



ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

ANÁLISIS Y RESOLUCIÓN NUMÉRICA DE PROBLEMAS DE PROGRAMACIÓN MATEMÁTICA CON RESTRICCIONES DE EQUILIBRIO Y COMPLEMENTARIEDAD (MPECS)

RESOLUCIÓN DE MPCCS Y MPECS CLÁSICOS CON MÉTODOS DE TIPO SEMISMOOTH NEWTON (SSN)

**TRABAJO DE INTEGRACIÓN CURRICULAR PRESENTADO COMO
REQUISITO PARA LA OBTENCIÓN DEL TÍTULO DE MATEMÁTICO**

DANNY MAURICIO MALLITASIG NACEVILLA

danny.mallitasig@epn.edu.ec

DIRECTOR: SERGIO GONZALEZ ANDRADE

sergio.gonzalez@epn.edu.ec

DMQ, AGOSTO 2023

CERTIFICACIONES

Yo, DANNY MAURICIO MALLITASIG NACEVILLA, declaro que el trabajo de integración curricular aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

Danny Mauricio Mallitasig Nacevilla

Certifico que el presente trabajo de integración curricular fue desarrollado por Danny Mauricio Mallitasig Nacevilla, bajo mi supervisión.

Sergio Gonzalez Andrade
DIRECTOR

DECLARACIÓN DE AUTORÍA

A través de la presente declaración, afirmamos que el trabajo de integración curricular aquí descrito, así como el(los) producto(s) resultante(s) del mismo, es(son) público(s) y estará(n) a disposición de la comunidad a través del repositorio institucional de la Escuela Politécnica Nacional; sin embargo, la titularidad de los derechos patrimoniales nos corresponde a los autores que hemos contribuido en el desarrollo del presente trabajo; observando para el efecto las disposiciones establecidas por el órgano competente en propiedad intelectual, la normativa interna y demás normas.

Danny Mauricio Mallitasig Nacevilla

Sergio Gonzalez Andrade

RESUMEN

El presente trabajo aborda el estudio de problemas de programación matemática con restricciones de complementariedad [MPCC] mediante una reformulación basada en la técnica de *lifting*. Luego, estudiamos el sistema de optimalidad de Lagrange que caracteriza los puntos estacionarios del problema reformulado con ayuda de una función Φ adecuada.

Posteriormente, se construye un algoritmo tipo Newton Semi-suave (SSN) (en primera instancia, con un alcance local y posteriormente se consigue una globalización), el cual posee un enfoque especial en los B-subdiferenciales y las propiedades que permiten una correcta incorporación de información de segundo orden en la matriz del sistema tipo Newton, construida a partir de los mencionados subdiferenciales. Con esto, se puede garantizar la convergencia superlineal del método.

Finalmente, realizamos la implementación del algoritmo en el lenguaje MATLAB y se proponen ejemplos que permiten evaluar el funcionamiento del código.

Palabras clave: MPCC, *lifting*, Newton Semi-suave, convergencia, MATLAB.

ABSTRACT

This paper deals with the study of mathematical programming problems with complementarity constraints [MPCC] by means of a reformulation based on the lifting technique. Then, we study the Lagrangian optimality system that characterizes the stationary points of the reformulated problem with the help of a suitable function Φ .

Subsequently, a Semi-Smooth Newton (SSN) type algorithm is constructed (in the first instance, with a local scope and subsequently globalization is achieved), which possesses a special focus on the B-subdifferentials and the properties that allow a correct incorporation of second order information into the Newton type system matrix, constructed from the mentioned subdifferentials. With this, the superlinear convergence of the method can be guaranteed.

Finally, we perform the implementation of the algorithm in MATLAB language and examples are proposed to evaluate the performance of the code.

Keywords: MPCC, lifting, Semi-Smooth Newton, convergence, MATLAB.

Índice general

1. Descripción del componente desarrollado	1
1.1. Objetivo general	1
1.2. Objetivos específicos	1
1.3. Alcance	2
1.4. Marco teórico	2
1.4.1. Conceptos preliminares	2
1.4.2. Problemas de minimización con restricciones de complementariedad (MPCCs)	3
1.4.3. Algunos conceptos básicos del MPCC y MPCC <i>lifted</i>	6
2. Metodología	15
2.1. Método de Newton Semi-suave (SSN) para el problema <i>lifted</i> MPCC	15
2.2. Globalización del método local	22
3. Resultados numéricos	33
3.1. Implementación	33
3.2. Ejemplos	33
3.3. Conclusiones	41
3.4. Recomendaciones	42

A. Título anexo

43

Bibliografía

51

Índice de figuras

1.1. Gráfica de la función $g(x) = (\text{máx}\{0, x\})^2$	4
3.1. Conjunto factible para el ejemplo 1.	34
3.2. Evolución del Error vs número de iteraciones para el Ejemplo 1, usando el Algoritmo local.	36
3.3. Evolución del Error vs número de iteraciones para el Ejemplo 1, usando el Algoritmo global con $M = 10^2$	37
3.4. Evolución del Error vs número de iteraciones para el Ejemplo 1, usando el Algoritmo global con $M = 10^5$	38
3.5. Evolución del Error vs número de iteraciones para el Ejemplo 2, usando el Algoritmo local.	40
3.6. Evolución del Error vs número de iteraciones para el Ejemplo 2, usando el Algoritmo global con $M = 10^5$	40

Capítulo 1

Descripción del componente desarrollado

1.1. Objetivo general

El objetivo general de este trabajo es estudiar problemas de programación matemática con restricciones de complementariedad o MPCCs (Mathematical Programs with Complementary Constraints), cuyos sistemas de optimalidad se pueden formular como ecuaciones vectoriales no lineales, y su solución numérica se puede alcanzar a través de algoritmos de tipo Newton semi-suaves.

1.2. Objetivos específicos

1. Estudiar los sistemas de optimalidad de los MPCCs como ecuaciones vectoriales no lineales, los cuales se obtienen a partir de la técnica de *lifting* propuesta en [3].
2. Estudiar un algoritmo semisuave de Newton, propuesto en [1], para los sistemas de optimalidad de los MPCCs obtenidos en el Oe1.
3. Implementar el algoritmo anterior en MATLAB y analizar sus propiedades.

1.3. Alcance

El alcance de este trabajo es generar un código en el lenguaje computacional MATLAB para la implementación eficiente de un algoritmo de Newton semi-suave para la resolución de MPCCs. Para esto, debemos estudiar y comprender en detalle la formulación de tipo *lifting* para los MPCCs, propuesta en [3]. Posteriormente, requerimos analizar la formulación de los sistemas de optimalidad como ecuaciones vectoriales no lineales, con funciones semisuaves asociadas. Una vez que contemos con los sistemas desarrollados, estudiamos en detalle el algoritmo de tipo Newton semi-suave para resolver los sistemas de optimalidad de los MPCCs, basando nuestro estudio en [1]. En particular, verificaremos las propiedades de convergencia superlineal que se obtiene con información generalizada de segundo orden.

1.4. Marco teórico

1.4.1. Conceptos preliminares

Sean $q, r \in \mathbb{N}$ y $u, v \in \mathbb{R}^q$; vamos a denotar por $\langle u, v \rangle$ al producto interno usual en \mathbb{R}^q entre los vectores u y v , además, por $\|\cdot\|$ a la norma asociada. Si $I \subset \{1, \dots, q\}$, entonces u_I representa el vector con componentes u_i , donde $i \in I$. Además, mediante $\text{diag}(u)$ definimos la matriz cuadrada de orden $q \times q$ con los componentes del vector u en la diagonal y ceros en otro caso. Para una matriz A cualquiera, denotamos por A^T su transpuesta.

Diremos que la función $\Phi : \mathbb{R}^q \rightarrow \mathbb{R}^r$ es localmente Lipschitz continua en el punto $\bar{u} \in \mathbb{R}^q$ si existe $l > 0$ tal que

$$\|\Phi(u) - \Phi(\bar{u})\| \leq l \|u - \bar{u}\|$$

para cada $u \in \mathbb{R}^q$ en una vecindad de \bar{u} .

1.4.2. Problemas de minimización con restricciones de complementariedad (MPCCs)

En el presente trabajo vamos a estudiar el siguiente problema de minimización con restricciones de complementariedad (MPCC¹)

$$\text{mín } f(x) \quad \text{s.a.} \quad G(x) \geq 0, H(x) \geq 0, \langle G(x), H(x) \rangle = 0, \quad (1.1)$$

donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^m$ son funciones suaves, y $n, m \in \mathbb{N}$. Vamos a abordar este problema mediante una reformulación basada en la idea del "lifting MPCC". Para ello, consideremos el siguiente conjunto

$$D = \{(u, v) \in \mathbb{R}^2 : u \geq 0, v \geq 0, u \cdot v = 0\},$$

así, vemos que el conjunto de puntos factibles del problema (1.1) coincide con el conjunto

$$\Omega = \{x \in \mathbb{R}^n : (G_j(x), H_j(x)) \in D, j \in K\}, \quad (1.2)$$

donde $K := \{1, \dots, m\}$. Ahora, notemos que D es "no suave", en el sentido de que tiene un pico en el origen. Si introducimos una variable artificial $w \in \mathbb{R}$, podemos considerar una curva suave J en el espacio \mathbb{R}^3 de variables (u, v, w) , tal que la proyección ortogonal de J sobre el plano coincida con D . Esto lo hacemos por construcción, tal como lo realiza Oliver Stein (2010) en [3].

Elegimos una función $g : \mathbb{R} \rightarrow \mathbb{R}$ con las siguientes propiedades:

(P1) $g(x) = 0$ para cada $x \leq 0$,

(P2) $g(x) > 0$ para cada $x > 0$,

(P3) $g|_{[0, +\infty)}$ es biyectiva,

(P4) g es al menos continuamente diferenciable en \mathbb{R} .

Un ejemplo de una función que satisface las propiedades anteriores

¹Mathematical program with complementary constraints.

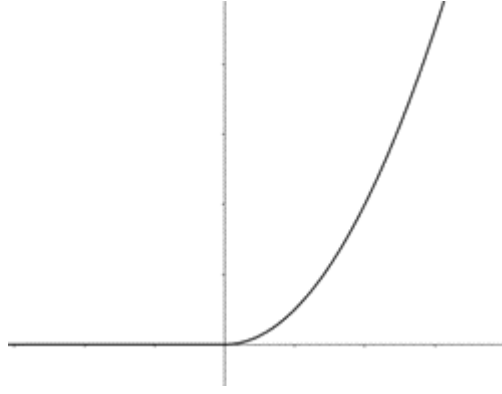


Figura 1.1: Gráfica de la función $g(x) = (\text{máx}\{0, x\})^2$

está definida por

$$g(x) = (\text{máx}\{0, x\})^s, \quad \text{con } s > 1.$$

La figura 1.1 nos muestra el caso para g con $s = 2$.

Por otra parte, definimos el conjunto

$$J = \{(u, v, w) \in \mathbb{R}^3 : u - g(-w) = 0, v - g(w) = 0\}.$$

Notemos que, J está descrito por dos restricciones de igualdad de clase C^1 gracias a la propiedad (P4), así, ésta es una curva suave. En lo que sigue, vamos a denotar con $\mathcal{P}_l(C)$ a la proyección ortogonal del conjunto $C \subset \mathbb{R}^n$ a \mathbb{R}^l , con $n > l$. Como podemos ver en [3], D es la proyección ortogonal de la curva J , es decir tenemos que

$$\mathcal{P}_2(J) = D. \tag{1.3}$$

La identidad anterior sugiere tratar el problema (1.1) mediante un enfoque *lifting*. Es así que definimos el siguiente conjunto

$$\begin{aligned} \Omega_p &= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : (G_j(x), H_j(x), y_j) \in J, \quad j \in K\} \\ &= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : G_j(x) = g(-y_j), H_j(x) = g(y_j), \quad j \in K\}, \end{aligned}$$

el cual nos conduce a la siguiente reformulación del problema (1.1)

$$\text{mín } f(x) \quad \text{s.a.} \quad (-\text{mín}\{0, y\})^s - G(x) = 0, (\text{máx}\{0, y\})^s - H(x) = 0, \tag{1.4}$$

donde las operaciones de tomar el máximo, mínimo y elevar a la potencia s son entendidas componente a componente. A fin de ver que resolver el problema (1.1) es equivalente a resolver (1.4), como se ve en [3], tenemos el siguiente lema:

Lema 1. $\mathcal{P}_n(\Omega_p) = \Omega$.

Demstración. Un punto x pertenece a $\mathcal{P}_n(\Omega_p)$ si y sólo si, existe $y \in \mathbb{R}^m$ tal que

$$(G_j(x), H_j(x), y_j) \in J, \quad \text{para cada } j \in K,$$

lo cual equivale a que el par $(G_j(x), H_j(x))$ pertenezca a $\mathcal{P}_2(S)$; y por la identidad (1.3), se sigue que $x \in \mathcal{P}_n(\Omega_p)$ si y sólo si

$$(G_j(x), H_j(x)) \in D, \quad \text{para cada } j \in K,$$

en otras palabras, $x \in \Omega$. ■

Observación 1. En la demostración del Lema 1 tenemos que el punto $y \in \mathbb{R}^m$, tal que $(x, y) \in \Omega_p$, está determinado de forma única por $x \in \mathbb{R}^n$. En efecto, gracias a la propiedad (P3) de g , podemos definir la función $w : D \rightarrow \mathbb{R}$ tal que, para cada $(u, v) \in D$

$$w(u, v) = \begin{cases} g^{-1}(v) & u = 0, v > 0 \\ -g^{-1}(u) & u \geq 0, v = 0, \end{cases}$$

y por tanto, $(u, v, w(u, v)) \in J$. Así, para cada $x \in \mathbb{R}^n$ y cada $j \in K$, con $(G_j(x), H_j(x)) \in D$, tenemos que si $y_j(x) = w(G_j(x), H_j(x))$, entonces se garantiza que $(G_j(x), H_j(x), y_j) \in J$ para cada $j \in K$; es decir, para cada $x \in \Omega$, existe un único $y := y(x) \in \mathbb{R}^m$ tal que $(x, y) \in \Omega_p$.

Por otra parte, el Lema 1 implica inmediatamente el siguiente resultado clásico, como se expone en [3]:

Proposición 1. Los minimizadores del problema (1.1) son las x -componentes de los minimizadores del problema (1.4).

En este trabajo vamos a considerar $s = 2$ en el problema (1.4), de modo

que nuestro propósito es resolver el siguiente problema

$$\min f(x) \quad \text{s.a.} \quad (\min\{0, y\})^2 - G(x) = 0, \quad (\max\{0, y\})^2 - H(x) = 0. \quad (1.5)$$

1.4.3. Algunos conceptos básicos del MPCC y MPCC *lifted*

Sea $\bar{x} \in \Omega$ (un punto factible del problema (1.1)), tal como lo hacen A. Izmailov, A. Pogosyan y M. Solodov en [1], consideremos lo siguiente: definimos los siguientes conjuntos de índices

$$\begin{aligned} I_G &= I_G(\bar{x}) = \{i \in K : G_i(\bar{x}) = 0\}, \\ I_H &= I_H(\bar{x}) = \{i \in K : H_i(\bar{x}) = 0\}, \\ I_0 &= I_G \cap I_H, \end{aligned}$$

donde I_0 es llamado el conjunto de índices degenerados. Notemos que, gracias a que $\bar{x} \in \Omega$, se tiene que $I_G \cup I_H = K$. Ahora, si $I_0 = \emptyset$ decimos que en \bar{x} se cumple la *complementariedad estricta de nivel inferior*. El MPCC-Lagrangiano especial para el problema (1.1) viene dado por

$$\mathcal{L}(x, \mu) = f(x) - \langle \mu_G, G(x) \rangle - \langle \mu_H, H(x) \rangle,$$

donde $x \in \mathbb{R}^n$ y $\mu = (\mu_G, \mu_H) \in \mathbb{R}^m \times \mathbb{R}^m$. En lo que sigue, consideremos la siguiente definición:

Definición 1. Un punto $\bar{x} \in \Omega$ se dice **débilmente estacionario** si existe $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H) \in \mathbb{R}^m \times \mathbb{R}^m$ tal que

$$\frac{\partial \mathcal{L}}{\partial x}(\bar{x}, \bar{\mu}) = 0, \quad (\bar{\mu}_G)_{I_H \setminus I_G} = 0, \quad (\bar{\mu}_H)_{I_G \setminus I_H} = 0. \quad (1.6)$$

Además, si se cumple también que

$$(\bar{\mu}_G)_{I_0} \geq 0, \quad (\bar{\mu}_H)_{I_0} \geq 0, \quad (1.7)$$

decimos que \bar{x} es **fuertemente estacionario**. En este caso, llamamos a $\bar{\mu}$ un **MPCC-multiplicador** asociado al punto fuertemente estacionario \bar{x} .

Por otro lado, recordemos que la condición de calificación de indepen-

dencia lineal MPCC [MPCC-LICQ] se cumple en un punto factible \bar{x} del problema (1.1) si, dados $i \in I_G$ y $j \in I_H$ entonces

$$\nabla G_i(\bar{x}), \nabla H_j(\bar{x})$$

son linealmente independientes.

Definimos el *cono crítico* para el problema (1.1) en el punto \bar{x} como el conjunto

$$C(\bar{x}) := \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} G'_{I_G \setminus I_H}(\bar{x})\xi = 0, H'_{I_H \setminus I_G}(\bar{x})\xi = 0, G'_{I_0}(\bar{x})\xi \geq 0, H'_{I_0}(\bar{x})\xi \geq 0, \\ \langle \nabla f(\bar{x}), \xi \rangle \leq 0 \end{array} \right\},$$

y también el cono

$$C_2(\bar{x}) := \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} G'_{I_G \setminus I_H}(\bar{x})\xi = 0, H'_{I_H \setminus I_G}(\bar{x})\xi = 0, G'_{I_0}(\bar{x})\xi \geq 0, H'_{I_0}(\bar{x})\xi \geq 0, \\ \langle \nabla f(\bar{x}), \xi \rangle \leq 0, \langle \nabla G_i(\bar{x}), \xi \rangle \langle \nabla H_i(\bar{x}), \xi \rangle = 0 \forall i \in I_0 \end{array} \right\},$$

de donde, evidentemente $C_2(\bar{x}) \subset C(\bar{x})$. A continuación, daremos una caracterización de estos conos, en el caso en que \bar{x} sea un punto fuertemente estacionario.

Lema 2. *Sea \bar{x} un punto fuertemente estacionario, entonces para cualquier MPCC-multiplicador $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H)$, se tiene que el cono crítico $C(\bar{x})$ coincide con el conjunto*

$$\left\{ \xi \in \mathbb{R}^n : \begin{array}{l} G'_{I_G \setminus I_H}(\bar{x})\xi = 0, H'_{I_H \setminus I_G}(\bar{x})\xi = 0, G'_{I_0}(\bar{x})\xi \geq 0, H'_{I_0}(\bar{x})\xi \geq 0, \\ (\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle = 0, (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle = 0, \forall i \in I_0 \end{array} \right\},$$

y también tenemos que

$$C_2(\bar{x}) = \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} G'_{I_G \setminus I_H}(\bar{x})\xi = 0, H'_{I_H \setminus I_G}(\bar{x})\xi = 0, G'_{I_0}(\bar{x})\xi \geq 0, H'_{I_0}(\bar{x})\xi \geq 0, \\ (\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle = 0, (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle = 0, \\ \langle \nabla G_i(\bar{x}), \xi \rangle \langle \nabla H_i(\bar{x}), \xi \rangle = 0, \forall i \in I_0 \end{array} \right\}.$$

Demostración. De la definición del MPCC-Lagrangiano para el problema (1.1), tenemos que

$$\frac{\partial \mathcal{L}}{\partial x}(x, \mu) = \nabla f(x) - G'(x)^T \bar{\mu}_G - H'(x)^T \bar{\mu}_H,$$

puesto que \bar{x} es fuertemente estacionario, de la primera identidad de (1.6) se tiene que, para cualquier MPCC-multiplicador $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H)$ asociado a \bar{x} , tenemos la siguiente identidad

$$G'(\bar{x})^T \bar{\mu}_G + H'(\bar{x})^T \bar{\mu}_H = \nabla f(\bar{x}). \quad (1.8)$$

Por otro lado, sea $\xi \in \mathbb{R}^n$. Notemos que, la condición

$$\langle \nabla f(\bar{x}), \xi \rangle \leq 0,$$

equivale, gracias a la identidad (1.8), a la desigualdad

$$\langle G'(\bar{x})^T \bar{\mu}_G, \xi \rangle + \langle H'(\bar{x})^T \bar{\mu}_H, \xi \rangle \leq 0,$$

en otras palabras

$$(\bar{\mu}_G)^T G'(\bar{x}) \xi + \bar{\mu}_H^T H'(\bar{x}) \xi \leq 0.$$

Ahora bien, si suponemos que $G'_{I_G \setminus I_H}(\bar{x}) \xi = 0$ y $H'_{I_H \setminus I_G}(\bar{x}) \xi = 0$ entonces la última desigualdad nos da que

$$\sum_{i \in I_H} (\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle + \sum_{i \in I_G} (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle \leq 0,$$

además, gracias a las dos últimas identidades de (1.6) se obtiene que

$$\sum_{i \in I_0} (\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle + \sum_{i \in I_0} (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle \leq 0.$$

Sin embargo, si $G'_{I_0}(\bar{x}) \xi \geq 0$ y $H'_{I_0}(\bar{x}) \xi \geq 0$, gracias a (1.7), vemos que la parte izquierda de la desigualdad anterior comprende la suma de términos no negativos; así obtenemos que

$$(\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle = 0, \quad (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle = 0,$$

para cada $i \in I_0$, lo que prueba que

$$C(\bar{x}) = \left\{ \xi \in \mathbb{R}^n : \begin{array}{l} G'_{I_G \setminus I_H}(\bar{x}) \xi = 0, H'_{I_H \setminus I_G}(\bar{x}) \xi = 0, G'_{I_0}(\bar{x}) \xi \geq 0, H'_{I_0}(\bar{x}) \xi \geq 0, \\ (\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle = 0, (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle = 0, \forall i \in I_0 \end{array} \right\}.$$

La última identidad de este lema se sigue inmediatamente de la definición de $C_2(\bar{x})$ y lo hecho anteriormente. ■

Por otro lado, introduzcamos la siguiente definición, la cual nos será útil en lo que sigue.

Definición 2. Se dice que la condición suficiente especial de segundo orden MPCC [MPCC-SOSC] se cumple en un punto fuertemente estacionario \bar{x} con un MPCC-multiplicador $\bar{\mu}$ asociado, si para cada $\xi \in C(\bar{x}) \setminus \{0\}$

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu})\xi, \xi \right\rangle > 0. \quad (1.9)$$

Mientras tanto, la condición suficiente de segundo orden por partes más débil consiste en que

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu})\xi, \xi \right\rangle > 0, \quad (1.10)$$

para cada $\xi \in C_2(\bar{x}) \setminus \{0\}$.

Observación 2. En realidad, como se puede ver en [6] la condición (1.10) (y, más aún (1.9)) es suficiente para la optimalidad local de \bar{x} .

En general, (1.10) es una condición más natural que la dada por (1.9). Esto se debe a que si se considera la desigualdad no estricta en la condición (1.10) obtenemos una condición necesaria para la optimalidad bajo [MPCC-LICQ], mientras que con (1.9) no podemos realizar algo similar.

Consideremos ahora la siguiente definición.

Definición 3. Sea \bar{x} un punto fuertemente estacionario con un MPCC-multiplicador $\bar{\mu}$ asociado. Decimos que la condición de complementariedad estricta de más alto nivel [ULSCC] se cumple en $\bar{\mu}$ si

$$(\bar{\mu}_G)_{I_0} > 0, \quad (\bar{\mu}_H)_{I_0} > 0. \quad (1.11)$$

Notemos que, bajo [ULSCC] se cumple que

$$K(\bar{x}) = C(\bar{x}) = C_2(\bar{x}),$$

donde

$$K(\bar{x}) = \{\xi \in \mathbb{R}^n : G'_{I_G}(\bar{x})\xi = 0, H'_{I_H}(\bar{x})\xi = 0\}. \quad (1.12)$$

En efecto, gracias al Lema 2, para $\xi \in C(\bar{x})$ se cumple que

$$(\bar{\mu}_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle = 0, \quad (\bar{\mu}_H)_i \langle \nabla H_i(\bar{x}), \xi \rangle = 0,$$

para cada $i \in I_0$, que junto a las desigualdades expuestas en (1.11), nos da que

$$\langle \nabla G_i(\bar{x}), \xi \rangle = 0, \quad \langle \nabla H_i(\bar{x}), \xi \rangle = 0,$$

para cada $i \in I_0$, lo que permite concluir que $\xi \in C_2(\bar{x})$, demostrando así que $C(\bar{x}) = C_2(\bar{x})$. Más aún, podemos escribir

$$G'_{I_0}(\bar{x})\xi = 0, \quad H'_{I_0}(\bar{x})\xi = 0,$$

y como $G'_{I_G \setminus I_H}(\bar{x})\xi = 0$ y $H'_{I_H \setminus I_G}(\bar{x})\xi = 0$, obtenemos que

$$G'_{I_G}(\bar{x})\xi = 0, \quad H'_{I_H}(\bar{x})\xi = 0,$$

es decir, $K(\bar{x}) = C(\bar{x}) = C_2(\bar{x})$. En este caso, las expresiones (1.9) y (1.10) son equivalentes y se escriben simplemente como

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu})\xi, \xi \right\rangle > 0, \quad (1.13)$$

para cada $\xi \in K(\bar{x}) \setminus \{0\}$.

Destinemos nuestra atención ahora a la reformulación *lifted* MPCC dada en (1.5). Notemos primero que, gracias a la Observación 1, el punto (\bar{x}, \bar{y}) es factible para el problema (1.5) si, y solo si

$$\bar{y}_{I_H \setminus I_G} = -(G_{I_H \setminus I_G}(\bar{x}))^{1/2}, \quad \bar{y}_{I_G \setminus I_H} = (H_{I_G \setminus I_H}(\bar{x}))^{1/2}, \quad \bar{y}_{I_0} = 0. \quad (1.14)$$

En efecto, de la Observación 1, tenemos que (\bar{x}, \bar{y}) es factible para el problema (1.5), si y solo si, para cada $j \in K$, $(G_j(\bar{x}), H_j(\bar{x})) \in D$ y además

$$\bar{y}_j(\bar{x}) = w(G_j(\bar{x}), H_j(\bar{x})),$$

donde, en este caso, la función w está definida por

$$w(u, v) = \begin{cases} v^{1/2} & u = 0, v > 0 \\ -u^{1/2} & u \geq 0, v = 0. \end{cases}$$

En consecuencia, si $j \in I_H \setminus I_G$, entonces $G_j(\bar{x}) > 0$, de donde

$$\bar{y}_j(\bar{x}) = -(G_j(\bar{x}))^{1/2},$$

probando así la primera identidad de (1.14); en cambio, si $j \in I_G \setminus I_H$, entonces $H_j(\bar{x}) > 0$, por lo que

$$\bar{y}_j(\bar{x}) = (H_j(\bar{x}))^{1/2},$$

lo que demuestra la segunda identidad de (1.14); la última se sigue inmediatamente pues, si $j \in I_0$, entonces $G_j(\bar{x}) = H_j(\bar{x}) = 0$.

Por otra parte, como se puede ver en [3] tenemos también que [MPCC-LICQ] se satisface en un punto $\bar{x} \in \Omega$ (factible para el problema (1.1)) si y sólo si se satisface la independencia lineal de los gradientes de las restricciones del problema *lifted* (1.5) en el punto (\bar{x}, \bar{y}) .

Definamos ahora el Lagrangiano usual del problema *lifted* (1.5):

$$L(x, y, \lambda) = f(x) + \langle \lambda_G, (\min\{0, y\})^2 - G(x) \rangle + \langle \lambda_H, (\max\{0, y\})^2 - H(x) \rangle,$$

donde $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ y $\lambda = (\lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^m$. Así, el sistema de optimalidad de Lagrange que caracteriza los puntos estacionarios del problema *lifted* (1.5) y los multiplicadores asociados, está dado por

$$\frac{\partial L}{\partial x}(x, y, \lambda) = 0, \quad \frac{\partial L}{\partial y}(x, y, \lambda) = 0, \quad (\min\{0, y\})^2 - G(x) = 0, \quad (1.15)$$

$$(\max\{0, y\})^2 - H(x) = 0, \quad (1.16)$$

donde

$$\frac{\partial L}{\partial x}(x, y, \lambda) = \frac{\partial \mathcal{L}}{\partial x}(x, \lambda), \quad (1.17)$$

y para cada $i \in K$

$$\frac{\partial L}{\partial y_i}(x, y, \lambda) = 2(\lambda_G)_i \min\{0, y_i\} + 2(\lambda_H)_i \max\{0, y_i\}. \quad (1.18)$$

Observe que, para cualquier $i \in K$, el lado derecho de la identidad anterior no es diferenciable en los puntos (x, y, λ) , tales que $y_i = 0$. Más adelante, veremos como abordar este problema de regularidad.

La correspondencia entre puntos estacionarios y multiplicadores para el problema original (1.1) y su reformulación *lifted* (1.5) viene dada por la siguiente proposición.

Proposición 2. *Sea $\bar{x} \in \mathbb{R}^n$ un punto factible del problema (1.1), tal que f, G y H son diferenciables en este punto. Si \bar{x} es un punto fuertemente estacionario de (1.1) y $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H)$ es un MPCC-multiplicador asociado, entonces el punto (\bar{x}, \bar{y}) , con \bar{y} dado por la expresión (1.14), es un punto estacionario del problema *lifted* (1.5), de modo que $\bar{\lambda} = \bar{\mu}$ es un multiplicador de Lagrange asociado.*

*Por otro lado, si (\bar{x}, \bar{y}) es un punto estacionario del problema *lifted* (1.5), entonces \bar{x} es un punto débilmente estacionario del problema (1.1). Además, si existe un multiplicador de Lagrange $\bar{\lambda} = (\bar{\lambda}_G, \bar{\lambda}_H)$ asociado a (\bar{x}, \bar{y}) y tal que*

$$(\bar{\lambda}_G)_{I_0} \geq 0, \quad (\bar{\lambda}_H)_{I_0} \geq 0,$$

entonces \bar{x} es un punto fuertemente estacionario del problema (1.1) y $\bar{\mu} = \bar{\lambda}$ es un MPCC-multiplicador asociado.

Demostración. Consideremos $\bar{x} \in \Omega$, un punto fuertemente estacionario del problema (1.1), con $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H)$ un MPCC-multiplicador asociado. Hemos visto que (\bar{x}, \bar{y}) es un punto factible del problema (1.5), donde \bar{y} está dado por (1.14). Por tanto, para probar que (\bar{x}, \bar{y}) es un punto estacionario del problema *lifted* (1.5), basta mostrar que

$$\frac{\partial L}{\partial x}(\bar{x}, \bar{y}, \bar{\mu}) = 0, \quad \frac{\partial L}{\partial y}(\bar{x}, \bar{y}, \bar{\mu}) = 0.$$

Puesto que \bar{x} es fuertemente estacionario y $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H)$ es su MPCC-

multiplicador asociado, por la identidad (1.17) se sigue que

$$\frac{\partial L}{\partial x}(\bar{x}, \bar{y}, \bar{\mu}) = 0.$$

Luego, utilizando la identidad (1.18), sabemos que para cada $i \in K$

$$\frac{\partial L}{\partial y_i}(\bar{x}, \bar{y}, \bar{\mu}) = 2(\bar{\mu}_G)_i \min\{0, \bar{y}_i\} + 2(\bar{\mu}_H)_i \max\{0, \bar{y}_i\}.$$

Así, podemos considerar los siguientes casos: primero, si $i \in I_H \setminus I_G$, como \bar{x} es fuertemente estacionario y $\bar{\mu} = (\bar{\mu}_G, \bar{\mu}_H)$ es su MPCC-multiplicador asociado, $(\bar{\mu}_G)_i = 0$. Además, por la definición de \bar{y} dada en la expresión (1.14), obtenemos

$$\bar{y}_i = -(G_i(\bar{x}))^{1/2} < 0,$$

de donde $\max\{0, \bar{y}_i\} = 0$, lo cual implica que

$$\frac{\partial L}{\partial y_i}(\bar{x}, \bar{y}, \bar{\mu}) = 0.$$

Ahora, la demostración de los casos $i \in I_G \setminus I_H$ e $i \in I_0$ resulta ser completamente análoga a los hecho anteriormente, gracias a la definición de \bar{y} . Así, terminamos la primera parte de la demostración de la Proposición 2.

Para la segunda parte, consideramos (\bar{x}, \bar{y}) un punto estacionario del problema *lifted* (1.5) con multiplicador de Lagrange $\bar{\lambda} = (\bar{\lambda}_G, \bar{\lambda}_H)$ asociado. Así, gracias a la identidad (1.17) tenemos que

$$\frac{\partial \mathcal{L}}{\partial x}(\bar{x}, \bar{\lambda}) = 0,$$

por lo que basta mostrar que

$$(\bar{\lambda}_G)_{I_H \setminus I_G} = 0, \quad (\bar{\lambda}_H)_{I_G \setminus I_H} = 0,$$

para obtener el resultado deseado.

Como (\bar{x}, \bar{y}) es un punto estacionario del problema (1.5), por la segunda identidad de la expresión (1.15) y la identidad (1.18) tenemos que

$$(\bar{\lambda}_G)_i \min\{0, y_i\} + (\bar{\lambda}_H)_i \max\{0, y_i\} = 0.$$

Vemos que, si $i \in I_H \setminus I_G$, nuevamente por la definición de \bar{y} dada en la expresión (1.14), tenemos que $y_i = -(G_i(\bar{x}))^{1/2} < 0$, de donde $(\bar{\lambda}_G)_i y_i = 0$, es decir, $(\bar{\lambda}_G)_i = 0$. De manera similar, si $i \in I_G \setminus I_H$, por la definición de \bar{y} tenemos que $y_i = (H_i(\bar{x}))^{1/2} > 0$, de donde $(\bar{\lambda}_H)_i y_i = 0$, es decir, $(\bar{\lambda}_H)_i = 0$.

Finalmente, si suponemos que

$$(\bar{\lambda}_G)_{I_0} \geq 0, \quad (\bar{\lambda}_H)_{I_0} \geq 0,$$

entonces \bar{x} es un punto fuertemente estacionario del problema (1.1), donde $\bar{\mu} = \bar{\lambda}$ es un MPCC-multiplicador asociado. ■

Capítulo 2

Metodología

2.1. Método de Newton Semi-suave (SSN) para el problema *lifted* MPCC

A continuación, vamos a construir un algoritmo tipo Newton semi-suave, con un enfoque en los B-subdiferenciables, daremos las condiciones suficientes que permitan garantizar la convergencia, y también obtendremos información sobre el radio de convergencia del método.

Sean $q, r \in \mathbb{N}$. Consideremos la función $\Phi : \mathbb{R}^q \rightarrow \mathbb{R}^r$, la cual es localmente Lipschitz continua en el punto $u \in \mathbb{R}^q$. Definimos el *B-diferencial* de Φ en el punto $u \in \mathbb{R}^q$ como el conjunto

$$\partial_B \Phi(u) = \{\Lambda \in \mathbb{R}^{r \times q} : \text{Existe } \{u_j\}_j \subset \mathcal{D}_\Phi \text{ tal que } u_j \rightarrow u, \Phi'(u_j) \rightarrow \Lambda\},$$

donde \mathcal{D}_Φ es el conjunto de puntos en los cuales Φ es diferenciable. Entonces, el *Jacobiano generalizado de Clarke* viene dado por

$$\partial \Phi(u) = \text{conv} \partial_B \Phi(u),$$

donde $\text{conv} X$ representa la envoltura convexa del conjunto X .

Además, se dice que Φ es **semi-suave** en $u \in \mathbb{R}^q$, si es localmente Lipschitz continua alrededor de u , admite todas las derivadas direccionales

en u , y satisface la condición

$$\sup_{\Lambda \in \partial\Phi(u)} \|\Phi(u+v) - \Phi(u) - \Lambda v\| = o(\|v\|).$$

Más aún, si se satisface la condición más fuerte

$$\sup_{\Lambda \in \partial\Phi(u)} \|\Phi(u+v) - \Phi(u) - \Lambda v\| = O(\|v\|^2),$$

entonces decimos que Φ es **fuertemente semi-suave en** u . Finalmente, recordemos también que $\Phi : \mathbb{R}^q \rightarrow \mathbb{R}^q$ es dicha **BD-regular** en un punto $\bar{u} \in \mathbb{R}^q$, si todas las matrices $\Lambda \in \partial_B\Phi(\bar{u})$ son no singulares.

Observación 3. Como menciona Liqun Qi(1993) en [2] la composición de funciones (fuertemente) semi-suaves también es (fuertemente) semi-suave; por ello, si $\Phi_1, \Phi_2 : \mathbb{R}^q \rightarrow \mathbb{R}^r$ son (fuertemente) semi-suaves en $u \in \mathbb{R}^q$, entonces también lo son los mapas $\Phi_1(\cdot) + \Phi_2(\cdot)$, $\langle \Phi_1(\cdot), \Phi_2(\cdot) \rangle$, $\min\{\Phi_1(\cdot), \Phi_2(\cdot)\}$, $\max\{\Phi_1(\cdot), \Phi_2(\cdot)\}$ y $(\Phi_1(\cdot), \Phi_2(\cdot))$.

El esquema iterativo SSN¹ para la ecuación

$$\Phi(u) = 0, \tag{2.1}$$

donde $\Phi : \mathbb{R}^q \rightarrow \mathbb{R}^q$, viene dado por

$$\Lambda_k(u^{k+1} - u^k) = -\Phi(u^k), \quad \Lambda_k \in \partial_B\Phi(u^k), \tag{2.2}$$

para cada $k \in \mathbb{N}_0$.

El siguiente es un resultado clásico de convergencia local usando SSN, cuya demostración se puede ver en [2].

Teorema 1. Sea $\Phi : \mathbb{R}^q \rightarrow \mathbb{R}^q$ semi-suave en $\bar{u} \in \mathbb{R}^q$. Supongamos que \bar{u} es una solución de la ecuación (2.1) tal que Φ es BD-regular en \bar{u} .

Entonces, para cualquier punto de partida $u^0 \in \mathbb{R}^q$ suficientemente cercano a \bar{u} , el esquema iterativo SSN definido por la expresión (2.2) define una única sucesión (u^k) que converge a \bar{u} . Además, la tasa de convergencia es superlineal, y es cuadrática si Φ es fuertemente semi-suave en \bar{u} .

¹Semismooth Newton Method.

A continuación, vamos a estudiar SSN aplicado al sistema de optimalidad de Lagrange de la reformulación *lifted* MPCC, a fin de hallar puntos estacionarios del problema (1.5). Teniendo en cuenta la identidad (1.17), este sistema toma la forma de la ecuación (2.1) si definimos $\Phi : \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m) \longrightarrow \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$ como

$$\Phi(u) = \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial x}(x, \lambda) \\ \frac{\partial \mathcal{L}}{\partial y}(x, y, \lambda) \\ (\min\{0, y\})^2 - G(x) \\ (\max\{0, y\})^2 - H(x) \end{pmatrix}, \quad (2.3)$$

donde $u = (x, y, \lambda)$, $\lambda = (\lambda_G, \lambda_H)$.

El siguiente lema nos va a permitir caracterizar a las matrices del B-diferencial de Φ en u .

Lema 3. Sean $\Phi : \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m) \longrightarrow \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$ definido como en la identidad (2.3) y $u = (x, y, \lambda) \in \text{Dom}(\Phi)$. El B-diferencial de Φ en u consiste de todas las matrices de la forma

$$\Lambda = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda) & 0 & -(G'(x))^T & -(H'(x))^T \\ 0 & 2A(y, \lambda) & 2B_{\min}(y) & 2B_{\max}(y) \\ -G'(x) & 2B_{\min}(y) & 0 & 0 \\ -H'(x) & 2B_{\max}(y) & 0 & 0 \end{pmatrix}, \quad (2.4)$$

donde

$$\begin{aligned} A(y, \lambda) &= \text{diag}(a(y, \lambda)), & B_{\min}(y) &= \text{diag}(\min\{0, y\}), \\ B_{\max}(y) &= \text{diag}(\max\{0, y\}), \end{aligned} \quad (2.5)$$

y el vector $a(y, \lambda) \in \mathbb{R}^m$ viene dado por

$$a_i = \begin{cases} (\lambda_G)_i & y_i \leq 0 \\ (\lambda_G)_i \circ (\lambda_H)_i & y_i = 0 \\ (\lambda_H)_i & y_i > 0, \end{cases} \quad (2.6)$$

para cada $i \in K$.

Demostración. Se tiene por un cálculo directo, la identidad (1.18) y la definición de Φ , dada en la expresión (2.3). ■

Sea \bar{x} un punto fuertemente estacionario estacionario del problema (1.1), $\bar{\mu}$ un MPCC-multiplicador asociado e \bar{y} definido por la expresión (1.14). Por el Lema 3 tenemos que el *B-diferencial* de Φ en el punto $\bar{u} = (\bar{x}, \bar{y}, \bar{\mu})$ consiste de todas las matrices de la forma

$$\Lambda = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu}) & 0 & -(G'(\bar{x}))^T & -(H'(\bar{x}))^T \\ 0 & 2A & 2B_{\text{mín}} & 2B_{\text{máx}} \\ -G'(\bar{x}) & 2B_{\text{mín}} & 0 & 0 \\ -H'(\bar{x}) & 2B_{\text{máx}} & 0 & 0 \end{pmatrix},$$

donde

$$A = \text{diag}(a), \quad B_{\text{mín}} = \text{diag}(b_{\text{mín}}), \quad B_{\text{máx}} = \text{diag}(b_{\text{máx}}),$$

con el vector $a \in \mathbb{R}^m$, cuyas componentes vienen dadas por

$$a_i = \begin{cases} 0 & i \in K \setminus I_0, \\ (\bar{\mu}_G)_i \circ (\bar{\mu}_H)_i & i \in I_0, \end{cases} \quad (2.7)$$

y los vectores $b_{\text{mín}}$ y $b_{\text{máx}}$ está definidos como

$$\begin{aligned} (b_{\text{mín}})_{I_H \setminus I_G} &= -(G_{I_H \setminus I_G}(\bar{x}))^{1/2}, & (b_{\text{mín}})_{I_G} &= 0, \\ (b_{\text{máx}})_{I_G \setminus I_H} &= (H_{I_G \setminus I_H}(\bar{x}))^{1/2}, & (b_{\text{máx}})_{I_H} &= 0. \end{aligned} \quad (2.8)$$

En efecto, gracias a la definición de $a(\bar{y}, \bar{\mu})$, dada en (2.6), de \bar{y} dada en (1.14) y la expresión (1.6), tenemos que $a_i = 0$ para $i \in K \setminus I_0$; en cambio, si $i \in I_0$, entonces $a_i = (\bar{\mu}_G)_i \circ a_i = (\bar{\mu}_H)_i$. Así mismo, por la definición de \bar{y} , se sigue de inmediato que los vectores $b_{\text{mín}}$ y $b_{\text{máx}}$ está definidos como en la expresión (2.8).

Así, tenemos que SSN para el sistema de optimalidad de Lagrange del problema *lifted* (1.5) se resume en el siguiente algoritmo:

Algoritmo 1. 1. Tomar $u^0 = (x^0, y^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$ y sea $k = 0$.

2. Calcular una matriz $\Lambda_k := \Lambda$, de acuerdo a las expresiones (2.4), (2.5) y (2.6), con $(x, y, \lambda) = (x^0, y^0, \lambda^0)$. Luego, calcular $u^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$ como solución del sistema lineal

$$\Lambda_k u = \Lambda_k u^k - \Phi(u^k), \quad (2.9)$$

donde Φ está definido como en la identidad (2.3).

3. Sea $k = k + 1$ y volver al paso 2.

De acuerdo con el Teorema 1, la convergencia local superlineal (respectivamente cuadrática) del Algoritmo 1 a la solución $\bar{u} = (\bar{x}, \bar{y}, \bar{\mu})$ de la ecuación (2.1) se conseguiría si podemos mostrar semi-suavidad (respectivamente fuerte) y BD-regularidad de Φ en \bar{u} .

El siguiente resultado nos proporciona las condiciones suficientes para la semi-suavidad (fuerte) de Φ en \bar{u}

Proposición 3. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciables alrededor de $\bar{x} \in \mathbb{R}^n$, tal que sus derivadas son semi-suaves en este punto. Entonces, para cualquier $\lambda = (\lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^m$ e $y \in \mathbb{R}^m$, la aplicación Φ definida en la identidad (2.3) es semi-suave en el punto (\bar{x}, y, λ) .

Demostración. Sean $\lambda = (\lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^m$ e $y \in \mathbb{R}^m$. Observemos que, gracias a la Observación 3, para mostrar que Φ es semi-suave en $u = (\bar{x}, y, \lambda)$, basta probar que cada componente lo es en u .

Sabemos que, para $\mu = (\mu_G, \mu_H) \in \mathbb{R}^m \times \mathbb{R}^m$, la derivada respecto de x del Lagrangeano del problema (1.1) viene dado por

$$\frac{\partial \mathcal{L}}{\partial x}(x, \mu) = \nabla f(x) - G'(x)^T \mu_G - H'(x)^T \mu_H.$$

Como f, G y H poseen derivadas semi-suaves en \bar{x} , por la Observación 3, tenemos que cada término de la parte derecha de la identidad anterior es semi-suave en u , con lo cual también lo es $\frac{\partial \mathcal{L}}{\partial x}(x, \mu)$. Por el mismo razonamiento, se tiene que

$$\frac{\partial L}{\partial y}(x, y, \lambda), \quad (\min\{0, y\})^2 - G(x), \quad (\max\{0, y\})^2 - H(x),$$

son semi-suaves en u , con lo que tenemos lo requerido. ■

Ahora bien, veamos las condiciones suficientes para asegurar que Φ es BD-regular en \bar{u} .

Proposición 4. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dos veces diferenciables en el punto $\bar{x} \in \mathbb{R}^n$, el cual es fuertemente estacionario para el problema (1.1), de modo que se satisface [MPCC-LICQ] y [MPCC-SOSC]. Sea $\bar{\mu}$ un MPCC-multiplicador asociado a \bar{x} , de modo que se cumple [ULSCC].

Entonces, Φ definido en (2.3) es BD-regular en $\bar{u} = (\bar{x}, \bar{y}, \bar{\mu})$, con \bar{y} definido en la expresión (1.14).

Demostración. Sean $\Lambda \in \partial\Phi_B(\bar{u})$ y $v = (\xi, \eta, \zeta_G, \zeta_H)$, con $\xi \in \mathbb{R}^n$, $\eta \in \mathbb{R}^m$, $\zeta_G \in \mathbb{R}^m$, $\zeta_H \in \mathbb{R}^m$, tales que $\Lambda v = 0$. Así, tenemos que

$$\begin{aligned} \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu})\xi - (G'(\bar{x}))^T \zeta_G - (H'(\bar{x}))^T \zeta_H &= 0, \\ \text{diag}(a)\eta + \text{diag}(b_{\min})\zeta_G + \text{diag}(b_{\max})\zeta_H &= 0, \\ -G'(\bar{x})\xi + 2\text{diag}(b_{\min})\eta &= 0, \\ -H'(\bar{x})\xi + 2\text{diag}(b_{\max})\eta &= 0. \end{aligned} \tag{2.10}$$

De la segunda identidad de (2.10), para cada $i \in K$

$$a_i \eta_i + (b_{\min})_i (\zeta_G)_i + (b_{\max})_i (\zeta_H)_i = 0,$$

de donde, podemos considerar los siguientes casos: si $i \in I_G \setminus I_H$, entonces por la definición de a y b_{\min} , dadas en (2.7) y (2.8) respectivamente, $a_i = (b_{\min})_i = 0$; y como $(b_{\max})_i = (H_i(\bar{x}))^{1/2} > 0$, obtenemos que $(\zeta_H)_i = 0$. Por un razonamiento análogo, se sigue que si $i \in I_H \setminus I_G$, entonces $(\zeta_G)_i = 0$; y por último, para $i \in I_0$, gracias a que se cumple [ULSCC] en $\bar{\mu}$, nos da que $\eta_i = 0$. En resumen

$$(\zeta_G)_{I_H \setminus I_G} = 0, \quad (\zeta_H)_{I_G \setminus I_H} = 0, \quad \eta_{I_0} = 0. \tag{2.11}$$

Luego, de la tercera identidad de (2.10), por la definición de b_{\min} , tenemos que

$$\begin{aligned} -\langle \nabla G_i(\bar{x}), \xi \rangle - 2(G_i(\bar{x}))^{1/2} \eta_i &= 0, \quad i \in I_H \setminus I_G, \\ G'_{I_G}(\bar{x})\xi &= 0, \end{aligned} \tag{2.12}$$

similarmente, de la cuarta identidad de (2.10) y la definición de b_{\max} , se sigue que

$$\begin{aligned} -\langle \nabla H_i(\bar{x}), \xi \rangle + 2(H_i(\bar{x}))^{1/2} \eta_i &= 0, \quad i \in I_G \setminus I_H, \\ H'_{I_H}(\bar{x})\xi &= 0. \end{aligned} \tag{2.13}$$

Ahora, de las segundas identidades de (2.12) y (2.13), concluimos que $\xi \in K(\bar{x})$. Así mismo, de la primera identidad de (2.10), tenemos que

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu})\xi, \xi \right\rangle - \langle (G'(\bar{x}))^T \zeta_G, \xi \rangle - \langle (H'(\bar{x}))^T \zeta_H, \xi \rangle = 0,$$

con lo cual podemos escribir

$$\langle (G'(\bar{x}))^T \zeta_G, \xi \rangle = \langle \zeta_G, (G'(\bar{x}))\xi \rangle = \sum_{i \in I_H \setminus I_G} (\zeta_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle + \sum_{i \in I_G} (\zeta_G)_i \langle \nabla G_i(\bar{x}), \xi \rangle.$$

Gracias a la primera identidad de (2.11) y como $\xi \in K(\bar{x})$ se obtiene

$$\langle (G'(\bar{x}))^T \zeta_G, \xi \rangle = 0,$$

análogamente, por la segunda identidad de (2.11), conseguimos

$$\langle (H'(\bar{x}))^T \zeta_H, \xi \rangle = 0,$$

de donde

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu}) \xi, \xi \right\rangle = 0,$$

y puesto que se cumple [MPCC-SOSC], la identidad (1.13) nos permite deducir que $\xi = 0$. Así, junto a las dos primeras identidades de (2.11) concluimos que la primera identidad de (2.10) es equivalente a

$$(G'_{I_G}(\bar{x}))^T (\zeta_G)_{I_G} + (H'_{I_H}(\bar{x}))^T (\zeta_H)_{I_H} = 0,$$

y, por último, como se satisface [MPCC-LICQ] en \bar{x} , se colige que

$$(\zeta_G)_{I_G} = (\zeta_H)_{I_H} = 0.$$

Por otra parte, como $\xi = 0$, de las dos primeras identidades de (2.12) y (2.13), obtenemos

$$(\eta)_{I_H \setminus I_G} = (\eta)_{I_G \setminus I_H} = 0.$$

Finalmente, las dos últimas expresiones junto a la tercera identidad de (2.11), nos dice que

$$\eta = \zeta = 0,$$

lo que implica que, $v = 0$; probando así que Λ es no singular. ■

El Teorema 1 y las Proposiciones 3 y 4 nos dan inmediatamente la con-

vergencia local y el radio de convergencia del Algoritmo 1, que reunidos nos dan el siguiente resultado.

Teorema 2. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dos veces diferenciables en una vecindad de un punto fuertemente estacionario $\bar{x} \in \mathbb{R}^n$, con MPCC-multiplicador $\bar{\mu}$ asociado, del problema (1.1). Suponemos que las segundas derivadas de f, G y H son continuas en \bar{x} , se satisfacen la condiciones [MPCC-LICQ] y [MPCC-SOSC] en \bar{x} , y se cumple [ULSCC] en $\bar{\mu}$.

Entonces, para cualquier punto $(x^0, y^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$ suficientemente cerca de $(\bar{x}, \bar{y}, \bar{\mu})$, el Algoritmo 1 define de manera única la sucesión $\{u^k\}$ que converge a $(\bar{x}, \bar{y}, \bar{\mu})$. También, la tasa de convergencia es superlineal, y si las segundas derivadas de f, G y H son localmente Lipschitz continuas con respecto a \bar{x} , entonces la tasa es cuadrática.

2.2. Globalización del método local

El principal problema del Algoritmo 1 es que el punto de inicialización debe estar suficientemente cercano a la solución de la ecuación (2.1). Nuestro propósito ahora es globalizar el SSN local del Algoritmo 1 introduciendo una búsqueda lineal para la función de mérito $\varphi : \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m) \rightarrow \mathbb{R}$ definida por

$$\varphi(u) = \frac{1}{2} \|\Phi(u)\|^2, \quad (2.14)$$

donde Φ está definida en (2.3).

Proposición 5. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dos veces diferenciables en el punto $x \in \mathbb{R}^n$. Entonces, para cualquier $y \in \mathbb{R}^m$ y $\lambda = (\lambda_G, \lambda_H) \in \mathbb{R}^m \times \mathbb{R}^m$, la función φ definida en (2.14) es diferenciable en $u = (x, y, \lambda)$. Más aún, se tiene que

$$\varphi'(u) = \Lambda \Phi(u), \quad (2.15)$$

para cada $\Lambda \in \partial_B \Phi(u)$. Además, si f, G y H son dos veces continuamente diferenciables en \mathbb{R}^n , entonces φ es continuamente diferenciable en $\mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$.

Demostración. Por la definición de φ podemos escribir

$$2\varphi(u) = \sum_{j=1}^n \left(\frac{\partial \mathcal{L}}{\partial x_j}(x, \lambda) \right)^2 + \sum_{i=1}^m \left(\frac{\partial L}{\partial y_i}(x, y, \lambda) \right)^2 \\ + \sum_{i=1}^m ((\text{mín}\{0, y_i\})^2 - G_i(x))^2 + \sum_{i=1}^m ((\text{máx}\{0, y_i\})^2 - H_i(x))^2.$$

Notemos que, las componentes del primer, tercer y cuarto sumando de la parte derecha de la identidad anterior, admiten todas sus derivadas parciales respecto a las variables x_j, y_i y λ_i , con $j \in \{1, \dots, n\}$ e $i \in K$, ya que f, G y H son dos veces diferenciables en x . Así mismo, por la identidad (1.18) se sigue que para cada $i \in K$

$$\left(\frac{\partial L}{\partial y_i}(x, y, \lambda) \right)^2 = 4(\lambda_G)_i^2 (\text{mín}\{0, y_i\})^2 + 4(\lambda_H)_i^2 (\text{máx}\{0, y_i\})^2 \\ + 8(\lambda_G)_i (\lambda_H)_i \text{mín}\{0, y_i\} \text{máx}\{0, y_i\} \\ = 4 \left((\lambda_G)_i^2 (\text{mín}\{0, y_i\})^2 + (\lambda_H)_i^2 (\text{máx}\{0, y_i\})^2 \right),$$

pues $\text{mín}\{0, t\} \text{máx}\{0, t\} = 0$ para cualquier $t \in \mathbb{R}$. Así, vemos que $\left(\frac{\partial L}{\partial y_i}(x, y, \lambda) \right)^2$ admite todas sus derivadas parciales respecto de x_j, y_i y λ_i , con $j \in \{1, \dots, n\}$ e $i \in K$. En resumen, φ es diferenciable en $u = (x, y, \lambda)$; más aún, de la última expresión tenemos que

$$\left(\frac{1}{2} \left\| \frac{\partial L}{\partial y} (x, y, \lambda) \right\|^2 \right)' = \left(\sum_{i=1}^m 2 \left((\lambda_G)_i^2 (\text{mín}\{0, y_i\})^2 + (\lambda_H)_i^2 (\text{máx}\{0, y_i\})^2 \right) \right)' \\ = \begin{pmatrix} 0_{\mathbb{R}^n} \\ 4 \left((\lambda_G)_1^2 \text{mín}\{0, y_1\} + (\lambda_H)_1^2 \text{máx}\{0, y_1\} \right) \\ \vdots \\ 4 \left((\lambda_G)_m^2 \text{mín}\{0, y_m\} + (\lambda_H)_m^2 \text{máx}\{0, y_m\} \right) \\ 4(\lambda_G)_1 (\text{mín}\{0, y_1\})^2 \\ \vdots \\ 4(\lambda_G)_m (\text{mín}\{0, y_m\})^2 \\ 4(\lambda_H)_1 (\text{máx}\{0, y_1\})^2 \\ \vdots \\ 4(\lambda_H)_m (\text{máx}\{0, y_m\})^2 \end{pmatrix},$$

es decir, gracias a las identidades (1.18), (2.6) y (2.5) podemos escribir

$$\left(\frac{1}{2} \left\| \frac{\partial L}{\partial y}(x, y, \lambda) \right\|^2 \right)' = \begin{pmatrix} 0 \\ 2A(y, \lambda) \frac{\partial L}{\partial y}(x, y, \lambda) \\ 2B_{\min}(y) \frac{\partial L}{\partial y}(x, y, \lambda) \\ 2B_{\max}(y) \frac{\partial L}{\partial y}(x, y, \lambda) \end{pmatrix}. \quad (2.16)$$

De manera similar, como

$$\frac{\partial \mathcal{L}}{\partial x}(x, \lambda) = \nabla f(x) - G'(x)^T \lambda_G - H'(x)^T \lambda_H,$$

entonces

$$\begin{aligned} \left(\frac{1}{2} \left\| \frac{\partial \mathcal{L}}{\partial x}(x, \lambda) \right\|^2 \right)' &= \frac{1}{2} \left(\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right)^2 \right)' \\ &= \begin{pmatrix} \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right) \cdot \\ \left(\frac{\partial^2 f}{\partial x_1 \partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial^2 G_j}{\partial x_1 \partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial^2 H_j}{\partial x_1 \partial x_i}(x) \right) \\ \vdots \\ \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right) \cdot \\ \left(\frac{\partial^2 f}{\partial x_n \partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial^2 G_j}{\partial x_n \partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial^2 H_j}{\partial x_n \partial x_i}(x) \right) \\ 0_{\mathbb{R}^m} \\ \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right) \cdot \\ \left(-\frac{\partial G_1}{\partial x_i}(x) \right) \\ \vdots \\ \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right) \cdot \\ \left(-\frac{\partial G_m}{\partial x_i}(x) \right) \\ \vdots \\ \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right) \cdot \\ \left(-\frac{\partial H_1}{\partial x_i}(x) \right) \\ \vdots \\ \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_G)_j \frac{\partial G_j}{\partial x_i}(x) - \sum_{j=1}^m (\lambda_H)_j \frac{\partial H_j}{\partial x_i}(x) \right) \cdot \\ \left(-\frac{\partial H_m}{\partial x_i}(x) \right) \end{pmatrix}, \end{aligned}$$

en otras palabras,

$$\left(\frac{1}{2} \left\| \frac{\partial \mathcal{L}}{\partial x}(x, \lambda) \right\|^2\right)' = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda) \frac{\partial \mathcal{L}}{\partial x}(x, \lambda) \\ 0 \\ -G'(x) \frac{\partial \mathcal{L}}{\partial x}(x, \lambda) \\ -H'(x) \frac{\partial \mathcal{L}}{\partial x}(x, \lambda) \end{pmatrix}. \quad (2.17)$$

De manera análoga a lo hecho anteriormente, tenemos que

$$\left(\frac{1}{2} \left\| (\min\{0, y\})^2 - G(x) \right\|^2\right)' = \begin{pmatrix} -G'(x)^T ((\min\{0, y\})^2 - G(x)) \\ 2B_{\min}(y) ((\min\{0, y\})^2 - G(x)) \\ 0 \\ 0 \end{pmatrix}, \quad (2.18)$$

y también

$$\left(\frac{1}{2} \left\| (\max\{0, y\})^2 - H(x) \right\|^2\right)' = \begin{pmatrix} -H'(x)^T ((\max\{0, y\})^2 - H(x)) \\ 2B_{\max}(y) ((\max\{0, y\})^2 - H(x)) \\ 0 \\ 0 \end{pmatrix}. \quad (2.19)$$

Juntando las identidades (2.16), (2.17), (2.18) y (2.19) conseguimos que

$$\varphi'(u) = \Lambda \Phi(u),$$

para cada $\Lambda \in \partial_B \Phi(u)$. Para terminar, de la identidad anterior se sigue de inmediato que si f, G y H son de clase C^2 , entonces φ es de clase C^1 . ■

Observación 4. Notemos que, la identidad (2.4) nos dice que todas las matrices $\Lambda \in \partial_B \Phi(u)$ son simétricas, en cualquier $u \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$. Así, por la identidad (2.15) obtenida en la Proposición 5, se sigue que para cualquier dirección $v \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$, calculada como la solución de la ecuación

$$\Lambda v = -\Phi(u), \quad (2.20)$$

tenemos que

$$\begin{aligned}\langle \varphi'(u), v \rangle &= \langle \Lambda \Phi(u), v \rangle \\ &= \langle \Phi(u), \Lambda v \rangle,\end{aligned}$$

que junto a la ecuación (2.20) nos da que

$$\langle \varphi'(u), v \rangle = - \|\Phi(u)\|^2 = -2\varphi(u). \quad (2.21)$$

En particular, si u no es una solución de la ecuación 2.1 o, de manera equivalente, no es un minimizador global de la función φ , entonces v es una dirección de descenso para φ en el punto u .

La observación anterior nos sugiere inmediatamente una estrategia de globalización natural para el SSN local en el Algoritmo (1). A continuación, enunciamos nuestro algoritmo globalizado.

Algoritmo 2. Elegir los siguientes parámetros: $\epsilon \in (0, 1/2)$, $\tau \in (0, 1)$, $M > 0$, $\theta > 0$ y un punto de inicio $u^0 = (x^0, y^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$. Sea $k = 0$.

1. Si $\Phi(u^k) = 0$, deténgase. De lo contrario, calcule alguna matriz $\Lambda_k = \Lambda$ de acuerdo con las expresiones (2.4), (2.5) y (2.6) con $u = u^k$. Luego, calcule $\tilde{u}^{k+1} \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$, como la solución de la ecuación (2.9) dado en el paso 2 del Algoritmo 1, con Φ definido en (2.3). Si \tilde{u}^{k+1} existe y cumple que

$$\|\tilde{u}^{k+1} - u^k\| \leq \max\{M, 1/(\varphi(u^k))^\theta\}, \quad (2.22)$$

entonces, tomamos

$$v^k = \tilde{u}^{k+1} - u^k;$$

caso contrario, consideramos

$$v^k = -\Lambda_k \Phi(u^k). \quad (2.23)$$

Si $v^k = 0$, parar.

2. Calcule el tamaño del paso α_k mediante la regla de Armijo: para ello,

sea $\alpha_k = \tau^j$, donde j es el entero no negativo más pequeño que satisface la desigualdad

$$\varphi(u^k + \tau^j v^k) \leq \varphi(u^k) + \epsilon \tau^j \langle \varphi'(u^k), v^k \rangle. \quad (2.24)$$

Consideramos $u^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}) = u^k + \alpha_k v^k$.

3. Sea $k = k + 1$ y volver al paso 1.

Observación 5. Dentro del paso 2 del Algoritmo 2, si \tilde{u}^{k+1} existiera, entonces

$$\begin{aligned} -\Phi(u^k) &= \Lambda_k(\tilde{u}^{k+1} - u^k) \\ &= \Lambda_k v^k, \end{aligned}$$

en otras palabras, v^k satisface la expresión (2.21), por lo que es una dirección de descenso que llamaremos **dirección de descenso de Newton**. Caso contrario, por la identidad (2.15) la dirección de descenso v^k coincide con $-\nabla\varphi$.

Por otra parte, los parámetros $M > 0$ y $\theta > 0$ deben elegirse lo suficientemente grandes a fin de admitir más direcciones de Newton.

El siguiente resultado nos da la validez de nuestro algoritmo globalizado.

Teorema 3. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $G, H : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dos veces continuamente diferenciables en \mathbb{R}^n . Entonces, el Algoritmo 2 está bien definido y cualquier punto de acumulación \bar{u} de cualquier sucesión $\{u^k\}$ generada por este algoritmo es un punto estacionario de la función φ , es decir

$$\varphi'(\bar{u}) = 0. \quad (2.25)$$

Demostración. Primero, observemos que la búsqueda lineal es siempre en una dirección de descenso; si la dirección de Newton satisface la identidad (2.21), o si la dirección es menos gradiente obtenemos que

$$\langle \varphi'(u^k), v^k \rangle < 0, \quad (2.26)$$

para todo $k \in \mathbb{N}$. Por otra parte, gracias a que f, G y H son dos veces

continuamente diferenciables, las sucesiones $\{\alpha_k\}_{k \in \mathbb{N}}$, generada por la búsqueda lineal de Armijo, y $\{u^k\}_{k \in \mathbb{N}}$ están bien definidas. Además, como se satisface la desigualdad (2.24) tenemos que la sucesión $\{\varphi(u^k)\}_{k \in \mathbb{N}}$ es decreciente. Como está acotada inferiormente (por cero), converge a $\inf_{k \in \mathbb{N}} \varphi(u^k)$. Entonces, por la desigualdad (2.24) tenemos que

$$\lim_{k \rightarrow +\infty} \alpha_k \langle \varphi'(u^k), v^k \rangle = 0. \quad (2.27)$$

En efecto, por la desigualdad (2.24) y como v^k es una dirección de descenso, tenemos que para cada $k \in \mathbb{N}$

$$\varphi(u^{k+1}) \leq \varphi(u^k) + \epsilon \alpha_k \langle \varphi'(u^k), v^k \rangle \leq \varphi(u^k),$$

de donde, tomando límite $k \rightarrow +\infty$, se consigue la identidad (2.27).

Por otra parte, sea \bar{u} un punto de acumulación de la sucesión $\{u^k\}_k$; así existe una subsucesión $\{u^{k_j}\}_j$ convergente a \bar{u} . Consideremos los dos casos posibles:

$$\limsup_{j \rightarrow +\infty} \alpha_{k_j} > 0 \quad \text{o} \quad \lim_{j \rightarrow +\infty} \alpha_{k_j} = 0. \quad (2.28)$$

En el primer caso, pasando a otra subsucesión si es necesario, podemos suponer que

$$\liminf_{j \rightarrow +\infty} \alpha_{k_j} > 0,$$

que, junto a la expresión (2.27), nos da que

$$\lim_{j \rightarrow +\infty} \langle \varphi'(u^{k_j}), v^{k_j} \rangle = 0. \quad (2.29)$$

Si dentro de esta subsucesión se usa la dirección de Newton para una cantidad infinita de índices j , por (2.21) y (2.29) tenemos que

$$\langle \varphi'(u^{k_j}), v^{k_j} \rangle = -2\varphi(u^{k_j}),$$

para una cantidad infinita de índices j . Entonces, como φ es de clase C^1 y u^{k_j} converge a \bar{u} , la identidad (2.29) implica que $\varphi(\bar{u}) = 0$, es decir, \bar{u} es un minimizador global de φ , lo que implica (2.25).

Ahora, si las direcciones de Newton se usan en una cantidad finita de índices j , entonces existe $N > 0$ suficientemente grande tal que

$$\langle \varphi'(u^{k_j}), v^{k_j} \rangle = - \|\varphi'(u^{k_j})\|^2,$$

para cada $j \geq N$; de donde, como φ es de clase C^1 y u^{k_j} converge a \bar{u} , por la identidad (2.29) se consigue que $\varphi'(\bar{u}) = 0$.

Ahora bien, debemos considerar el segundo caso en la expresión (2.27). Supongamos primero que la sucesión $\{v^{k_j}\}$ no está acotada. Así, existe una cantidad infinita de índices j donde se utilizan las direcciones de Newton (de lo contrario, existe $N > 0$ suficientemente grande, de modo que para todo $j \geq N$, $v^{k_j} = -\varphi'(u^{k_j})$; y por lo tanto, como u^{k_j} converge a \bar{u} y φ es de clase C^1 , v^{k_j} converge a $-\varphi'(\bar{u})$, lo cual contradice que la sucesión $\{v^{k_j}\}$ no está acotada).

Sin embargo, sabemos que las direcciones de Newton satisfacen la condición

$$\|v^{k_j}\| \leq \max\{M, 1/(\varphi(u^{k_j}))^\theta\} \quad (2.30)$$

para cada $j \in \mathbb{N}$, lo cual implica que la sucesión $\{\varphi(u^{k_j})\}_{k \in \mathbb{N}}$ converge al $\inf_{j \in \mathbb{N}} \varphi(u^{k_j}) = 0$, pues de lo contrario, si $\inf_{j \in \mathbb{N}} \varphi(u^{k_j}) > 0$ entonces

$$\max\left\{M, \frac{1}{(\varphi(u^{k_j}))^\theta}\right\} \leq M + \frac{1}{(\inf_{j \in \mathbb{N}} \varphi(u^{k_j}))^\theta},$$

lo cual implica, junto a la desigualdad (2.30), que $\{v^{k_j}\}$ está acotada, contradiciendo nuestra suposición. Así, como φ es de clase C^1 y u^{k_j} converge a \bar{u} se consigue que $\varphi(\bar{u}) = 0$, es decir, \bar{u} es un minimizador global de φ , lo que implica nuevamente la identidad (2.25).

Finalmente, supongamos que la sucesión $\{v^{k_j}\}$ está acotada. Por el Teorema de Bolzano-Weierstrass, $\{v^{k_j}\}$ posee una subsucesión, notada de la misma forma, que converge a algún \bar{v} . Por el paso 2 del Algoritmo 2, para cada $j \in \mathbb{N}$, $\alpha_{k_j} = \tau^m$, donde $m = m(j) \in \mathbb{N}$ es el natural más pequeño que satisface la desigualdad (2.24), así

$$\frac{\varphi(u^{k_j} + (\alpha_{k_j}/\tau)v^{k_j}) - \varphi(u^{k_j})}{\alpha_{k_j}/\tau} > \epsilon \langle \varphi'(u^{k_j}), v^{k_j} \rangle.$$

Empleando el Teorema del Valor Medio y el hecho de que α_{k_j} converge a 0 cuando $j \rightarrow +\infty$ obtenemos que

$$\langle \varphi'(\bar{u}), \bar{v} \rangle \geq \epsilon \langle \varphi'(\bar{u}), \bar{v} \rangle. \quad (2.31)$$

En efecto, por el Teorema de Valor Medio, para cada $j \in \mathbb{N}$, existe $\beta_j \in [0, 1]$ tal que

$$\begin{aligned} \frac{\varphi(u^{k_j} + (\alpha_{k_j}/\tau)v^{k_j}) - \varphi(u^{k_j})}{\alpha_{k_j}/\tau} &= \frac{\alpha_{k_j}/\tau \langle \varphi'(u^{k_j} + \beta_j(\alpha_{k_j}/\tau)), v^{k_j} \rangle}{\alpha_{k_j}/\tau} \\ &= \langle \varphi'(u^{k_j} + \beta_j(\alpha_{k_j}/\tau)), v^{k_j} \rangle, \end{aligned}$$

y como α_{k_j} converge a 0 cuando $j \rightarrow +\infty$ y la sucesión $\{\beta_j\}$ está acotada, se consigue la expresión (2.31). Luego, junto al hecho que $\epsilon \in (0, 1/2)$, se sigue que

$$\langle \varphi'(\bar{u}), \bar{v} \rangle \geq 0.$$

Combinando la última desigualdad con la expresión (2.26), obtenemos que

$$\langle \varphi'(\bar{u}), \bar{v} \rangle = 0.$$

Considerando, como antes, los dos casos en los que el número de veces que se ha utilizado la dirección de Newton es infinito o finito, la última relación implica que la identidad (2.25) se cumple. ■

Observación 6. De acuerdo con la demostración del Teorema 3, si durante la construcción de una subsucesión de $\{u^k\}$, la cual converge a $\bar{u} = (\bar{x}, \bar{y}, \bar{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m \times (\mathbb{R}^m \times \mathbb{R}^m)$, la dirección de Newton se ha utilizado una cantidad infinita de veces, entonces $\Phi(\bar{u}) = 0$. Así, como Φ está definida en la expresión (2.3) tenemos que (\bar{x}, \bar{y}) es un punto estacionario del problema (1.5) y $\bar{\lambda}$ es un multiplicador de Lagrange asociado. Por la Proposición 2, se sigue que \bar{x} es un punto débilmente estacionario del problema (1.1).

Además, notemos que la convergencia a un punto estacionario de φ del cual no es su minimizador global, sólo puede ocurrir cuando las direcciones de Newton no se han usado a lo largo de la subsucesión correspondiente desde algún punto. Pero incluso en ese caso, por la Proposición 5 sabemos que para cualquier punto estacionario \bar{u} de la función φ se cumple que $\Lambda\Phi(\bar{u}) = 0$ para cada matriz $\Lambda \in \partial_B\Phi(\bar{u})$. Si al menos una de ellas es no singular obtenemos inmediatamente que $\Phi(\bar{u}) = 0$. Por lo tanto, podemos esperar una convergencia global del Algoritmo 2 a puntos débilmente estacionarios del problema (1.1).

Finalmente, mostremos que el Algoritmo 2 preserva la rápida convergencia local del Algoritmo 1 bajo ciertas hipótesis adicionales.

Teorema 4. Sean f, G y H dos veces continuamente diferenciables en \mathbb{R}^n , y una sucesión $\{u^k\}$ generada por el Algoritmo 2 tal que posee un punto de acumulación $\bar{u} = (\bar{x}, \bar{y}, \bar{\mu})$, donde \bar{x} es un punto fuertemente estacionario del problema (1.1), \bar{y} está dada por la expresión (1.14) y $\bar{\mu}$ es un MPCC-multiplicador asociado a \bar{x} . Supongamos además que se satisfacen las condiciones [MPCC-LICQ] y [MPCC-SOSC] en \bar{x} y [ULSCC] en $\bar{\mu}$.

Entonces toda la sucesión $\{u^k\}$ converge a \bar{u} , y la tasa de convergencia es superlineal. En particular, si las segundas derivadas de f, G y H son localmente Lipschitz continuas con respecto a \bar{x} , entonces es cuadrática.

Demostración. Sea u^k suficientemente cerca de \bar{u} . Según el Teorema 2, podemos hallar \tilde{u}^{k+1} , calculado como en el paso 1 del Algoritmo 2, el cual está bien definido pues satisface la identidad (2.9) y adicional tenemos que

$$\|\tilde{u}^{k+1} - \bar{u}\| = o(\|u^k - \bar{u}\|). \quad (2.32)$$

En consecuencia, \tilde{u}^{k+1} verifica la desigualdad (2.22). Además, de acuerdo con la Proposición 4, bajo los supuestos establecidos, la función Φ es *BD-regular* en \bar{u} . Como menciona Pang y Qi (1993) en [5] la BD-regularidad implica el error

$$\|u - \bar{u}\| = O(\|\Phi(u)\|). \quad (2.33)$$

Como \bar{u} satisface la ecuación (2.1) entonces

$$\begin{aligned}\varphi(\tilde{u}^{k+1}) &= \frac{1}{2} \|\Phi(\tilde{u}^{k+1})\|^2 \\ &= \frac{1}{2} \|\Phi(\tilde{u}^{k+1}) - \Phi(\bar{u})\|^2.\end{aligned}$$

Teniendo en cuenta la Lipschitz continuidad de Φ en una vecindad de \bar{u} , la identidad (2.32) y la expresión anterior, obtenemos que

$$\begin{aligned}\varphi(\tilde{u}^{k+1}) &= O\left(\|\tilde{u}^{k+1} - \bar{u}\|^2\right) \\ &= o\left(\|u^k - \bar{u}\|^2\right),\end{aligned}$$

de donde, empleando la cota de error (2.33) conseguimos

$$\begin{aligned}\varphi(\tilde{u}^{k+1}) &= o\left(\|\Phi(u^k)\|^2\right) \\ &= o(\varphi(u^k)).\end{aligned}$$

Por otro lado, consideremos $v^k = \tilde{u}^{k+1} - u^k$; como $\epsilon \in (0, 1/2)$, la relación anterior implica que si u^k está lo suficientemente cerca de \bar{u} entonces

$$\begin{aligned}\varphi(u^k + v^k) &= \varphi(\tilde{u}^{k+1}) \\ &\leq (1 - 2\epsilon)\varphi(u^k),\end{aligned}$$

de donde, como v^k satisface la identidad (2.21), se tiene que

$$\begin{aligned}\varphi(u^k + v^k) &\leq \varphi(u^k) - \epsilon \|\Phi(u^k)\|^2 \\ &= \varphi(u^k) + \epsilon \langle \varphi'(u^k), v^k \rangle.\end{aligned}$$

Por lo tanto, $\alpha_k = 1$ será aceptado por la regla de Armijo, es decir, la desigualdad (2.24) se cumple con $j = 0$. Esto muestra que las iteraciones del Algoritmo 2 a partir de la iteración k se reducen a las del Algoritmo 1. Así, las afirmaciones de este teorema ahora se deducen del Teorema 2. ■

Capítulo 3

Resultados numéricos

3.1. Implementación

Vamos a utilizar el lenguaje de programación MATLAB para implementar los Algoritmo 1 y 2. En el Capítulo de Anexos se puede encontrar el código empleado en este trabajo.

3.2. Ejemplos

Ejemplo 1. Para el problema (1.1) consideremos $n = 2$, $m = 1$ y las siguientes funciones

$$f(x) = (x_1 + 1)^2 + (x_2 - 2)^2$$

$$G(x) = x_2 - x_1$$

$$H(x) = x_2.$$

En primer lugar, gracias a la expresión (1.2), el conjunto factible para este problema es

$$\mathcal{M} = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 - x_1 \geq 0, x_2 \geq 0, (x_2 - x_1)x_2 = 0\},$$

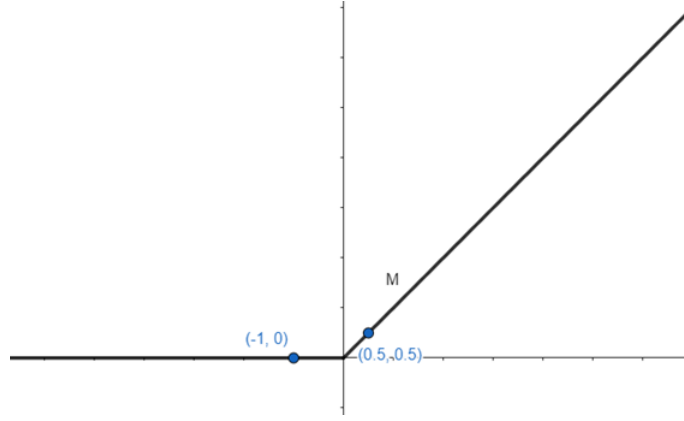


Figura 3.1: Conjunto factible para el ejemplo 1.

cuya gráfica la podemos ver en la Figura 3.1, donde también se encuentran ubicados los puntos estacionarios (según [4]) $z^* = (-1, 0)$ y $x^* = \left(\frac{1}{2}, \frac{1}{2}\right)$ de la función f .

En este caso, tenemos que la función Φ , definida como en la expresión (2.3), viene dada por

$$\Phi(u) = \begin{pmatrix} 2(x_1 + 1) + \lambda_G \\ 2(x_2 - 2) - (\lambda_G + \lambda_H) \\ 2(\lambda_G(\min\{0, y\}) + \lambda_H(\max\{0, y\})) \\ (\min\{0, y\})^2 - x_2 + x_1 \\ (\max\{0, y\})^2 - x_2 \end{pmatrix},$$

con $u = (x, y, \lambda) \in \mathbb{R}^2 \times \mathbb{R} \times (\mathbb{R} \times \mathbb{R})$ y $\lambda = (\lambda_G, \lambda_H)$. Además, gracias al Lema 3 cualquier $\Lambda \in \partial_B \Phi(u)$ se ve como sigue

$$\Lambda = \begin{pmatrix} 2 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & -1 & -1 \\ 0 & 0 & 2a(y, \lambda) & 2 \min\{0, y\} & 2 \max\{0, y\} \\ 1 & -1 & 2 \min\{0, y\} & 0 & 0 \\ 0 & -1 & 2 \max\{0, y\} & 0 & 0 \end{pmatrix}, \quad (3.1)$$

donde $a(y, \lambda)$ está definido en la expresión (2.6).

Vamos a ejecutar los Algoritmos 1 y 2 a fin de determinar cuál es el más eficiente para este problema. Consideramos los siguientes valores

para los parámetros iniciales del Algoritmo 2:

$$\epsilon = 10^{-4}, \quad \tau = \frac{1}{2}, \quad \theta = 1, \quad M \in \{10^2, 10^5\},$$

y los siguientes puntos iniciales $u = (x, y, \lambda_G, \lambda_H)$

$$\begin{aligned} u_1 &= (1, 1, 2, 4, 5), \\ u_2 &= (-1, -1, 1, 8, 5), \\ u_3 &= \left(-\frac{1}{2}, \frac{1}{2}, -\frac{4}{5}, 0, -3\right), \\ u_4 &= (30, -15, -5, 0, 20), \\ u_5 &= (70, -15, 10, -1, 100), \\ u_6 &= (-30, 70, -15, -10, 210), \end{aligned}$$

Al utilizar nuestro programa, se obtuvieron los siguientes resultados, resumidos en las tablas 3.1, 3.2 y 3.3, donde el *error* considera la norma euclídeana de la diferencia del punto al que converge el algoritmo y su valor teórico más cercano; además, se presentan los gráficos de convergencia en las figuras 3.2, 3.3 y 3.4, donde se puede ver la convergencia superlineal que esperábamos.

Punto de inicio	Número de iteraciones	Tiempo de ejecución (s)	Solución \bar{x}	Error
u_1	6	0,064293	$\begin{pmatrix} 0,500000000004 \\ 0,500000000004 \end{pmatrix}$	$5,6498 \times 10^{-12}$
u_2	5	0,005918	$\begin{pmatrix} 0,500000000003 \\ 0,500000000003 \end{pmatrix}$	$4,8484 \times 10^{-12}$
u_3	3	0,001372	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	0
u_4	6	0,003478	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	0
u_5	9	0,002462	$\begin{pmatrix} 0,500000000005 \\ 0,500000000005 \end{pmatrix}$	$7,0381 \times 10^{-11}$
u_6	7	0,004044	$\begin{pmatrix} -1,00000001 \\ 0 \end{pmatrix}$	$1,2060 \times 10^{-8}$

Cuadro 3.1: Resultados para el Ejemplo 1 usando el Alg. 1.

Vemos que, en efecto, solamente cuando se elige un punto de inicialización cercano a z^* el algoritmo local converge a éste punto estacionario (en este caso, vemos que $y < 0$ y $\lambda_G = 0$ nos permiten estar en una

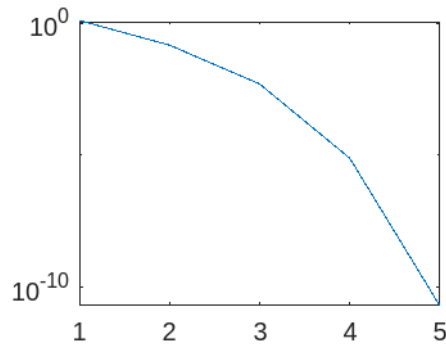


Figura 3.2: Evolución del Error vs número de iteraciones para el Ejemplo 1, usando el Algoritmo local.

vecindad suficientemente cercana); en las demás experiencias, converge siempre a x^* .

Por otro lado, veamos lo que sucede cuando ejecutamos el algoritmo global, considerando los parámetros $M = 10^2$ y $M = 10^5$.

Punto de inicio	Número de iteraciones	Tiempo de ejecución (s)	Solución \bar{x}	Error
u_1	6	0,009321	$\begin{pmatrix} 0,500000000004 \\ 0,500000000004 \end{pmatrix}$	$5,6498 \times 10^{-12}$
u_2	5	0,003589	$\begin{pmatrix} 0,500000000003 \\ 0,500000000003 \end{pmatrix}$	$4,8484 \times 10^{-12}$
u_3	3	0,000557	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	0
u_4	6	0,000236	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	0
u_5	9	0,006725	$\begin{pmatrix} 0,5000000003 \\ 0,5000000003 \end{pmatrix}$	$7,0381 \times 10^{-11}$
u_6	No converge			

Cuadro 3.2: Resultados para el Ejemplo 1 usando el Alg. 2 y $M = 10^2$.

De las tablas 3.1 y 3.2 vemos que se obtiene una mejora en los tiempos de ejecución al utilizar el Algoritmo global en lugar del local.

De la tabla 3.3 notemos que, al utilizar el Algoritmo global con un valor del parámetro M más grande se obtiene una mejora en los tiempos de ejecución. De hecho, tal como se menciona en la Observación 5, se consiguen mejores resultados al considerar valores de M suficientemente grandes ya que admiten más direcciones de Newton; en efecto, si $M = 10^2$ el algoritmo global tiene problemas en la convergencia con el punto inicial

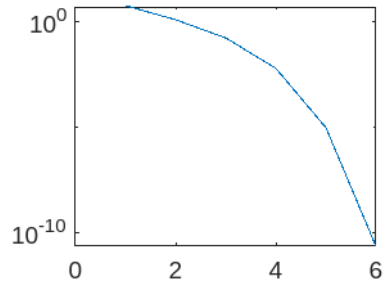


Figura 3.3: Evolución del Error vs número de iteraciones para el Ejemplo 1, usando el Algoritmo global con $M = 10^2$.

Punto de inicio	Número de iteraciones	Tiempo de ejecución (s)	Solución \bar{x}	Error
u_1	6	0,006668	$\begin{pmatrix} 0,500000000004 \\ 0,500000000004 \end{pmatrix}$	$5,6498 \times 10^{-12}$
u_2	5	0,003332	$\begin{pmatrix} 0,500000000003 \\ 0,500000000003 \end{pmatrix}$	$4,8484 \times 10^{-12}$
u_3	3	0,000546	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	0
u_4	6	0,000374	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$	0
u_5	9	0,000517	$\begin{pmatrix} 0,5000000003 \\ 0,5000000003 \end{pmatrix}$	$7,0381 \times 10^{-11}$
u_6	7	0,000243	$\begin{pmatrix} -1,00000001 \\ 0 \end{pmatrix}$	$1,2060 \times 10^{-8}$

Cuadro 3.3: Resultados para el Ejemplo 1 usando el Alg. 2 y $M = 10^5$.

u_6 , mientras tanto para $M = 10^3$ no hay inconvenientes.

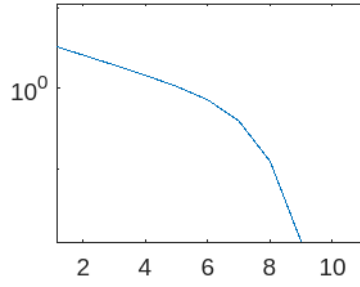


Figura 3.4: Evolución del Error vs número de iteraciones para el Ejemplo 1, usando el Algoritmo global con $M = 10^5$.

Ejemplo 2. Para el problema (1.1) consideremos $n = 3$, $m = 2$ y las siguientes funciones

$$f(x) = 0,2x_1^2 + 0,1x_2^2 + 0,8x_3^3,$$

$$G(x) = \begin{pmatrix} x_2 \\ x_1 + x_2 - x_3 - 1 \end{pmatrix},$$

$$H(x) = \begin{pmatrix} x_1 \\ x_1 + x_2 + x_3 - 1 \end{pmatrix}.$$

En este caso, tenemos que la función Φ , definida como en la expresión (2.3), viene dada por

$$\Phi(u) = \begin{pmatrix} 0,4x_1 - (\lambda_G)_2 - ((\lambda_H)_1 + (\lambda_H)_2) \\ 0,2x_2 - ((\lambda_G)_1 + (\lambda_G)_2) - (\lambda_H)_2 \\ 2,4x_3^2 + (\lambda_G)_2 - (\lambda_H)_2 \\ 2((\lambda_G)_1(\min\{0, y_1\}) + (\lambda_H)_1(\min\{0, y_1\})) \\ 2((\lambda_G)_2(\min\{0, y_2\}) + (\lambda_H)_2(\min\{0, y_2\})) \\ (\min\{0, y_1\})^2 - x_2 \\ (\min\{0, y_2\})^2 - x_1 - x_2 + x_3 + 1 \\ (\max\{0, y_1\})^2 - x_1 \\ (\max\{0, y_2\})^2 - x_1 - x_2 - x_3 + 1 \end{pmatrix},$$

con $u = (x, y, \lambda) \in \mathbb{R}^3 \times \mathbb{R}^2 \times (\mathbb{R}^2 \times \mathbb{R}^2)$ y $\lambda = (\lambda_G, \lambda_H)$. Además, tenemos que

$$\frac{\partial^2 \mathcal{L}}{\partial x^2}(x, \lambda) = \begin{pmatrix} 0,4 & 0 & 0 \\ 0 & 0,2 & 0 \\ 0 & 0 & 4,8x_3 \end{pmatrix},$$

$$G'(x) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix},$$

$$H'(x) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

con lo que cada $\Lambda \in \partial_B \Phi(u)$ está plenamente determinado.

Consideramos los siguientes valores para los parámetros iniciales del Algoritmo 2:

$$\epsilon = 10^{-4}, \quad \tau = \frac{1}{2}, \quad \theta = 1, \quad M \in \{10^2, 10^5\},$$

y los siguientes puntos iniciales $u = (x, y, \lambda_G, \lambda_H)$

$$u_1 = (1, 2, 3, 1, 1, 2, 4, 1, 2),$$

$$u_2 = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{5}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{5}, \frac{1}{5} \right),$$

Al utilizar nuestro programa, se obtuvieron los siguientes resultados, resumidos en las tablas 3.4, 3.6 y 3.5, acompañados también de las gráficas de convergencia en las figuras 3.5 y 3.6.

Punto de inicio	Número de iteraciones	Tiempo de ejecución (s)	Solución \bar{x}	Error
u_1	12	1,759753	$\begin{pmatrix} 1,500000003 \\ 0 \\ -0,500000003 \end{pmatrix}$	$3,6917 \times 10^{-9}$
u_2	5	2,112173	$\begin{pmatrix} 0 \\ 0,99999999999 \\ 0,0000000001 \end{pmatrix}$	$1,4106 \times 10^{-10}$

Cuadro 3.4: Resultados para el Ejemplo 2 usando el Alg. 1.

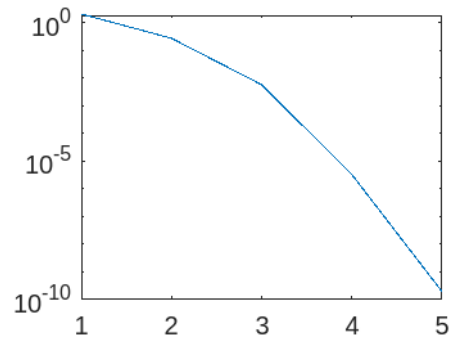


Figura 3.5: Evolución del Error vs número de iteraciones para el Ejemplo 2, usando el Algoritmo local.

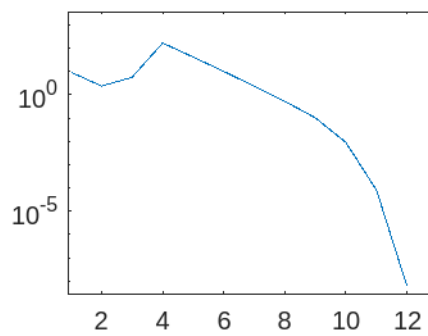


Figura 3.6: Evolución del Error vs número de iteraciones para el Ejemplo 2, usando el Algoritmo global con $M = 10^5$.

Punto de inicio	Número de iteraciones	Tiempo de ejecución (s)	Solución \bar{x}	Error
u_1	12	1,409228	$\begin{pmatrix} 1,500000003 \\ 0 \\ -0,500000003 \end{pmatrix}$	$3,6917 \times 10^{-9}$
u_2	5	1,927910	$\begin{pmatrix} 0 \\ 0,9999999999 \\ 0,0000000001 \end{pmatrix}$	$1,4106 \times 10^{-10}$

Cuadro 3.5: Resultados para el Ejemplo 2 usando el Alg. 2 con $M = 10^2$

Punto de inicio	Número de iteraciones	Tiempo de ejecución (s)	Solución \bar{x}	Error
u_1	12	0,798144	$\begin{pmatrix} 1,500000003 \\ 0 \\ -0,500000003 \end{pmatrix}$	$3,6917 \times 10^{-9}$
u_2	5	1,393111	$\begin{pmatrix} 0 \\ 0,9999999999 \\ 0,0000000001 \end{pmatrix}$	$1,4106 \times 10^{-10}$

Cuadro 3.6: Resultados para el Ejemplo 2 usando el Alg. 2 con $M = 10^5$.

3.3. Conclusiones

Las dificultades en el tratamiento teórico y numérico de los problemas de programación matemática con restricciones de complementariedad [MPCC], que vienen ligados a la forma de sus restricciones, motivan el uso de técnicas que permitan una reformulación "suave" del conjunto factible del problema. En este sentido, al utilizar la reformulación por elevación (*lifting approach*), obtuvimos un problema equivalente donde las nuevas restricciones son más sencillas de estudiar; así mismo los puntos factibles preservan las condiciones de calificación estándar.

Por otra parte, el método de Newton semi-suave [SSN] permitió el abordaje del sistema de optimalidad Lagrangeano del problema reformulado, donde aprovechamos el enfoque en los B-subdiferenciales para mostrar resultados de convergencia superlineal del método bajo condiciones de calificación estándar y condiciones razonables de regularidad de la función objetivo y sus restricciones.

Finalmente, implementamos el algoritmo [SSN] local y global en un ambiente MATLAB, de donde se mostró la mejora en los tiempos de eje-

cución con el global. Sin embargo, vemos que el algoritmo local se comporta bastante bien. Además, comprobamos la importancia de elegir unos parámetros adecuados para obtener mejores resultados.

3.4. Recomendaciones

Se sugiere aplicar el algoritmo a problemas de gran escala para analizar su comportamiento y, de ser necesario, realizar los cambios pertinentes para preservar la convergencia superlineal del método.

Además, se propone discutir otras técnicas de globalización del algoritmo SSN o usar otros métodos de búsqueda lineal a fin de obtener resultados más eficientes.

Capítulo A

Título anexo

Presentamos el código en ambiente MATLAB correspondiente a la implementación de los Algoritmos 1 y 2. El *script* principal es el siguiente, donde están programados ambos algoritmos:

```
1  %Codigo SSN para MPECS, paper Izmailov-Solodov
2  clc
3  clear
4  err = 1;
5  verr=[];
6  iter=0;
7
8  %Parametros iniciales
9  M = 1e5;
10 th = 1;
11 tau=0.5;
12 ep=1e-4;
13
14 %Elefir funcion objetivo
15 l=input('Funcion obj=');
16
17 %Punto de inicializacion del Algoritmo
18 x = input('x=');
19 y = input('y=');
```

```

20 lamG = input('lambdaG=');
21 lamH = input('lambdaH=');
22
23 tic
24 u=[x;y;lamG;lamH];
25 n=length(x);
26 m=length(y);
27
28 %Elegir algoritmo local o global
29 q=input('Algoritmo');
30 if q==1
31     while (err>1e-6 && iter<=200)
32
33     A=Lambda(l,x,y,lamG,lamH);
34     Auk = A*u;
35     Phi = phi(l,u);
36
37     %Resolucion eq. (2.9)
38     u=A\ (Auk-Phi);
39
40     x=u(1:n);
41     y=u(n+1:n+m);
42     lamG = u(n+m+1:n+2*m);
43     lamH= u(n+2*m+1:n+3*m);
44
45     err = norm(phi(l,u));
46     iter = iter + 1;
47     verr=[verr;err];
48 end
49 else
50     while (err>1e-6 && iter<=200)
51
52     A=Lambda(l,x,y,lamG,lamH);
53     Auk = A*u;
54     Phi = phi(l,u);
55

```

```

56     %Vectores auxiliares del Paso 2
57     ut=A\ (Auk-Phi);
58     dir=norm(u-ut);
59
60     varphi = 0.5*norm(Phi)^2;
61
62     if dir <= max(M,1/(varphi^th))
63         v=ut-u;
64     else
65         v=-A*Phi;
66     end
67
68     if v==0
69         break
70     end
71
72     [alfa]=armijo_opt(l,u,v,tau,ep);
73     u = u+alfa*v;
74
75     x=u(1:n);
76     y=u(n+1:n+m);
77     lamG = u(n+m+1:n+2*m);
78     lamH= u(n+2*m+1:n+3*m);
79
80     err = norm(phi(l,u));
81     iter = iter + 1;
82     verr=[verr;err];
83 end
84 end
85
86 toc
87 f=norm(funobj(l,x));
88
89 if iter >= 200
90     fprintf('Fails to converge within maximum number of
           iterations\n');

```



```

91     fprintf('The iterate returned has relative residual %e\n', f);
92 else
93     fprintf('The method converged at iteration %i with a residual %e\n', iter, f);
94 end
95 %Graficas de la evolucion del error
96
97 figure(1)
98 semilogy(verr)

```

Además, hemos usado los siguientes programas auxiliares, cuyo objetivo principal es construir las matrices del B -diferencial de Φ en $u = (x, y, \lambda)$.

```

1  function [A]=Lambda(l, x, y, lamG, lamH)
2  n=length(x);
3  m=length(y);
4
5     Aylam=diag(aylam(y, lamG, lamH));
6     Bmin=diag(min(0, y));
7     Bmax=diag(max(0, y));
8     derivadaG=derivG(l);
9     derivadaH=derivH(l);
10
11     A=[segderivL(l, x) zeros(n, m) derivadaG derivadaH;
12        zeros(m, n) 2*Aylam 2*Bmin 2*Bmax;
13        derivadaG.' 2*Bmin zeros(m) zeros(m);
14        derivadaH.' 2*Bmax zeros(m) zeros(m)];
15 end

```

Así mismo, para obtener las componentes $\frac{\partial^2 \mathcal{L}}{\partial x^2}(\bar{x}, \bar{\mu})$, $-(G'(\bar{x}))^T \mathbf{y}$ y $-(H'(\bar{x}))^T$ de cada matriz $\Lambda \in \partial_B \Phi(u)$ utilizamos los siguientes programas

```

1  function [aylam]=aylam(y, lamG, lamH)
2  aylam=zeros(length(y), 1);
3  for i=1:length(y)
4     if y(i)<=0
5         aylam(i)=lamG(i);

```

```

6     elseif y(i)>0
7         aylam(i)=lamH(i);
8     end
9 end

```

```

1 function dH = derivH(l)
2 if l==1
3     dH=-[0;1];
4 elseif l==2
5     dH=-[1 1;
6         0 1;
7         0 1];
8 end

```

```

1 function dG = derivG(l)
2 if l==1
3     dG=-[-1;1];
4 elseif l==2
5     dG=-[0 1;
6         1 1;
7         0 -1];
8 end

```

```

1 function dL2 = segderivL(l,x)
2 if l==1
3     dL2=[2 0;
4         0 2];
5 elseif l==2
6     dL2=[0.4 0 0;
7         0 0.2 0;
8         0 0 4.8*x(3)];
9 end

```

Finalmente, la búsqueda lineal con la regla de Armijo, la función objetivo f y la función Φ vienen dadas en las siguientes líneas de código.

```

1 % BUsqueda lineal con la regla de Armijo

```

```

2  %
3  function [alfa]=armiyo_opt(l,u,v,tau,ep)
4  countarm=0;
5
6  % Evaluacion de grad f en u y <grad f,p>
7  Phi=phi(l,u);
8  varphi=0.5*norm(Phi)^2;
9  gradvarphiv=-2*varphi;
10 %
11 % Valores auxiliares iniciales
12 %
13 alfa=1;
14 armiyo=1;
15
16 while armiyo > ep*alfa*gradvarphiv
17     alfa=tau^countarm;
18     u1=u+alfa*v;
19     Phi1=phi(l,u1);
20     Phi=phi(l,u);
21     armiyo=Phi1-Phi;
22     countarm=countarm+1;
23 end
24 end

```

```

1  function [phi] = phi(l,u)
2
3  if l==1
4  x1 = u(1);
5  x2 = u(2);
6  y = u(3);
7  lamG = u(4);
8  lamH = u(5);
9
10 phi1=2*(x1+1)+lamG;
11 phi2=2*(x2-2)-(lamG+lamH);
12 phi3=2*(lamG*min(0,y)+lamH*max(0,y));

```

```

13 phi4=min(0,y)^2-x2+x1;
14 phi5=max(0,y)^2-x2;
15
16 phi=[phi1;phi2;phi3;phi4;phi5];
17 elseif l==2
18 x1=u(1);
19 x2=u(2);
20 x3=u(3);
21 y1=u(4);
22 y2=u(5);
23 lamG1=u(6);
24 lamG2=u(7);
25 lamH1=u(8);
26 lamH2=u(9);
27
28     phi1=0.4*x1-lamG2-(lamH1+lamH2);
29     phi2=0.2*x2-(lamG1+lamG2)-lamH2;
30     phi3=2.4*x3^2+lamG2-lamH2;
31     phi4=2*(lamG1*min(0,y1)+lamH1*max(0,y1));
32     phi5=2*(lamG2*min(0,y2)+lamH2*max(0,y2));
33     phi6=min(0,y1)^2-x2;
34     phi7=min(0,y2)^2-(x1+x2-x3-1);
35     phi8=max(0,y1)^2-x1;
36     phi9=max(0,y2)^2-(x1+x2+x3-1);
37 phi=[phi1;phi2;phi3;phi4;phi5;phi6;phi7;phi8;phi9];
38 end

```

```

1 function [f] = funobj(l,x)
2 if l==1
3     x1=x(1);
4     x2=x(2);
5     f=(x1+1)^2+(x2-2)^2;
6 elseif l==2
7     x1=x(1);
8     x2=x(2);
9     x3=x(3);

```

```
10     f=0.2*(x1^2)+0.1*(x2^2)+0.8*(x3^3);  
11 end  
12 end
```

Referencias bibliográficas

- [1] A. Izmailov M. Solodov, A. Pogosyan. Semismooth newton method for the lifted reformulation of mathematical programs with complementary constraints. *Comput Optim Appl*, 51:199–221, 2010.
- [2] Liqun Qi. Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of Operations Research*, 18(1):227–244, 1993.
- [3] Oliver Stein. Lifting mathematical programs with complementary constraints. *Math. Program., Ser. A*, 131:71–94, 2010.
- [4] Houyuan Jiang y Daniel Ralph. Qpecgen, a matlab generator for mathematical programs with quadratic objectives and affine variational inequality constraints. *Comput Optim Appl*, 13:25–59, 1999.
- [5] J. Pang y L. Qi. Nonsmooth equations: Motivation and algorithms. *SIAM J. Optim.*, 3:443–465, 1993.
- [6] A. Izmailov y M. Solodov. An active-set newton method for the lifted reformulation of mathematical programs with complementary constraints. *SIAM J. Optim.*, 19:1003–1027, 2008.