzow*) [29], *dimensión* $n = 10$ y $n = 256$.

Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por

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

donde

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

Para $n = 10$, se tiene que la solución para el problema de complementariedad lineal anterior es:

$$\mathbf{x}^* = (0.3660 \ 0.4641 \ 0.4904 \ 0.4974 \ \dots \ 0.4974 \ 0.4904 \ 0.4641 \ 0.3660)^T.$$

Para el caso en que $n = 256$, la solución es:

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

Problema 4.8. Ahn (*Ahn*) [9]. $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por,

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

donde

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

Para $n = 10$, se tiene que la solución para el problema de complementariedad lineal anterior es:

$$\mathbf{x}^* = (0.4081 \quad 0.3162 \quad 0.3366 \quad 0.3312 \cdots \\ 0.3308 \quad 0.3272 \quad 0.3197 \quad 0.3031 \quad 0.2660 \quad 0.1835)^T.$$

Para el caso en que $n = 100$, la solución es:

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

Problema 4.9. Función casi Lineal de Brown (Brown) [60] dimensión $n = 100, 200, 400, 600$ y $n = 1000$.

Sea $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por,

$$F_i(\mathbf{x}) = \begin{cases} f_i(\mathbf{x}) - f_i(\mathbf{x}^*) + 1, & \text{si } i \text{ es impar,} \\ f_i(\mathbf{x}) - f_i(\mathbf{x}^*), & \text{en otro caso.} \end{cases}$$

donde,

$\mathbf{x}^* = (0 \ 1 \ 0 \ 1 \ \dots)^T$, y la función vectorial $f = (f_1, f_2, \dots, f_n)$ está definida

como sigue:

$$f_i(\mathbf{x}) = \begin{cases} -(n+1) + \mathbf{x}_i + \sum_{j=1}^n \mathbf{x}_j, & i = 1, 2, 3, \dots, n-1, \\ -1 + \prod_{j=1}^n \mathbf{x}_j, & i = n. \end{cases}$$

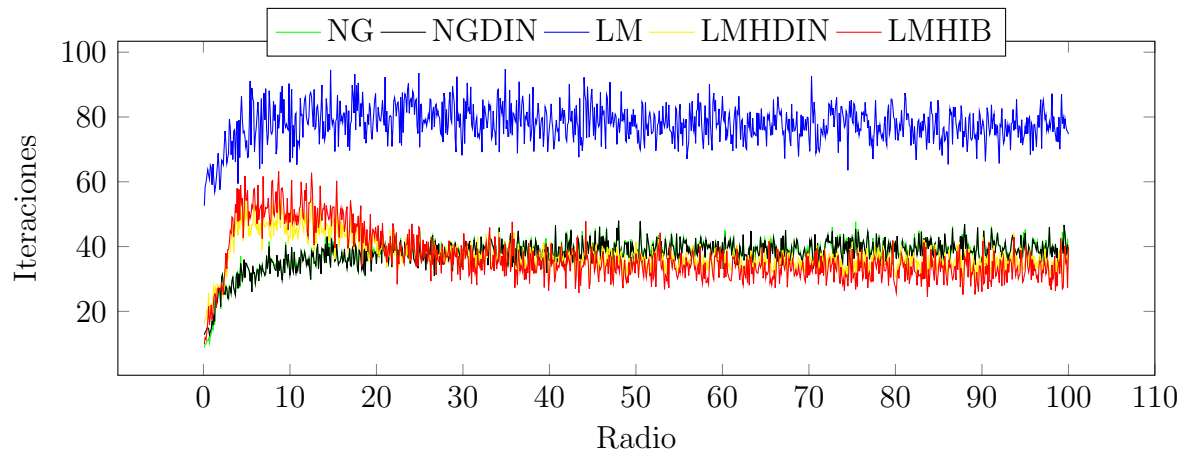
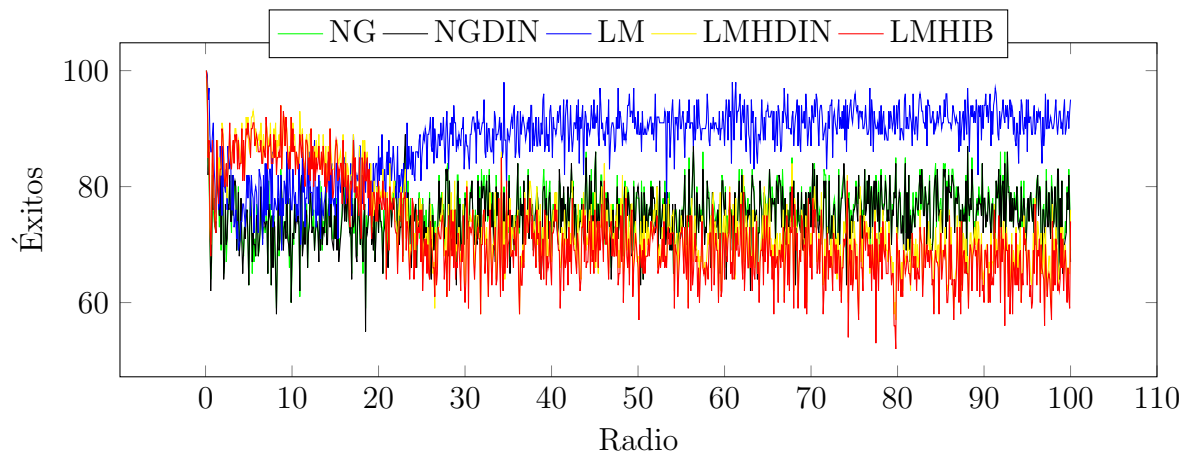
Se tiene que el vector \mathbf{x}^* es una solución no degenerada para el problema.

Para dar mayor claridad a los resultados numéricos obtenidos, declaramos convergencia si $\|\Psi(\mathbf{x}_k)\|_2 \leq 10^{-6}$ y divergencia si el número de iteraciones excedió las 200. Además, cuando haya convergencia se dice que hubo éxito; en esos casos medimos el tiempo y número de iteraciones.

Para la implementación de los algoritmos, es importante mencionar que para cada radio $r = 0.1, 0.2, \dots, 100$, se resolvieron 100 experimentos. Es decir, se realizaron 100.000 experimentos por cada problema.

Las gráficas siguientes son el resultado de nuestros experimentos, en los cuales estudiaremos el promedio de: número de éxitos, número de iteraciones y tiempo en el que el algoritmo determina la solución del problema (tiempo de ejecución). Así como también, diremos que un algoritmo es efectivo cuando alcanza un porcentaje alto de éxitos en un número mínimo de iteraciones, y eficiente cuando determina la solución en un tiempo relativamente bajo.

Es por ello que analizaremos en conjunto: el número de éxitos respecto al número de iteraciones y el número de éxitos con el tiempo en el que tarda cada método en encontrar la solución del problema (tiempo de ejecución).

Figura 4.1: Problema Khojima-Shindo, $n=4$, $\exp=100$ Figura 4.2: Problema Khojima-Shindo, $n=4$, $\exp=100$

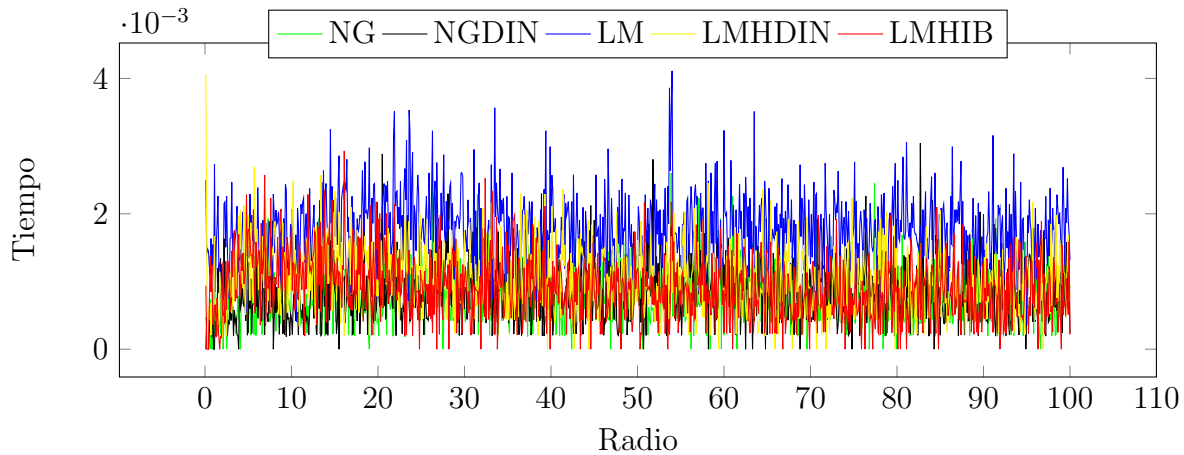
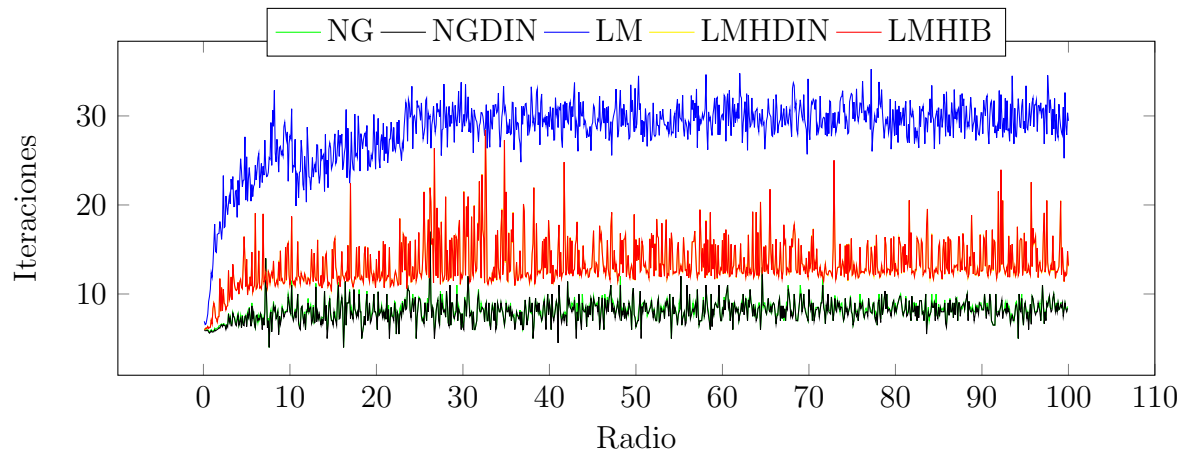
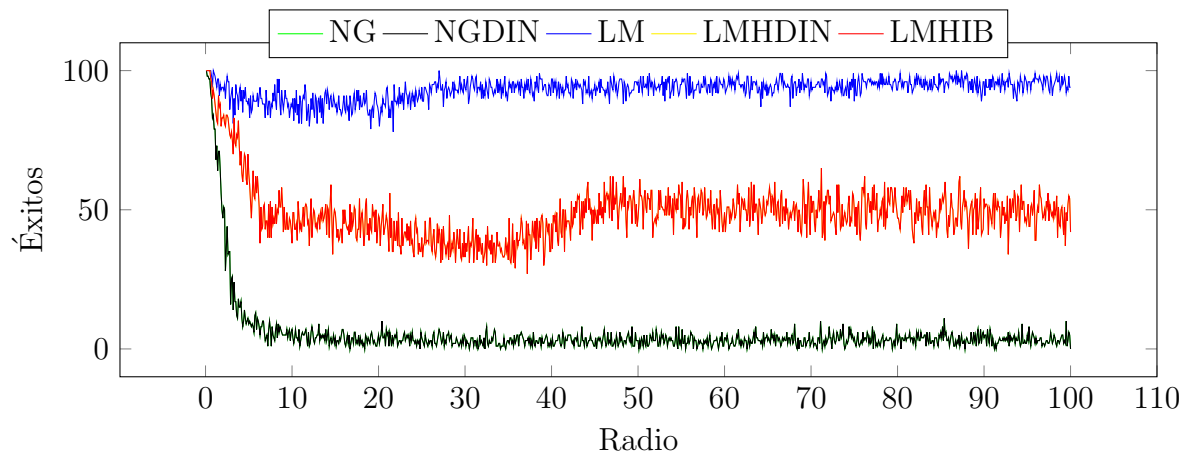


Figura 4.3: Problema Khojima-Shindo, $n=4$, $\text{exp}=100$

De las Figuras 4.1 y 4.2 podemos observar el comportamiento de los algoritmos **NG**, **NGDIN**, **LMHIB** y **LMHDIN**, donde su porcentaje de éxitos supera el 60 %, y el promedio de iteraciones oscila entre 30 y 45 iteraciones a medida que el radio aumenta. Además, a pesar de que el método **LM** tiene un porcentaje alto de iteraciones, se observa que a medida que el radio aumenta el promedio de éxitos se mantiene por encima de 80 %, superando al resto de métodos. Es importante mencionar que el parámetro de λ dinámico no aporta mucho, pues no refleja menor número iteraciones ni mayor número de éxitos, para este problema. Para finalizar, nuestro algoritmo **LMHIB** a pesar de que el número de iteraciones disminuye, el porcentaje de éxitos se mantiene en un 60 % y 70 %, es decir por debajo del resto de métodos.

En cuanto al tiempo de ejecución, analizando la figura 4.3, nos permite concluir lo eficiente de los métodos **NG** y **NGDIN**, sin descartar el método **LM**, pues el tiempo de convergencia es pequeño. Respecto al algoritmo **LMHIB** se puede ver lo eficiente que es cuando este encuentra la solución.

Figura 4.4: Problema Khojima-Josephy, $n=4$, $\exp=100$ Figura 4.5: Problema Khojima-Josephy, $n=4$, $\exp=100$

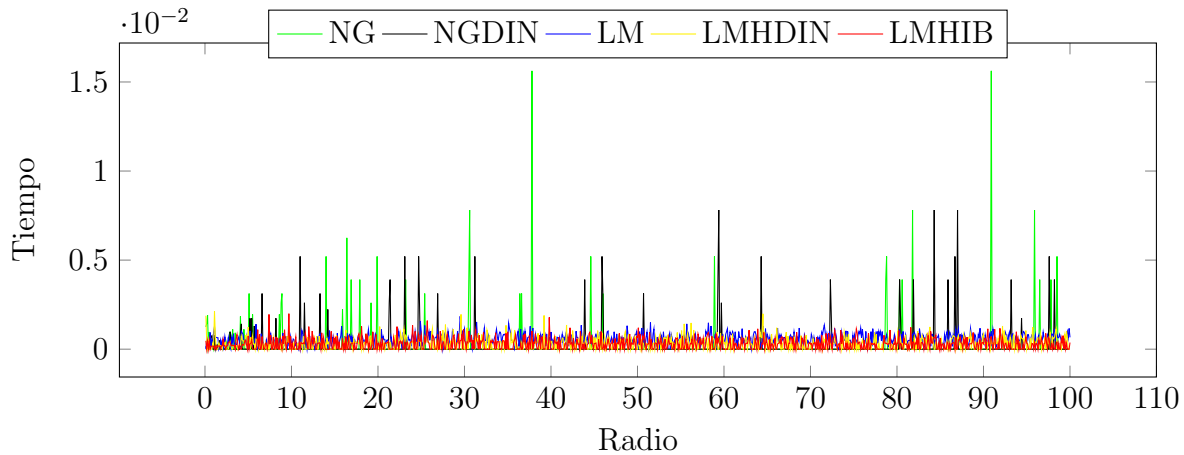


Figura 4.6: Problema Khojima-Josephy, $n=4$, $\text{exp}=100$

En las Figuras 4.4 y 4.5, se ve como los algoritmos **NG** y **NGDIN**, tiene problema de convergencia, esto debido a la singularidad de las matrices H_k . Además, a pesar que el algoritmos **LM** tiene un promedio de iteraciones entre 20 y 30, el alto porcentaje de convergencia, lo hace efectivo respecto al resto de algoritmos. Por otro lado, se tiene que los algoritmos **LMHIB** y **LMHDIN** para radios entre 5 y 30, el promedio de éxitos no supera el 55 por ciento y para radios mayores a 40, se tiene que el promedio de éxitos se mantiene entre 45 y 55 por ciento.

En cuanto al tiempo en determinar la solución, se observa como el algoritmo **LM** es eficiente. Los métodos **LMHIB** y **LMHDIN** convergen, lo hacen más rápido que el algoritmo **LM**.

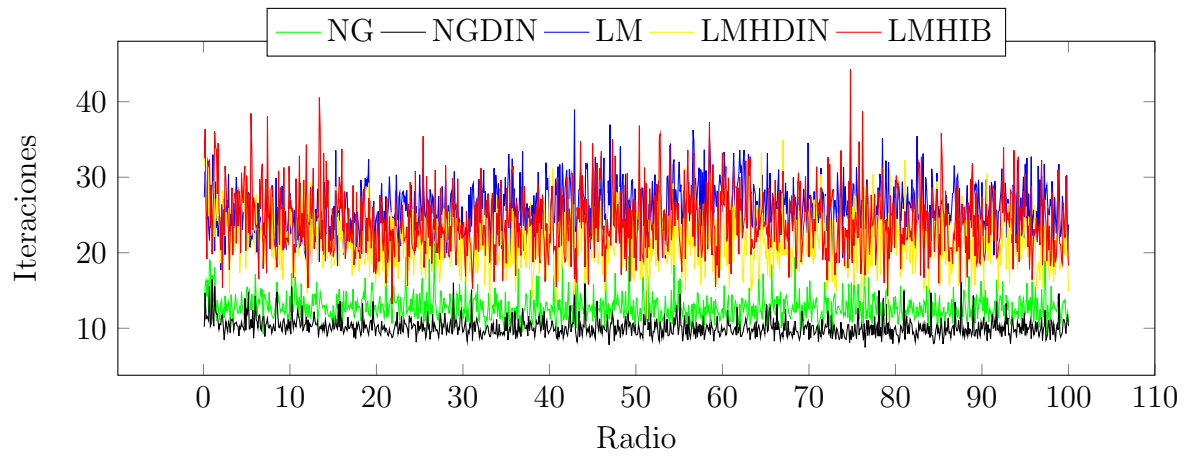


Figura 4.7: Problema Mathiesen, n=4, exp=100

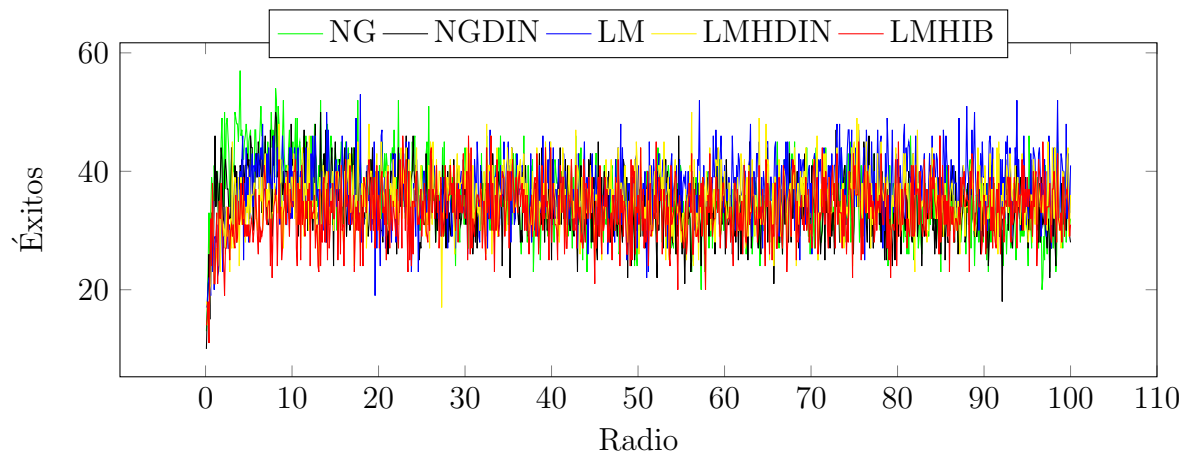


Figura 4.8: Problema Mathiesen, n=4, exp=100

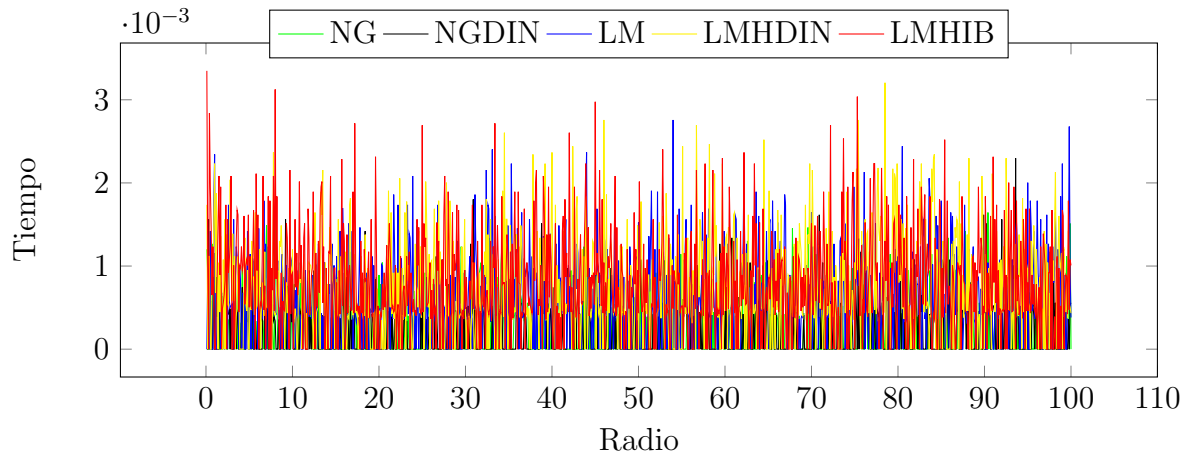
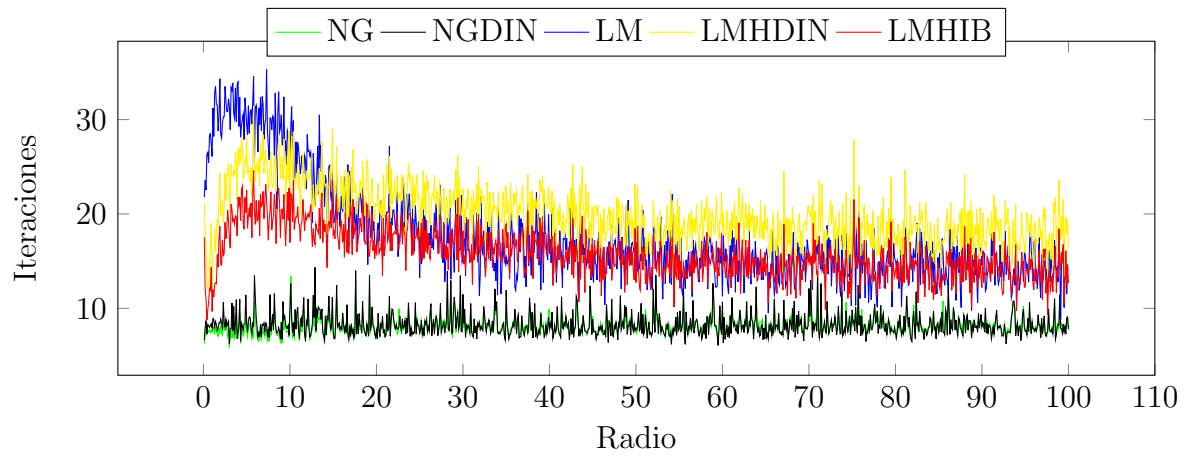
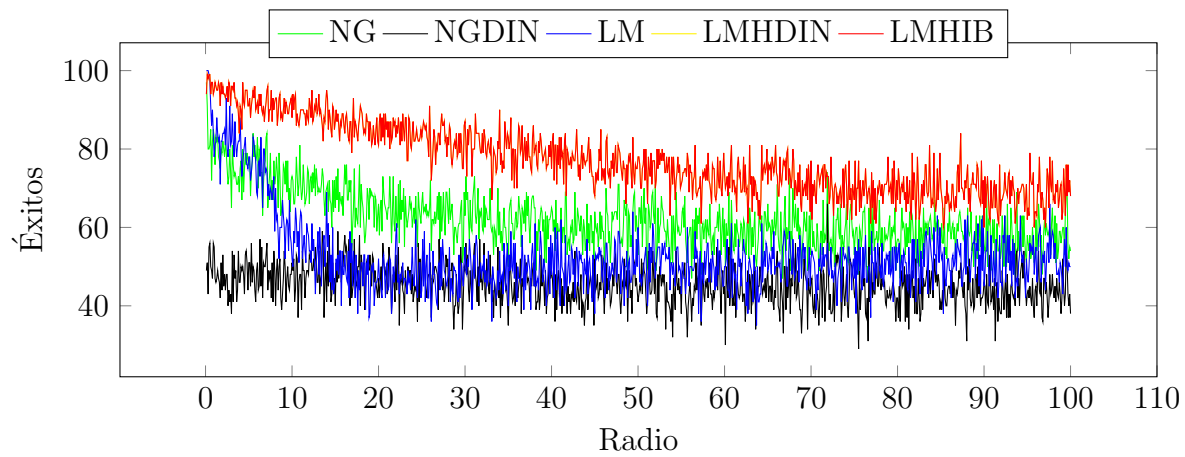


Figura 4.9: Problema Mathiesen, $n=4$, $\exp=100$

De las Figuras 4.7 y 4.8, queremos resaltar que los algoritmos no superan el 60 % de éxitos. Además, los algoritmos **LMHIB** y **LMHDIN** tienen un comportamiento similar en cuanto al número de éxitos respecto al número de iteraciones; es menor el algoritmo **LMHDIN**. Por otro lado, los algoritmos **NG** y **NGDIN** son los que mejor se posicionan en las gráficas, pues el promedio de iteraciones oscila entre 10 y 15 iteraciones, respecto al número de éxitos, se mantiene por encima del 35 %, para todos los radios.

Queremos aclarar que durante los experimentos numéricos, para diversos radios, la matriz H resultó mal condicionada, para dichos casos se perturbó la matriz.

Por último, el tiempo de convergencia: en la Figura 4.9 se puede observar que no hay una diferencia relevante, pero quienes mejores se comportan son los algoritmos que solo usan búsqueda lineal, es decir **NG** y **NGDIN**.

Figura 4.10: Problema Mathiesen Modificado, $n=4$, $\exp=100$ Figura 4.11: Problema Mathiesen Modificado, $n=4$, $\exp=100$

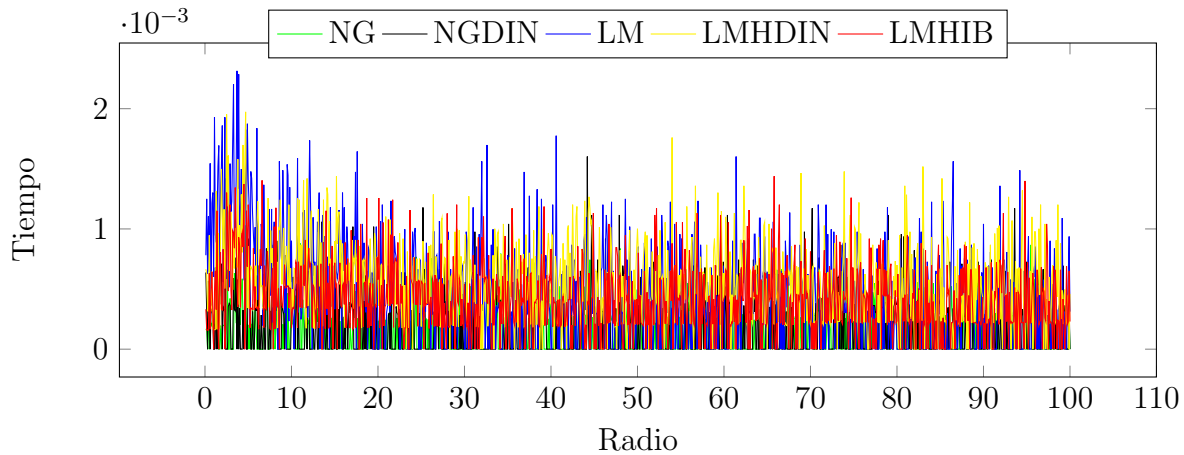
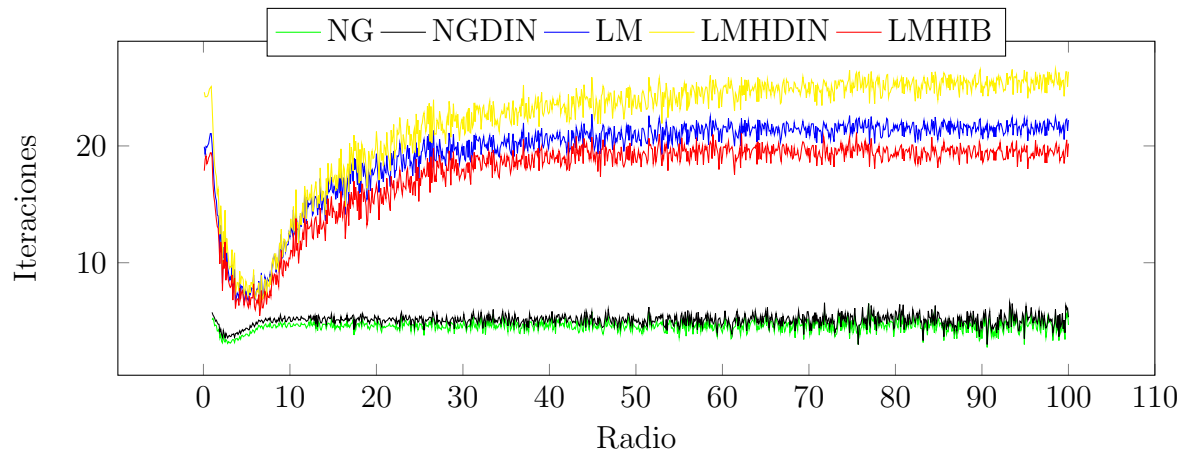
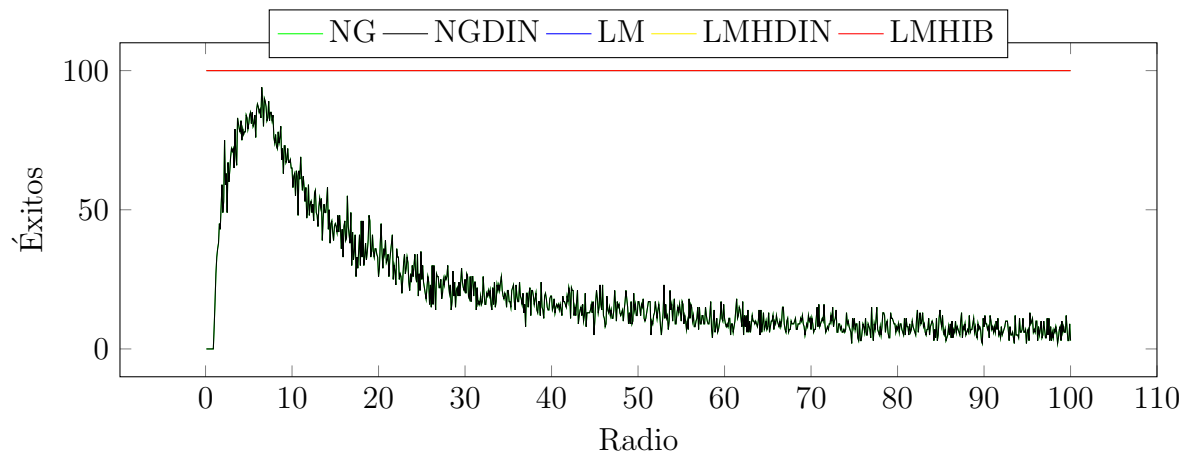


Figura 4.12: Problema Mathiesen Modificado, $n=4$, $\exp=100$

De la Figura 4.11, podemos observar que el número de éxitos de los algoritmos, disminuye a medida que el radio aumenta. El porcentaje de éxitos del algoritmo que solo usa región de confianza (**LM**) no supera el 60%, así como también se puede ver que el algoritmo **LMHIB** y **LMHDIN** son los que mejor se comportan, respecto al número de éxitos, de otro lado se puede observar que en la Figura 4.10 y 4.11, como el parámetro λ dinámico hace que los algoritmos desmejoren en cuanto al porcentaje de número de éxitos y número de iteraciones.

Para finalizar, podemos observar en este problema, que los métodos **LMHIB** y **LMHDIN** tiene un mejor comportamiento que los métodos **NG**, **NGDIN** y **LM**; esto, sin desmeritar el hecho que los métodos **NG** y **NGDIN** cuando hallan la solución lo hacen con pocas iteraciones.

Respecto al tiempo de convergencia, como se observa en la Figura 4.12, no hay una diferencia relevante, pero destacamos los métodos **LMHIB** y **LMHDIN** por su porcentaje de éxitos.

Figura 4.13: Problema Billups, $n=1$, $\exp=100$ Figura 4.14: Problema Billups, $n=1$, $\exp=100$

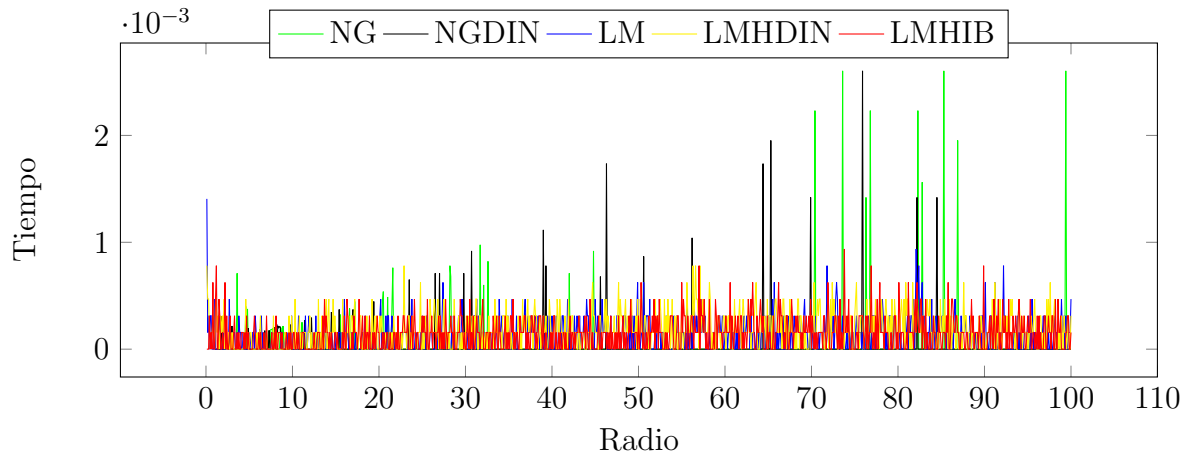
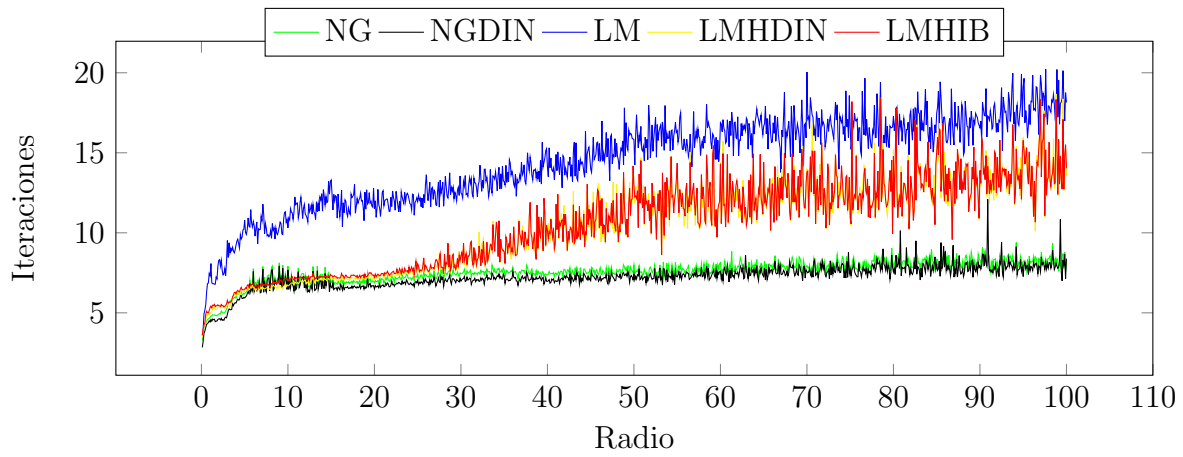
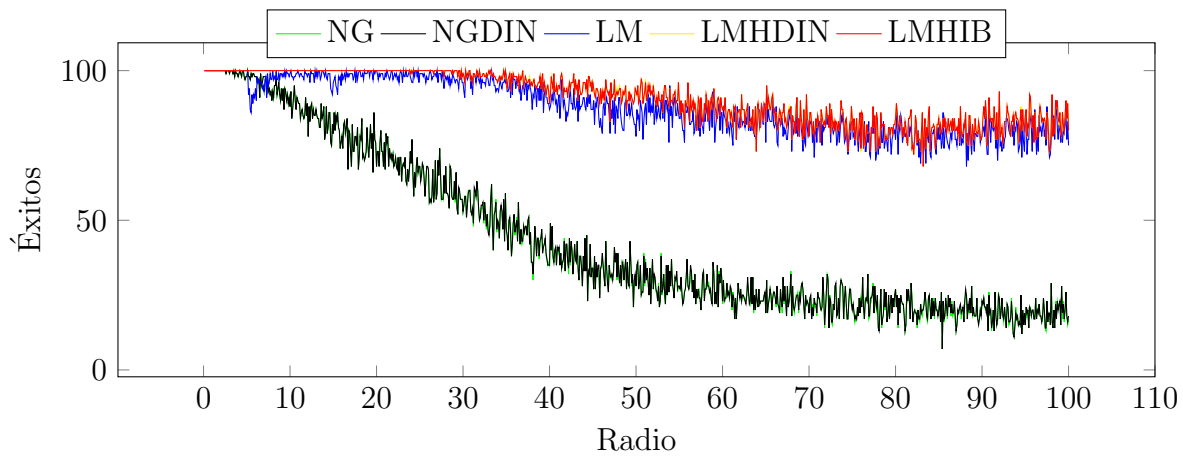


Figura 4.15: Problema Billups, $n=1$, $\text{exp}=100$

La Figura 4.14 nos permite observar que los algoritmos **NG** y **NGDIN** tiene un bajo porcentaje de éxitos, en cuanto a los otros tres métodos, se puede observar que tiene un 100% de éxito, por tanto para determinar que método es mejor para este problema, podemos considera la Figuras 4.13 y 4.15: primero se observa como el algoritmo de **LMHIB** es quien tiene menor número de iteraciones. Así como también, sigue siendo el algoritmo que tarda menos tiempo en determinar la solución al problema, esto respecto a los otros dos algoritmos restantes. Sin embargo, no implica que los otros sean malos, pues gráficamente se puede ver que estos dos métodos **LMHIB** y **LMHDIN** resuelven el problema 4.5 de manera eficiente y efectiva.

Figura 4.16: Problema Nash-Cournot, $n=5$, $\exp=100$ Figura 4.17: Problema Nash-Cournot, $n=5$, $\exp=100$

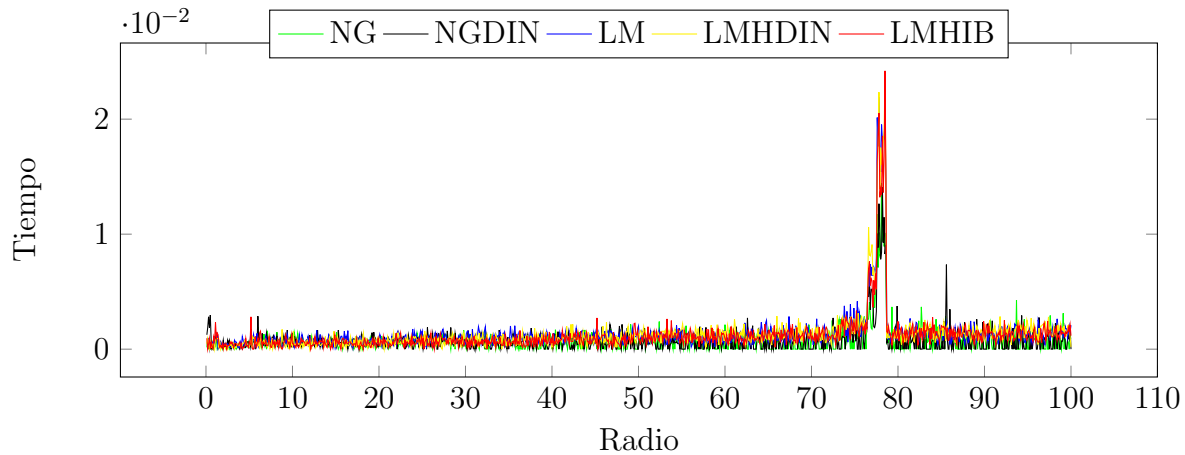
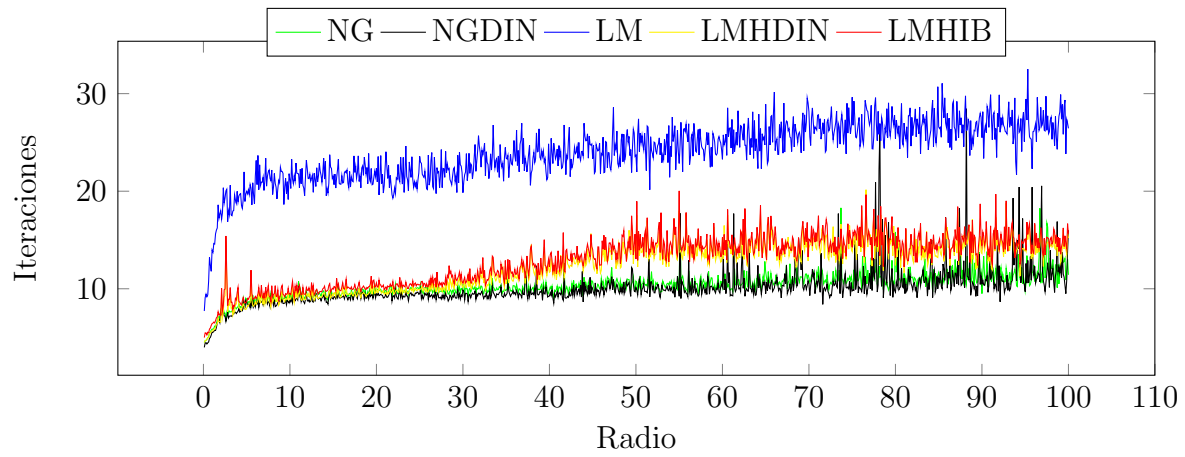
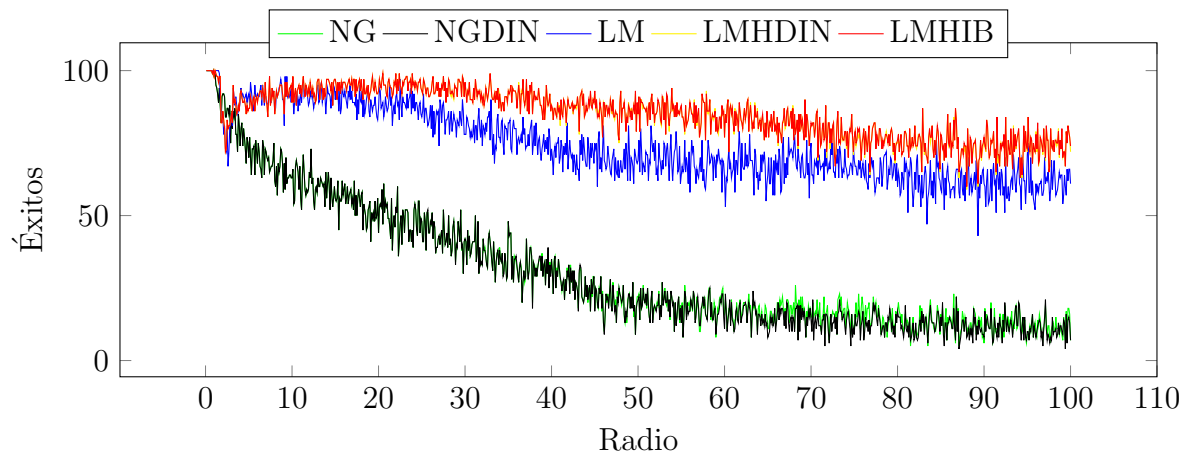


Figura 4.18: Problema Nash-Cournot, $n=5$, $\text{exp}=100$

Podemos observar por las Figuras 4.16 y 4.17, los algoritmos **NG** y **NGDIN** son los que tienen menor número de iteraciones, pero su porcentaje de éxitos disminuye considerablemente cuando el radio aumenta. Por otro lado, los algoritmos **LMHDIN** y **LMHIB** son los que tienen un mejor comportamiento, alcanzado un 100% de éxito para radios r inferiores a 40, y el algoritmo de **LM** también tiene un buen comportamiento en cuanto a éxitos, pero claramente por la Figura 4.18 tiene mayor número de iteraciones.

En cuanto al tiempo de convergencia, no hay una diferencia relevante que se deba tener en consideración, a pesar de que la gráfica tiene algunos saltos, no hacen gran diferencia, es por ello que para este problema destacamos el buen comportamiento de los algoritmos **LMHDIN** y **LMHIB**.

Figura 4.19: Problema Nash-Cournot, $n=10$, $\exp=100$ Figura 4.20: Problema Nash-Cournot, $n=10$, $\exp=100$

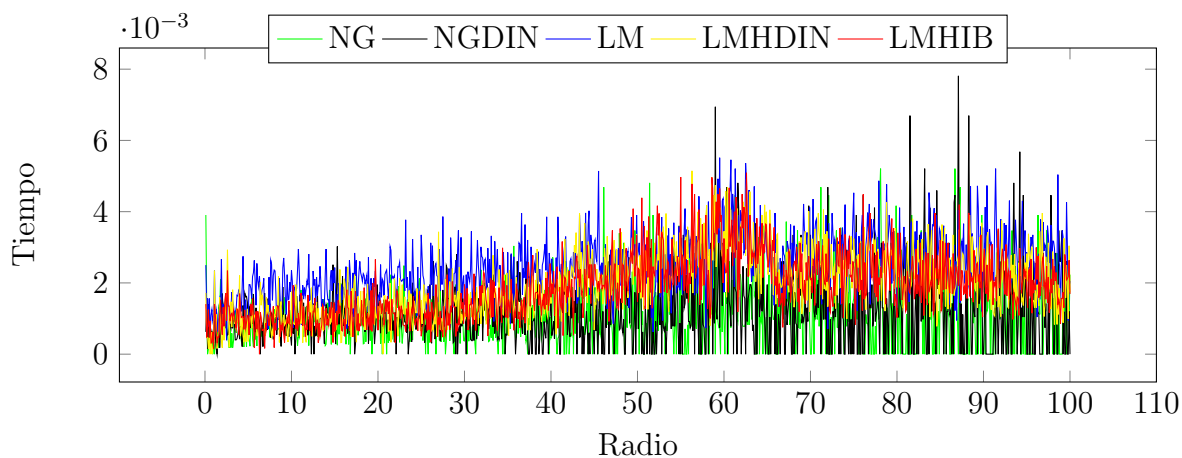
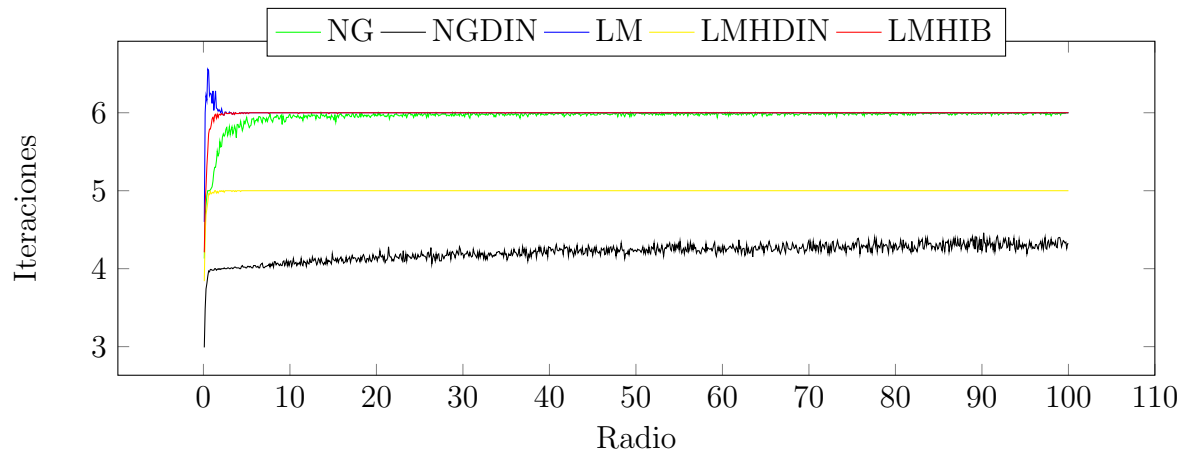
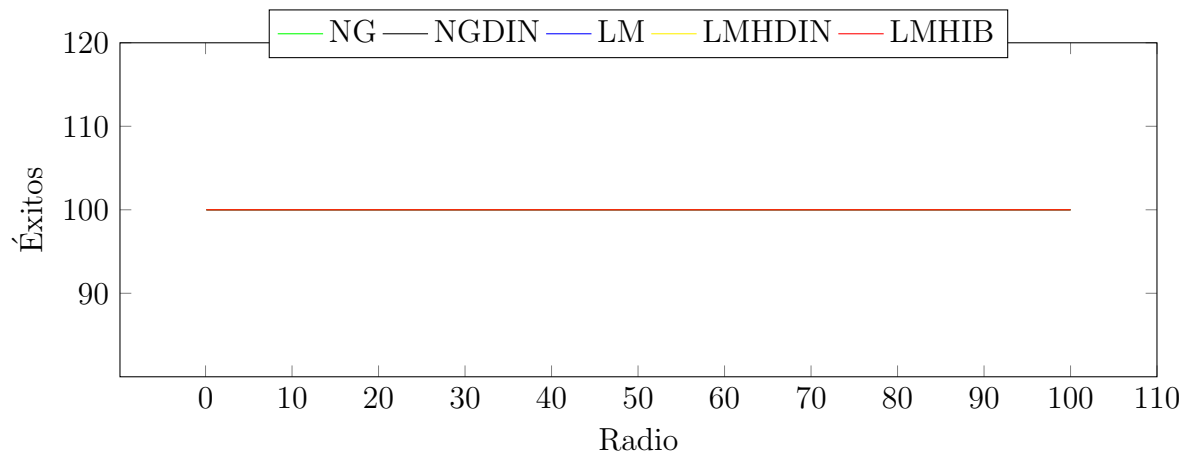


Figura 4.21: Problema Nash-Cournot, $n=10$, $\exp=100$

Para el problema de *Nash-Cournot* de dimensión $n = 10$, cuando el radio aumenta. Las Figuras 4.19, 4.20 y 4.21 nos permiten ver el buen comportamiento de los algoritmos **LMHDIN** y **LMHIB** respecto al resto de algoritmos. Donde, para los mencionados algoritmos es visible que tienen un mayor porcentaje de éxito y un menor número de iteraciones.

Además el tiempo de convergencia es menor que el algoritmo LM, respecto a los algoritmos **NG** y **NGDIN** podemos decir que es eficiente cuando converge a la solución del problema, pero en general tiene un bajo porcentaje de éxitos.

Figura 4.22: Problema Geiger and Kanzow, $n=10$, $\exp=100$ Figura 4.23: Problema Geiger and Kanzow, $n=10$, $\exp=100$

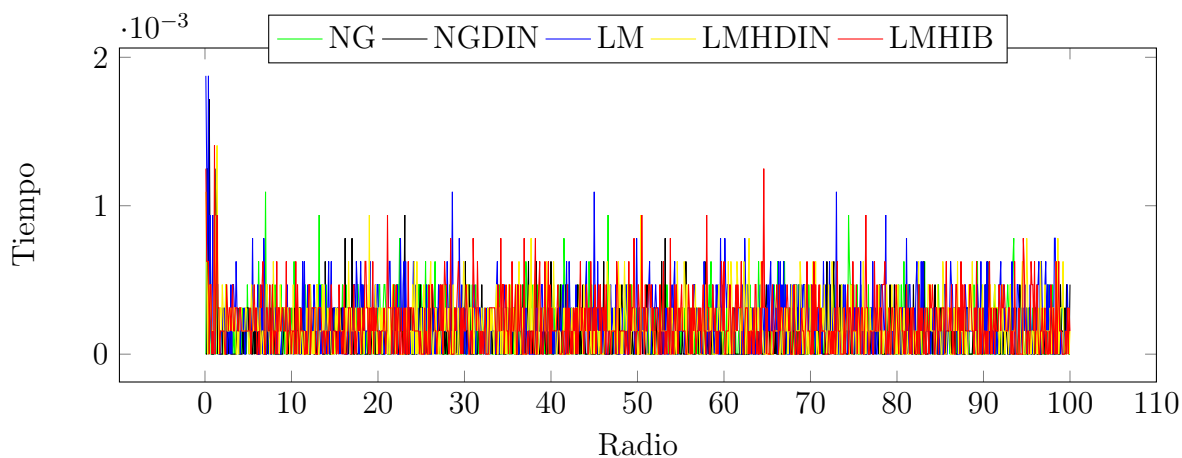
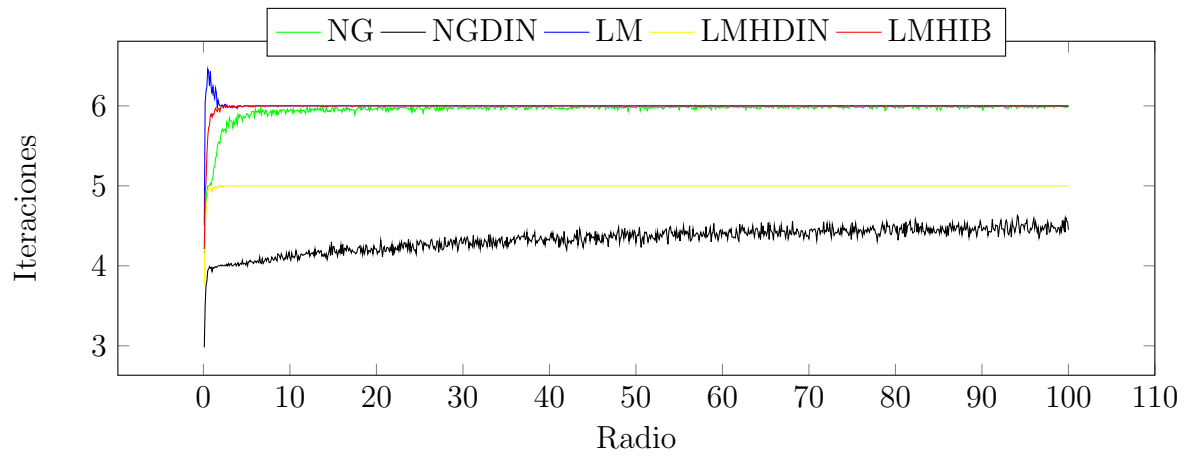
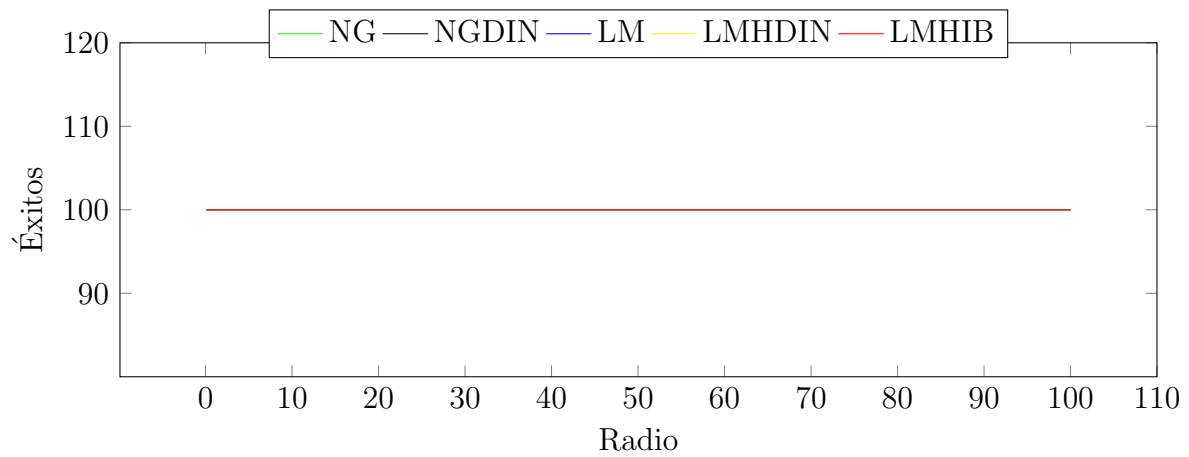


Figura 4.24: Problema Geiger and Kanzow, $n=10$, $\exp=100$

Para este problema de complementariedad lineal, podemos ver por la figura 4.23 que los métodos tienen un 100 % de éxito, es por ello que nos centraremos en la figura 4.22 y 4.24.

Primero destacamos que todos los métodos resuelven el problema en un número pequeño de iteraciones, además el parámetro λ dinámico permite que los algoritmos **NGDIN** y **LMHDIN** tengan menor número de iteraciones, de otro lado, respecto al tiempo que se demoran en determinar la solución del problema claramente se observa que todos los métodos son eficientes.

Figura 4.25: Problema Geiger and Kanzow, $n=128$, $\exp=100$ Figura 4.26: Problema Geiger and Kanzow, $n=128$, $\exp=100$

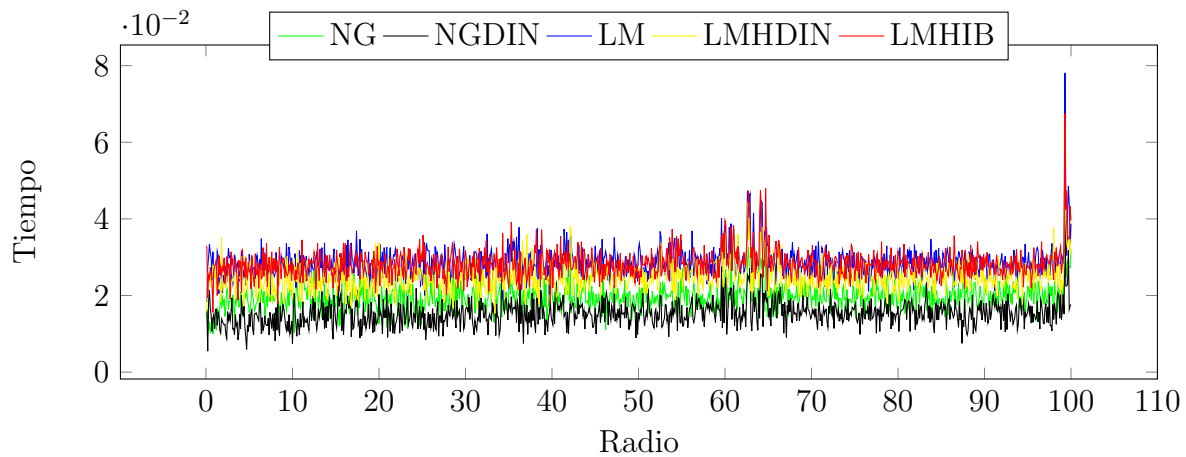
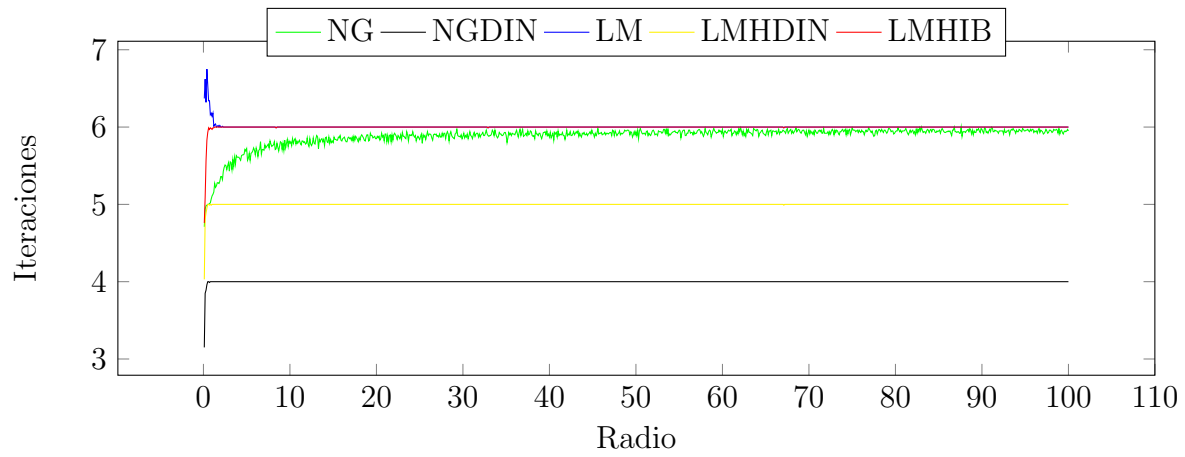
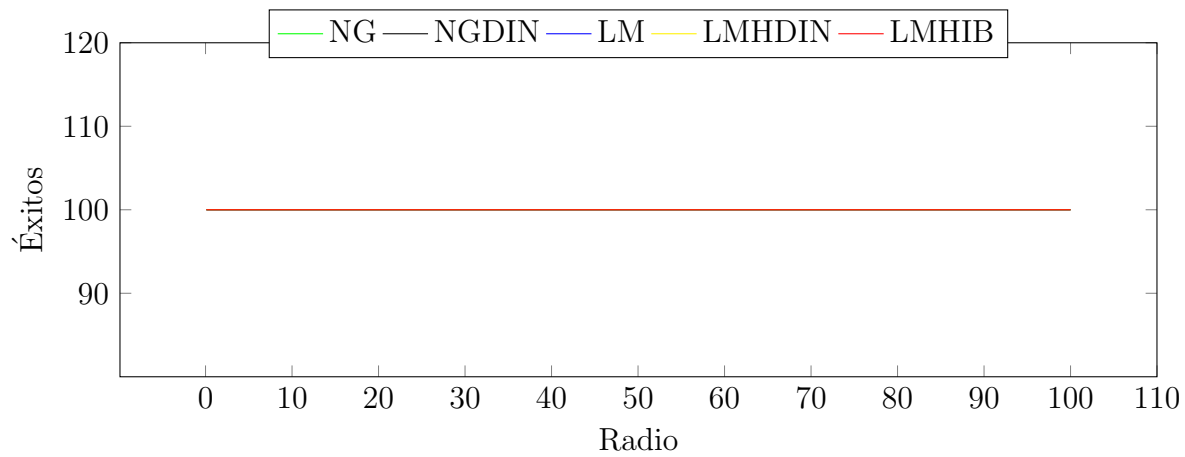


Figura 4.27: Problema Geiger and Kanzow, $n=128$, $\text{exp}=100$

Respecto a las Figuras 4.25, 4.26 y 4.27 se puede deducir que cuando aumentamos la dimensión para el problema de complementariedad lineal, vemos que el comportamiento en cuanto al número de éxitos e iteraciones se mantiene igual que para la dimensión $n = 10$.

Respecto al tiempo en que tarda los algoritmos en determinar la solución, se ve con mayor claridad el buen comportamiento de los métodos **NG** y **NGDIN**, además se puede observar que el parámetro λ dinámico ayuda a disminuir el número de iteraciones y el tiempo de convergencia.

Figura 4.28: Problema Ahn, $n=10$, $\exp=100$ Figura 4.29: Problema Ahn, $n=10$, $\exp=100$

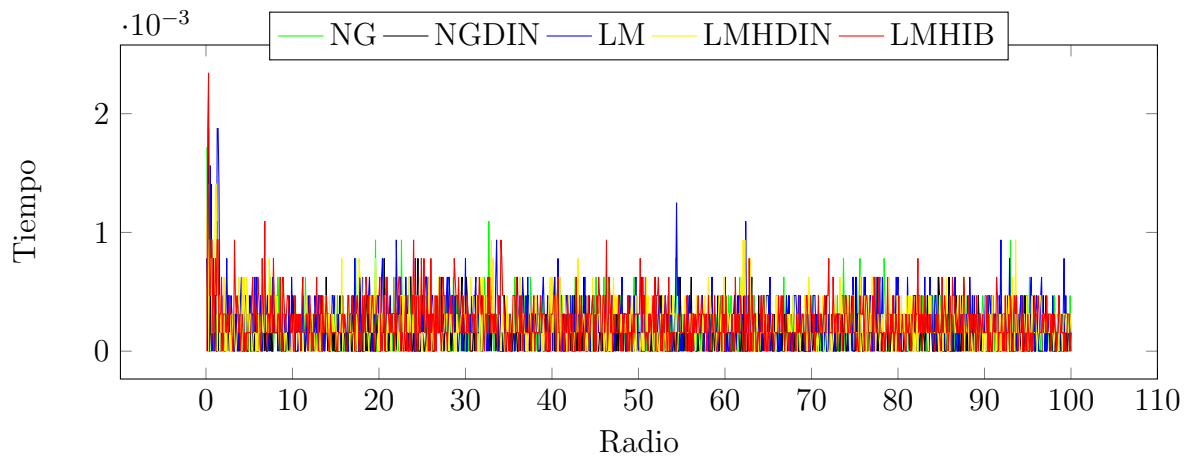
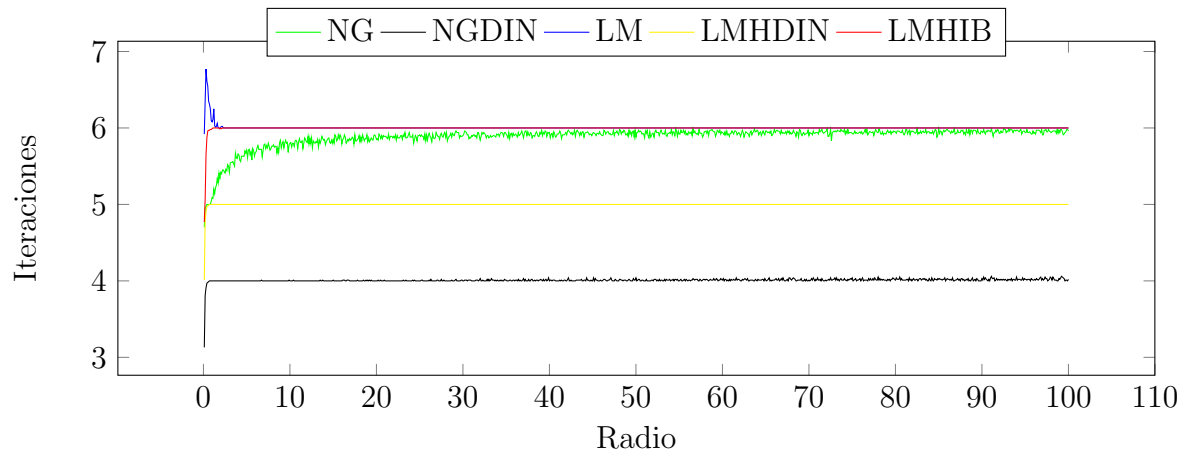
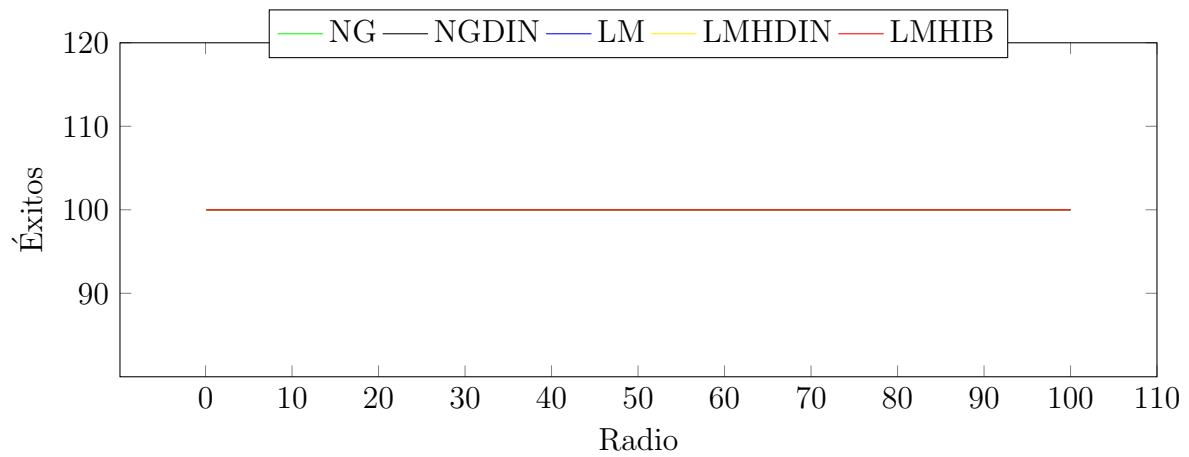


Figura 4.30: Problema Ahn, $n=10$, $\text{exp}=100$

Los resultados numéricos para este problema de complementariedad lineal, al igual que el anterior problema, nos permiten inferir el buen comportamiento de todos los algoritmos respecto porcentaje de número de éxitos, número de iteraciones y promedio de tiempo.

Figura 4.31: Problema Ahn, $n=100$, $\exp=100$ Figura 4.32: Problema Ahn, $n=100$, $\exp=100$

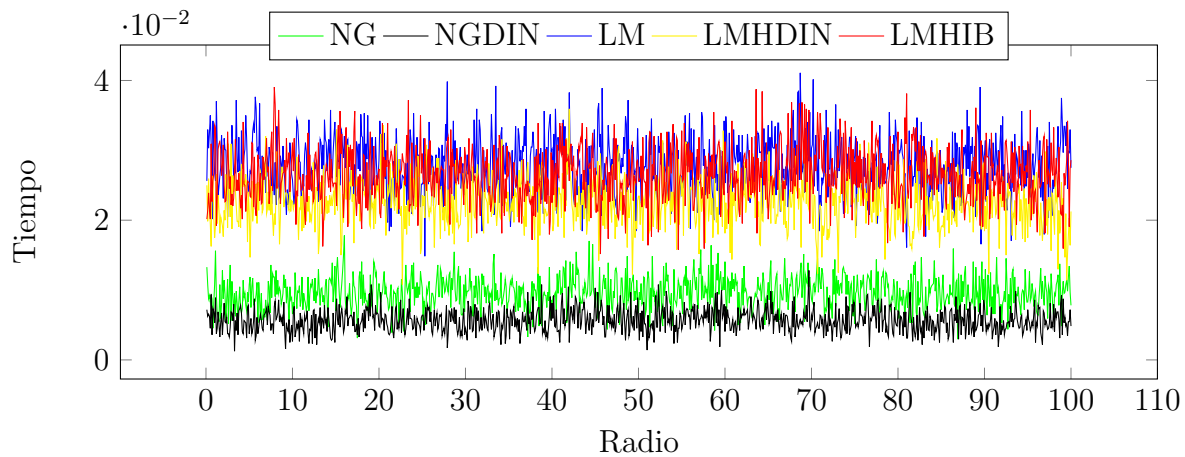


Figura 4.33: Problema Ahn , $n=100$, $\exp=100$

Cuando aumentamos la dimensión para el problema Ahn , podemos ver dos cambios relevantes, el primero fue el tiempo en que tarda los algoritmos en determinar la solución, pero que se ajusta dado al tamaño de la matriz que define el problema de complementariedad lineal, el segundo lo efectivo que son los métodos (**NG** y **NGDIN**).

4.1. Resumen

Observando los resultados numéricos que se obtuvieron para los nueve problemas presentados en este capítulo, se evidencia lo eficiente de los métodos que usan la estrategia de búsqueda lineal; es decir, cuando el algoritmo determina una solución para el PCNL, el tiempo de ejecución es considerablemente pequeño; en contraste, los algoritmos que usan región de confianza nos proporciona mayor número de éxitos, por lo tanto el usar la combinación de las dos estrategias (región de confianza y búsqueda lineal) permitió conseguir un equilibrio entre eficiencia y eficacia.

El bajo porcentaje de éxitos de los algoritmos que usaron solo búsqueda lineal en la mayoría de los casos se debe a la singularidad de la matriz H_k .

Por otro lado, el costo computacional que genera la estrategia de región de confianza, se produce por la veces que se repite el proceso de resolver un sistema lineal cada vez que se rechaza un iterante.

Respecto a los algoritmos (**NG** y **NGDIN**) los cuales usaron el paso de Newton, esperábamos que obtuvieran un mayor rapidez de convergencia, respecto a los los algoritmos que usaron el paso de L.M (**LMHIB** y **LMHDIN**), pero esta no fue en gran proporción. Adicional a esto, se observo que el método **LM** genera un alto número de iteraciones pero con un buen porcentaje de éxitos, para la mayoría de problemas nuestro algoritmo híbrido(**LMHIB**) ayudo a disminuir el número de iteraciones y aumento el porcentaje de éxitos.

Por último, se propuso usar el parámetro λ dinámico con el fin de observar mejor compartimiento de los algoritmo **NG** y **LMHIB** en cuanto al número de éxitos, pero los resultados numéricos obtenidos no presentan diferencias significativas en cuanto tiempo, número de iteraciones o éxitos con respecto a $\lambda = 2$, posiblemente una nueva propuesta para variar el λ dinámicamente puede generar mejores resultados.

4.2. Comparación numérica

En esta sección, vamos a hacer una comparación numérica entre el **Algoritmo 1** y el algoritmo propuesto por *Fan* [20] el cual llamaremos **NMLMHIB**, observemos que para $\lambda = 2$ el algoritmo **Algoritmo 1** es la versión monótona del algoritmo **NMLMHIB** (**Algoritmo 3**).

Algoritmo 3 Algoritmo Híbrido no monótono de Levenberg-M.

Entrada: Sea $\mathbf{x}_0 \in \mathbb{R}^n$, $0 < \rho_1 < \rho_2 < 1$, $0 < \tau_1 < 1 < \tau_2$, $\beta_0 = \mu_0 > \Delta_{min} > 0$,
 $\delta \in (0, 1)$, $t \in (0, 1)$, $\lambda = 2$ $\eta_k = 0.15$ $k := 0$.

1: $\nabla\Psi_{FB}(\mathbf{x}_k) \leftarrow H_k\Phi_{FB}(\mathbf{x}_k)$, donde $H_k \in \partial\Phi_{FB}(\mathbf{x}_k)$.

2: $\mathbf{d}_k \leftarrow -(H_k^T H_k + \mu_k I_n)^{-1} \nabla\Psi_{FB}(\mathbf{x}_k)$.

3: **mientras** $\|\nabla\Psi_\lambda(\mathbf{x}_k)\| < 10^{-6}$ y $k < 200$ **hacer**

4: $Pred_k \leftarrow -\nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k - \frac{1}{2} \mathbf{d}_k^T H_k^T H_k \mathbf{d}_k$
 $Ared_k \leftarrow C_k - \Psi_{FB}(\mathbf{x}_k + \mathbf{d}_k)$

donde,

$$C_k \leftarrow \begin{cases} \Psi_{FB}(\mathbf{x}_k) & \text{si } k = 0 \\ (\eta_k Q_{k-1} C_{k-1} + \Psi_{FB}(\mathbf{x}_k))/Q_k & \text{si } k \geq 1 \end{cases}$$

$$Q_k \leftarrow \begin{cases} 1 & \text{si } k = 0 \\ (\eta_k Q_{k-1} + 1)Q_k & \text{si } k \geq 1 \end{cases}$$

$\tilde{\rho}_k \leftarrow Ared_k/Pred_k$.

5: **si** $\tilde{\rho}_k \geq \rho_1$ **entonces**

6: $\mathbf{x}_k \leftarrow \mathbf{x}_k + \mathbf{d}_k$

7: **si no**

8: Calcular $\alpha_k := \max\{t^l | l = 0, 1, \dots\}$ tal que

$$\Psi_\lambda(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq \Psi_\lambda(\mathbf{x}_k) + \delta \alpha_k \nabla\Psi_\lambda(\mathbf{x}_k)^T \mathbf{d}_k$$

9: $\mathbf{x}_k \leftarrow \mathbf{x}_k + \alpha_k \mathbf{d}_k$

10: **fin si**

11: actualizar,

$$\beta_{k+1} \leftarrow \begin{cases} \tau_2 \beta_k & \text{si } \tilde{\rho}_k \leq \rho_1 \\ \beta_k & \text{si } \tilde{\rho}_k \in (\rho_1, \rho_2) \\ \max\{\tau_1 \beta_k, \Delta_{min}\} & \text{si } \tilde{\rho}_k \geq \rho_2 \end{cases}$$

12: $\mu_k \leftarrow \beta_k \|\Phi_\lambda(\mathbf{x}_k)\|^2$

13: $\nabla\Psi_{FB}(\mathbf{x}_k) \leftarrow H_k\Phi_{FB}(\mathbf{x}_k)$,

14: $\mathbf{d}_k \leftarrow -(H_k^T H_k + \mu_k I_n)^{-1} \nabla\Psi_{FB}(\mathbf{x}_k)$.

15: $k \leftarrow k + 1$

16: **fin mientras**

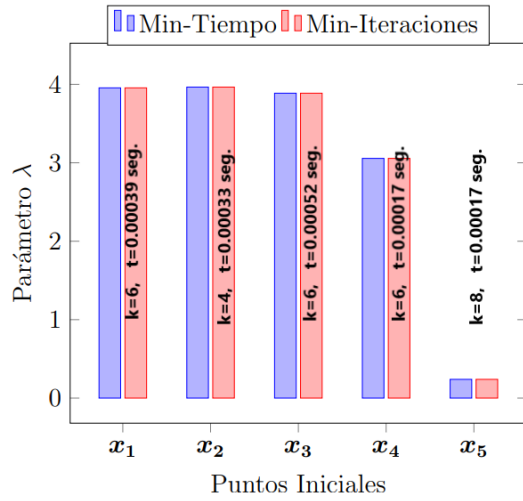
17: **Salida:** \mathbf{x}_{k+1} .

Es importante mencionar que este **Algoritmo 3** es eficiente y robusto, por ello, con base en los resultado de la anterior sección, consideramos que se debe buscar un $\lambda \in (0, 4)$ apropiado para que nuestro algoritmo sea competitivo. Para ello, realizamos pruebas respecto a los **Problemas 4.1, 4.3, 4.6, 4.8, 4.9**, para distintos λ desde $\lambda = 10^{-3}$ hasta $\lambda = 3.998$ con incrementos de 10^{-3} y registramos la convergencia, número de iteraciones y tiempo en hallar la solución del problema; con estos resultados se quiere determinar un $\lambda \in (0, 4)$, que llamaremos λ_{opt} para el cual se obtiene el menor número de iteraciones y menor tiempo de ejecución.

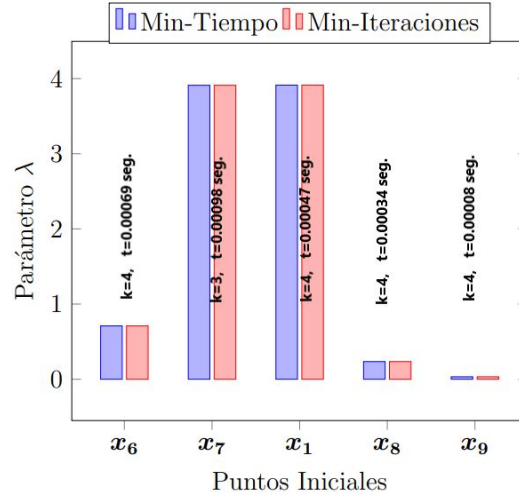
En cuanto a los vectores iniciales, consideramos los tomados en [20], según el experimento numérico a realizar. Con el objetivo de hacer referencia rápida, los enumeramos a continuación:

$$\begin{array}{ll}
 \mathbf{x}_1 = (2 \ 2 \ 2 \ 2)^T & \mathbf{x}_9 = (-5 \ -5 \ -5 \ -5)^T \\
 \mathbf{x}_2 = (5 \ 0 \ 5 \ 0)^T & \mathbf{x}_{10} = (5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5)^T \\
 \mathbf{x}_3 = (10 \ 10 \ 10 \ 10)^T & \mathbf{x}_{11} = (10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10 \ 10)^T \\
 \mathbf{x}_4 = (10^3 \ 10^3 \ 10^3 \ 10^3)^T & \mathbf{x}_{12} = (1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10)^T \\
 \mathbf{x}_5 = (-10 \ -10 \ -10 \ -10)^T & \mathbf{x}_{13} = (5 \ 4 \ 3 \ 2 \ 1 \ 6 \ 7 \ 8 \ 9 \ 10)^T \\
 \mathbf{x}_6 = (0.5 \ 0.5 \ 0.5 \ 0.5)^T & \mathbf{x}_{14} = (7 \ 4 \ 3 \ 1 \ 18 \ 4 \ 1 \ 6 \ 3 \ 2)^T \\
 \mathbf{x}_7 = (1 \ 2 \ 1 \ 2)^T & \mathbf{x}_0^n = (0 \ \dots \ 0)^T \\
 \mathbf{x}_8 = (100 \ 100 \ 100 \ 100)^T & \mathbf{x}_1^n = (0.5 \ \dots \ 0.5)^T
 \end{array}$$

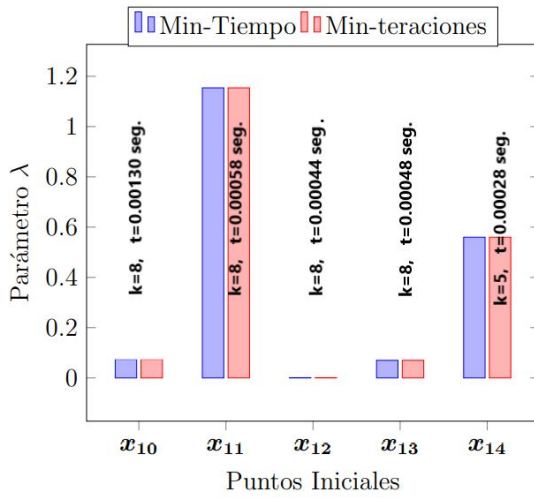
Los experimentos que se muestran en la Figura 4.34, nos permiten observar que para algunos puntos iniciales, en relación con el menor tiempo y menor número de iteraciones, el parámetro λ es el mismo, a dicho parámetro le llamaremos λ_{opt} .



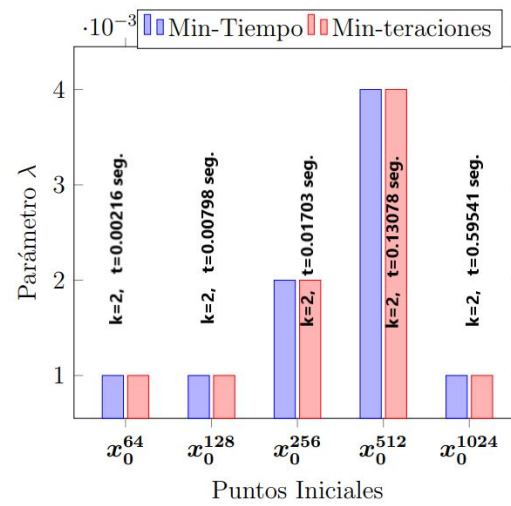
(a) Problema *Koj-Shi*



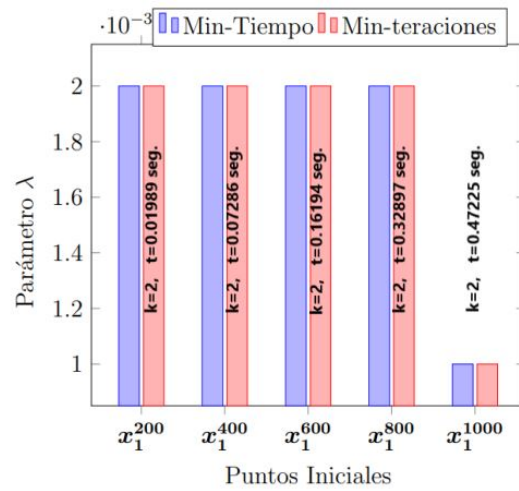
(b) Problema *Math mod*



(c) Problema *Nash Co*



(d) Problema *Ahn*



(e) Problema *Brown*

Figura 4.34: λ_{opt} para el PCNL(F)

En la siguiente tabla, encontramos los resultados obtenidos al resolver los problemas anteriormente mencionados con los algoritmos **LMHIB** y **NMLMHIB**, donde cada una de las columnas contiene la siguiente información:

Problema: Nombre del problema.

λ_{opt} : λ optimo.

n: Dimensión del problema.

\mathbf{x}_0 : Punto inicial.

Método: Nombre del algoritmo.

k : Número de iteraciones.

$\Psi_\lambda(\mathbf{x}^*)$ la función Ψ_λ evaluada en la solución del problema.

Tiempo: Tiempo CPU (en segundos).

Problema	λ_{opt}	n	\mathbf{x}_0	Método	k	$\Psi_\lambda(\mathbf{x}^*)$	Tiempo
<i>Koj-Shi</i>	3.955	4	\mathbf{x}_1	LMHIB	6	9.395e-15	0.00039
	2			NMLMHIB	16*	1.605e-16	0.00120
<i>Koj-Shi</i>	3.965	4	\mathbf{x}_2	LMHIB	4	3.029e-14	0.00033
	2			NMLMHIB	7	2.043e-24	0.00039
<i>Koj-Shi</i>	3.887	4	\mathbf{x}_3	LMHIB	6	2,089e-21	0.00052
	2			NMLMHIB	17*	1.015e-16	0.00036
<i>Koj-Shi</i>	3.056	4	\mathbf{x}_4	LMHIB	6	2,641e-16	0.00017
	2			NMLMHIB	13	5.39e-23	0.00036
<i>Koj-Shi</i>	0.239	4	\mathbf{x}_5	LMHIB	8	4,125e-25	0.00017
	2			NMLMHIB	14*	6.953e-15	0.00027

Problema	λ_{opt}	n	\mathbf{x}_0	Método	k	$\Psi_\lambda(\mathbf{x}^*)$	Tiempo
<i>Math mod</i>	0.710	4	\mathbf{x}_6	LMHIB	4*	9.000e-30	0.00069
	2			NMLMHIB	4*	7.572e-24	0.00034
<i>Math mod</i>	3.911	4	\mathbf{x}_7	LMHIB	3	5.920e-14	0.00098
	2			NMLMHIB	5	1.734e-17	0.00028
<i>Math mod</i>	3.913	4	\mathbf{x}_1	LMHIB	4*	5.423e-16	0.00047
	2			NMLMHIB	4	6.393e-16	0.00028
<i>Math mod</i>	0.235	4	\mathbf{x}_8	LMHIB	4*	5,106e-14	0.00034
	2			NMLMHIB	6*	1.570e-16	0.00038
<i>Math mod</i>	0.032	4	\mathbf{x}_9	LMHIB	4*	1,135e-13	0.00008
	2			NMLMHIB	7*	0.100e-24	0.00022
<i>Nash Co</i>	0.074	10	\mathbf{x}_{10}	LMHIB	8	9.910e-28	0.00130
	2			NMLMHIB	9	3.0310e-18	0.00183
<i>Nash Co</i>	1.154	10	\mathbf{x}_{11}	LMHIB	8	1.212e-25	0.00058
	2			NMLMHIB	13	1.494e-17	0.00117
<i>Nash Co</i>	0.001	10	\mathbf{x}_{12}	LMHIB	8	6.060e-28	0.00044
	2			NMLMHIB	10	5.810e-28	0.00073
<i>Nash Co</i>	0.070	10	\mathbf{x}_{13}	LMHIB	8	2.970e-28	0.00048
	2			NMLMHIB	9	1.331e-16	0.00053
<i>Nash Co</i>	0.560	10	\mathbf{x}_{14}	LMHIB	5	2.408e-24	0.00028
	2			NMLMHIB	6	4.915e-026	0.00038
<i>Ahn</i>	0.001	64	\mathbf{x}_0^{56}	LMHIB	2	3.889e-21	0.00216
	2			NMLMHIB	5	7.520e-23	0.00380
<i>Ahn</i>	0.001	128	\mathbf{x}_0^{128}	LMHIB	2	7.842e-21	0.00798
	2			NMLMHIB	5	1.516e-22	0.01603
<i>Ahn</i>	0.002	256	\mathbf{x}_0^{256}	LMHIB	2	9.431e-19	0.01703
	2			NMLMHIB	5	3.043e-22	0.04070
<i>Ahn</i>	0.004	512	\mathbf{x}_0^{512}	LMHIB	2	1.167e-16	0.13078
	2			NMLMHIB	5	6.100e-22	0.26298
<i>Ahn</i>	0.001	1024	\mathbf{x}_0^{1024}	LMHIB	2	6.319e-20	0.59541
	2			NMLMHIB	5	1.2222e-21	1.48756
<i>Brown</i>	0.002	200	\mathbf{x}_1^{200}	LMHIB	2	1.711e-20	0.01989
	2			NMLMHIB	5	5.182e-25	0.06695
<i>Brown</i>	0.002	400	\mathbf{x}_1^{400}	LMHIB	2	3.447e-20	0.07286
	2			NMLMHIB	5	8.038e-26	0.17856
<i>Brown</i>	0.002	600	\mathbf{x}_1^{600}	LMHIB	2	4.782e-20	0.16194
	2			NMLMHIB	5	2.684e-23	0.35405
<i>Brown</i>	0.002	800	\mathbf{x}_1^{800}	LMHIB	2	1.711e-20	0.32897
	2			NMLMHIB	5	1.963e-24	0.71559
<i>Brown</i>	0.001	1000	\mathbf{x}_1^{1000}	LMHIB	2	1.886e-21	0.47225
	2			NMLMHIB	5	2.727e-22	1.21139

Tabla 4.1: Resultados numéricos de los algoritmos LMHIB y NMLMHIB

El asterisco (*) denota que el punto límite generado por los algoritmos es la solución degenerada, de lo contrario es la solución no degenerada.

Con el objetivo de comparar el desempeño global del **algoritmo 1**, lo comparamos con el algoritmo **NMLMHIB** propuesto en [20]. Para cada punto inicial se realizaron 1000 experimento y el tiempo (CPU) es su tiempo promedio en segundos.

Los resultados consignados en la tabla 4.1 muestran que para todos los problemas evaluados en los distintos puntos iniciales, tanto el número de iteraciones como el tiempo de ejecución del **Algoritmo 1 (LMHIB)** es considerablemente menor que los reportados por el algoritmo **NMLMHIB** [20]. Otro aspecto importante a tener en cuenta es que el **Algoritmo 1**, para todos los problemas, usa en la mayoría de pasos región de confianza.

Por otro lado, destacamos los resultados que obtuvo el algoritmo **LMHIB** para el problema de *Ahn*, pues el número de iteraciones y el tiempo en el cual determina la solución del problema son notablemente pequeño, respecto al número de iteraciones es menor que el algoritmo **NMLMHIB**.

4.3. Comentarios finales

El enfoque tradicional para para resolver el PCNL(F) (1.1) implica reformular el problema, como un problema de optimización[37, 28, 63, 1] o como un sistema no lineal, no diferenciable [46, 68, 67]. Nuevos métodos surgen para resolver el problema de optimización, los cuales se basan en la clase de métodos de región de confianza [39, 34] y también, métodos híbridos de región de confianza y búsqueda lineal [20, 44, 45].

Motivados por los resultados obtenidos en [20], quienes resolvieron el PCNL (1.1), a través de la reformulación como un problema de minimi-

zar la función de mérito 2.23 usando la función de complementariedad de *Fischer-Burmeister* en la reformulación del sistema no lineal (2.3), en este trabajo de investigación, propusimos un algoritmo híbrido que combina las técnicas de búsqueda lineal y región de confianza, donde cada paso se usa el método de *Levenberg-Marquardt*, además para la reformulación como un problema de minimización (2.24), usamos la familia uniparamétrica de funciones de complementariedad (2.2). Desarrollamos la teoría de convergencia del método y realizamos pruebas numéricas para analizar su desempeño.

Los experimentos nos llevaron a proponer un λ_{opt} para el cual se obtuvieron ventajas notorias respecto al número de éxitos y en cuanto al tiempo de ejecución del método. Además, para posteriores trabajos creemos que es posible incorporar una estrategia que permita elegir el parámetro λ con el objetivo de mejorar la efectividad del algoritmo 1.

Para trabajos futuros, creemos que la combinación no monótona de la búsqueda lineal y región de confianza (**NMLMHIB**) proporcionará una alta efectividad en cuanto a la convergencia del algoritmo, es por ello que el **Algoritmo 1** pueda ser mejorado usando la técnica no monótona, además si se escoge un λ_{opt} proporcionaría mayor rapidez de convergencia.

Bibliografía

- [1] P. Alvarez and P. Edith, *Solución de sistemas de ecuaciones restringidas mediante método de puntos interiores*, Master's thesis, Universidad Nacional de Trujillo (2017).
- [2] M. Anitescu, J. Cremer, and F. Potra, *On the existence of solutions to complementarity formulations of contact problems with friction*, *Complementary and Variational Problems* (1997), 12–21.
- [3] F. Arenas, *Métodos secantes de cambio mínimo para el problema de complementariedad no lineal*, Master's thesis, Universidad del Cauca (2013).
- [4] F. Arenas, H.J. Martínez, and R. Pérez, *Least change secant update methods for nonlinear complementarity problem*, *Ingeniería y Ciencia* **11** (2015), no. 21, 11–36.
- [5] F. Arenas, R. Pérez, and H. Vivas, *A neural network model for nonlinear complementarity problems*, *Revista Integración* **34** (2016), no. 2, 169–185.
- [6] C Arias, *Un algoritmo cuasi Newton global para problemas de complementariedad no lineal*, Master's thesis, Universidad del Cauca (2014).

-
- [7] S. Billups, *Algorithms for complementarity problems and generalized equations*, Tech. report, University of Wisconsin at Madison, 1995.
- [8] C. Broyden, J.E Dennis, and J. Moré, *On the local and superlinear convergence of quasi-Newton methods*, IMA Journal of Applied Mathematics **12** (1973), no. 3, 223–245.
- [9] A. Byong-Hun, *Iterative methods for linear complementarity problems with upperbounds on primary variables*, Mathematical Programming **26** (1983), no. 3, 295–315.
- [10] A. Chen, J. Oh, D. Park, and W. Recker, *Solving the bicriteria traffic equilibrium problem with variable demand and nonlinear path costs*, Applied Mathematics and Computation **217** (2010), no. 7, 3020–3031.
- [11] F. Clarke, *Necessary conditions for nonsmooth problems in optimal control and the calculus of variations*, Ph.D. thesis, University of Washington, 1973.
- [12] ———, *Generalized gradients and applications*, Transactions of the American Mathematical Society **205** (1975), 247–262.
- [13] ———, *Optimization and nonsmooth analysis*, vol. 5, Siam, 1990.
- [14] R. W Cottle, *Nonlinear programs with positively bounded jacobians*, SIAM Journal on Applied Mathematics **14** (1966), no. 1, 147–158.
- [15] T. De Luca, F. Facchinei, and C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Mathematical programming **75** (1996), no. 3, 407–439.
- [16] J. Dennis and R. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, vol. 16, Siam, 1996.
- [17] S. Du and Y. Gao, *Convergence analysis of nonmonotone levenberg-marquardt algorithms for complementarity problem*, Applied mathematics and computation **216** (2010), no. 5, 1652–1659.

- [18] H. Escobar, *Optimización no lineal y dinámica*, Universidad Nacional de Colombia, 2001.
- [19] F. Facchinei and J. Soares, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM Journal on Optimization **7** (1997), no. 1, 225–247.
- [20] B. Fan, C. Ma, A. Wu, and C. Wu, *A levenberg–marquardt method for nonlinear complementarity problems based on nonmonotone trust region and line search techniques*, Mediterranean Journal of Mathematics **15** (2018), no. 3, 118.
- [21] J. Fan, *A modified levenberg-marquardt algorithm for singular system of nonlinear equations*, Journal of Computational Mathematics **21** (2003), 628–630.
- [22] J. Fan and J. Pan, *A note on the levenberg–marquardt parameter*, Applied Mathematics and Computation **207** (2009), no. 2, 351–359.
- [23] JY Fan and YX Yuan, *On the convergence of a new levenberg-marquardt method*, Technical Report, AMSS, Chinese Academy of Sciences (2001), 1–11.
- [24] M. Ferris and J Pang, *Complementarity and variational problems*, SIAM, 1997.
- [25] M. Ferris and J. Pang, *Engineering and economic applications of complementarity problems*, State of the art. **39** (1997), no. 4, 669–713.
- [26] A. Fischer, *A special Newton-type optimization method*, Optimization **24** (1992), no. 3-4, 269–284.
- [27] ———, *Solution of monotone complementarity problems with locally lipschitzian functions*, Mathematical Programming **76** (1997), no. 3, 513–532.

-
- [28] S. A Gabriel and J. Pang, *A trust region method for constrained nonsmooth equations*, Large Scale Optimization, Springer, 1994, pp. 155–181.
- [29] C. Geiger and C. Kanzow, *On the resolution of monotone complementarity problems*, Computational Optimization and Applications **5** (1996), no. 2, 155–173.
- [30] G. Golub and C. Van Loan, *Matrix computations*, vol. 3, JHU press, 2013.
- [31] L. Grippo, F. Lampariello, and S. Lucidi, *A nonmonotone line search technique for newton’s method*, SIAM Journal on Numerical Analysis **23** (1986), no. 4, 707–716.
- [32] P. Harker, *Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities*, Mathematical Programming **41** (1988), no. 1-3, 29–59.
- [33] P. Harker and B. Xiao, *Newton’s method for the nonlinear complementarity problem: A b -differentiable equation approach*, Mathematical Programming **48** (1990), no. 1-3, 339–357.
- [34] H. Jiang, M. Fukushima, L. Qi, and D. Sun, *A trust region method for solving generalized complementarity problems*, SIAM Journal on Optimization **8** (1998), no. 1, 140–157.
- [35] C. Kanzow, *Some equation-based methods for the nonlinear complementarity problem*, Optimization Methods and Software **3** (1994), no. 4, 327–340.
- [36] ———, *Some noninterior continuation methods for linear complementarity problems*, SIAM Journal on Matrix Analysis and Applications **17** (1996), no. 4, 851–868.

- [37] C. Kanzow, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 4, 851–868.
- [38] C. Kanzow and H. Kleinmichel, *A new class of semismooth Newton-type methods for nonlinear complementarity problems*, Computational Optimization and Applications **11** (1998), no. 3, 227–251.
- [39] C. Kanzow and M. Zupke, *Inexact trust-region methods for nonlinear complementarity problems*, Reformulation: Nonsmooth, piecewise smooth, semismooth and smoothing methods, Springer, 1998, pp. 211–233.
- [40] Yamashita N. Kanzow, C. and Masao Fukushima, *New ncp-functions and their properties*, Journal of Optimization Theory and Applications **94** (1997), no. 1, 115–135.
- [41] D. Li and M. Fukushima, *A derivative-free line search and global convergence of broyden-like method for nonlinear equations*, Optimization Methods and Software **13** (2000), no. 3, 181–201.
- [42] ———, *Globally convergent broyden-like methods for semismooth equations and applications to vip, ncp and mcp*, Annals of Operations Research **103** (2001), no. 1-4, 71–97.
- [43] V. Lopes, J. Martinez, and R. Pérez, *On the local convergence of quasi-newton methods for nonlinear complementarity problems*, Applied Numerical Mathematics **30** (1999), no. 1, 3–22.
- [44] Changfeng Ma, *A new smoothing quasi-Newton method for nonlinear complementarity problems*, Applied mathematics and computation **171** (2005), no. 2, 807–823.
- [45] ———, *A new smoothing and regularization newton method for p - 0 -ncp*, Journal of Global Optimization **48** (2010), no. 2, 241–261.

- [46] O. Mangasarian, *Equivalence of the complementarity problem to a system of nonlinear equations*, SIAM Journal on Applied Mathematics **31** (1976), no. 1, 89–92.
- [47] L. Mathiesen, *An algorithm based on a sequence of linear complementarity problems applied to a walrasian equilibrium model: An example*, Mathematical programming **37** (1987), no. 1, 1–18.
- [48] J. Moré, *The levenberg-marquardt algorithm: implementation and theory*, Numerical analysis, Springer, 1978, pp. 105–116.
- [49] J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.
- [50] J. Nocedal and Y. Yuan, *Combining trust region and line search techniques*, Advances in nonlinear programming, Springer, 1998, pp. 153–175.
- [51] J. Pang, *A b -differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems*, Mathematical Programming **51** (1991), no. 1-3, 101–131.
- [52] J. Pang and L. Qi, *Nonsmooth equations: motivation and algorithms*, SIAM Journal on optimization **3** (1993), no. 3, 443–465.
- [53] R Pérez, F Arenas, J.M. Hector, and C. Arias, *El problema de complementariedad no lineal: Teoría, aplicaciones y nuevos algoritmos para su solución*, Programa Editorial Univalle, 2019.
- [54] R. Pérez and T. Díaz, *Minimización sin restricciones*, Editorial Universidad del Cauca (2010).
- [55] M. Powell, *Convergence properties of a class of minimization algorithms*, Nonlinear programming 2, Elsevier, 1975, pp. 1–27.

- [56] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations*, Mathematics of operations research **18** (1993), no. 1, 227–244.
- [57] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical programming **58** (1993), no. 1-3, 353–367.
- [58] W. Rudin, M. Sánchez, and L. Aguirre, *Principios de análisis matemático*, no. 515 R835 1981., McGraw-Hill, 1966.
- [59] A. Sherman, *On Newton-iterative methods for the solution of systems of nonlinear equations*, SIAM Journal on Numerical Analysis **15** (1978), no. 4, 755–771.
- [60] E. Spedicato and Z. Huang, *Numerical experience with newton-like methods for nonlinear algebraic systems*, Computing **58** (1997), no. 1, 69–89.
- [61] X. Tong and S. Zhou, *Combining trust region and line search methods for equality constrained optimization*, Numerical Functional Analysis and Optimization **24** (2003), no. 1-2, 143–162.
- [62] H. Wang and J. Fan, *Convergence rate of the levenberg-marquardt method under hölderian local error bound*, Optimization Methods and Software **35** (2020), no. 4, 767–786.
- [63] H. Wang and D. Pu, *A kind of nonmonotone filter method for nonlinear complementarity problem*, Journal of Applied Mathematics and Computing **36** (2011), no. 1, 27–40.
- [64] N. Yamashita and M. Fukushima, *On the rate of convergence of the levenberg-marquardt method*, (2001), 239–249.
- [65] L. Yong, *Nonlinear complementarity problem and solution methods*, International Conference on Artificial Intelligence and Computational Intelligence, Springer, 2010, pp. 461–469.

-
- [66] Ya-xiang Yuan, *A new trust region algorithm with trust region radius converging to zero*, (2001), 1–5.
- [67] H. Zheng and L. Liu, *The sign-based methods for solving a class of nonlinear complementarity problems*, *Journal of Optimization Theory and Applications* **180** (2019), no. 2, 480–499.
- [68] L. Zheng, H. Liu, *A two-step modulus-based matrix splitting iteration method for solving nonlinear complementarity problems*, *Computational and Applied Mathematics* **37** (2018), no. 4, 5410–5423.