

UNIVERSIDAD AUTONOMA DE NUEVO LEON
FACULTAD DE CIENCIAS FISICO-MATEMATICAS
DIVISION DE ESTUDIOS DE POSTGRADO



PROBLEMAS DE COMPLEMENTARIEDAD Y
DESIGUALDADES VARIACIONALES: NUEVOS
METODOS DE SOLUCION

TESIS

propuesta para la obtención del Grado de Doctor en Ingeniería
Física Industrial del

Lic. Aarón Arévalo Franco

Asesor: Dra. Nataliya Kalashnykova
Co-asesor: Dr. Vyacheslav Kalashnikov

San Nicolás de los Garza, Nuevo León, México Junio de 2010

Agradecimientos

El presente escrito está dedicado a todos aquellos que hicieron posible éste proyecto. Agadezco en Dios por iluminar éste camino que se emprendió hace cuatro y medio años y que ahora se encuentre concluído en la manera esperada.

A mi familia que han seguido mi desarrollo desde los primeros pasos, por las dificultades que se atravesaron, por su paciencia y por todo el apoyo que me dieron. A mis padres Magdalena y José por permitirme realizar éste sueño, por haberme regalado la vida y por todo su amor incondicional. A mi hermano Adrián, por tu amistad y por lo que compartimos juntos desde niños.

A mi esposa e hija que han sido mi círculo de amor, y porque 'we are one team'. A Maria por estar siempre ahí, por tu soporte, por ser como eres, porque contigo la soledad no existe, por caminar siempre conmigo durante todo éste tiempo. Mi hija Ave Maria que desde pequeña nos diste esa alegría que aún sigue y que día a día la incrementas, por tu apoyo y ayuda que desde tus cinco días de nacida me diste para terminar de escribir esta tesis, o como tu dirías 'erle'.

A mis asesores Dra. Nataliya Kalashnykova y Dr. Vyacheslav Kalashnikov que con todo su apoyo, dedicación y confianza pude sacar adelante éste trabajo, por enseñarme el camino de la investigación y por su amor a las matemáticas, gracias.

Al Dr. Stephan Dempe, por su co-asesoría y que durante mi estancia en

Alemania me dió todo el apoyo y confianza. Su forma de dirigir el proyecto me ayudo mucho a comprender el campo de la investigación.

A la Universität Bergakademie Freiberg por haberme auspiciado en mi estancia en Alemania.

Índice general

1. Introducción	1
2. La Convergencia Global: El Método de Newton Aproximado para Resolver el Problema Implícito de Complementariedad No Lineal	7
2.1. Revisión de la Literatura	7
2.2. Medida del Error y Perturbaciones	11
2.3. Convergencia global del método de Newton	20
2.4. El método de Newton aproximado	38
3. Solución de Problemas Implícitos de Complementariedad con un Algoritmo de Perturbación	56
3.1. Revisión de la Literatura	56
3.2. Existencia de soluciones a problemas...	58
3.3. Taza de convergencia	62
3.4. Resolviendo Problemas Perturbados	66
4. Un Nuevo Algoritmo para Resolver el Problema de Progra-	

mación Binivel Entero Mixto	72
4.1. Revisión de la Literatura	72
4.2. Formulación General	75
4.3. Modelo Matemático	76
4.4. Geometría del problema	78
4.5. Un Algoritmo de Aproximación	84
4.6. Un ejemplo numérico	90
5. Conclusión	94
6. Trabajo futuro	96

Capítulo 1

Introducción

La teoría de la complementariedad trata problemas de optimización en donde el objetivo principal se centra en encontrar un vector $x \in R^n$ tal que para una función $f(x)$ se cumple lo siguiente:

$$x \geq 0, \quad f(x) \geq 0 \quad \text{y} \quad x^T f(x) = 0, \quad (1.1)$$

en donde $f : R^n \rightarrow R^n$ es una aplicación continua y las desigualdades son por componentes. A éste último problema se le conoce como problema estándar de complementariedad. Si decimos además que la función f es lineal, entonces el problema (1.1) es llamado el problema lineal de complementariedad.

Sin embargo, podemos decir que en cierto grado (vea [22]) un problema de complementariedad es análogo al concepto de punto estacionario en problemas de extremos, es decir, si el punto $x = 0$ es un mínimo (local) de una función real diferenciable F definida sobre la mitad positiva $R_+ = [0, +\infty)$

entonces la desigualdad $F'(x) \geq 0$, junto con la condición de complementariedad $x^T F'(x) = 0$ es su condición necesaria. Ahora bien, si el punto en cuestión x es estrictamente positivo entonces la condición del mínimo necesaria se reduce a la igualdad $F'(x) = 0$. Estos dos casos pueden resumirse como sigue:

$$x \geq 0, \quad F'(x) \geq 0 \quad \text{y} \quad xF'(x) = 0.$$

En forma más general, cuando una función diferenciable F está definida sobre un espacio vectorial R^n , y tenemos que encontrar un punto $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ tal que sea un mínimo local de la función F sobre el ortante no negativo R_+^n , las condiciones necesarias para que esto se cumpla son:

$$\frac{\partial F(\bar{x})}{\partial x_i} \geq 0, \quad \text{si } \bar{x}_i = 0, \quad \text{y} \quad \frac{\partial F(\bar{x})}{\partial x_i} = 0, \quad \text{si } \bar{x}_i > 0,$$

para todo $i = 1, \dots, n$. Ahora bien, si ponemos $f_i(x) = \frac{\partial F(x)}{\partial x_i}$, podemos reescribir la última relación en la forma siguiente:

$$f_i(\bar{x}) \geq 0, \quad x_i \geq 0, \quad i = 1, \dots, n, \quad \sum_{i=1}^n \bar{x}_i f_i(\bar{x}) = 0. \quad (1.2)$$

Si nos olvidamos del origen de las funciones f_i y las consideramos arbitrarias, éste último problema (1.2) puede ser llamado el problema de complementariedad no lineal.

El estudio de los problemas de complementariedad es un área de las matemáticas que se ha tornado más que interesante con muchas aplicaciones en la optimización, teoría de juegos, control óptimo estocástico, ingeniería, economía, etc. Además, tiene estrecha relación con dominios de

matemática fundamental como lo son la teoría del punto fijo, análisis no lineal, teoría del grado topológico, desigualdades variacionales así como modelación matemática y análisis numérico.

Los problemas de complementariedad se remontan al estudio hecho por Lemke y Howson en 1963 en donde el cálculo del punto de equilibrio de Nash en la teoría de juegos para dos personas, es equivalente a un problema lineal de complementariedad. Y para 1968 Cottle y Dantzig unificaron la programación lineal y cuadrática con los juegos bimatriciales.

En fin, el desarrollo y producción de nuevas técnicas para su solución han incrementado el número de publicaciones al respecto, así como también el número de aplicaciones. Enseguida se presenta una de esas aplicaciones.

La teoría de la complementariedad tiene una gran variedad de aplicaciones en la vida real, por ejemplo en el campo de la física como lo es la lubricación elasto-electrodinámica [31] en la cual Michael Kostreva lo formula como un problema de complementariedad no lineal.

La lubricación elasto-hidrodinámica es un tipo de lubricación en la cual la deformación elástica de las superficies juega un papel importante en el proceso de lubricación hidrodinámica, implica la presencia de una película de fluido entre las superficies y no existe contacto alguno entre ellas.

Se considera el caso del material puro (sin desplazamiento) en condiciones isotérmicas. Las soluciones de estado serán consideradas constantes. El lubricante entre los dos cilindros en línea elástica de contacto se considera incompresible. Se supondrá que la viscosidad del lubricante llega de su valor de equilibrio de forma instantánea de acuerdo con la ley de la presión

$v = v_0 e^{\alpha p}$, donde v denota la viscosidad, p la presión y α y v_0 son lubricantes parametricos dependientes.

Cualquier carga aplicada externamente será equilibrada por la presión generada en la película lubricante. El espesor de la película lubricante se asumirá positivo sobre todo el dominio de interés.

Entre el punto de entrada denotado por x_a , y el punto de salida x_b ($x_b > x_a$), la fórmula de Reynolds cuantifica la relación entre la presión, espesor de la película y la viscosidad. La entrada de punto x_a será conocida, mientras que el punto de salida (sin límite) x_b se determinará como parte de la solución.

La ecuación de Reynolds, la ecuación de la elasticidad lineal, la ecuación de la presión, de viscosidad y la restricción de carga no son de dimensión correspondiente por $x = \frac{\bar{x}}{a}$, $h = \frac{\bar{h}}{\delta}$ y $p = \frac{\bar{p}}{p_{max}}$.

Además introducimos la variable k que es el equivalente dimensional a la suma de los términos (h_0 , + constante) en la formulación de Dowson y Higginson. Esto lleva al siguiente problema de frontera libre para la distribución de la presión $p(x)$ y el espesor de la película $h(x)$.

Dados los parámetros α y λ y un punto de entrada x_a , encontrar $p(x)$ y $h(x)$, la frontera libre x_b y la variable global k satisfaciendo lo siguiente:

$$\frac{d}{dx} \left(\frac{h^3(x)}{e^{\alpha p}} \frac{dp}{dx} \right) = \lambda \frac{dh}{dx} \text{ en } [x_a, x_b]$$

$$h(x) = x^2 + k - \frac{2}{\pi} \int_{x_a}^{x_b} p(s) \ln|x-s| ds \text{ en } [x_a, \infty)$$

$$\frac{2}{\pi} \int_{x_a}^{x_b} p(s) ds = 1$$

con las condiciones de frontera:

$$p(x_a) = 0 \text{ y } p(x_b) = 0 = \frac{dp}{dx}(x_b).$$

Para que el problema tenga un sentido físico busquemos soluciones $p(x)$ que satisfagan $p(x) \geq 0$ en $[x_a, x_b]$ y $p(x) = 0$ en $[x_b, \infty)$. Para darle éste enfoque es necesario aplicar la Teoría de la Complementariedad. Entonces el anterior problema puede ser reformulado como sigue:

$$R(p, h, k) = 0, \quad \text{en } [x_a, x_b]$$

$$R(p, h, k) \geq 0, \quad \text{en } [x_b, x_F]$$

$$S(p, h, k) = 0, \quad \text{en } [x_a, x_F]$$

$$T(p, h, k) \geq 0, \quad \text{en } [x_b, x_F]$$

sujeto a: $p(x) \geq 0$, en $[x_a, x_b]$, $p(x) = 0$, en $[x_b, x_F]$, $h(x) > 0$ en $[x_a, x_F]$, k sin restricción de signo, junto con las condiciones de frontera. Aquí $R(p, h, k) = \lambda \frac{dh}{dx} - \frac{d}{dx} \left(\frac{h^3(x)}{e^{\alpha p}} \frac{dp}{dx} \right)$, $S(p, h, k) = h(x) - x^2 - k + \frac{2}{\pi} \int_{x_a}^{x_b} p(s) \ln|x-s| ds$ y $T(p, h, k) = 1 - \frac{2}{\pi} \int_{x_a}^{x_b} p(s) ds$.

En el Capítulo 2 se presenta el problema implícito de complementariedad para el cual primero establecemos medidas del error y algunas estimaciones para problemas perturbados, después mediante un cambio de variable pasamos del problema implícito de complementariedad al problema estándar, inmediatamente después se aplican los métodos de Newton exacto e inexacto para su solución y para obtener la tasa de convergencia del punto solución del problema auxiliar al problema original.

En el Capítulo 3 resolvemos el problema implícito de complementariedad mediante un algoritmo de perturbación. Primeramente establecemos condiciones de existencia y unicidad de la solución con respecto a la perturbación, enseguida mostramos la tasa de convergencia de la solución con respecto a la tasa de convergencia de la secuencia de perturbaciones a cero. Y finalmente, trabajamos con un algoritmo para resolver problemas implícitos de complementariedad perturbados.

En el Capítulo 4 se presenta un caso particular de problemas de dos niveles en el que la variable del lado derecho de las restricciones del seguidor funciona como parámetro. Además, consideramos el caso de las variables enteras-mixtas. El presente trabajo lo considera debido a dos razones. La primera viene del hecho ya demostrado de que todo problema de optimización puede ser reformulado como un problema de complementariedad; dicho lo anterior el problema descrito en Capítulo 4, uno de los dos niveles puede ser reformulado como un problema de complementariedad. La segunda razón que ya es considerada en una cantidad creciente de artículos es la de los problemas MPEC (por sus siglas en inglés Mathematical Programs with Equilibrium Constraints) que tratan Problemas de Optimización con restricciones de complementariedad.

Capítulo 2

La Convergencia Global: El

Método de Newton

Aproximado para Resolver el

Problema Implícito de

Complementariedad No Lineal

2.1. Revisión de la Literatura

La Teoría de la Complementariedad es un área relativamente nueva en las Matemáticas Aplicadas, que ha sido rápidamente desarrollada y extendida en los últimos años. Se relaciona con modelación matemática, teoría de

optimización, ingeniería, mecánica estructural, elasticidad, teoría de la lubricación, economía, cálculo variacional, teoría del equilibrio, control óptimo estocástico, etc.

Los artículos publicados hasta ahora sobre éste tema soportan la idea de que éste dominio está en los procesos de desarrollo ambos intensivo y extensivo, y el rango de sus aplicaciones está realmente creciendo.

La Teoría de la Complementariedad ha sido generalizada y extendida a muchas diferentes direcciones. El Problema Implícito de Complementariedad, que es una de tales generalizaciones del problema estándar de complementariedad, fue planteado en la vida de algunos problemas especiales de control óptimo estocástico en [3], [4] y [29], entre muchos otros.

El Problema Implícito de Complementariedad (ICP por sus siglas en inglés) puede ser introducido como sigue: Encontrar un vector $x \in R^n$ tal que:

$$f(x) \geq 0, \quad g(x) \geq 0, \quad \text{y} \quad [g(x)]^T f(x) = 0,$$

en donde $f, g : R^n \rightarrow R^n$ son aplicaciones continuas, y las desigualdades son por componentes.

El ICP ha sido estudiado por un abundante número de autores, *cf.* [6], [8], [18], [19], [21], [25], [26], [28], [35], [36], [37], [39], [40]. Recientemente, un nuevo método ha sido introducido en el estudio de problemas de complementariedad. Éste método se basa en el concepto de una excepcional familia de elementos (EFE) para una aplicación continua. La noción de la EFE fue introducida con el fin de usar la teoría de grado topológico cuando probamos la existencia de las soluciones a los problemas de com-

plementariedad en [25], basada sobre la idea de una secuencia excepcional definida en [45]. Haciendo uso de éste concepto, muchos nuevos resultados han sido obtenidos, vea [6], [7], [21], [23], [25] entre otros.

En [6] y [21], fueron obtenidos interesantes teoremas alternativos relacionados a la existencia de soluciones para problemas implícitos de complementariedad.

En [26], fue demostrado que por una pequeña modificación del concepto de la excepcional familia de elementos (EFE) para un par de funciones y reemplazando los resultados de grado topológico por la Alternativa de Leray-Schauder, podemos obtener una alternativa general, implicando un teorema de existencia para problemas implícitos de complementariedad.

Los autores del artículo [51] introdujeron un nuevo tipo de excepcional familia de elementos (EFE) para resolver problemas implícitos de complementariedad.

En [1], los autores consideran el problema implícito de complementariedad en el marco de la celosía de Hilbert. Ha sido demostrado que éste problema puede ser formulado como un problema del Punto Fijo con adecuados cambios de variable. Por otra parte, ellos mostraron la existencia y unicidad de las soluciones de problemas implícitos de complementariedad haciendo uso de esta formulación.

En [49], es introducido un problema nuevo llamado el *problema implícito de complementariedad generalizado*, que generaliza todas las formulaciones de problemas de complementariedad que han sido estudiados hasta ahora. También muestran aplicaciones del problema implícito de complementariedad.

En [30], es considerada una clase de problemas implícitos de complementariedad con dependencia de parámetros, teniendo una solución única para cada valor del parámetro (control) de un conjunto dado. Entonces un problema de optimización puede ser formulado, tal que un problema implícito de complementariedad surge como una restricción. Ellos usan herramientas de análisis no diferencial y proponen una aproximación a su solución numérica usando un método de optimización no diferenciable y una variante del método de Newton no diferenciable.

Los autores de [32] trataron el problema implícito de complementariedad paramétrico, y [44] trabajó con el problema implícito de complementariedad no lineal.

En [24], fue introducida la noción de una infinitesimal excepcional familia de elementos para un par ordenado de aplicaciones. Usando una aplicación de inversión especial, se presenta una dualidad entre la excepcional familia de elementos para un par ordenado de aplicaciones y la infinitesimal excepcional familia de elementos para un par ordenado de aplicaciones. Usando esta dualidad y la noción de derivadas escalares, se presentan teoremas de existencia para problemas implícitos de complementariedad en espacios de Hilbert.

Soluciones numéricas de problemas de complementariedad pueden ser realizadas mediante varias clases de algoritmos, tales como métodos de Newton y quasi-Newton, métodos de regularización, etc. (*cf.*, por ejemplo, [39], [40], [41], [42], entre otros).

El capítulo está basado en [27] que trabaja el problema de complemen-

tariedad estándar no lineal. De esta manera, hacemos uso del método de Newton y una variante del algoritmo de Newton inexacto para resolver el problema implícito de complementariedad no lineal.

2.2. Medida del Error y Perturbaciones

Consideramos el Problema general de Complementariedad Implícito (no lineal) (ICP por sus siglas en inglés) como sigue: Encontrar un vector $x \in R^n$ tal que:

$$g(x) \geq 0, \quad f(x) \geq 0, \quad \text{y} \quad [g(x)]^T[f(x)] = 0, \quad (2.1)$$

en donde $f, g : R^n \rightarrow R^n$ y $f, g \in C^1(R^n)$. El método de Newton para resolver el problema (2.1) consiste en lo siguiente [42]: generar una secuencia de puntos x^k : seleccionar una iteración inicial $x^0 \in R^n$, y al asumir que el valor x^k es conocido, determinar x^{k+1} como la solución exacta del Problema Implícito de Complementariedad Lineal:

$$z \equiv g(x^k) + g'(x^k)(x - x^k) \geq 0, \quad w \equiv f(x^k) + f'(x^k)(x - x^k) \geq 0, \quad \text{y} \quad z^T w = 0. \quad (2.2)$$

Bajo ciertas suposiciones [41], la secuencia $\{x^k\}$ está definida correctamente y converge localmente a la solución del problema (2.1) a una tasa cuadrática.

Con el fin de examinar el comportamiento del método de Newton, obtendremos algunas estimaciones para una función de medida del error (una función de mérito) $\|h(x)\|$, en donde $h(x) = \min(g(x), f(x))$ es el mínimo

por componentes de los vectores $g(x)$ y $f(x)$.

Definición 2.2.1 Una aplicación $f : R^n \rightarrow R^n$ es llamada fuertemente monótona con respecto a otra aplicación $g : R^n \rightarrow R^n$, si existe un escalar $\beta > 0$ tal que:

$$\langle f(x) - f(y), g(x) - g(y) \rangle \geq \beta \|g(x) - g(y)\|^2, \quad (2.3)$$

para cualquier $x, y \in R^n$.

Definición 2.2.2 Una aplicación $f : R^n \rightarrow R^n$ es llamada Lipschitz-continua, si existe un escalar $L > 0$ tal que:

$$\|f(x) - f(y)\| \leq L \|x - y\|, \quad (2.4)$$

para cualquier $x, y \in R^n$.

Definición 2.2.3 Una aplicación $f : R^n \rightarrow R^n$ es llamada g -Lipschitz continua con respecto a otra aplicación $g : R^n \rightarrow R^n$, si existe un escalar $L > 0$ tal que:

$$\|f(x) - f(y)\| \leq L \|g(x) - g(y)\|, \quad (2.5)$$

para cualquier $x, y \in R^n$.

De aquí en adelante, $\langle \cdot, \cdot \rangle$ denota el producto punto euclideo en R^n .

Lema 2.2.4 Sea $f \in C^1(R^n)$ una aplicación g -fuertemente monótona. Por otra parte, sea g una aplicación fuertemente monótona con un parámetro $\beta_g > 0$, y Lipschitz continua con un parámetro $L_g > 0$.

Entonces, para cualquier conjunto compacto y convexo $V \subset R^n$, existen escalares: $\lambda_V > 0$ y $\delta_V > 0$ tales que para cualquier $x, y \in V$, la siguiente desigualdad está cumplida:

$$\lambda_V \|x - y\| \leq \|h(x) - h(y)\| \leq \delta_V \|x - y\|. \quad (2.6)$$

Demostración. Primeramente, establecemos la siguiente desigualdad a demostrar:

$$\bar{\lambda}_V \|g(x) - g(y)\| \leq \|h(x) - h(y)\| \leq \bar{\delta}_V \|g(x) - g(y)\|. \quad (2.7)$$

Definimos $\chi(x) = g(x) - h(x) \geq 0$, $\tau(x) = f(x) - h(x) \geq 0$ para cualquier punto $x \in R^n$. Además, es evidente que $\chi^T \tau = 0$ para cualquier x . Ahora bien, consideramos el siguiente producto punto:

$$\begin{aligned} & \langle f(x) - h(x) - f(y) + h(y), g(x) - h(x) - g(y) + h(y) \rangle = \\ & = \langle f(x) - f(y), g(x) - g(y) \rangle - \langle h(x), g(x) - g(y) \rangle + \\ & \quad + \langle h(y), g(x) - g(y) \rangle - \langle f(x) - f(y), h(x) \rangle + \\ & + \langle f(x) - f(y), h(y) \rangle + \|h(x)\|^2 + \|h(y)\|^2 - 2 \langle h(x), h(y) \rangle = \\ & = \langle f(x) - f(y), g(x) - g(y) \rangle - \langle h(x) - h(y), g(x) - g(y) \rangle - \\ & \quad - \langle f(x) - f(y), h(x) - h(y) \rangle + \|h(x) - h(y)\|^2. \end{aligned} \quad (2.8)$$

Ahora, ya que f es g -fuertemente monótona, i. e. (2.3) se cumple, entonces tenemos lo siguiente:

$$\langle f(x) - h(x) - f(y) + h(y), g(x) - h(x) - g(y) + h(y) \rangle \geq$$

$$\begin{aligned} &\geq \beta \|g(x) - g(y)\|^2 - \langle h(x) - h(y), g(x) - g(y) \rangle - \\ &\quad - \langle f(x) - f(y), h(x) - h(y) \rangle + \|h(x) - h(y)\|^2. \end{aligned}$$

Por otro lado, el producto punto (2.8) puede ser reescrito como sigue:

$$\begin{aligned} &\langle \tau(x) - \tau(y), \chi(x) - \chi(y) \rangle = \\ &= \chi^T(x)\tau(x) - \chi^T(x)\tau(y) - \chi^T(y)\tau(x) + \chi^T(y)\tau(y) = \\ &= -\chi^T(x)\tau(y) - \chi^T(y)\tau(x) \leq 0; \end{aligned}$$

combinando las dos cadenas, obtenemos:

$$\begin{aligned} &\beta \|g(x) - g(y)\|^2 + \|h(x) - h(y)\|^2 \leq \\ &\leq \langle h(x) - h(y), g(x) - g(y) \rangle + \langle f(x) - f(y), h(x) - h(y) \rangle \\ &\leq \|h(x) - h(y)\| \|g(x) - g(y)\| + \|f(x) - f(y)\| \|h(x) - h(y)\|. \quad (2.9) \end{aligned}$$

Por otra parte, con la condición $f \in C^1(R^n)$, podemos encontrar para cualquier conjunto compacto $V \in R^n$, $\hat{L}_V = \max_{x \in \bar{V}} \|f'(x)\| > 0$. Entonces f es Lipschitz continua con la constante \hat{L}_V , i. e.:

$$\|f(x) - f(y)\| \leq \hat{L}_V \|x - y\|, \forall x, y \in V,$$

por otro lado, si asumimos que g es fuertemente monótona con una constante β_g , i.e.

$$\langle g(x) - g(y), x - y \rangle \geq \beta_g \|x - y\|^2, \quad (2.10)$$

y también (2.3) se cumple, entonces f es g -Lipschitz continua con constante $L_V = \frac{\hat{L}_V}{\beta_g}$, i. e.:

$$\|f(x) - f(y)\| \leq L_V \|g(x) - g(y)\|, \forall x, y \in V. \quad (2.11)$$

Combinando (2.11) con la cadena (2.9), la siguiente desigualdad está cumplida:

$$\|h(x) - h(y)\|^2 - (L_V + 1) \|g(x) - g(y)\| \|h(x) - h(y)\| + \beta \|g(x) - g(y)\|^2 \leq 0,$$

para cualquier $x, y \in V$. De aquí que

$$\frac{\|h(x) - h(y)\|^2}{\|g(x) - g(y)\|^2} - (L_V + 1) \frac{\|h(x) - h(y)\|}{\|g(x) - g(y)\|} + \beta \leq 0.$$

Sea $t = \frac{\|h(x) - h(y)\|}{\|g(x) - g(y)\|}$, entonces tenemos la siguiente desigualdad cuadrática:

$$t^2 - (L_V + 1)t + \beta \leq 0,$$

directamente llegamos a la siguiente estimación necesaria

$$\frac{\beta}{L_V + 1} \|g(x) - g(y)\| \leq \|h(x) - h(y)\| \leq (L_V + 1) \|g(x) - g(y)\|.$$

Sea $\bar{\lambda}_V = \frac{\beta}{L_V + 1}$ y $\bar{\delta}_V = L_V + 1$, esto implica

$$\bar{\lambda}_V \|g(x) - g(y)\| \leq \|h(x) - h(y)\| \leq \bar{\delta}_V \|g(x) - g(y)\|. \quad (2.12)$$

Ahora bien, ya que g es fuertemente monótona, la siguiente desigualdad se cumple:

$$\bar{\lambda}_V \|g(x) - g(y)\| \geq \bar{\lambda}_V \beta_g \|x - y\|. \quad (2.13)$$

Por otro lado, si la aplicación g es Lipschitz continua con la constante $L_g > 0$, i.e.

$$\|g(x) - g(y)\| \leq L_g \|x - y\|, \quad (2.14)$$

se sigue directamente la siguiente desigualdad

$$\bar{\delta}_V \|g(x) - g(y)\| \leq \bar{\delta}_V L_g \|x - y\|. \quad (2.15)$$

Sea $\lambda_V = \bar{\lambda}_V \beta_g$ y $\delta_V = \bar{\delta}_V L_g$. Finalmente, combinando (2.12), (2.13) y (2.15) obtenemos la siguiente desigualdad

$$\lambda_V \|x - y\| \leq \|h(x) - h(y)\| \leq \delta_V \|x - y\|.$$

□

Comentario 2.2.5 *La condición de que la aplicación f sea g -Lipschitz continua no es exótica, ya que esta puede ser prevista, por ejemplo, asumiendo que f es Lipschitz continua y que g es fuertemente monótona. Primero, si g es fuertemente monótona con constante $\beta_g > 0$, i.e. (2.10) es válida y usando la desigualdad de Cauchy-Buniakovski i.e.*

$$\langle g(x) - g(y), x - y \rangle \leq \|g(x) - g(y)\| \|x - y\|,$$

llegamos a la siguiente desigualdad

$$\frac{1}{\beta_g} \|g(x) - g(y)\| \geq \|x - y\|. \quad (2.16)$$

Por otro lado, si la aplicación f es Lipschitz continua con constante $\hat{L}_V > 0$, i.e., la siguiente desigualdad está cumplida

$$\|f(x) - f(y)\| \leq \hat{L}_V \|x - y\|, \quad \forall x, y \in V,$$

juntandola con (2.16), obtenemos

$$\|f(x) - f(y)\| \leq \frac{\hat{L}_V}{\beta_g} \|g(x) - g(y)\|,$$

esto implica que, f es g -fuertemente monótona con la constante $L_V = \frac{\hat{L}_V}{\beta_g} > 0$.

Corolario 2.2.6 *Si x^* es la solución del ICP (2.1) con la aplicación g fuertemente monótona y además $f \in C^1(\mathbb{R}^n)$, entonces para cualquier conjunto convexo y acotado $V \subset \mathbb{R}^n$ que contiene al punto x^* , existen escalares λ_V y δ_V tales que para cualquier $x \in V$, la siguiente desigualdad está cumplida:*

$$\lambda_V \|x - x^*\| \leq \|h(x)\| \leq \delta_V \|x - x^*\|. \quad (2.17)$$

La demostración sigue directamente del Lema 2.2.4, ya que $h(x^*) = 0$.

Ahora, consideramos el Problema Lineal de Complementariedad Perturbado (PLCP por sus siglas en inglés): Encontrar un vector $x \in \mathbb{R}^n$ tal que:

$$\bar{g}(x) = Bx + d - q \geq 0, \quad \bar{f}(x) = Ax + b - p \geq 0, \quad \text{y} \quad [\bar{g}(x)]^T [\bar{f}(x)] = 0, \quad (2.18)$$

en donde p y q son dichas perturbaciones, $A, B \in \mathbb{R}^{n \times n}$, $b, d, p, q \in \mathbb{R}^n$, donde la matriz B se asume que tiene inversa. El problema (2.18) es equivalente al siguiente problema: Encontrar un vector $y \in \mathbb{R}^n$ tal que:

$$\begin{aligned} y \geq 0, \quad \phi(y) &\equiv AB^{-1}y - AB^{-1}d + AB^{-1}q + b - p \geq 0 \\ \text{y} \quad y^T \phi(y) &= 0, \end{aligned} \quad (2.19)$$

si y sólo si $y = Bx + d - q$.

Lema 2.2.7 *Sea ϕ una aplicación fuertemente monótona, i.e., la matriz AB^{-1} es positiva definida.*

Entonces para cualquier par de vectores perturbación $p, q \in \mathbb{R}^n$, existe el único vector $y(p, q)$ que resuelve al problema (2.19). Además, si x_1 resuelve

(2.18) con las perturbaciones p_1, q_1 , y x_2 hace lo mismo para p_2, q_2 , entonces la siguiente desigualdad está cumplida:

$$\|x_1 - x_2\| \leq \frac{1}{\eta\rho} \|(p_1 - p_2) + AB^{-1}(q_2 - q_1)\|, \quad (2.20)$$

en donde $\rho = \sqrt{\lambda_{\min}(B^T B)}$, y η es la constante de la definidad positiva de AB^{-1} .

Demostración. Para cualquier punto fijo $x \in R^n$ y para cualquier $u \in R^n$ podemos escribir lo siguiente:

$$\begin{aligned} u^T \phi'(x)u &= \lim_{t \rightarrow 0} \frac{1}{t} (\phi(x + tu) - \phi(x), u) = \\ &= \lim_{t \rightarrow 0} \frac{1}{t^2} (\phi(x + tu) - \phi(x), tu) \geq \eta \|u\|^2, \end{aligned}$$

entonces, como para cualquier $x \in R^n$ la matriz $\phi'(y(x, p, q)) = AB^{-1}$ es positiva definida, de aquí que la existencia de la única solución $y(x, p, q)$ del problema (2.19) esté garantizada para cualquier $x, p, q \in R^n$.

Ahora bien, consideremos $y_1 = y(x, p_1, q_1), y_2 = y(x, p_2, q_2)$ como soluciones del problema con perturbaciones p_1, q_1 y p_2, q_2 , tomando lugar en las siguientes condiciones:

$$\begin{aligned} y_1 \geq 0, Py_1 &\equiv AB^{-1}y_1 - AB^{-1}d + b \geq p_1 - AB^{-1}q_1, y_1 Py_1 = y_1[p_1 - AB^{-1}q_1], \\ y_2 \geq 0, Py_2 &\equiv AB^{-1}y_2 - AB^{-1}d + b \geq p_2 - AB^{-1}q_2, y_2 Py_2 = y_2[p_2 - AB^{-1}q_2], \end{aligned} \quad (2.21)$$

por otro lado, las relaciones (2.21) implican las siguientes desigualdades:

$$-y_1 Py_2 \leq -y_1 [p_2 - AB^{-1}q_2], \quad -y_2 Py_1 \leq -y_2 [p_1 - AB^{-1}q_1],$$

que dan lugar a las siguientes desigualdades:

$$(y_1 - y_2)(Py_1 - Py_2) \leq (y_1 - y_2) [(p_1 - AB^{-1}q_1) - (p_2 - AB^{-1}q_2)],$$

$$(y_1 - y_2)AB^{-1}(y_1 - y_2) \leq (y_1 - y_2) [(p_1 - p_2) - AB^{-1}(q_2 - q_1)],$$

y ya que AB^{-1} es positiva definida con el parámetro η , se tiene

$$\eta \|y_1 - y_2\|^2 \leq \|y_1 - y_2\| \|(p_1 - p_2) - AB^{-1}(q_2 - q_1)\|$$

$$\|y_1 - y_2\| \leq \frac{1}{\eta} \|(p_1 - p_2) - AB^{-1}(q_2 - q_1)\|.$$

Sin embargo, si regresamos a nuestras variables originales, tenemos

$$y_1 - y_2 = B(x_1 - x_2)$$

y entonces

$$\|B(x_1 - x_2)\| \leq \frac{1}{\eta} \|(p_1 - p_2) + AB^{-1}(q_2 - q_1)\|.$$

Ahora analicemos el lado izquierdo, sabemos que:

$$\|B(x_1 - x_2)\|^2 = \langle B(x_1 - x_2), B(x_1 - x_2) \rangle$$

$$= \langle B^T B(x_1 - x_2), x_1 - x_2 \rangle > 0.$$

Ahora, sea $\rho^2 = \lambda_{\min}(B^T B) > 0$, entonces

$$\|B(x_1 - x_2)\|^2 \geq \rho^2 \|x_1 - x_2\|^2$$

y ya que

$$\|B(x_1 - x_2)\|^2 \leq \frac{1}{\eta^2} \|(p_1 - p_2) + AB^{-1}(q_2 - q_1)\|^2,$$

por lo tanto

$$\|x_1 - x_2\| \leq \frac{1}{\eta\rho} \|(p_1 - p_2) + AB^{-1}(q_2 - q_1)\|.$$

□

2.3. Convergencia global del método de Newton

Reescribamos el problema implícito de complementariedad lineal (2.2) en la siguiente forma

$$F(x) \geq 0, \quad G(x) \geq 0, \quad \text{y} \quad [F(x)]^T[G(x)] = 0;$$

aquí, $F(x) = Ax + b$ y $G(x) = Mx + c$, en donde $A = \nabla f(x^k)$, $b = f(x^k) - \nabla f(x^k)x^k$, $M = \nabla g(x^k)$, $c = g(x^k) - \nabla g(x^k)x^k$, para cualquier x^k .

Entonces, podemos reformular el problema como sigue:

$$F(x) = \nabla f(x^k)x + (f(x^k) - \nabla f(x^k)x^k)$$

$$G(x) = \nabla g(x^k)x + (g(x^k) - \nabla g(x^k)x^k)$$

haciendo un cambio de variable del siguiente modo

$$y = \nabla g(x^k)x + (g(x^k) - \nabla g(x^k)x^k),$$

despejamos:

$$x = [\nabla g(x^k)]^{-1} y - [\nabla g(x^k)]^{-1} (g(x^k) - \nabla g(x^k)x^k).$$

Y ahora sustituyendo en $F(x)$, obtenemos lo siguiente:

$$\begin{aligned} \Phi(y) = \nabla f(x^k) \left([\nabla g(x^k)]^{-1} y - [\nabla g(x^k)]^{-1} (g(x^k) - \nabla g(x^k)x^k) \right) + \\ + (f(x^k) - \nabla f(x^k)x^k). \end{aligned}$$

Además, en cada paso del método de Newton (exacto), en vez del problema implícito de complementariedad linealizado (2.2), podemos considerar el problema lineal de complementariedad estándar: Encontrar un vector $y^k \in R^n$, tal que

$$y^k \geq 0, \quad \Phi(y^k) \geq 0, \quad \text{y} \quad (y^k)^T[\Phi(y^k)] = 0, \quad \forall k = 0, 1, \dots \quad (2.22)$$

Ahora, deduciremos condiciones que garanticen la convergencia del método de Newton iniciando en cualquier punto inicial $x^0 \in R^n$.

De aquí en adelante diremos que $\|D\| = \sqrt{\lambda_{\max}(D^T D)}$.

Definición 2.3.1 *A una matriz cuadrada $A \in L(R^n)$ se le llama M -matriz, si $a_{ij} \leq 0$ para $i \neq j$, A es no singular, y $A^{-1} \geq 0$.*

Teorema 2.3.2 *Sea $g = Bx + d$ con la matriz B positiva definida y $\Phi : R^n \rightarrow R^n$ una aplicación fuertemente monótona y tal que cada una de sus componentes es una función convexa, y en cualquier punto $y \in R_+^n$ la matriz $\Phi'(y)$ es una M -matriz.*

Entonces, para cualquier punto inicial $y^0 = g(x^0) \in R^n$, la secuencia $\{y^k\}$ generada por el método de Newton converge a $y^ = g(x^*)$, y además, el proceso (2.2) converge a $x^* = g^{-1}(y^*)$, que es la única solución del problema (2.1). Si, además, $f \in C^2(R^n)$, entonces la tasa de convergencia es cuadrática, es decir, existe una constante $C > 0$, tal que la siguiente desigualdad está cumplida:*

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2, \quad k = 0, 1, 2, \dots \quad (2.23)$$

Demostración. Considerando un punto inicial x^0 cualquiera, podemos construir el sistema:

$$y = \nabla g(x^0)x - (g(x^0) - \nabla g(x^0)x^0),$$

$$\Phi(y) = \nabla f(x^0) [\nabla g(x^0)]^{-1} y + \nabla f(x^0) \left(- [\nabla g(x^0)]^{-1} (g(x^0) - \nabla g(x^0)x^0) \right) + f(x^0) - \nabla f(x^0)x^0.$$

Encontramos y^1 como una solución al problema:

$$y \geq 0,$$

$$w = \nabla f(x^0) [\nabla g(x^0)]^{-1} y + \nabla f(x^0) \left(- [\nabla g(x^0)]^{-1} (g(x^0) - \nabla g(x^0)x^0) \right) + f(x^0) - \nabla f(x^0)x^0 \geq 0 \quad \text{y} \quad y^T w = 0,$$

por esa razón, las siguientes desigualdades son válidas

$$y^1 \geq 0, \quad \Phi(y^1) \geq \nabla f(x^0) [\nabla g(x^0)]^{-1} y^1 + \nabla f(x^0) \left[- [\nabla g(x^0)]^{-1} (g(x^0) - \nabla g(x^0)x^0) \right] + f(x^0) - \nabla f(x^0)x^0 \geq 0.$$

Ahora, ya que todas las componentes de la aplicación Φ son convexas, podemos probar que las siguientes desigualdades son válidas:

$$y^k \geq 0 \quad \text{y} \quad \Phi(y^k) \geq 0, \tag{2.24}$$

para todo $k = 1, 2, 3, \dots$. Ahora bien, consideramos y^{k+1} para cualquier $k \geq 0$ y construimos el conjunto $\alpha_k = \{i : 1 \leq i \leq n, y_i^{k+1} \geq 0\}$. Si $\alpha_k = \emptyset$, y ya que (2.24) es válida, el proceso termina; $y^{k+1} = 0 = y^*$, y $x^* = \nabla^{-1}g(x^k)y^k - \nabla^{-1}g(x^k)(g(x^k) - \nabla g(x^k)x^k)$, esto es

$$x^* = -\nabla^{-1}g(x^k)(g(x^k) - \nabla g(x^k)x^k),$$

es la solución del problema (2.1).

Por otro lado, si $\alpha_k \neq \emptyset$, ya que y^{k+1} es la solución del problema (2.22), la existencia y unicidad siguen directamente de la monotonía fuerte de Φ , entonces las siguientes igualdades se cumplen:

$$y_j^{k+1} = 0 \quad j \notin \alpha_k$$

$$w_i^{k+1} \equiv f_i(x^k) + \nabla f(x^k)[\nabla g(x^k)]^{-1}(y^{k+1} - g(x^k)) = 0, \quad i \in \alpha_k. \quad (2.25)$$

Por comodidad denotamos $\alpha = \alpha_k$, $\bar{\alpha} = \{1, \dots, n\} \setminus \alpha$, y reescribimos (2.25) en su forma vectorial, esto es

$$y_{\bar{\alpha}}^{k+1} = 0_{\bar{\alpha}};$$

$$\begin{aligned} w_{\alpha}^{k+1} &\equiv f_{\alpha}(x^k) + \nabla f_{\alpha\alpha}(x^k)[\nabla g_{\alpha\alpha}(x^k)]^{-1}(y_{\alpha}^{k+1} - g_{\alpha}(x^k)) + \\ &+ \nabla f_{\alpha\bar{\alpha}}(x^k)[\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1}(y_{\bar{\alpha}}^{k+1} - g_{\bar{\alpha}}(x^k)) = 0_{\alpha}. \end{aligned} \quad (2.26)$$

Es bien sabido (vea [5], Capítulo 6. Teoremas 2.3 y 2.4) que cualquier submatriz principal de una M-matriz es M-matriz. Entonces las siguientes desigualdades se cumplen:

$$[\nabla f_{\alpha\alpha}(x^k)[\nabla g_{\alpha\alpha}(x^k)]^{-1}]^{-1} \geq 0, \quad \nabla f_{\alpha\bar{\alpha}}(x^k)[\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1} \leq 0. \quad (2.27)$$

Multiplicando ambas partes de (2.26) por $[\nabla f_{\alpha\alpha}(x^k)[\nabla g_{\alpha\alpha}(x^k)]^{-1}]^{-1}$ en el lado izquierdo y tomando en cuenta que $y_{\bar{\alpha}} = 0_{\bar{\alpha}}$, obtenemos la siguiente igualdad:

$$\begin{aligned} 0_{\bar{\alpha}} &= [\nabla f_{\alpha\alpha}(x^k)[\nabla g_{\alpha\alpha}(x^k)]^{-1}]^{-1} f_{\alpha}(x^k) + (y_{\alpha}^{k+1} - g_{\alpha}(x^k)) - \\ &- [\nabla f_{\alpha\alpha}(x^k)[\nabla g_{\alpha\alpha}(x^k)]^{-1}]^{-1} \nabla f_{\alpha\bar{\alpha}}(x^k)[\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1} g_{\bar{\alpha}}(x^k). \end{aligned}$$

De esta igualdad encontramos y_α^{k+1} , en donde

$$\begin{aligned} y_\alpha^{k+1} &= g_\alpha(x^k) - \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} f_\alpha(x^k) + \\ &+ \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \nabla f_{\alpha\bar{\alpha}}(x^k) [\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1} g_{\bar{\alpha}}(x^k). \end{aligned} \quad (2.28)$$

De (2.24), (2.27) y (2.28) tenemos $y_\alpha^{k+1} < g_\alpha(x^k)$, para $k \geq 1$. Entonces $y_\alpha^{k+1} < y_\alpha^k - \nabla g(x^k)(x - x^k)$ implica $y_\alpha^{k+1} < y_\alpha^k$ y ya que $y_\alpha^{k+1} = 0_{\bar{\alpha}} \leq y_{\bar{\alpha}}^k$, entonces las siguientes desigualdades están cumplidas:

$$0 \leq y^{k+1} \leq y^k \leq \dots \leq y^1,$$

para $k = 1, 2, \dots$. Por lo tanto $\{y^k\}_{k=1}^\infty$ es monotonamente no creciente y acotada; además tiene el punto límite $y^* \geq 0$. Para investigar la tasa de convergencia, linealizamos $f(x)$ para cualquier k , obteniendo $F(x) = f(x^k) + \nabla f(x^k)(x - x^k)$ y sea $G(x)$ lineal, es decir, $G(x) = g(x) = Bx + d$. Ahora consideremos el producto punto

$$\langle w^{k+1} - w^*, y^{k+1} - y^* \rangle = -(y^{k+1})^T w^* - (y^*)^T w^{k+1} \leq 0.$$

Por otro lado

$$\begin{aligned} &\langle w^{k+1} - w^*, y^{k+1} - y^* \rangle = \\ &= (y^{k+1} - y^*)^T \left(f(x^k) - f(x^*) + \nabla f(x^k) [\nabla g(x^k)]^{-1} (y^{k+1} - g(x^k)) \right) = \\ &= (y^{k+1} - y^*)^T (f(x^{k+1}) - f(x^*) - f(x^{k+1}) + f(x^k) + \\ &+ \nabla f(x^k) [\nabla g(x^k)]^{-1} (y^{k+1} - g(x^k))) = (y^{k+1} - y^*)^T (f(x^{k+1}) - f(x^*)) - \\ &- (y^{k+1} - y^*)^T \left(f(x^{k+1}) - f(x^k) - \nabla f(x^k) [\nabla g(x^k)]^{-1} (y^{k+1} - g(x^k)) \right). \end{aligned}$$

Ahora, ya que $g(x) = Bx + d = y$, entonces $y^{k+1} = Bx^{k+1} + d$ y $y^* = Bx^* + d$. De aquí obtenemos que $y^{k+1} - y^* = g(x^{k+1}) - g(x^*)$. Además, ya que (2.3) es válida, tenemos la siguiente desigualdad

$$[g(x^{k+1}) - g(x^*)]^T [f(x^{k+1}) - f(x^*)] \geq \beta \|g(x^{k+1}) - g(x^*)\|^2.$$

Pero también sabemos que

$$\beta \|g(x^{k+1}) - g(x^*)\|^2 = \beta \|B(x^{k+1} - x^*)\|^2.$$

Con esto llegamos a la siguiente igualdad

$$\beta \|B(x^{k+1} - x^*)\|^2 = \beta (x^{k+1} - x^*)^T B^T B (x^{k+1} - x^*),$$

si B es no singular, entonces $B^T B$ es positiva definida, i.e., existe un escalar $\eta > 0$ tal que $u^T B^T B u \geq \eta \|u\|^2$, entonces usando la desigualdad de Cauchy-Buniakovski obtenemos lo siguiente

$$\begin{aligned} & [g(x^{k+1}) - g(x^*)]^T [f(x^{k+1}) - f(x^*)] \leq \\ & \|g(x^{k+1}) - g(x^*)\| \cdot \|f(x^{k+1}) - f(x^k) - \nabla f(x^k)[\nabla g(x^k)]^{-1}(y^{k+1} - g(x^k))\|, \end{aligned}$$

eso significa que,

$$\begin{aligned} & \beta \eta \|x^{k+1} - x^*\|^2 \leq \|g(x^{k+1}) - g(x^*)\| \times \\ & \times \|f(x^{k+1}) - f(x^k) - \nabla f(x^k)[\nabla g(x^k)]^{-1}(y^{k+1} - g(x^k))\| \leq \\ & \leq \|B\| \|x^{k+1} - x^*\| \|f(x^{k+1}) - f(x^k) - \nabla f(x^k)[\nabla g(x^k)]^{-1}(y^{k+1} - g(x^k))\| \\ & \|x^{k+1} - x^*\| \leq \frac{\|B\|}{\beta \eta} \|f(x^{k+1}) - f(x^k) - \nabla f(x^k)[\nabla g(x^k)]^{-1}(y^{k+1} - g(x^k))\| \end{aligned}$$

$$\|x^{k+1} - x^*\| \leq \frac{\|B\|}{\beta\eta} \|f(x^{k+1}) - f(x^k) - \nabla f(x^k)[\nabla g(x^k)]^{-1}B(x^{k+1} - x^k)\|.$$

Ahora bien, ya que $y^{k+1} = Bx^{k+1} + d$ y $g(x^k) = Bx^k + d$, entonces $y^{k+1} - g(x^k) = B(x^{k+1} - x^k)$. Por otro lado $\nabla g(x^k) = B$, entonces obtenemos la siguiente estimacion:

$$\|x^{k+1} - x^*\| \leq \frac{\|B\|}{\beta\eta} \|f(x^{k+1}) - f(x^k) - \nabla f(x^k)(x^{k+1} - x^k)\|. \quad (2.29)$$

Sabemos que si $f \in C^2$, entonces para cualquier conjunto convexo y acotado $V \subset R^n$, existe un escalar $r_v > 0$ tal que

$$\|f(x) - f(y) - f'(y)(x - y)\| \leq r_v \|x - y\|^2, \quad \forall x, y \in V, \quad (2.30)$$

esto implica que

$$\|f(x^{k+1}) - f(x^k) - \nabla f(x^k)(x^{k+1} - x^k)\| \leq r_v \|x^{k+1} - x^k\|^2,$$

de aquí y de (2.29)

$$\|x^{k+1} - x^*\| \leq \frac{\|B\| r_v}{\beta\eta} \|x^{k+1} - x^k\|^2.$$

Sea $C = \frac{\|B\| r_v}{\beta\eta}$ entonces llegamos a la siguiente desigualdad

$$\|x^{k+1} - x^*\| \leq C \|x^{k+1} - x^k\|^2. \quad (2.31)$$

De la monotonocidad de la secuencia $\{y^k\}_{k=1}^\infty$, tenemos $0 \leq y^{k+1} \leq y^k \leq \dots \leq y^1$ para $k = 1, 2, \dots$. Ahora, regresando a la variable original x , tenemos que

$$0 \leq Bx^{k+1} + d \leq Bx^k + d \leq \dots \leq Bx^1 + d.$$

Por otro lado, es claro que

$$\begin{aligned} \|(Bx^{k+1} + d) - (Bx^k + d)\| &\leq \|(Bx^k + d) - (Bx^* + d)\| \\ \|B(x^{k+1} - x^k)\| &\leq \|B(x^k - x^*)\| \end{aligned}$$

esto implica

$$\|x^{k+1} - x^k\| \leq q \|x^k - x^*\| \quad \text{para cierto valor } q > 0.$$

De aquí y de (2.31), la desigualdad (2.23) está cumplida, por lo tanto la tasa de convergencia es cuadrática.

□

Definición 2.3.3 *Decimos que dos soluciones x_1 y x_2 del problema (2.1) son funcionalmente equivalentes si $f(x_1) = f(x_2)$ y $g(x_1) = g(x_2)$. Además, decimos que la solución del problema (2.1) es única, si todos los $x \in R^n$ que resuelven al problema (2.1) son funcionalmente equivalentes.*

Definición 2.3.4 *(vea Definición 4.2.1 en [22] Capítulo 4.2) Decimos que una familia de elementos $\{x^r\}_{r>0} \subset R^n$ es una excepcional familia de elementos para el par (f, g) con respecto a R_+^n si las siguientes condiciones están cumplidas*

1. $\|x^r\| \rightarrow +\infty$ cuando $r \rightarrow +\infty$,
2. $g(x^r) \geq 0$ para cada $r > 0$,
3. para cada $r > 0$, existe $\mu_r > 0$ tal que para $i = 1, 2, \dots, n$ tenemos
 - (i) $f_i(x^r) = -\mu_r g_i(x^r)$, si $g_i(x^r) > 0$,
 - (ii) $f_i(x^r) \geq 0$ si $g_i(x^r) = 0$.

Definición 2.3.5 (vea Definición 5.2.5 en [22] Capítulo 5.2) Decimos que el par de aplicaciones (f, g) que van de $R^n \rightarrow R^n$ es (ρ, g) -copositivo, con respecto a un cono cerrado convexo $K \subset R^n$, si ahí existe $\rho > 0$ y $\alpha > 0$ tales que para cada $x \in K$, con $\|x\| > \rho$ tenemos

1. $\langle g(x), f(x) \rangle \geq 0$,
2. $\langle g(x), x \rangle \geq \alpha \|x\|^2$.

Definición 2.3.6 (vea Definición 5.2.2 en [22] Capítulo 5.2) Decimos que el par de aplicaciones (f, g) satisfacen la condición (θ_g) con respecto a K si existe $\rho > 0$ tal que para cualquier $x \in K$ con $\|x\| > \rho$, existe $y \in K$ tal que:

1. $\langle g(x) - y, f(x) \rangle \geq 0$ y
2. $\langle g(x) - y, x \rangle > 0$.

Proposición 2.3.7 (vea Proposición 5.2.10 en [22] Capítulo 5.2) Si el par de aplicaciones (f, g) que van de $R^n \rightarrow R^n$ es (ρ, g) -copositivo, con respecto a un cono cerrado convexo $K \subset R^n$, entonces el par (f, g) satisface la condición (θ_g) .

Teorema 2.3.8 (vea Teorema 5.24 en [22] Capítulo 5.2) Sea $(R^n, \langle \cdot, \cdot \rangle)$ un espacio Euclideano, $K \subset R^n$ un cono cerrado convexo puntajeado y $f, g : R^n \rightarrow R^n$ un par de aplicaciones. Si el par (f, g) satisface la condición (θ_g) , entonces el par (f, g) está sin una excepcional familia de elementos (EFE) con respecto a K .

Corolario 2.3.9 (vea Corolario 5.2.3 en [22]) Sea $(R^n, \langle \cdot, \cdot \rangle)$ el espacio Euclideano n -dimensional, $K \subset R^n$ un cono cerrado convexo y $f, g : R^n \rightarrow$

R^n un par de aplicaciones continuas. Si el par (f, g) está sin EFEs, entonces el ICP (f, g, K) tiene solución.

Ahora, podemos describir la obtención de otro teorema de convergencia para el método de Newton exacto como sigue. Sea $x^0 \in R^n$ cualquier punto; si tenemos $x^k \in R^n$, entonces encontramos $x^{k+1} \in R^n$ como una solución exacta del problema implícito de complementariedad lineal:

$$z_k(x) \equiv f(x^k) + f'(x^k)(x - x^k) \geq 0,$$

$$w_k(x) \equiv g(x^k) + g'(x^k)(x - x^k) \geq 0, \quad \text{y} \quad [z_k(x)]^T [w_k(x)] = 0.$$

Teorema 2.3.10 Sean $f(x) = Ax + \varphi(x)$ y $g(x) = Bx + \sigma(x)$, donde $A, B \in L(R^n)$, la matriz B sea positiva definida con un parámetro $\bar{\eta} > 0$, y la matriz $B^T A$ sea positiva definida con parámetro $\mu > 0$ (i.e., $u^T B u \geq \bar{\eta} \|u\|^2$ y $u^T B^T A u \geq \mu \|u\|^2$ para todo $u \in R^n$). Además, asumimos que las aplicaciones $\varphi : R^n \rightarrow R^n$, $\sigma : R^n \rightarrow R^n$ son continuamente diferenciables con derivadas acotadas:

$$\|\varphi'(x)\| \leq q\mu,$$

$$\text{y} \quad \|\sigma'(x)\| \leq \min\{p\mu, \bar{\eta} - \gamma\}, \quad \text{en donde} \quad 0 < \gamma < \bar{\eta}, \quad \forall x \in R^n, \quad (2.32)$$

y también

$$q \|B\| + p \|A\| + 3pq\mu < 1/3. \quad (2.33)$$

Entonces, para cualquier punto inicial $x^0 \in R^n$, la secuencia $\{x^k\}$ construida por el método de Newton (exacto), converge linealmente a x^* , la única solución del problema (2.1). Además, si $\varphi, \sigma \in C^2(R^n)$, entonces la tasa de

convergencia es cuadrática, i.e., la siguiente desigualdad se cumple para alguna constante C :

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2, \quad k = 1, 2, \dots \quad (2.34)$$

Demostración. No es difícil demostrar que las condiciones (2.32) son suficientes para que la aplicación f sea g -fuertemente monótona. En verdad, para cualquier $x, y \in R^n$ la siguiente cadena de relaciones está cumplida:

$$\begin{aligned} & \langle g(x) - g(y), f(x) - f(y) \rangle = \\ & = (x - y)^T B^T A(x - y) + (x - y)^T B^T (\varphi(x) - \varphi(y)) + \\ & + (\sigma(x) - \sigma(y))^T A(x - y) + (\sigma(x) - \sigma(y))^T (\varphi(x) - \varphi(y)) \geq \\ & \geq \mu \|x - y\|^2 - \|x - y\| \|B^T (\varphi(x) - \varphi(y))\| - \\ & - \|(\sigma(x) - \sigma(y))A\| \|x - y\| - \|(\sigma(x) - \sigma(y))\| \|(\varphi(x) - \varphi(y))\| \geq \\ & \geq \mu \|x - y\|^2 - q\mu \|B\| \|x - y\|^2 - p\mu \|A\| \|x - y\|^2 - pq\mu^2 \|x - y\|^2 = \\ & = (\mu - q\mu \|B\| - p\mu \|A\| - pq\mu^2) \|x - y\|^2. \end{aligned}$$

Por otro lado, mediante la desigualdad del valor medio

$$\|g(x) - g(y)\|^2 \leq (\|B\| + p\mu)^2 \|x - y\|^2,$$

sea

$$N = \frac{\mu - \mu q \|B\| - \mu p \|A\| - \mu^2 pq}{(\|B\| + p\mu)^2},$$

aquí garantizamos que $N > 0$ debido a que se cumple la siguiente desigualdad

$$1 - q \|B\| - p \|A\| - pq\mu > 0,$$

que viene directamente de (2.33). Con esto se concluye que

$$\langle g(x) - g(y), f(x) - f(y) \rangle \geq N \|g(x) - g(y)\|^2, \quad \forall x, y \in R^n.$$

De aquí, y de ([22] Capítulo 6, Teorema 6.10), sigue la existencia de una solución $x^* \in R^n$.

Con el fin de probar la unicidad de la solución en el sentido de la Definición 2.3.3, consideramos $x_1, x_2 \in R^n$ como dos diferentes soluciones del LICP, suponemos que $g(x_1) \neq g(x_2)$, y ya que la aplicación f es g -fuertemente monótona tenemos la siguiente cadena:

$$\begin{aligned} 0 &\geq -[g(x_1)]^T[f(x_2)] - [g(x_2)]^T[f(x_1)] = \\ &= [g(x_1)]^T[f(x_1)] - [g(x_1)]^T[f(x_2)] - [g(x_2)]^T[f(x_1)] - [g(x_2)]^T[f(x_2)] = \\ &= [g(x_1) - g(x_2)]^T [f(x_1) - f(x_2)] \geq N \|g(x_1) - g(x_2)\|^2 > 0. \end{aligned}$$

Esto nos da una contradicción, entonces $g(x_1) = g(x_2)$ para soluciones x_1, x_2 en donde $x_1 \neq x_2$. Del mismo modo g siendo f -fuertemente monótona, concluimos que $f(x_1) = f(x_2)$ para soluciones x_1, x_2 en donde $x_1 \neq x_2$. Entonces con esto demostramos la unicidad en sentido de la Definición 2.3.3.

Ahora probaremos que es única en el sentido comúnmente conocido. Sabemos que la matriz B es positiva definida con el parámetro $\bar{\eta}$, es decir la siguiente desigualdad está cumplida

$$u^T B u \geq \bar{\eta} \|u\|^2.$$

Utilizando esta definidad positiva, las desigualdades (2.32) y el Teorema del Valor Medio, demostraremos la monotonía fuerte de la aplicación g . Verdaderamente,

$$\begin{aligned}
& \langle g(x_1) - g(x_2), x_1 - x_2 \rangle = \\
& = \langle B(x_1 - x_2), x_1 - x_2 \rangle + \langle \sigma(x_1) - \sigma(x_2), x_1 - x_2 \rangle \geq \\
& \geq \bar{\eta} \|x_1 - x_2\|^2 - \|\sigma'(x_1 + \theta(x_2 - x_1))\| \|x_1 - x_2\|^2 \geq \\
& \geq \gamma \|x_1 - x_2\|^2,
\end{aligned}$$

para $x_1, x_2 \in R^n$. Esto último implica que la solución es única: al suponer el contrario, es decir, que hay dos soluciones del problema original, tales que $x_1 \neq x_2$ pero $g(x_1) = g(x_2)$, inmediatamente llegaríamos a una desigualdad imposible $0 > 0$.

Ahora, aplicamos el método de Newton como sigue:

Sea $x^0 \in R^n$ un punto cualquiera; si tenemos $x^k \in R^n$, entonces encontramos $x^{k+1} \in R^n$ como una solución exacta del Problema Implícito de Complementariedad Lineal (LICP) como sigue:

$$z_k(x) \equiv Bx + \sigma(x^k) + \sigma'(x^k)(x - x^k) \geq 0,$$

$$w_k(x) \equiv Ax + \varphi(x^k) + \varphi'(x^k)(x - x^k) \geq 0, \text{ y } [z_k(x)]^T [w_k(x)] = 0. \quad (2.35)$$

Con el fin de demostrar que la secuencia $\{x^k\}$ generada por el método de Newton en las últimas expresiones está bien definido, debemos mostrar que el (LICP) una tiene solución. Según las suposiciones, tenemos que la matriz B es positiva definida con una constante $\bar{\eta} > 0$ y $\|\sigma'(x)\| \leq \bar{\eta} - \gamma$,

$\forall x \in R^n$. Anteriormente ya demostramos que en este caso la matriz $B + \sigma'(x)$ es positiva definida $\forall x \in R^n$, con la constante $\gamma > 0$.

También podemos probar que la matriz $(B + \sigma'(x))^T(A + \varphi'(x))$ es positiva definida $\forall x \in R^n$, con una constante $\bar{\mu} > 0$ ya que

$$\begin{aligned}
x^T(B + \sigma'(x))^T(A + \varphi'(x))x &= \\
&= x^T B^T A x + x^T B^T \varphi'(x)x + x^T \sigma'(x)^T A x + x^T \sigma'(x)^T \varphi'(x)x \leq \\
&\geq \mu \|x\|^2 - \|B\| \|\sigma'(x)\| \|x\|^2 - \|A\| \|\varphi'(x)\| \|x\|^2 - \|\sigma'(x)\| \|\varphi'(x)\| \|x\|^2 \geq \\
&\geq \mu \|x\|^2 - \|B\| q\mu \|x\|^2 - \|A\| p\mu \|x\|^2 - pq\mu^2 \|x\|^2 = \\
&= \mu(1 - q\|B\| - p\|A\| - pq\mu) \|x\|^2.
\end{aligned}$$

Sea $\bar{\mu} = \mu(1 - q\|B\| - p\|A\| - pq\mu)$ y porque $(1 - q\|B\| - p\|A\| - pq\mu) > 0$ deducida directamente de (2.33), entonces $\bar{\mu} > 0$.

Ahora, mostraremos que el par (w, z) es (ρ, z) -cpositivo. Primeramente, consideramos la siguiente cadena de relaciones

$$\begin{aligned}
\langle z(x), x \rangle &= x^T(B + \sigma'(x^k))^T x + (\sigma(x^k) - \sigma'(x^k)x^k)^T x \geq \\
&\geq \gamma \|x\|^2 - \|\sigma(x^k) - \sigma'(x^k)x^k\| \|x\|.
\end{aligned}$$

Ahora bien, para algún $\alpha = \gamma - \epsilon$, donde $0 < \epsilon < \gamma$, la siguiente desigualdad se cumple

$$\gamma \|x\|^2 - \|\sigma(x^k) - \sigma'(x^k)x^k\| \|x\| \geq \alpha \|x\|^2,$$

si

$$\epsilon \|x\|^2 - \|\sigma(x^k) - \sigma'(x^k)x^k\| \|x\| \geq 0,$$

que está cumplida cuando $\|x\| \geq \frac{\|\sigma(x^k) - \sigma'(x^k)x^k\|}{\epsilon}$, además, $\rho_1^k = \frac{\|\sigma(x^k) - \sigma'(x^k)x^k\|}{\epsilon}$.

Por otro lado, consideramos la siguiente cadena de relaciones

$$\begin{aligned} \langle z(x), w(x) \rangle &= x^T (B + \sigma'(x^k))^T (A + \varphi'(x^k))x + \\ &+ x^T (B + \sigma'(x^k))^T (\varphi(x^k) - \varphi'(x^k)x^k) + (\sigma(x^k) - \sigma'(x^k)x^k)^T (A + \varphi'(x^k))x + \\ &+ (\sigma(x^k) - \sigma'(x^k)x^k)^T (\varphi(x^k) - \varphi'(x^k)x^k) \geq \bar{\mu} \|x\|^2 - \\ &- \|B + \sigma'(x^k)\| \|\varphi(x^k) - \varphi'(x^k)x^k\| \|x\| - \|A + \varphi'(x^k)\| \|\sigma(x^k) - \sigma'(x^k)x^k\| \times \\ &\times \|x\| - \|\sigma(x^k) - \sigma'(x^k)x^k\| \|\varphi(x^k) - \varphi'(x^k)x^k\| \geq 0. \end{aligned}$$

Sea $\hat{a} = \bar{\mu}$, $\hat{b} = -\|B + \sigma'(x^k)\| \|\varphi(x^k) - \varphi'(x^k)x^k\| - \|A + \varphi'(x^k)\| \times$
 $\times \|\sigma(x^k) - \sigma'(x^k)x^k\|$ y $\hat{c} = -\|\sigma(x^k) - \sigma'(x^k)x^k\| \|\varphi(x^k) - \varphi'(x^k)x^k\|$. Entonces tenemos

$$\hat{a} \|x\|^2 + \hat{b} \|x\| + \hat{c} \geq 0,$$

que es válida cuando

$$\|x\| \geq \frac{-\hat{b} + \sqrt{\hat{b}^2 - 4\hat{a}\hat{c}}}{2\hat{a}},$$

aquí, $\hat{b}^2 - 4\hat{a}\hat{c} > 0$ ya que $\hat{a} > 0$ y $\hat{c} < 0$. Sea $\rho_2^k = \frac{-\hat{b} + \sqrt{\hat{b}^2 - 4\hat{a}\hat{c}}}{2\hat{a}}$ y tomamos $\rho^k = \max(\rho_1^k, \rho_2^k)$. Por lo tanto, el par (w, z) es (ρ^k, z) -copositivo.

Ahora, por la Proposición 2.3.7 el par (w, z) satisface la condición (θ_z) , por otro lado por el Teorema 2.3.8 el par (w, z) está sin una excepcional familia de elementos con respecto a $K \subset R^n$, significando por el Corolario 2.3.9 que el $\text{ICP}(w, z, K)$ tiene una solución.

Para la solución x^* del problema (2.1) las siguientes condiciones se cumplen:

$$z^* \equiv Bx^* + \sigma(x^*) \geq 0, \quad w^* \equiv Ax^* + \varphi(x^*) \geq 0, \quad \text{y } (z^*)^T w^* = 0.$$

Por simplicidad, denotamos $w^{k+1} = w(x^{k+1})$ y $z^{k+1} = z(x^{k+1})$. Entonces calculamos $(z^{k+1} - z^*)^T(w^{k+1} - w^*) = -(z^{k+1})^T w^* - (z^*)^T w^{k+1} \leq 0$. Por otro lado,

$$\begin{aligned}
& (z^{k+1} - z^*)^T(w^{k+1} - w^*) = \\
& = [B(x^{k+1} - x^*) + \sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)]^T \times \\
& \times [A(x^{k+1} - x^*) + \varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)] = \\
& = (x^{k+1} - x^*)^T B^T A(x^{k+1} - x^*) + (x^{k+1} - x^*)^T B^T \times \\
& \quad \times [\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)] + \\
& \quad + [\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)]^T A(x^{k+1} - x^*) + \\
& + [\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)]^T [\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)] \geq \\
& \geq \mu \|x^{k+1} - x^*\|^2 + (x^{k+1} - x^*)^T B^T [\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)] + \\
& \quad + [\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)]^T A(x^{k+1} - x^*) + \\
& + [\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)]^T [\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)].
\end{aligned}$$

Si juntamos las últimas dos cadenas, obtenemos:

$$\begin{aligned}
\mu \|x^{k+1} - x^*\|^2 & \leq (x^{k+1} - x^*)^T B^T [\varphi(x^*) - \varphi(x^k) - \varphi'(x^k)(x^{k+1} - x^k)] + \\
& \quad + [\sigma(x^*) - \sigma(x^k) - \sigma'(x^k)(x^{k+1} - x^k)]^T A(x^{k+1} - x^*) + \\
& - [\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)]^T [\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)].
\end{aligned}$$

Usando la desigualdad de Cauchy-Buniakovski obtenemos:

$$\mu \|x^{k+1} - x^*\|^2 \leq \|x^{k+1} - x^*\| \|B\| \|\varphi(x^*) - \varphi(x^k) - \varphi'(x^k)(x^{k+1} - x^k)\| +$$

$$\begin{aligned}
& + \|\sigma(x^*) - \sigma(x^k) - \sigma'(x^k)(x^{k+1} - x^k)\| \|A\| \|x^{k+1} - x^*\| + \\
& + \|\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(x^{k+1} - x^k)\| \times \\
& \times \|\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(x^{k+1} - x^k)\|. \tag{2.36}
\end{aligned}$$

Ahora, usando la desigualdad del triángulo y las condiciones (2.32):

$$\begin{aligned}
\|\varphi(x^*) - \varphi(x^k)\| &\leq q\mu \|x^k - x^*\|, \quad \|\sigma(x^*) - \sigma(x^k)\| \leq p\mu \|x^k - x^*\|, \\
\|\varphi'(x^k)(x^{k+1} - x^k)\| &\leq q\mu \|x^{k+1} - x^k\|, \\
\|\sigma'(x^k)(x^{k+1} - x^k)\| &\leq p\mu \|x^{k+1} - x^k\|,
\end{aligned}$$

y considerando que:

$$\|x^{k+1} - x^k\| \leq \|x^{k+1} - x^*\| + \|x^k - x^*\|,$$

y además usando (2.36) llegamos a la siguiente desigualdad:

$$\begin{aligned}
\|x^{k+1} - x^*\|^2 &\leq 4pq\mu \|x^k - x^*\|^2 + \\
& + (2q\|B\| + 2p\|A\| + 4pq\mu) \|x^{k+1} - x^*\| \|x^k - x^*\| + \\
& + (q\|B\| + p\|A\| + pq\mu) \|x^{k+1} - x^*\|^2.
\end{aligned}$$

Sean $a = 4pq\mu$, $b = 2q\|B\| + 2p\|A\| + 4pq\mu$ y $c = q\|B\| + p\|A\| + pq\mu$, entonces, haciendo manipulaciones algebraicas evidentes obtenemos:

$$\|x^{k+1} - x^*\| \leq \left(\frac{2a}{-b + \sqrt{b^2 - 4a(c-1)}} \right) \|x^k - x^*\|, \quad k = 0, 1, \dots, \tag{2.37}$$

donde $b^2 - 4a(c - 1) = 4q^2 \|B\|^2 + 4p^2 \|A\|^2 + 8pq \|A\| \|B\| + 16pq\mu > 0$. Entonces (2.37) implica la convergencia lineal de la secuencia $\{x^k\}$ a x^* , ya que

$$\left(\frac{2a}{-b + \sqrt{b^2 - 4a(c - 1)}} \right) < 1,$$

cuando (2.33) se cumple.

Entonces, si $\varphi, \sigma \in C^2(R^n)$, para cualquier conjunto convexo y acotado $V \in R^n$, existen escalares $r_V > 0$ y $s_V > 0$ tales que:

$$\|\varphi(x) - \varphi(y) - \varphi'(y)(x - y)\| \leq r_v \|x - y\|^2,$$

$$\|\sigma(x) - \sigma(y) - \sigma'(y)(x - y)\| \leq s_v \|x - y\|^2, \quad \forall x, y \in V.$$

Sin pérdida de generalidad, consideramos el conjunto V como una bola centrada en x^* con radio $\rho = \|x^0 - x^*\|$. Entonces de (2.37), sabemos que $x^k \in V$, $k = 0, 1, \dots$

Entonces

$$\|\varphi(x^*) - \varphi(x^k) - \varphi'(x^k)(x^* - x^k)\| \leq r_V \|x^k - x^*\|^2$$

y

$$\|\sigma(x^*) - \sigma(x^k) - \sigma'(x^k)(x^* - x^k)\| \leq s_V \|x^k - x^*\|^2,$$

$k = 0, 1, \dots$, y de (2.36) obtenemos

$$\begin{aligned} 0 &\leq (s_V r_V) \|x^k - x^*\|^4 + \\ &+ (r_V \|B\| + s_V \|A\| + s_V q\mu + r_V p\mu) \|x^{k+1} - x^*\| \|x^k - x^*\|^2 + \\ &+ (-\mu + q\mu \|B\| + p\mu \|A\| + pq\mu^2) \|x^{k+1} - x^*\|^2. \end{aligned}$$

Por simplicidad sea $\bar{a} = s_V r_V$, $\bar{b} = r_V \|B\| + s_V \|A\| + s_V q\mu + r_V p\mu$ y $\bar{c} = -\mu + q\mu \|B\| + p\mu \|A\| + pq\mu^2$, con esto llegamos a la siguiente estimación:

$$\|x^{k+1} - x^*\| \leq \left(\frac{2\bar{a}}{-\bar{b} + \sqrt{\bar{b}^2 - 4\bar{a}\bar{c}}} \right) \|x^k - x^*\|^2,$$

sea $C = \left(\frac{2\bar{a}}{-\bar{b} + \sqrt{\bar{b}^2 - 4\bar{a}\bar{c}}} \right)$, donde $C < 1$ cuando la siguiente desigualdad se cumple

$$q \|B\| + p \|A\| + pq\mu < 1,$$

que siempre se cumple cuando (2.33) es válida. También garantiza que $\bar{c} < 0$ y ya que $\bar{a} > 0$ entonces $\bar{b}^2 - 4\bar{a}\bar{c} > 0$.

□

Comentario 2.3.11 *También podemos encontrar una constante $N_1 > 0$ tal que la siguiente desigualdad se cumple:*

$$[g(x) - g(y)]^T [f(x) - f(y)] \geq N_1 \|f(x) - f(y)\|^2, \quad \forall x, y \in R^n.$$

2.4. El método de Newton aproximado

Describamos una versión inexacta del método de Newton como sigue. Sea $x^k \in R^n$ la iteración actual con $k \geq 0$. Construimos x^{k+1} mediante la siguiente regla

$$\|h_k(x^{k+1})\| \leq \gamma_k = \xi_k \|h(x^k)\|, \quad \xi_k > 0, \quad (2.38)$$

en donde $h_k(x) = \min(g(x^k) + g'(x^k)(x - x^k), f(x^k) + f'(x^k)(x - x^k))$, y $\{\xi_k\}$ es una secuencia de números positivos seleccionados anteriormente.

$$\begin{aligned} \text{Sean } \bar{\mu} &= \mu(1 - q\|B\| - p\|A\| - pq\mu), L_{fg} = \frac{(\|A\| + q\mu)^2}{\bar{\mu}}, N^k = \frac{\mu(1-q)}{\|B + \sigma'(x^k)\|^2}, \\ L_{wz}^k &= \frac{\|A + \varphi'(x^k)\|^2}{\mu(1-p)}, N_z^k = \bar{\eta} - \|\sigma'(x^k)\|, \tilde{a} = \mu - \|B\| \|\varphi'(x^k)\| - \|A\| \|\sigma'(x^k)\| - \\ &\|\sigma'(x^k)\| \|\varphi'(x^k)\|, \tilde{b} = -q\mu\|B\| - p\mu\|A\| - \|\varphi'(x^k)\| \|B\| - \|\sigma'(x^k)\| \|A\| - \\ &p\mu \|\varphi'(x^k)\| - q\mu \|\sigma'(x^k)\| - 2\|\sigma'(x^k)\| \|\varphi'(x^k)\|, \tilde{c} = -(p\mu + \|\sigma'(x^k)\|)(q\mu + \\ &\|\varphi'(x^k)\|), \bar{\bar{b}} = -2q\mu\|B\| - 2p\mu\|A\| - 4pq\mu^2 \text{ y } \bar{\bar{c}} = -4pq\mu^2. \end{aligned}$$

Teorema 2.4.1 *Sea $x^* \in R_+^n$ la solución del problema (2.1) con las aplicaciones $g(x) = Bx + \sigma(x)$ y $f(x) = Ax + \varphi(x)$, satisfaciendo las condiciones del Teorema 2.3.10. Entonces para cualquier punto inicial $x^0 \in R^n$, la secuencia $\{x^k\}$ generada por la regla (2.38):*

(a) *Converge a x^* , si*

$$\frac{L_{fg}(L_{wz}^k + 1)}{N^k N_z^k} \xi_k \leq 1 - \epsilon + \left(\tilde{b} - \frac{\sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}} \right), \quad k = 0, 1, \dots, \quad (2.39)$$

para algún $0 < \epsilon < \min \left\{ 1 + \left(\frac{\bar{\bar{b}} - \sqrt{\bar{\bar{b}}^2 - 4\bar{\mu}\bar{\bar{c}}}}{2\bar{\mu}} \right), 1 + \left(\tilde{b} - \frac{\sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}} \right) \right\}$.

(b) *Por otro lado, si $\varphi, \sigma \in C^2(R^n)$ y $\xi_k \rightarrow 0$ cuando $k \rightarrow \infty$, entonces la convergencia es superlineal, i.e.*

$$\lim_{k \rightarrow \infty} \left(\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \right) = 0. \quad (2.40)$$

(c) *Si iniciando en algún $k_0 \in N$, la siguiente condición se cumple $\xi_k \leq \vartheta \|h(x^k)\|$, $\vartheta > 0$, $k \geq k_0$, entonces la convergencia es cuadrática, i.e., la desigualdad (2.23) se cumple para algún $C > 0$.*

Demostración. Consideramos la secuencia $\{x^k\}$ generada por la regla (2.38), de un punto inicial $x^0 \in R^n$. Sean x^{k+1} y ξ_k satisfaciendo las condiciones (2.38) y (2.39). Denotamos y^{k+1} como la solución exacta del siguiente problema lineal de complementariedad:

$$z_k(y) \equiv By + \sigma(x^k) + \sigma'(x^k)(y - x^k) \geq 0,$$

$$w_k(y) \equiv Ay + \varphi(x^k) + \varphi'(x^k)(y - x^k) \geq 0, \text{ y } [z_k(y)]^T w_k(y) = 0. \quad (2.41)$$

Podemos ver que por la condición (2.32) que la aplicación f es g -fuertemente monótona con constante $N = \frac{\bar{\mu}}{(\|B\| + p\mu)^2}$ y también es g -Lipschitz continua con constante $L_{fg} = \frac{(\|A\| + q\mu)^2}{\bar{\mu}}$. Además, la aplicación g es Lipschitz-continua con constante $L_g = \|B\| + p\mu$ y es fuertemente monótona con constante $N_g = \bar{\eta} - p\mu$.

Análogamente, la aplicación $w_k(y)$ es z_k -fuertemente monótona con el parámetro $N^k = \frac{\mu(1-q)}{\|B + \sigma'(x^k)\|^2}$ y es z_k -Lipschitz continua con constante $L_{wz}^k = \frac{\|A + \varphi'(x^k)\|^2}{\mu(1-p)}$. Además, z es Lipschitz-continua con parámetro $L_z^k = \|B + \varphi'(x^k)\|$ y fuertemente monótona con constante $N_z^k = \bar{\eta} - \|\sigma'(x^k)\|$.

Por eso, podemos aceptar el Corolario 2.2.6 del Lema 2.2.4 al problema (2.41) y obtener la siguiente desigualdad

$$\|x^{k+1} - y^{k+1}\| \leq \frac{L_{wz}^k + 1}{N^k N_z^k} \|h_k(x^{k+1})\|. \quad (2.42)$$

De (2.38), (2.42) y el Corolario 2.2.6 del Lema 2.2.4 obtenemos la siguiente cadena de desigualdades

$$\|y^{k+1} - x^{k+1}\| \leq \frac{L_{wz}^k + 1}{N^k N_z^k} \xi_k \|h(x^k)\| \leq$$

$$\leq \frac{L_{fg} (L_{wz}^k + 1)}{N^k N_z^k} \xi_k \|x^k - x^*\|. \quad (2.43)$$

Por otro lado, usando (2.36) tenemos la siguiente desigualdad

$$\begin{aligned} \mu \|y^{k+1} - x^*\|^2 &\leq \|y^{k+1} - x^*\| \|B\| \|\varphi(x^*) - \varphi(x^k) - \varphi'(x^k)(y^{k+1} - x^k)\| + \\ &\quad + \|\sigma(x^*) - \sigma(x^k) - \sigma'(x^k)(y^{k+1} - x^k)\| \|A\| \|y^{k+1} - x^*\| + \\ &\quad + \|\sigma(x^k) - \sigma(x^*) + \sigma'(x^k)(y^{k+1} - x^k)\| \|\varphi(x^k) - \varphi(x^*) + \varphi'(x^k)(y^{k+1} - x^k)\|. \end{aligned}$$

De nuevo, por la condición (2.32), tenemos que

$$\begin{aligned} \|\varphi(x^*) - \varphi(x^k)\| &\leq q\mu, \quad \|\sigma(x^*) - \sigma(x^k)\| \leq p\mu, \\ \|\varphi'(x^k)(y^{k+1} - x^k)\| &\leq \|\varphi'(x^k)\| \|x^k - y^{k+1}\|, \\ \|\sigma'(x^k)(y^{k+1} - x^k)\| &\leq \|\sigma'(x^k)\| \|x^k - y^{k+1}\|, \end{aligned}$$

con esto llegamos a la siguiente condición:

$$\begin{aligned} \mu \|y^{k+1} - x^*\|^2 &\leq \|y^{k+1} - x^*\| \times \\ &\quad \times \|B\| [\|\varphi(x^k)\| (\|x^k - x^*\| + \|y^{k+1} - x^*\|) + q\mu \|x^k - x^*\|] + \\ &\quad + [\|\sigma(x^k)\| (\|x^k - x^*\| + \|y^{k+1} - x^*\|) + p\mu \|x^k - x^*\|] \|x^k - x^*\| \|A\| \times \\ &\quad \times \|y^{k+1} - x^*\| + [\|\sigma(x^k)\| (\|x^k - x^*\| + \|y^{k+1} - x^*\|) + p\mu \|x^k - x^*\|] \times \\ &\quad \times [\|\varphi(x^k)\| (\|x^k - x^*\| + \|y^{k+1} - x^*\|) + q\mu \|x^k - x^*\|]. \end{aligned}$$

Esta última desigualdad puede ser reescrita como sigue

$$\tilde{a} \|y^{k+1} - x^*\|^2 + \tilde{b} \|y^{k+1} - x^*\| \|x^k - x^*\| + \tilde{c} \|x^k - x^*\|^2 \leq 0,$$

$$\begin{aligned}
& \text{en donde } \tilde{a} = \mu - \|B\| \|\varphi'(x^k)\| - \|A\| \|\sigma'(x^k)\| - \|\sigma'(x^k)\| \|\varphi'(x^k)\|, \\
\tilde{b} &= -q\mu \|B\| - p\mu \|A\| - \|\varphi'(x^k)\| \|B\| - \|\sigma'(x^k)\| \|A\| - p\mu \|\varphi'(x^k)\| - \\
& - q\mu \|\sigma'(x^k)\| - 2 \|\sigma'(x^k)\| \|\varphi'(x^k)\|, \\
\tilde{c} &= -(p\mu + \|\sigma'(x^k)\|)(q\mu + \|\varphi'(x^k)\|).
\end{aligned}$$

Esta desigualdad implica que

$$\|y^{k+1} - x^*\| \leq \left(\frac{-\tilde{b} + \sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}} \right) \|x^k - x^*\|, \quad (2.44)$$

aquí $\tilde{b}^2 - 4\tilde{a}\tilde{c} > 0$ ya que $\tilde{a} > 0$ y $\tilde{c} < 0$. Ahora, usando la desigualdad del triángulo en (2.43) y (2.44) y la condición (2.39) acerca de ξ_k , llegamos a la siguiente estimación:

$$\begin{aligned}
\|x^{k+1} - x^*\| &\leq \|x^{k+1} - y^{k+1}\| + \|y^{k+1} - x^*\| \leq \\
&\leq \left[1 - \epsilon + \left(\tilde{b} - \frac{\sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}} \right) + \left(\frac{-\tilde{b} + \sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}}{2\tilde{a}} \right) \right] \|x^k - x^*\| = \\
&= (1 - \epsilon) \|x^k - x^*\|, \quad (2.45)
\end{aligned}$$

de aquí sigue la convergencia de $\{x^k\}$ a x^* .

Nosotros sabemos que si $\varphi, \sigma \in C^2(R^n)$, para cualquier conjunto convexo y acotado $V \in R^n$, existen escalares $r_V > 0$ y $s_V > 0$ tales que:

$$\|\varphi(x) - \varphi(y) - \varphi'(y)(x - y)\| \leq r_v \|x - y\|^2,$$

$$\|\sigma(x) - \sigma(y) - \sigma'(y)(x - y)\| \leq s_v \|x - y\|^2, \quad \forall x, y \in V.$$

Sin pérdida de generalidad, consideramos el conjunto V como una bola centrada en x^* con radio $\rho = \|x^0 - x^*\|$. Entonces de (2.45), nosotros sabemos

que $x^k \in V$, $k = 0, 1, \dots$. Entonces la desigualdad (2.34) se cumple, i.e.

$$\|x^{k+1} - x^*\| \leq \left(\frac{2\bar{a}}{-\bar{b} + \sqrt{\bar{b}^2 - 4\bar{a}\bar{c}}} \right) \|x^k - x^*\|^2,$$

en donde $\bar{a} = s_V r_V$, $\bar{b} = r_V \|B\| + s_V \|A\| + s_V q\mu + r_V p\mu$ y $\bar{c} = -\mu + q\mu \|B\| + p\mu \|A\| + pq\mu^2$. Pero, si $\xi_k \rightarrow \infty$ cuando $k \rightarrow \infty$, entonces, ya que $\{\|\varphi'(x^k)\|\}_{k=0}^\infty$ y $\{\|\sigma'(x^k)\|\}_{k=0}^\infty$ son acotados, la magnitud

$$\zeta_k = \frac{L_{fg} (L_{wz}^k + 1)}{N^k N_z^k} \xi_k$$

tiende a cero cuando $k \rightarrow \infty$. Entonces

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \leq \left(\frac{2\bar{a}}{-\bar{b} + \sqrt{\bar{b}^2 - 4\bar{a}\bar{c}}} \right) \|x^k - x^*\| + \zeta_k \rightarrow 0$$

si $k \rightarrow \infty$, aquí $\bar{a} = s_V r_V$, $\bar{b} = r_V \|B\| + s_V \|A\| + s_V q\mu + r_V p\mu$ y $\bar{c} = -\mu + q\mu \|B\| + p\mu \|A\| + pq\mu^2$.

Ahora bien, si $\xi_k \leq \vartheta \|h(x^k)\|$, $\vartheta > 0$ para $k \geq k_0$, entonces, usando las desigualdades para $\|h(x^k)\|$ del Corolario 2.2.6 del Lema 2.2.4, obtenemos la siguiente desigualdad

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2, \quad k \geq k_0,$$

en donde podemos considerar C como sigue

$$C = \left(\frac{2\bar{a}}{-\bar{b} + \sqrt{\bar{b}^2 - 4\bar{a}\bar{c}}} \right) + \frac{L_{fg}^2 (\overline{L_{wz}} + 1)}{\overline{N^k} \overline{N_z^k}}$$

en donde, $\overline{L_{wz}^k} = \frac{(\|A\| + q\mu)^2}{\bar{\mu}}$, $\overline{N^k} = \frac{\bar{\mu}}{(\|B\| + p\mu)^2}$ y $\overline{N_z^k} = \bar{\eta} - p\mu$.

□

En seguida, consideramos otra variación del método de Newton inexacto. Sean $\{\varepsilon_k\}$ y $\{\xi_k\}$ secuencias de números positivos. Entonces, empezando en cualquier punto $x^0 \in R^n$, generamos la secuencia de puntos $\{x^k\}$ del siguiente modo: teniendo $x^k \in R^n$ encontramos x^{k+1} satisfaciendo la siguiente condición:

$$\|h_k(x^{k+1})\| \leq \gamma_k = \min(\varepsilon_k, \xi_k \|h(x^k)\|). \quad (2.46)$$

Nosotros sabemos que si la matriz AB^{-1} es positiva definida, entonces, el siguiente problema lineal de complementariedad:

$$z \equiv p + Bx \geq 0, \quad w \equiv q + Ax, \quad y \quad z^T w = 0 \quad (2.47)$$

tiene una única solución.

Para investigar esta variación del método de Newton inexacto, demostramos la dependencia continua de la solución a la matriz AB^{-1} , p y q .

Lema 2.4.2 Sean $AB^{-1}, [A_k B_k]^{-1} \in L(R^n), k = 1, 2, \dots$, las matrices positivas definidas con parámetros $\eta > 0$ y $\eta_k > 0$ respectivamente, y sean $p, p_k, q, q_k, k = 1, 2, \dots$ vectores en R^n tales que $[A_k B_k]^{-1} \rightarrow AB^{-1}, p_k \rightarrow p$ y $q_k \rightarrow q$ cuando $k \rightarrow \infty$.

Entonces, si x^* es la solución del problema (2.47) y x^k son las soluciones de los siguientes problemas lineales de complementariedad

$$z_k \equiv p_k + B_k x \geq 0, \quad w_k \equiv q_k + A_k x \geq 0 \quad y \quad z_k^T w_k = 0, \quad (2.48)$$

entonces $x^k \rightarrow x^*$ cuando $k \rightarrow \infty$.

Demostración. No es difícil ver que x^* es la solución para el siguiente problema lineal de complementariedad

$$z \equiv B_k x - (B_k - B)x^* - (p_k - p) + p_k \geq 0,$$

$$w \equiv A_k x - (A_k - A)x^* - (q_k - q) + q_k \geq 0, \text{ y } z^T w = 0,$$

siendo el problema (2.48) con perturbaciones $q_z = (B_k - B)x^* + (q_k - q)$ y $p_w = (A_k - A)x^* + (p_k - p)$ respectivamente.

Entonces, podemos aceptar el Lema 2.2.7 y así, obtener la siguiente estimación:

$$\|x^k - x^*\| \leq \frac{1}{\eta_k \rho} \|p_w - AB^{-1}q_z\|,$$

en donde $\rho = \sqrt{\lambda_{\min}(B^T B)}$, de esta desigualdad sigue que

$$\|x^k - x^*\| \leq \frac{1}{\eta_k \rho} \times$$

$$\times (\|B_k - B\| \|x^*\| + \|p_k - p\| + \|AB^{-1}\| \|A_k - A\| \|x^*\| + \|AB^{-1}\| \|q_k - q\|).$$

Ya que podemos seleccionar η y η_k tal que $\eta_k \rightarrow \eta$ cuando $k \rightarrow \infty$, entonces existe k_0 tal que para $k \geq k_0$ toma lugar la siguiente estimación:

$$\|x^k - x^*\| \leq \frac{2}{\eta \rho} \times$$

$$\times (\|B_k - B\| \|x^*\| + \|p_k - p\| + \|AB^{-1}\| \|A_k - A\| \|x^*\| + \|AB^{-1}\| \|q_k - q\|),$$

y de aquí sigue la demostración del Lema.

□

Teorema 2.4.3 Sea $f(x) = Ax + \phi(x)$ una aplicación g -fuertemente monótona con parámetro $\beta > 0$ y $g(x) = Bx + \sigma(x)$ una aplicación fuertemente monótona con el parámetro $\beta_g > 0$. Además, sea cada componente de la aplicación f una función convexa con respecto a cada variable, y continuamente diferenciable. Sean ∇g y ∇f aplicaciones que satisfacen la condición de Lipschitz con constantes $Q_1 > 0$ y $Q_2 > 0$ respectivamente, y en cualquier punto $x \in R^n$ el producto $\nabla g(x) [\nabla f(x)]^{-1}$ es una M -matriz.

Sea $\{x^k\}$ una secuencia generada por la regla (2.46) iniciando en cualquier punto $x_0 \in R^n$, satisfaciendo las siguientes condiciones:

(a) existe un número positivo S tal que:

$$\sum_{k=1}^{\infty} \xi_k^2 \left[\frac{\|\nabla f(x^k)\| + 1}{\beta\beta_g} \right]^2 \leq S$$

(b) para todo $k = 0, 1, \dots$, la siguiente condición se cumple:

$$\xi_k (\eta\rho + \|1 - \nabla f(x^k)\| + \|AB^{-1}\| \|1 - \nabla g(x^k)\|) \leq \xi,$$

en donde $\eta > 0$ es el parámetro de la definidad positiva de AB^{-1} , $0 < \xi < \frac{\eta\rho}{L_g(1+L)}$; aquí $L = \sup_{x \in W} \|f'(x)\|$ y el conjunto $W \subset R^n$ es una bola centrada en el origen y radio $\bar{\rho} = \|y^1\| + \frac{Q_2 S}{\eta} + \frac{\sqrt{S}}{\sqrt{\eta}}$ (el vector y^1 esta definido como la solución única y exacta del problema (2.2) para $k = 0$). En este caso, las soluciones $\{x^k\}$ convergen a $x^* \in R_+^n$, la solución única del problema (2.1). Además, si $\xi_k \rightarrow 0$ cuando $k \rightarrow \infty$, entonces la tasa de convergencia es más que lineal, esto es, la igualdad (2.40) se cumple. Si, para valores de k suficientemente grandes la siguiente condición se cumple:

$$\xi_k \leq \vartheta \|h(x^k)\|, \vartheta > 0,$$

entonces la convergencia es cuadrática, i.e., la desigualdad (2.23) se cumple con alguna constante $C > 0$.

Demostración. Investiguemos el paso k del proceso. Sea $x^k \in R^n$, $k \geq 0$. Consideremos el siguiente Problema Lineal de Complementariedad:

$$\begin{aligned} z &\equiv g(x^k) + \nabla g(x^k)(v - x^k) \geq 0, \\ w &\equiv f(x^k) + \nabla f(x^k)(v - x^k) \geq 0, \text{ y } [z]^T [w] = 0. \end{aligned} \quad (2.49)$$

Este problema es equivalente a (2.22), el siguiente problema de complementariedad estándar:

$$y \geq 0, \quad \Phi(y) \geq 0, \text{ y } [y]^T [\Phi(y)] = 0, \quad (2.50)$$

aquí,

$$\begin{aligned} y &= \nabla g(x^k)v + g(x^k) - \nabla g(x^k)x^k \\ \Phi(y) &= \nabla f(x^k) [\nabla g(x^k)]^{-1} y - \nabla f(x^k) [\nabla g(x^k)]^{-1} g(x^k) + f(x^k). \end{aligned}$$

La solución aproximada de este problema como lo describe (2.46) es el punto x^{k+1} y a la solución exacta la denotamos como $y^{k+1} \geq 0$, regresando a las variables originales la solución del problema (2.49) es $v^{k+1} = [\nabla g(x^k)]^{-1} [y^{k+1} - g(x^k)] + x^k$ (la solución existe y es única ya que el producto $\nabla f(x^k) [\nabla g(x^k)]^{-1}$ es positiva definida). Ahora, ya que f es convexa, entonces

$$f(y^{k+1}) \geq f(x^k) + \nabla f(x^k)(y^{k+1} - x^k) \geq 0.$$

Aceptando el Corolario 2.2.6 del Lema 2.2.4 al problema (2.49) para cada $k = 0, 1, \dots$ obtenemos la siguiente desigualdad:

$$\frac{\beta\beta_g}{\|\nabla f(x^k)\| + 1} \|x^{k+1} - y^{k+1}\| \leq \|h_k(x^{k+1})\|,$$

y ya que la regla (2.46) se cumple, i.e.,

$$\|h_k(x^{k+1})\| \leq \xi_k \quad k = 0, 1, \dots$$

Finalmente obtenemos la estimación:

$$\|x^{k+1} - y^{k+1}\| \leq \xi_k \frac{\|\nabla f(x^k)\| + 1}{\beta\beta_g} \text{ para todo } k = 0, 1, \dots \quad (2.51)$$

Ahora consideramos el conjunto $\alpha_k = \alpha \subseteq \{1, 2, \dots, n\}$ tal que, $y_j^{k+1} = 0$ para $j \notin \alpha$ y $y_i^{k+1} > 0, i \in \alpha$. Y tomamos (2.28) de la demostración del Teorema 2.3.2,

$$\begin{aligned} y_\alpha^{k+1} &= g_\alpha(x^k) - \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} f_\alpha(x^k) + \\ &+ \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \nabla f_{\alpha\bar{\alpha}}(x^k) [\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1} g_{\bar{\alpha}}(x^k). \end{aligned}$$

Ahora bien, reescribimos la última igualdad como sigue

$$\begin{aligned} y_\alpha^{k+1} &= g_\alpha(y^k) - \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} f_\alpha(y^k) + \\ &+ \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \nabla f_{\alpha\bar{\alpha}}(x^k) [\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1} g_{\bar{\alpha}}(y^k) + R_\alpha(x^k, y^k), \end{aligned} \quad (2.52)$$

en donde

$$\begin{aligned} R_\alpha(x^k, y^k) &= g_\alpha(x^k) - g_\alpha(y^k) + \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} [f_\alpha(y^k) - f_\alpha(x^k)] - \\ &- \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \nabla f_{\alpha\bar{\alpha}}(x^k) [\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1} (g_{\bar{\alpha}}(y^k) - g_{\bar{\alpha}}(x^k)), \end{aligned}$$

de esta forma

$$R_\alpha(x^k, y^k) = \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \times \\ \times \{f_\alpha(y^k) - f_\alpha(x^k) - [\nabla f(x^k)(y^k - x^k)]_\alpha\} \geq 0_\alpha.$$

Ya que, cada componente de f es convexa y la matriz $\nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1}$ es no negativa. También, de (2.52) y el hecho que $\nabla f_{\alpha\bar{\alpha}}(x^k) [\nabla g_{\alpha\bar{\alpha}}(x^k)]^{-1}$ es no positiva; y $f_\alpha(y^k)$ y $g_\alpha(y^k)$ son no negativas ($\forall k \geq 1$), sigue la desigualdad $y_\alpha^{k+1} \leq g_\alpha(y^k) + R_\alpha(x^k, y^k)$. Por otro lado, es evidente que:

$$y_{\bar{\alpha}}^{k+1} = 0_{\bar{\alpha}} \leq y_{\bar{\alpha}}^k \quad \forall k \geq 1.$$

Por esa razón, la siguientes estimaciones se cumplen:

$$0 \leq y^{k+1} \leq y^k + R(x^k, y^k) \quad k = 1, 2, \dots,$$

en donde

$$R(x^k, y^k) = \begin{bmatrix} R_\alpha(x^k, y^k) \\ 0_{\bar{\alpha}} \end{bmatrix}.$$

De aquí obtenemos las siguientes desigualdades:

$$0 \leq y^{k+1} \leq y^1 + \sum_{l=1}^k R(x^l, y^l). \quad (2.53)$$

Ya que, ∇f satisface la condición de Lipschitz con la constante $Q_2 > 0$, tenemos que

$$\|R_\alpha(x^k, y^k)\| \leq \left\| \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \right\| \times \\ \times \|f_\alpha(y^k) - f_\alpha(x^k) - [\nabla f(x^k)(y^k - x^k)]_\alpha\| \leq$$

$$\leq Q_2 \left\| \nabla g_{\alpha\alpha}(x^k) [\nabla f_{\alpha\alpha}(x^k)]^{-1} \right\| \|x^k - y^k\|^2.$$

Es bien sabido que cualquier submatriz principal $A_{\alpha\alpha} [B_{\alpha\alpha}]^{-1}$ de la matriz positiva definida AB^{-1} también es positiva definida. Por otro lado, ya que la matriz AB^{-1} es positiva definida con el parámetro $\eta > 0$ (i.e., $u^T AB^{-1}u \geq \eta \|u\|^2$) la siguiente desigualdad se cumple $\|AB^{-1}\| \leq \frac{1}{\eta}$.

Entonces, de la matriz $\nabla f(x^k) [\nabla g(x^k)]^{-1}$ positiva definida y de (2.51) sigue que:

$$\|R(x^k, y^k)\| = \|R_{\alpha}(x^k, y^k)\| \leq \frac{Q_2}{\eta} \xi_{k-1} \left[\frac{\|\nabla f(x^{k-1})\| + 1}{\beta\beta_g} \right]^2.$$

Entonces, si la serie

$$\sum_{k=1}^{\infty} \xi_k^2 \left[\frac{\|\nabla f(x^k)\| + 1}{\beta\beta_g} \right]^2 \quad (2.54)$$

converge, de la fórmula (2.53) sigue la acotación de la secuencia $\{y^k\}$, y así, la acotación de la secuencia $\{x^k\}$. En verdad, de (2.51) y de la tendencia a cero de la serie convergente (2.54) sigue que,

$$\|x^k - y^k\| \rightarrow 0 \quad \text{cuando} \quad k \rightarrow \infty. \quad (2.55)$$

Ahora seleccionamos cualquier subsecuencia convergente $\{y^{k_l}\}_{l=1}^{\infty}$, y denotamos $\bar{x}^1 = \lim_{l \rightarrow \infty} y^{k_l}$. Simultáneamente de la fórmula (2.55), tenemos $\lim_{l \rightarrow \infty} x^{k_l} = \bar{x}^1 \geq 0$. Entonces, por la continuidad de f , g , ∇f y ∇g obtenemos

$$\lim_{l \rightarrow \infty} f(x^{k_l}) = f(\bar{x}^1), \quad \lim_{l \rightarrow \infty} \nabla f(x^{k_l}) = \nabla f(\bar{x}^1)$$

$$\lim_{l \rightarrow \infty} g(x^{k_l}) = g(\bar{x}^1), \quad \lim_{l \rightarrow \infty} \nabla g(x^{k_l}) = \nabla g(\bar{x}^1).$$

Pero, y^{k_l} es la solución exacta del Problema Lineal de Complementariedad:

$$y \geq 0, \quad \Phi(y) = \Phi(y) = \nabla f(x^{k_l}) [\nabla g(x^{k_l})]^{-1} y - \\ - \nabla f(x^{k_l}) [\nabla g(x^{k_l})]^{-1} g(x^{k_l}) + f(x^{k_l}) \geq 0, \text{ y } [y]^T [\Phi(y)] = 0.$$

Entonces, del Lema 2.4.2 $\{y^{k_l}\}$ converge cuando $l \rightarrow \infty$ a $\bar{x}^2 \geq 0$ que es la solución del último problema:

$$z \equiv g(\bar{x}^1) + \nabla g(\bar{x}^1)(x - \bar{x}^1) \geq 0, w \equiv f(\bar{x}^1) + \nabla f(\bar{x}^1)(x - \bar{x}^1) \geq 0, \text{ y } z^T w = 0.$$

Entonces, por la fórmula (2.36) $x^{k_{l+1}} \rightarrow \bar{x}^2$ cuando $l \rightarrow \infty$. Análogamente, siguiendo adelante obtenemos por medio de los elementos de las soluciones $\{y^k\}$ y $\{x^k\}$ un conjunto numerable de las subsecuencias $\{\{y^{k_l+m}\}_{l=1}^{\infty}\}_{m=0}^{\infty}$ y $\{\{x^{k_l+m}\}_{l=1}^{\infty}\}_{m=0}^{\infty}$. Para un $m \in \mathbb{N}$ fijo las subsecuencias $\{y^{k_l+m}\}$ y $\{x^{k_l+m}\}$ convergen cuando $l \rightarrow \infty$ to \bar{x}^{m+1} siendo la solución del Problema Lineal de Complementariedad:

$$z \equiv g(\bar{x}^m) + \nabla g(\bar{x}^m)(x - \bar{x}^m) \geq 0, w \equiv f(\bar{x}^m) + \nabla f(\bar{x}^m)(x - \bar{x}^m) \geq 0, \text{ y } z^T w = 0.$$

Pero del Teorema 2.3.2, sabemos que la secuencia $\{\bar{x}^m\}$ obtenida del método de Newton exacto, converge a x^* que es la solución única del problema (2.1) cuando $m \rightarrow \infty$, entonces para cualquier punto inicial $x^0 \in R^n$ y cualquier vecindad V del punto x^* existe un número $k_0 \in \mathbb{N}$ tal que $x^{k_0} \in V$.

Para terminar la demostración del teorema, usaremos los resultados de [41], construimos tal vecindad V en la que pertenece x^* , que garantiza la convergencia de la secuencia $\{x^k\}$ a x^* . Para esto, consideramos nuevamente el paso k , i.e., construimos el punto x^{k+1} satisfaciendo la regla (2.46).

Denotamos $u^{k+1} = h_k(x^{k+1})$, $v^{k+1} = x^{k+1} - u^{k+1}$, y vemos que v^{k+1} satisface a las siguientes condiciones:

$$\begin{aligned} z^{k+1} &\equiv g(x^k) - (I - \nabla g(x^k)) u^{k+1} + \nabla g(x^k)(v^{k+1} - x^k) \geq 0, \\ w^{k+1} &\equiv f(x^k) - (I - \nabla f(x^k)) u^{k+1} + \nabla f(x^k)(v^{k+1} - x^k) \geq 0, \\ &\text{y } [z^{k+1}]^T [w^{k+1}] = 0. \end{aligned}$$

Análogamente, si introducimos la notación:

$$t_1^k = g(x^*) - g(x^k) - \nabla g(x^k)(x^* - x^k), \quad t_2^k = f(x^*) - f(x^k) - \nabla f(x^k)(x^* - x^k),$$

entonces es evidente que x^* satisface a las siguientes condiciones:

$$\begin{aligned} z^* &\equiv g(x^k) + t_1^k + \nabla g(x^k)(x^* - x^k) \geq 0, \\ w^* &\equiv f(x^k) + t_2^k + \nabla f(x^k)(x^* - x^k) \geq 0, \text{ y } [z^*]^T [w^*] = 0. \end{aligned}$$

En otras palabras, usando la notación del Lema 2.2.7, la siguiente desigualdad está definida correctamente:

$$\begin{aligned} &\|v^{k+1} - x^*\| \leq \frac{1}{\eta\rho} \times \\ &\times \|(I - \nabla f(x^k)) u^{k+1} + t_2^k + AB^{-1}(-t_1^k - (I - \nabla g(x^k)) u^{k+1})\|. \end{aligned} \quad (2.56)$$

Por la desigualdad del triángulo, tenemos:

$$\|v^{k+1} - x^*\| = \|x^{k+1} - x^* - u^{k+1}\| \geq \|x^{k+1} - x^*\| - \|u^{k+1}\|.$$

De aquí y de (2.56) siguen las desigualdades:

$$\|x^{k+1} - x^*\| \leq \|u^{k+1}\| + \|v^{k+1} - x^*\| \leq$$

$$\begin{aligned} &\leq \frac{\eta\rho + \|1 - \nabla f(x^k)\| + \|AB^{-1}\| \|1 - \nabla g(x^k)\|}{\eta\rho} \|u^{k+1}\| + \\ &\quad + \frac{1}{\eta\rho} (\|t_2^k\| + \|AB^{-1}\| \|t_1^k\|). \end{aligned}$$

Ya que, los gradientes ∇f y ∇g satisfacen a la condición de Lipschitz-continua obtenemos que $\|t_1^k\| \leq Q_1 \|x^k - x^*\|^2$ y $\|t_2^k\| \leq Q_2 \|x^k - x^*\|^2$ y además de la regla (2.46) llegamos a la siguiente desigualdad:

$$\|u^{k+1}\| \leq \gamma_k \leq \xi_k \|h(x^k)\|, k = 0, 1, \dots$$

Inmediatamente de aquí sigue la estimación:

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \frac{1}{\eta\rho} (Q_2 + Q_1 \|AB^{-1}\|) \|x^k - x^*\|^2 + \\ &+ \left(\frac{\eta\rho + \|1 - \nabla f(x^k)\| + \|AB^{-1}\| \|1 - \nabla g(x^k)\|}{\eta\rho} \right) \gamma_k. \end{aligned} \quad (2.57)$$

Siguiendo adelante, de la fórmula (2.53) y de la convergencia de la serie (2.54), concluimos que la secuencia $\{y^k\}$ está dentro de una bola con centro en el punto 0 y de radio $\rho_0 = \|y^1\| + \frac{Q_2 S}{\eta}$. Entonces como sigue de (2.55), a partir de un número $k_1 \in \mathbb{N}$, la secuencia $\{x^k\}$ está dentro de $W \in R^n$, definido en (b) por las condiciones del teorema. Entonces, usando las constantes introducidas anteriormente $\widehat{L} > 0, L_g > 0$, podemos escribir la siguiente desigualdad:

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \frac{1}{\eta\rho} (Q_2 + Q_1 \|AB^{-1}\|) \|x^k - x^*\|^2 + \\ &+ \left(\frac{\eta\rho + 1 + \widehat{L} + \|AB^{-1}\| (1 + L_g)}{\eta\rho} \right) \gamma_k, \end{aligned}$$

esta puede ser reescrita como sigue

$$\|x^{k+1} - x^*\| \leq \Lambda \|x^k - x^*\|^2 + \Gamma \gamma_k,$$

en donde

$$\Lambda = \frac{1}{\eta\rho} (Q_2 + Q_1 \|AB^{-1}\|), \text{ y } \Gamma = \left(\frac{\eta\rho + 1 + \widehat{L} + \|AB^{-1}\| (1 + L_g)}{\eta\rho} \right).$$

Podemos ver que por la fórmula (2.46) y el Corolario 2.2.6 del Lema 2.2.4:

$$\gamma_k \leq \xi_k \|h(x^k)\| \leq \xi_k L_g (L + 1), \quad k = 0, 1, \dots$$

Es por eso que si $\xi_k (\eta\rho + \|1 - \nabla f(x^k)\| + \|AB^{-1}\| \|1 - \nabla g(x^k)\|) \leq \xi$, $\forall k = 0, 1, \dots$. Entonces de (2.57) sigue evidentemente la convergencia de la secuencia $\{x^k\}$ al punto x^* si para algún número $m_0 \in \mathbb{N}$ satisface la desigualdad $\|x^{m_0} - x^*\| < \delta$, en donde $\delta > 0$ satisface a la siguiente condición:

$$\frac{1}{\eta\rho} (Q_2 + Q_1 \|AB^{-1}\|) \delta + \xi L_g \left(\frac{L + 1}{\eta\rho} \right) = v < 1.$$

Primero, es claro que $\delta > 0$ existe y satisface a la condición de arriba, y también que

$$\xi < \frac{\eta\rho}{L_g (L + 1)};$$

y segundo, como demostramos anteriormente, para cualquier valor fijo $\delta > 0$, el número m_0 tal que $\|x^{m_0} - x^*\| < \delta$ es válida debe existir.

Si pedimos como condición extra, que $\xi_k \rightarrow 0$ cuando $k \rightarrow \infty$, entonces la tasa de convergencia es superlineal, (esto viene de (2.57)). Además, si para valores bastante grandes de k , sigue la condición: $\xi_k \leq \vartheta \|h(x^k)\|$, $\vartheta > 0$,

entonces obtenemos la convergencia cuadrática (en este caso tomamos como constante C la siguiente magnitud):

$$C = \frac{1}{\eta\rho} (Q_2 + Q_1 \|AB^{-1}\|) + \\ + \vartheta \left(\frac{\eta\rho + 1 + \widehat{L} + \|AB^{-1}\| (1 + L_g)}{\eta\rho} \right) [L_g(L + 1)]^2.$$

□

Capítulo 3

Solución de Problemas

Implícitos de

Complementariedad con un

Algoritmo de Perturbación

3.1. Revisión de la Literatura

En el capítulo anterior presentamos el problema implícito de complementariedad no lineal: encontrar un $x \in R^n$ tal que:

$$f(x) \geq 0, \quad g(x) \geq 0, \quad \text{y} \quad [g(x)]^T [f(x)] = 0, \quad (3.1)$$

en donde $f, g : R^n \rightarrow R^n$ y $f, g \in C^1(R^n)$; al cual aplicamos el Método de Newton para su solución.

Ahora bien, en este capítulo se presenta uno de los enfoques fructíferos a los problemas de complementariedad con una clase de algoritmos de perturbación. Un método de perturbación resuelve un problema de complementariedad estándar, es decir, con $g(x) = x$, fue presentado en [34] en la forma siguiente: para un vector positivo $z \in R_{++}^n$ consideramos una ecuación auxiliar no lineal

$$x_i f_i(x) = z_i, \quad i = 1, \dots, n. \quad (3.2)$$

Si el vector perturbación z tiende a cero, uno puede esperar una secuencia de soluciones $x = x(z)$ de la ecuación (3.2) a obtener una solución x^* del problema original de complementariedad como su punto límite.

Extendemos el algoritmo de perturbación antes mencionado al caso del problema implícito de complementariedad. Esto es, en la sección 3.2, probamos la existencia de soluciones al problema implícito de complementariedad perturbado: Para un vector perturbación positivo $z \in R^n$, $z > 0$, encontrar un punto $x \in R^n$ que satisfaga las siguientes condiciones:

$$g(x) \geq 0, \quad g_i(x) f_i(x) = z_i, \quad i = 1, \dots, n. \quad (3.3)$$

En la sección 3.3, mostramos la convergencia de las soluciones del problema perturbado a una solución del problema original, y examinamos la tasa de convergencia. En la sección 3.4 describimos un algoritmo resolviendo el problema perturbado en cada paso y probamos su convergencia también.

3.2. Existencia de soluciones a problemas perturbados

Definición 3.2.1 Una aplicación $f : R^n \rightarrow R^n$ es llamada localmente fuertemente monótona con respecto a otra aplicación $g : R^n \rightarrow R^n$ (o localmente g -fuertemente monótona) sobre un conjunto $X \subseteq R^n$ si existe una función escalar $\beta : R_+^1 \rightarrow R_+^1$ tal que $\beta > 0$ para todo $t > 0$, y:

$$[g(x) - g(y)]^T [f(x) - f(y)] \geq \beta(t) \|g(x) - g(y)\|^2. \quad (3.4)$$

para cualquier $x, y \in X \cap g^{-1}[B_t(0)]$; aquí $B_t(0) = \{v \in R^n \mid \|v\| \leq t\}$ es una bola cerrada de radio t centrada en el origen.

Note primeramente que la función $\beta(t)$ es necesariamente no creciente, y segundo, cualquier aplicación localmente g -monótona es estrictamente g -monótona.

Teorema 3.2.2 Si la aplicación $g : R^n \rightarrow R_{++}^n$ es un homeomorfismo, la aplicación f es continua y local g -fuertemente monótona, y la función β tiene la propiedad:

$$\lim_{t \rightarrow +\infty} \beta(t)t = +\infty, \quad (3.5)$$

entonces el problema (3.4) tiene una única solución $x = x_z$ para cada $z > 0$.

Demostración. Fijemos un vector positivo $\hat{\epsilon} = (\epsilon, \dots, \epsilon)^T > 0$ y consideremos una aplicación $F^\epsilon : R^n \rightarrow R^n$ definida como sigue:

$$F_i^\epsilon(x) = f_i(x) + \epsilon - \frac{z_i}{g_i(x) + \epsilon}, \quad i = 1, \dots, n. \quad (3.6)$$

Es fácil verificar que F^ϵ es localmente g -fuertemente monótona; en verdad,

$$\begin{aligned}
[g(x) - g(y)]^T [F^\epsilon(x) - F^\epsilon(y)] &= [g(x) - g(y)]^T [f(x) - f(y)] - \\
&\quad - \sum_{i=1}^n [g_i(x) - g_i(y)] z_i \left[\frac{1}{g_i(x) + \epsilon} - \frac{1}{g_i(y) + \epsilon} \right] \geq \\
&\geq \beta(t) \|g(x) - g(y)\|^2 + \sum_{i=1}^n \frac{z_i}{[g_i(x) + \epsilon][g_i(y) + \epsilon]} [g_i(x) - g_i(y)]^2 \geq \\
&\geq \beta(t) \|g(x) - g(y)\|^2
\end{aligned}$$

para cualquier x, y en el conjunto $g^{-1}[R_+^n \cap B_t(0)]$. Ahora bien, el Teorema 6.10 y la Proposición 7.4 en [22] implican que existe una única solución $x = x(\epsilon)$ del siguiente problema de complementariedad: Encontrar un $x \in R^n$ tal que:

$$g(x) \geq 0, \quad F^\epsilon(x) \geq 0, \quad g(x)^T F^\epsilon(x) = 0. \quad (3.7)$$

Tomando en cuenta (3.6), reescribimos (3.7) como sigue:

$$\begin{aligned}
g_i \geq 0, \quad f_i(x) + \epsilon &\geq \frac{z_i}{g_i(x) + \epsilon}, \\
g_i(x)[f_i(x) + \epsilon] &= \frac{g_i(x)z_i}{g_i(x) + \epsilon}, \quad i = 1, \dots, n.
\end{aligned} \quad (3.8)$$

Esta última ecuación implica claramente:

$$g_i(x)[f_i(x) + \epsilon] \leq z_i \quad i = 1, \dots, n. \quad (3.9)$$

Ahora consideramos una secuencia decreciente de parámetros $\{\epsilon_k\}_{k \in \mathbb{N}}$ y demostramos que la secuencia de puntos $\{g(x(\epsilon_k))\}_{k=1}^\infty$ es acotada. Por el contrario, sin pérdida de generalidad, suponemos que:

$$\|g(x(\epsilon_k))\| \rightarrow +\infty \quad \text{cuando } k \rightarrow \infty. \quad (3.10)$$

Entonces, por un lado, (3.8) y (3.9) implican:

$$A_k \equiv [g(x(\epsilon_k)) - g(x(\epsilon_1))]^T [f(x(\epsilon_k)) - f(x(\epsilon_1))] \leq 2 \sum_{i=1}^n z_i, \quad k = 1, 2, \dots; \quad (3.11)$$

por otro lado, debido a la condición de que f es local g -fuertemente monótona, tenemos las relaciones:

$$A_k \geq \beta(t_k) \|g(x(\epsilon_k)) - g(x(\epsilon_1))\|^2, \quad k = 1, 2, \dots, \quad (3.12)$$

en donde $t_k = \max \{g(x(\epsilon_k)), g(x(\epsilon_1))\}$. En seguida, (3.10) garantiza que $t_k \rightarrow +\infty$ cuando $k \rightarrow \infty$. Entonces tomando en cuenta (3.5), obtenemos de (3.12) que $A_k \rightarrow +\infty$ cuando $k \rightarrow \infty$ que contradice (3.11), y esto prueba que la secuencia $\{x(\epsilon_k)\}$ es también acotada. Otra vez sin pérdida de generalidad, podemos considerar un punto límite $x^* = \lim_{k \rightarrow \infty} x(\epsilon_k)$. Note primero que ninguno de los componentes del vector $g(x^*)$ pueden ser iguales a cero: esto último se sigue de las relaciones (3.8) y de la continuidad de f con respecto a x . Ahora haciendo uso de (3.8) y de la continuidad de f y g podemos obtener el límite de las desigualdades e igualdades:

$$g_i(x^*) \geq 0, \quad g_i(x^*)f_i(x^*) = z_i, \quad i = 1, \dots, n,$$

que prueba la existencia de soluciones al problema (3.3). La unicidad de x^* se sigue de estricta g -monotonocidad de f y del hecho de que g es un homeomorfismo. De verdad, habiendo otra solución y^* del problema (3.3) distinta de x^* (por lo tanto, con $g(x^*) \neq g(y^*)$), pudimos obtener la siguiente serie de relaciones

$$0 \leq [g(x^*) - g(y^*)]^T [f(x^*) - f(y^*)] =$$

$$= \sum_{i=1}^n \left[2z_i - \frac{z_i}{g_i(x^*)} g_i(y^*) - \frac{z_i}{g_i(y^*)} g_i(x^*) \right] < 0,$$

implicando la desigualdad imposible $0 < 0$. Esta contradicción completa la demostración.

□

Comentario 3.2.3 *La continuidad de la solución $x = x_z$ del problema (3.3) con respecto a la perturbación $z \in R_{++}^n$ está garantizada por la continuidad de las aplicaciones f y g y la condición límite (3.5).*

Corolario 3.2.4 *Bajo las suposiciones del **Teorema 3.2.2**, la solución $x = x_z$ del problema (3.3) es continua con respecto a la perturbación $z \in R_{++}^n$.*

Demostración. En verdad, consideremos una secuencia de perturbaciones $\{z^k\}$ convergiendo a $z \in R_{++}^n$. Primeramente, demostramos que la secuencia $\{x(z^k)\}_{k=0}^{\infty}$ de soluciones al problema (3.3) está acotada. Tenemos las desigualdades

$$g_i(x(z^k))f_i(x(z^k)) = z_i^k, \quad i = 1, \dots, n,$$

implicando la serie de relaciones:

$$\begin{aligned} A_k &\equiv [g(x(z^k)) - g(x(z^0))]^T [f(x(z^k)) - f(x(z^0))] = \\ &\sum_{i=1}^n \left[z_i^k + z_i^0 - z_i^0 \frac{g_i(x(z^k))}{g_i(x(z^0))} - z_i^k \frac{g_i(x(z^0))}{g_i(x(z^k))} \right] \leq \\ &\leq \sum_{i=1}^n (z_i^k + z_i^0) \leq \sum_{i=1}^n (z_i + z_i^0) + \sigma, \end{aligned} \quad (3.13)$$

para algún $\sigma > 0$ y todo $k = 0, 1, \dots$. Por otro lado,

$$A_k \geq \beta(t_k) \|g(x(z^k)) - g(x(z^0))\|^2,$$

donde $t_k = \max \{\|g(x(z^k))\|, \|g(x(z^0))\|\}$. Consecuentemente, A_k podría tomar valores arbitrariamente grandes si la secuencia $\{x(z^k)\}$ (y por tanto, la secuencia $\{g(x(z^k))\}$) no estaría acotada, como sigue de (3.5). Además, como la secuencia $\{A_k\}$ está acotada por (3.13), lo último implica que la secuencia $\{x(z^k)\}$ está acotada también. Ahora bien, la continuidad de f y g nos permite concluir que cualquier punto límite de esta secuencia resuelve el problema (3.3) con el vector perturbación z . Como la solución de (3.3) está determinada única, el punto límite de la secuencia $\{x(z^k)\}$ es también único implicando la continuidad de la aplicación $z \rightarrow x(z)$ y así se completa la demostración.

□

3.3. Taza de convergencia

Considerando nuevamente (3.3) con un homeomorfismo $g : R^n \rightarrow R_{++}^n$ y una aplicación continua f siendo también g -localmente fuertemente monótona sobre R^n . Como una consecuencia del Teorema 3.2.2, el problema: Encontrar un $x \in R^n$ tal que:

$$g_i(x)f_i(x) = z_i, \quad i = 1, \dots, n, \quad (3.14)$$

tiene solución únicamente para cualquier $z > 0$ si la condición límite (3.5) se cumple: $\lim_{t \rightarrow +\infty} \beta(t)t = +\infty$. Fijemos dos vectores perturbación $v > 0$

y $w > 0$, y denotamos por x y y a las soluciones del problema (3.14) en caso de $z = v$ y $z = w$, respectivamente.

Lema 3.3.1 *Sea $f : R^n \rightarrow R^n$ una aplicación continua y g -local fuertemente monótona, y la condición límite (3.5) es válida. Entonces, para soluciones x y y del problema (3.14) para vectores perturbación $v > 0$ y $w > 0$, respectivamente, la siguiente estimación es válida:*

$$\|g(x) - g(y)\| \leq \frac{1}{\sqrt{\beta(t)}} \|\sqrt{v} - \sqrt{w}\|, \quad (3.15)$$

en donde $t = \max\{\|g(x)\|, \|g(y)\|\}$, y $\sqrt{z} = (\sqrt{z_1}, \dots, \sqrt{z_n})^T$ para $z > 0$.

Demostración. Debido a (3.4) y (3.14), la siguiente serie de relaciones es válida:

$$\begin{aligned} \beta(t) \|g(x) - g(y)\|^2 &\leq [g(x) - g(y)]^T [f(x) - f(y)] = \\ &= \sum_{i=1}^n [g_i(x)f_i(x) + g_i(y)f_i(y) - g_i(x)f_i(y) - g_i(y)f_i(x)] = \\ &= \sum_{i=1}^n \left(v_i + w_i - w_i \frac{g_i(x)}{g_i(y)} - v_i \frac{g_i(y)}{g_i(x)} \right) \leq \\ &\leq \sum_{i=1}^n (v_i + w_i - 2\sqrt{v_i w_i}) = \sum_{i=1}^n (\sqrt{v_i} - \sqrt{w_i})^2. \end{aligned}$$

Extrayendo la raíz cuadrada del primer y último terminos de la serie, obtenemos la desigualdad

$$\sqrt{\beta(t)} \|g(x) - g(y)\| \leq \|\sqrt{v} - \sqrt{w}\|,$$

implica (3.15) completando la demostración.

□

Ahora, asumimos que la aplicación f es g -Lipschitz continua sobre cada subconjunto compacto $S \subset R^n$, esto es, para algunos escalares positivos $L = L(S)$ y cualquier $x, y \in g^{-1}(S)$ la siguiente desigualdad está cumplida:

$$|f_i(x) - f_i(y)| \leq L \|g(x) - g(y)\|, i = 1, \dots, n. \quad (3.16)$$

Nuevamente fijamos perturbaciones $v > 0$ y $w > 0$ en el problema (3.14) y denotamos por $x = x(v)$ y $y = y(w)$ a las soluciones correspondientes del problema.

Lema 3.3.2 *La siguiente estimación es válida:*

$$\begin{aligned} \|v - w\|_1 &\leq \sqrt{n} [L(S) \|g(x) - g(y)\| + \|f(y)\| + L(S) \|g(y)\|] \times \\ &\quad \times \|g(x) - g(y)\|, \end{aligned} \quad (3.17)$$

en donde $\|p\|_1 = \sum_{i=1}^n |p_i|$ y $\|p\|_2 = \sqrt{\sum_{i=1}^n p_i^2}$ para $p \in R^n$.

Demostración. Consideramos un subconjunto compacto arbitrario $S \subset R^n$ que comprende ambas x y y . Enseguida escribimos las relaciones

$$\begin{aligned} v_i - w_i &= g_i(x)f_i(x) - g_i(y)f_i(y) = [g_i(x) - g_i(y)][f_i(x) - f_i(y)] + \\ &\quad + g_i(y)f_i(x) + g_i(x)f_i(y) - 2g_i(y)f_i(y), \quad i = 1, \dots, n, \end{aligned}$$

que, junto con (3.16), implican

$$\begin{aligned} |v_i - w_i| &\leq |g_i(x) - g_i(y)| |f_i(x) - f_i(y)| + |g_i(x) - g_i(y)| |f_i(y)| + \\ &\quad + |g_i(y)| |f_i(x) - f_i(y)| \leq L(S) |g_i(x) - g_i(y)| \|g(x) - g(y)\| + \end{aligned}$$

$$\begin{aligned}
& + |g_i(x) - g_i(y)| |f_i(y)| + L(S) |g_i(y)| \|g(x) - g(y)\| \leq \\
\leq & \|g(x) - g(y)\| [L(S) |g_i(x) - g_i(y)| + |f_i(y)| + L(S) |g_i(y)|], \quad i = 1, \dots, n.
\end{aligned}$$

Resumiendo estas desigualdades para todo $i = 1, \dots, n$, llegamos a

$$\|v - w\|_1 \leq \|g(x) - g(y)\| [L(S) \|g(x) - g(y)\|_1 + \|f(y)\|_1 + L(S) \|g(y)\|_1]$$

que nos da el límite superior esperado (3.17), ya que $\|p\|_1 \leq \sqrt{n} \|p\|_2$ para todo $p \in R^n$. La demostración está completa.

□

Intercambiando los roles de v y w en ambas partes de (3.17) sigue directamente la estimación:

$$\begin{aligned}
\|v - w\|_1 \leq & \sqrt{n} [L(S) \|g(x) - g(y)\| + \|f(x)\| + L(S) \|g(x)\|] \times \\
& \times \|g(x) - g(y)\|, \tag{3.18}
\end{aligned}$$

Juntando las estimaciones (3.17) y (3.18) llegamos a la siguiente desigualdad (simétrica)

$$\begin{aligned}
\|v - w\|_1 \leq & \sqrt{n} [L(S) \|g(x) - g(y)\| + \\
& + \text{mín} \{ \|f(x)\| + L(S) \|g(x)\|, \|f(y)\| + L(S) \|g(y)\| \}] \|g(x) - g(y)\|.
\end{aligned}$$

Reemplazando y en (3.16) y (3.17) con la solución x^* del problema inicial (3.1), obtenemos las respectivas estimaciones:

$$\|g(x) - g(x^*)\| \leq \frac{1}{\sqrt{\beta(t)}} \|\sqrt{z}\| = \frac{1}{\sqrt{\beta(t)}} \sqrt{\|z\|_1} \tag{3.19}$$

en donde $t = \|g(x^*)\| + \|g(x) - g(x^*)\|$, y

$$\begin{aligned} \|z\|_1 &= \sqrt{n} [L(S) \|g(x) - g(x^*)\| + \|f(x^*)\| + \\ &\quad + L(S) \|g(x^*)\| \|g(x) - g(x^*)\|]. \end{aligned} \quad (3.20)$$

aquí $S = g^{-1}[B_t(0)]$. De esos límites por ambos lados (3.19) y (3.20) y el hecho que $\|g(\hat{x}) - g(x^*)\| \leq \|g(x) - g(x^*)\|$ implica que \hat{x} pertenece a la bola $g^{-1}[B_t(0)]$ junto con x y x^* , obtenemos el siguiente enunciado.

Si asumimos que g^{-1} es una aplicación localmente Lipschitziana, entonces la secuencia de las soluciones $x(z^k)$ del problema (3.14) converge a la solución x^* del problema implícito de complementariedad (3.1) a una velocidad igual a la mitad de la convergencia de las perturbaciones z^k tendientes a cero. Por ejemplo, en el caso de una convergencia cuadrática de z^k a cero, $x(z^k)$ convergerá a x^* al mínimo lineal (en la taza geométrica).

3.4. Resolviendo Problemas Perturbados

Fijemos un vector perturbación $z \in R_{++}^n$ y consideremos el problema perturbado: Encontrar un $x \in R^n$ tal que:

$$g(x) \geq 0, \quad f(x) \geq 0 \quad \text{y} \quad g_i(x)f_i(x) = z_i, \quad i = 1, \dots, n. \quad (3.21)$$

Como $z > 0$, este problema puede ser reescrito como sigue:

$$F(x) = 0 \quad (3.22)$$

en donde

$$F_i(x) = f_i(x) - \frac{z_i}{g_i(x)}, \quad i = 1, \dots, n. \quad (3.23)$$

Por conveniencia asumimos que el problema perturbado tiene soluciones siendo g un homeomorfismo de R^n en sí mismo, y representamos F como $F(x) = f(x) + H(x)$, con:

$$H_i(x) = -\frac{z_i}{g_i(x)}, \quad i = 1, \dots, n.$$

Es fácil verificar que la aplicación $H : R_{++}^n \rightarrow R^n$ es g -monótona. Además, si la aplicación f es continua y g -local fuertemente monótona sobre R^n , podemos aplicar un método de Peaceman-Rachford modificado [38] para resolver el problema (3.21).

Teniendo una iteración actual $g(x^k)$, primeramente encontramos el vector $g(x^{k+\frac{1}{2}})$ como una solución positiva de la ecuación:

$$\lambda g(x^{k+\frac{1}{2}}) + H(x^{k+\frac{1}{2}}) = \lambda g(x^k) - f(x^k); \quad (3.24)$$

después de haber obtenido $g(x^{k+\frac{1}{2}})$, construimos el vector $g(x^{k+1})$ como una solución a la ecuación:

$$\lambda g(x^{k+1}) + f(x^{k+1}) = \lambda g(x^{k+\frac{1}{2}}) - H(x^{k+\frac{1}{2}}). \quad (3.25)$$

Aquí $\lambda > 0$ es un parámetro fijo.

Uno puede probar la convergencia de este algoritmo por haber adaptado la prueba del teorema de Peaceman-Rachford global en [38].

Teorema 3.4.1 *Sea $f : R^n \rightarrow R^n$ una aplicación g -local fuertemente monótona y g -Lipschitz continua sobre cada subconjunto en R^n . Entonces para cualquier inicial $x^0 \in R^n$, y un fijo $\lambda > 0$, la secuencia $\{x^k\}$ está bien definida por las formulas (3.24) - (3.25) y converge a la solución única $x^* = x^*(z)$ del problema (3.1).*

Demostración. La existencia y unicidad de la solución x^* está garantizada por el Teorema 3.2.2. El punto auxiliar $x^{k+1/2}$ está bien definido debido a la forma de las ecuaciones cuadráticas con respecto a las componentes de $g(x_i^{k+1/2})$, inducidos por la ecuación (3.24). Adelante, la aplicación $(f + \lambda g)$ es continua y g -fuertemente monótona en R^n y por lo tanto es un homeomorfismo de R^n en sí misma (Teorema 6.4.4 en [38]). Lo anterior implica que el sistema de ecuaciones (3.25) tiene una única solución x^{k+1} y justifica el proceso de iteración (3.24)-(3.25). Escribámoslo en la forma vectorial:

$$x^{k+1} = Px^k, \quad k = 0, 1, \dots, \quad (3.26)$$

en donde

$$Px = (\lambda g + f)^{-1}(\lambda g - H)(\lambda g + H)_+^{-1}(\lambda g - f)x;$$

aquí $(\lambda g + H)_+^{-1}(\lambda g - f)$ es el operador que asocia el vector $x^k \in R^n$ con tal solución de la ecuación (3.24) que $g(x^{k+1/2}) > 0$. Las relaciones

$$F(x^*) = f(x^*) + H(x^*) = 0 \quad \text{y} \quad g(x^*) > 0$$

implican las igualdades

$$(\lambda g + H)_+^{-1}(\lambda g - f)x^* = x^* \quad \text{y} \quad (\lambda g + f)_+^{-1}(\lambda g - H)x^* = x^*.$$

Además, x^* es un punto fijo de la aplicación P .

Inversamente, si x^* es un punto fijo de la aplicación P y $g(x^*) > 0$, entonces x^* resuelve la ecuación (3.22), esto es, $f(x^*) + H(x^*) = 0$. En verdad, si $x^* = Px^*$, denotamos por

$$y^* = (\lambda g + H)_+^{-1}(\lambda g - f)x^*,$$

y obtenemos la igualdad

$$(\lambda g + f)x^* = (\lambda g - H)y^*. \quad (3.27)$$

Por otro lado, la definición de y^* implica que

$$(\lambda g + H)y^* = (\lambda g - f)x^*. \quad (3.28)$$

Escribamos las ecuaciones (3.27)-(3.28) en la forma siguiente

$$\lambda [g(x^*) - g(y^*)] = -f(x^*) - H(y^*),$$

y

$$\lambda [g(x^*) - g(y^*)] = f(x^*) + H(y^*),$$

respectivamente. Con esto llegamos inmediatamente a $g(x^*) = g(y^*)$, y por tanto

$$f(x^*) + H(x^*) = 0.$$

Además, sea $S \subset R^n$ un subconjunto compacto arbitrario. Entonces el subconjunto $\widehat{S} = (\lambda g + f)^{-1}S$ es también compacto, y para cualquier $u, v \in \widehat{S}$ uno tiene

$$[g(u) - g(v)]^T [f(u) - f(v)] \geq \beta(t) \|g(u) - g(v)\|^2, \quad (3.29)$$

donde $t = \max \{ \|w\| \mid w \in \widehat{S} \}$. En seguida, denotamos por ν el coeficiente g -Lipschitz de la aplicación f sobre \widehat{S} , y para arbitrarios $x, y \in S, x \neq y$, sea $u = (\lambda g + f)^{-1}x, v = (\lambda g + f)^{-1}y$. Entonces las siguientes relaciones son

válidas para la aplicación $P_1 = (\lambda g - f)(\lambda g + f)^{-1}$ (aquí $\|\cdot\| = \|\cdot\|_2$ es la norma Euclidea en R^n):

$$\begin{aligned} & \left[\frac{\|P_1x - P_1y\|}{\|x - y\|} \right]^2 = \\ &= \frac{\lambda^2 \|g(u) - g(v)\|^2 - 2\lambda [g(u) - g(v)]^T [f(u) - f(v)] + \|f(u) - f(v)\|^2}{\lambda^2 \|g(u) - g(v)\|^2 + 2\lambda [g(u) - g(v)]^T [f(u) - f(v)] + \|f(u) - f(v)\|^2} \leq \\ &\leq \frac{[\lambda^2 - 2\lambda\beta(t)] \|g(u) - g(v)\|^2 + \|f(u) - f(v)\|^2}{[\lambda^2 + 2\lambda\beta(t)] \|g(u) - g(v)\|^2 + \|f(u) - f(v)\|^2} \leq \frac{\lambda^2 - 2\lambda\beta(t) + v^2}{\lambda^2 + 2\lambda\beta(t) + v^2} < 1, \end{aligned}$$

en donde la segunda desigualdad no estricta está basada en el hecho de la monotonocidad de la función $\frac{a+\tau}{b+\tau}$ sobre $\tau \geq 0$ si $b \geq a \geq 0$. Así obtenida, la estimación muestra que la aplicación P_1 es una contracción estricta sobre S .

Ahora consideramos la aplicación $P_2 = (\lambda g - H)(\lambda g + H)_+^{-1}$ y mostramos que satisface la condición

$$\|P_2x - P_2y\| \leq \|x - y\| \quad (3.30)$$

para cualquier $x, y \in R^n$. De verdad, notamos primero que la aplicación H es g -monótona y g -Lipschitz continua sobre cada subconjunto compacto $S \subset R_{++}^n$. Por otra parte, para cada par $x, y \in R^n$ los puntos $u = (\lambda g + H)_+^{-1}x$ y $v = (\lambda g + H)_+^{-1}y$ pertenecen a R_{++}^n , y la fórmula de la ecuación cuadrática implica que el conjunto $\widehat{S} = (\lambda g + H)_+^{-1}S$ es un subconjunto compacto de R_{++}^n siempre que $S \subset R^n$ es un compacto. Ahora bien, es claro que la validez de la relación (3.30) para todo $x, y \in R^n$ puede ser probada repitiendo exactamente los argumentos usados al demostrar la contractividad de P_1 .

Así, la composición P_2P_1 es una aplicación contractiva sobre cada subconjunto compacto en R^n . Debido a la identidad

$$P = (\lambda g - f)^{-1}P_2P_1(\lambda g + f)$$

y el hecho que P tiene el (único) punto fijo x^* tal que $g(x^*) > 0$, la aplicación P_2P_1 tiene el punto fijo $y^* = (\lambda g + f)x^*$. Lo último nos permite concluir ([38], Teorema 12.1.4) que la secuencia $\{y^k\}$, construida por la recurrencia $y^{k+1} = P_2P_1y^k$, $y^0 = (\lambda g + f)x^0$, $k = 0, 1, \dots$, convergerá a y^* . Sin embargo, esto último implica ([38], Teorema 12.1.6) que la secuencia $\{x^k\}$ convergerá a x^* . La demostración está completa.

□

Comentario 3.4.2 *El punto $g(x^{k+1/2})$ que resuelve la ecuación (3.24) es calculada fácilmente con las fórmulas de la ecuación cuadrática. Esto es,*

$$g_i(x^{k+1/2}) = \frac{g_i(x^k)}{2} - \frac{f_i(x^k)}{2\lambda} + \frac{\sqrt{[\lambda g_i(x^k) - f_i(x^k)]^2 + 4\lambda z_i}}{2\lambda}, \quad i = 1, \dots, n.$$

Consecuentemente, en cada paso del algoritmo descrito, uno sólo necesita resolver la ecuación no lineal

$$\begin{aligned} \lambda g_i(x^{k+1}) + f_i(x^{k+1}) = & \lambda \frac{g_i(x^k)}{2} - \frac{f_i(x^k)}{2} + \frac{\sqrt{[\lambda g_i(x^k) - f_i(x^k)]^2 + 4\lambda z_i}}{2} + \\ & + \frac{z_i}{g_i(x^{k+1/2})}, \end{aligned}$$

$i = 1, \dots, n$. Sin embargo, la aplicación $(\lambda g + f)$ es g -fuertemente monótona, por tanto el último sistema de ecuaciones puede ser tratado eficientemente por algoritmos tipo Newton.

Capítulo 4

Un Nuevo Algoritmo para Resolver el Problema de Programación Binivel Entero Mixto

4.1. Revisión de la Literatura

La investigación de problemas de programación binivel está fuertemente motivada por sus aplicaciones (en el mundo real). Tales problemas surgen en los juegos de Stackelberg que tratan la economía de mercado [46], en donde los distintos tomadores de decisión tratan de comprender mejor las decisiones en el mercado con respecto a sus propios objetivos generalmente diferentes. Sin embargo, a menudo son capaces de realizar sus decisiones de

forma independiente, pero se ven obligados a actuar de acuerdo a una cierta jerarquía. Vamos a considerar el caso más simple de tal situación donde sólo hay dos tomadores de decisión. El líder es el que puede manejar el mercado de forma independiente, mientras que el seguidor tiene que actuar de una manera dependiente.

Es obvio que, si un tomador de decisiones es capaz de asumir una posición independiente (y por tanto para observar y utilizar las reacciones del dependiente tomador de decisiones a las decisiones del líder), y luego a tratar de hacer una ventaja de esto.

Los Problemas de Programación Binivel son más generales que los juegos de Stackelberg en el sentido de que ambos conjuntos factibles pueden depender sobre la decisión de otro tomador de decisiones.

En términos matemáticos, el conjunto de variables es particionado en dos variables vectoriales, x y y , en donde $y \in R^m$ es la variable del líder y $x \in R^n$ es la variable del seguidor. Usando y como un parámetro, el seguidor resuelve un problema de optimización paramétrico, y los valores $x = x(y)$ están determinados por el seguidor conociendo de antemano la elección y del líder. El líder tiene que determinar la mejor elección de y conociendo la reacción (óptima) $x = x(y)$ del seguidor a la decisión del líder. Tenemos un líder (nivel superior) que elige primero su decisión con el objetivo de minimizar una cierta función $f(x(y), y)$, y un seguidor (en el nivel inferior) que responde óptimamente a esta decisión.

Un caso particular del problema de programación binivel está presentado por el siguiente modelo entero-mixto surgiendo del problema de minimizar

los costos de penalización de una compañía cargadora de gas natural. Éste es un problema de la vida real (*vea* [12]). El problema en cuestión se genera cuando un cargador (de gas) firma un contrato con una compañía de gasoductos a entregar cierta cantidad de gas a cierta distancia de la entrega. Lo que realmente es enviado puede ser superior o inferior a la cantidad que inicialmente se había acordado (éste fenómeno es llamado un *desbalance*). Cuando ocurre tal desequilibrio, el gasoducto penaliza al cargador por el que se establece una política de penalización. Dado que esta penalización es una función de la operación de los desbalances diarios, un problema importante para los cargadores, es cómo llevar a cabo sus desbalances diarios, así como minimizar la penalización impuesta.

Nuestro principal objetivo es proponer un algoritmo eficiente para resolver el Problema de Programación Binivel Entero Mixto (MIBLPP por sus siglas en inglés). El término *entero mixto* significa que el problema tiene ambas variables, continuas y discretas. Por otra parte, en nuestro caso x es la variable vectorial continua y y puede incluir variables discretas. Sabiendo que este problema es difícil de resolver, se propone un algoritmo de aproximación que nos conduce a una solución global.

Muchos autores han trabajado en las diferentes formulaciones de los problemas de programación binivel. En [9] se propuso un algoritmo para resolver el problema de programación lineal binivel (BLPP) utilizando el método simplex con variables adicionales en el conjunto de base y la teoría de subgradientes. En [2] se obtuvieron límites superiores para las funciones objetivo en ambos niveles. Así se generó una secuencia no decreciente de límites inferi-

ores de la función objetivo en el nivel superior, que, bajo ciertas condiciones, converge a la solución del BLPP general para funciones continuamente diferenciables.

En [11] se presentan varias alternativas para resolver el MIBLPP con las condiciones de integralidad. En [47] se resolvió el MIBLPP con el algoritmo branch and bound. El análisis de sensibilidad para MIBLPP también fue considerado en [48].

4.2. Formulación General

Consideramos el problema jerárquico de optimización binivel: la toma de decisiones en el nivel superior está gobernada por las restricciones que están definidas en parte por un problema de optimización (paramétrico) en el nivel inferior. Sea éste segundo problema definido como sigue:

$$\min_x \{f(x, y) \mid g(x, y) \leq 0, h(x, y) = 0\}, \quad (4.1)$$

en donde $f : R^n \times R^m \rightarrow R$, $g : R^n \times R^m \rightarrow R^p$ y $h : R^n \times R^m \rightarrow R^q$ con $g(x, y) = -(g_1(x, y), \dots, g_p(x, y))^T$ y $h(x, y) = (h_1(x, y), \dots, h_q(x, y))^T$. Éste problema es llamado el de nivel inferior o el problema del seguidor. Denotemos como $\psi(y)$ al conjunto solución del problema (4.1) para un parámetro fijo $y \in R^m$.

Ahora podemos formular el problema binivel como sigue:

$$\min_{x,y} \{F(x, y) \mid y \in Y, x \in \psi(y)\}, \quad (4.2)$$

en donde $F : R^n \times R^m \rightarrow R$, y Y es un subconjunto cerrado de R^m . Éste es el problema de nivel superior o el problema del líder. El problema (4.1) – (4.2) es el *problema de programación binivel*.

Con el fin de garantizar que el problema de programación binivel está bien definido, asumimos lo siguiente:

1. El conjunto $M = \{(x, y) \mid g(x, y) \leq 0, h(x, y) = 0\}$ es no vacío.
2. Ambos $F(x, y)$ y $f(x, y)$ están acotadas por debajo sobre M .

Definición 4.2.1 *Un par ordenado (x, y) se le llama factible al problema de programación lineal binivel si éste satisface $y \in Y$ y $x \in \psi(y)$.*

Definición 4.2.2 *Un par ordenado factible (x', y') es llamado solución óptima al problema de programación binivel si $F(x', y') \leq F(x, y)$ para todas las soluciones factibles (x, y) .*

4.3. Modelo Matemático

El Problema de Programación Lineal Binivel Entero Mixto con un parámetro en el lado derecho del nivel inferior está formulado como sigue:

$$\min_{x,y} \{ \langle a, x \rangle + \langle b, y \rangle \mid Gy = d, x \in \psi(y), y \in Z_+^m \}, \quad (4.3)$$

que representa el nivel superior en donde $a, x \in R^n$, $b, y \in R^m$, G es una matriz $r \times m$, $d \in R^r$. Note que aquí usamos la versión optimística del problema de programación binivel, *vea* [10]. De aquí en adelante, $\langle \cdot, \cdot \rangle$

representa el producto interno Euclideo. La función de valor objetivo del nivel superior (Problema del Cargador) representa a los costos de producción más la penalización que intentamos minimizar. Aquí $\psi(y)$ está definida como sigue:

$$\psi(y) = \underset{x}{\text{Argmin}} \{ \langle c, x \rangle \mid Ax = y, x \geq 0 \}, \quad (4.4)$$

que describe al conjunto de soluciones óptimas del problema de nivel inferior (el conjunto de las reacciones racionales). Aquí $c, x \in R^n$, A es una matriz $m \times n$ con $m \leq n$.

Determinemos ahora la función de valor óptimo del problema de nivel inferior (Problema del Gasoducto) como sigue:

$$\varphi(y) = \underset{x}{\text{mín}} \{ \langle c, x \rangle \mid Ax = y, x \geq 0 \}. \quad (4.5)$$

Suponemos que el conjunto factible del problema (4.4) es no vacío. Aquí y es el vector paramétrico que puede representar los valores de diferentes desbalances diarios, la cantidad de gas transportada o también la utilidad del cargador en el caso cuando necesitamos trabajar con números enteros. El nivel inferior (Problema del Gasoducto), en función de nuestros objetivos puede intentar minimizar el desbalance, los volúmenes de gas transportados, o el valor absoluto de las transacciones entre el líder y el seguidor.

En éste trabajo, consideramos una reformulación de (4.3)–(4.5) basada sobre una aproximación reportada en la literatura (*vea* [50] o [10]) como un problema clásico de optimización no diferenciable. Si tomamos en cuenta la función de valor óptimo del nivel inferior (4.5), entonces el problema (4.3)–

(4.5) puede ser reemplazado por:

$$\min_{x,y} \{ \langle a, x \rangle + \langle b, y \rangle \mid Gy = d, \langle c, x \rangle \leq \varphi(y), Ax = y, x \geq 0, y \in Z_+^m \}. \quad (4.6)$$

Nuestro trabajo está concentrado en la función de valor objetivo del nivel inferior (4.5). Por esta razón, mostramos algunas características importantes (vea [17] o [13]) que serán de ayuda para resolver el problema (4.6).

4.4. Geometría del problema

Consideremos el problema lineal de programación paramétrico (4.5):

$$\varphi(y) = \min_x \{ \langle c, x \rangle \mid Ax = y, x \geq 0 \}.$$

Con el fin de resolver éste problema, usamos el método simplex dual, como se hizo en [13]. Sea un vector fijo $y = y^*$ y x^* una solución básica óptima para $y = y^*$ con su correspondiente matriz básica B , que es una submatriz cuadrada y del mismo rango de A , y tal que $x^* = (x_B^*, x_N^*)^T$, con $x_B^* = B^{-1}y$ y $x_N^* = 0$. Además, fijemos el valor del parámetro del nivel superior $y = y^*$. Entonces, podemos decir que $x^*(y^*) = (x_B^*(y^*), x_N^*(y^*))^T = (B^{-1}y^*, 0)^T$ es una solución básica óptima del problema (4.5) para un parámetro fijo y^* . Y si la siguiente desigualdad es válida:

$$B^{-1}y \geq 0,$$

entonces $x^*(y) = (x_B^*(y), x_N^*(y))^T = (B^{-1}y, 0)^T$ es también óptima para el vector paramétrico y .

Es posible perturbar y^* de modo que B siga siendo una matriz básica óptima [17]. Denotamos por $\mathfrak{R}(B)$ a un conjunto que llamaremos la *región de estabilidad* de B , que está definida como:

$$\mathfrak{R}(B) = \{y \mid B^{-1}y \geq 0\}.$$

Para todo $y \in \mathfrak{R}(B)$, el punto $x^*(y) = (x_B^*(y), x_N^*(y))^\top = (B^{-1}y, 0)^\top$ es una solución básica óptima al problema (4.5).

Esta región es no vacía ya que $y^* \in \mathfrak{R}(B)$. Además, esta región es cerrada pero no necesariamente acotada. Si $\mathfrak{R}(B)$ y $\mathfrak{R}(B')$ son dos diferentes regiones de estabilidad con $B \neq B'$, entonces sólo uno de los siguientes casos es posible.

1. $\mathfrak{R}(B) \cap \mathfrak{R}(B') = \{0\}$.
2. $\mathfrak{R}(B) \cap \mathfrak{R}(B')$ contiene la frontera común de las regiones $\mathfrak{R}(B)$ y $\mathfrak{R}(B')$.
3. $\mathfrak{R}(B) = \mathfrak{R}(B')$.

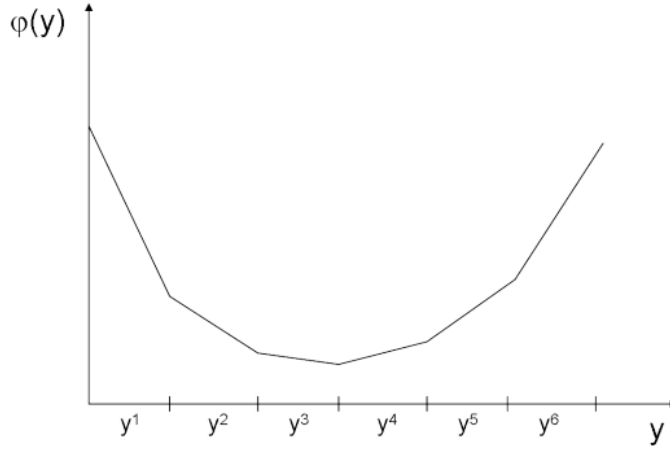


Figura 1. Representación de φ en 1 dimensión

Por otra parte, $\mathfrak{R}(B)$ es un conjunto poliédrico convexo, en la que la función de valor óptimo del nivel inferior es una función lineal y finita. Para determinar una descripción explícita de la función φ consideremos el problema dual al problema (4.5). Si $\varphi(y)$ es finita, entonces

$$\varphi(y) = \max\{\langle y, u \rangle : A^T u \leq c\}.$$

Ahora u^1, u^2, \dots, u^s denotan los vértices del conjunto poliédrico $\{u : A^T u \leq c\}$. Entonces,

$$\varphi(y) = \max\{\langle y, u^1 \rangle, \langle y, u^2 \rangle, \dots, \langle y, u^s \rangle\},$$

siempre que $\varphi(y)$ es finita.

Por dualidad, para alguna matriz básica B_i con $y \in \mathfrak{R}(B_i)$ tenemos $B_i^T u =$

c_{B_i} o $u = (B_i^T)^{-1} c_{B_i}$ y así,

$$\langle y, u^i \rangle = \langle y, (B_i^T)^{-1} c_{B_i} \rangle = \langle (B_i)^{-1} y, c_{B_i} \rangle.$$

Poniendo $x^i(y) = ((B_i)^{-1} y, 0)$ llegamos a:

$$\varphi(y) = \text{máx} \{ \langle c, x^1(y) \rangle, \langle c, x^2(y) \rangle, \dots, \langle c, x^q(y) \rangle \}.$$

Como podemos ver en la figura 1, las regiones de estabilidad están representadas por los segmentos sobre el eje de las y . La función φ es no diferenciable, esto hace a este tipo de problemas difíciles de resolver.

Ahora, introducimos la siguiente definición [50].

Definición 4.4.1 *Sea el par (x^*, y^*) tal que resuelva al problema (4.6). Entonces (4.6) es llamado partially calm en (x^*, y^*) si existe una constante $\mu > 0$ y una vecindad U de $(x^*, y^*, 0) \in R^n \times R^m \times R$, tal que para todo $(x, y, u) \in U$ factible al problema:*

$$\min_{x, y, u} \{ \langle a, x \rangle + \langle b, y \rangle \mid Gy = d, \langle c, x \rangle - \varphi(y) + u = 0, Ax = y, x \geq 0, y \in Z_+^m \}, \quad (4.7)$$

tenemos

$$\langle a, x \rangle + \langle b, y \rangle - \langle a, x^* \rangle - \langle b, y^* \rangle + \mu |u| \geq 0.$$

Aquí $|u|$ representa el valor absoluto de u .

Teorema 4.4.2 *Sea un par (x^*, y^*) que resuelva al problema (4.3)–(4.5), entonces (4.6) es partially calm en (x^*, y^*) .*

Demostración. Fijamos un valor arbitrario $\delta > 0$ y suponemos que $(x', y', u) \in (x^*, y^*, 0) + \delta B(0, 1)$, en donde $B(0, 1)$ es una bola unitaria en R^{n+m+m} centrada en el origen, es factible al problema (4.7), i.e.,

$$\langle c, x' \rangle - \varphi(y') + u = 0 \quad (4.8)$$

$$Ax' - y' = 0 \quad (4.9)$$

Sea $x(y')$ una solución al problema lineal de nivel inferior, i.e.,

$$\langle c, x(y') \rangle - \varphi(y') = 0 \quad (4.10)$$

$$Ax(y') - y' = 0 \quad (4.11)$$

Entonces $(x(y'), y')$ es factible para el Problema Binivel de Programación Lineal. Por la optimalidad de (x^*, y^*) y usando la desigualdad de Cauchy-Buniakovski tenemos:

$$\langle a, x' \rangle + \langle b, y' \rangle - \langle a, x^* \rangle - \langle b, y^* \rangle \geq \langle a, x' \rangle + \langle b, y' \rangle - \langle a, x(y') \rangle - \langle b, y' \rangle = \quad (4.12)$$

$$= \langle a, x' \rangle - \langle a, x(y') \rangle \geq -\|a\|_\infty \|x' - x(y')\|_\infty, \quad (4.13)$$

en donde $\|\cdot\|_\infty$ representa a la norma ∞ , i.e. $\|v\|_\infty = \max\{|v_i|, i = 1, \dots, n\}$. Por la definición de la función de valor óptimo en el nivel inferior, ya que $(x', y', 0)$ es factible para el problema (4.7), u debe ser no positiva.

Si la solución $x = x(y')$ del problema de nivel inferior no es única, entonces podemos escoger la más cercana al punto x' , de modo que:

$$\begin{aligned} & \|x' - x(y')\|_\infty = \\ & = \min_{\varepsilon, x} \{ \varepsilon \mid -\varepsilon e \leq x' - x \leq \varepsilon e, \langle c, x \rangle - \varphi(y') = 0, Ax - y' = 0, x \geq 0 \}. \end{aligned}$$

Aquí $e = (1, 1, \dots, 1)^T \in R^n$, y las desigualdades son por componentes.

Ahora consideramos el dual del último problema de programación lineal:

$$\begin{aligned} w = & \max_{\xi_1 \in R_+^n, \xi_2 \in R_+^n, \xi_3 \geq 0, \xi_4 \in R_+^m} \{ \langle \xi_1 - \xi_2, x' \rangle + \xi_3(-\varphi(y')) + \langle \xi_4, (-y') \rangle \mid \\ & \xi_1 - \xi_2 - \xi_3 c - A^T \xi_4 = 0, \langle e, \xi_1 \rangle + \langle e, \xi_2 \rangle = 1 \} = \end{aligned} \quad (4.14)$$

$$\begin{aligned} & = \max \{ \xi_3(\langle c, x' \rangle - \varphi(y')) + \langle \xi_4, Ax' - y' \rangle \mid \\ & \langle e, \xi_1 \rangle + \langle e, \xi_2 \rangle = 1, \xi_i \geq 0, i = 1, 2, 3, 4 \}. \end{aligned} \quad (4.15)$$

Por dualidad de programación lineal tenemos:

$$\begin{aligned} \|x' - x(y')\|_\infty & = \xi_3(x', y', 0) [\langle c, x' \rangle - \varphi(y')] + \langle \xi_4(x', y', 0), Ax' - y' \rangle = \\ & = \xi_3(x', y', 0) [\langle c, x' \rangle - \varphi(y')], \end{aligned}$$

en donde $(\xi_1(x', y', 0), \xi_2(x', y', 0), \xi_3(x', y', 0), \xi_4(x', y', 0)) \in R_+^n \times R_+^n \times R_+ \times R_+^m$ es una solución del problema de maximización (4.14). Entonces tenemos las siguientes relaciones:

$$\begin{aligned} \|x' - x\|_\infty & = \xi_3(x', y', 0) [\langle c, x' \rangle - \varphi(y')] = \\ & = \xi_3(x', y', 0)(-u) = \xi_3(x', y', 0) |u|. \end{aligned} \quad (4.16)$$

Ya que $\xi_3(x', y', 0)$ puede ser escogido como una componente de un vértice solución del problema de maximización (4.14), y ya que la región factible es independiente de $(x', y', 0)$ y tiene un número finito de vértices, llegamos a:

$$\xi_3(x', y', 0) \leq L \tag{4.17}$$

en donde

$$L := \text{máx} \{ \xi_3 \mid (\xi_1, \xi_2, \xi_3, \xi_4) \text{ es un vértice del conjunto de restricciones en (4.14)} \}.$$

Entonces, combinando (4.13), (4.16) y (4.17), concluimos que el problema (4.6) es partially calm con $\mu = \|a\|_\infty L$, ya que

$$\langle a, x' \rangle + \langle b, y' \rangle - \langle a, x^* \rangle - \langle b, y^* \rangle + \mu |u| \geq 0.$$

□

4.5. Un Algoritmo de Aproximación

Las bases para empezar a describir el algoritmo están dadas en las secciones anteriores. Las dificultades que surgen al trabajar con la función de valor objetivo (4.5) se deben al simple hecho de que no la tenemos en una forma explícita. Éste algoritmo trata en aproximar (4.5) con un número infinito de iteraciones. También es bien sabido que (4.5) es no diferenciable, es por eso que en algunos trabajos como [14] y [50] trabajan con cálculo subdiferencial que se basan sobre la restricción de cualificación no suave de Mangasarian-Fromowitz.

Las herramientas que utilizaremos están basadas principalmente en el hecho que (4.5) es lineal por partes y convexa. También, las bases para desarrollar un buen algoritmo están dados en los siguientes teoremas bien conocidos, importantes para guardar la convexidad en cada nivel de aproximación.

Definición 4.5.1 *La intersección de todos los conjuntos convexos que contienen un subconjunto dado S de R^m es llamado el casco convexo de S y se denota por $\text{conv } S$.*

Teorema 4.5.2 Teorema de Carathéodory. *Sea S cualquier conjunto de puntos en R^m , y sea $C = \text{conv } S$. Entonces $y \in C$ si y sólo si y puede ser expresada como una combinación convexa de $m + 1$ (no necesariamente distintos) puntos en S . De hecho C es la unión de todos los d -dimensionales símlices generalizados cuyos vértices pertenecen a S , en donde $d = \dim C$.*

Corolario 4.5.3 *Sea $\{C_i \mid i \in I\}$ una colección arbitraria de conjuntos convexos en R^m , y sea C el casco convexo de la unión de la colección. Entonces cada punto de C puede ser expresado como una combinación convexa de $m+1$ o menos puntos afinamente independientes, cada uno perteneciendo a alguno de los C_i .*

Teorema 4.5.4 *La intersección de una colección arbitraria de conjuntos convexos es convexa.*

El casco convexo $\text{conv } S$ es un conjunto convexo por el Teorema 4.5.4, el único más pequeño que contiene S .

Los detalles y las demostraciones de los Teoremas 4.5.2, 4.5.4, y del Corolario 4.5.3 pueden ser encontrados en [43].

Ahora, describimos el algoritmo propuesto como sigue:

Step 0. Inicialización. Sea la lista inicial de problemas incluyendo sólomente la Aproximación del Problema Entero (APE) construida del siguiente modo:

Consideramos el problema (4.6):

$$\min_{x,y} \{ \langle a, x \rangle + \langle b, y \rangle : Gy = d, \langle c, x \rangle \leq \varphi(y), Ax = y, x \geq 0, y \in Z_+^m \}.$$

Ahora, consideramos el politopo Y compuesto como un casco convexo de las estrategias del líder en el nivel superior: $Y = \{y \mid Gy = d, y \geq 0\}$, y seleccionamos $\hat{m} + 1$ puntos afinamente independientes y^i tales que $Y \subset \text{conv} \{y^1, \dots, y^{\hat{m}+1}\} \subset \{y : |\varphi(y)| < \infty\}$. Aquí $\hat{m} = m - \text{rank}(G)$, y $y^2 - y^1, y^3 - y^1, \dots, y^{\hat{m}+1} - y^1$ forman un sistema linealmente independiente. Denotamos este conjunto de vértices como $V = \{y^1, \dots, y^{\hat{m}+1}\}$. También consideramos un valor de tolerancia $\epsilon > 0$. Entonces, resolvemos el problema de programación lineal del nivel inferior (4.5) en cada vértice, i.e., encontrar $\varphi(y^1), \dots, \varphi(y^{\hat{m}+1})$.

Ahora bien, construimos la primer aproximación de la función de valor óptimo como sigue:

$$\Phi(y) = \sum_{i=1}^{\hat{m}+1} \lambda_i \varphi(y^i), \quad (4.18)$$

definido sobre

$$y = \sum_{i=1}^{\hat{m}+1} \lambda_i y^i, \quad (4.19)$$

con $\lambda_i \geq 0$ $i = 1, \dots, \hat{m} + 1$ y

$$\sum_{i=1}^{\hat{m}+1} \lambda_i = 1. \quad (4.20)$$

En (4.18) tenemos una expresión con la variable λ , que nos conduce a la variable y usando (4.19) y (4.20). Ahora, ya que la función φ es convexa,

$$\langle c, x \rangle \leq \varphi(y) \leq \Phi(y),$$

la condición $\langle c, x \rangle \leq \varphi(y)$ puede ser reemplazada en (4.6) con la siguiente desigualdad explícita:

$$\langle c, x \rangle \leq \Phi(y).$$

Con esto obtenemos un nuevo problema de optimización que puede ser resuelto con el método símplex clásico. La Aproximación del Problema Entero (APE) es descrita como sigue:

$$\min_{x,y} \{ \langle a, x \rangle + \langle b, y \rangle : Gy = d, \langle c, x \rangle \leq \Phi(y), Ax = y, x \geq 0, y \in Z_+^m \}, \quad (4.21)$$

sea $t = 1$, y $z_t = +\infty$, en donde z_t es el valor objetivo incumbente. Ponemos éste problema dentro de la lista. Por definición, éste problema corresponde al poliedro convexo Y . Ir al Paso 1.

Paso 1. *Terminación.* Detener el algoritmo si la lista de problemas es vacía. De otra forma seleccionar arbitrariamente y remover un problema de la lista. Ir al paso 2.

Paso 2. Resolver el problema tomado de la lista usando métodos comunes para programación entera como branch and bound para resolver la restricción

de integralidad. Denotamos al conjunto de las soluciones optimas básicas como $S = \{(x^{1R}; y^{1R}), \dots\}$ y z^R como la función de valor objetivo. Si el problema no tiene soluciones factibles, o si esta función de valor óptimo es mayor que z_t , entonces cortamos esta rama. Asignar $z_{t+1} = z_t$, $t = t + 1$ y vamos al Paso 1. De otra forma vamos al Paso 3.

Paso 3. Si las componentes y de todas las soluciones pertenecientes a S son vértices de V , entonces guardamos las soluciones, asignamos $z_{t+1} = z^R$, $t = t + 1$ y vamos al Paso 1 (para tales valores de y , el punto (x, y) es factible para el problema (4.6)). De otra manera, considerando la solución $(x^{jR}; y^{jR})$ de S tal que la componente y^{jR} es diferente de todos los vértices de V , añadimos y^{jR} a V , asignamos $z_{i+1} = z_i$, $t = t + 1$ y vamos al Paso 4.

Step 4. Subdivisión. Hacer una subdivisión del conjunto Y correspondiente a éste problema. Por construcción, el problema (4.21) corresponde a un conjunto de $\hat{m} + 1$ puntos afinamente independientes, los cuales sin pérdida de generalidad asumimos sean los puntos $y^1, \dots, y^{\hat{m}+1}$. Añadiendo el punto y^{jR} a éste conjunto, estos vienen a ser afinamente independientes. Excluyendo un elemento del conjunto resultante, la independencia afina puede ser obtenida eventualmente (esto está garantizado si algún elemento correcto es eliminado). Cuando uno hace uso de esta aproximación, a lo más $\hat{m} + 1$ nuevos conjuntos afinamente independientes surgen, cada uno correspondiendo a nuevas aproximaciones lineales de la función objetivo del nivel inferior sobre el casco convexo de esos puntos. Si un tal símplex T es un subconjunto de alguna región de estabilidad: $T \subset \mathfrak{R}(B_i)$, los puntos factibles (x, y) del problema

(4.21) son también factibles para el problema (4.6). El objetivo de éste paso es encontrar esas simples subdivisiones subsecuentes del conjunto Y . Estos problemas son entonces añadidos a la lista de problemas.

Para calcular la nueva aproximación de la función de valor óptimo del nivel inferior procederemos como sigue: Primero calcular $\varphi(y^{jR})$. Entonces construir un conjunto de puntos afinamente independientes como se describió anteriormente, i.e., eliminar uno de los puntos, digamos y^l , y calcular

$$\Phi_l(y) = \sum_{i=1, i \neq l}^{\widehat{m}+1} \lambda_i \varphi(y^i) + \mu \varphi(y^{jR}),$$

definido sobre

$$y = \sum_{i=1, i \neq l}^{\widehat{m}+1} \lambda_i y^i + \mu y^{jR}, \quad (4.22)$$

con $\lambda_i \geq 0$ $i = 1, \dots, \widehat{m} + 1$ y

$$\sum_{i=1, i \neq l}^{\widehat{m}+1} \lambda_i + \mu = 1, \quad (4.23)$$

para $l \in \{1, \dots, \widehat{m} + 1\}$. Así construimos a lo más $\widehat{m} + 1$ nuevos problemas:

$$(P^l) \quad \min_{x, y} \{ \langle a, x \rangle + \langle b, y \rangle : Gy = d, \langle c, x \rangle \leq \Phi_l(y), Ax = y, x \geq 0, y \in Z_+^m \},$$

y los añadimos a la lista de problemas. Vamos al Paso 1.

Como un criterio de paro, usamos la condición:

$$\| (x^s, y^s) - (x^{s-1}, y^{s-1}) \| < \epsilon;$$

si es válida, seleccionamos el punto $(x^s; y^s)$ como una solución aproximada del problema de programación entera (4.6).

4.6. Un ejemplo numérico

Consideremos el siguiente problema de programación binivel lineal paramétrico, en donde el nivel superior está planteado como sigue:

$$\min_{x_1, x_2, x_3, y_1, y_2} \{3x_1 + 2x_2 + 6x_3 + 2y_1 \mid 4y_1 + y_2 = 10, x \in \psi(y), y_1, y_2 \in Z_+\},$$

en donde

$$\begin{aligned} \psi(y_1, y_2) = \operatorname{Argmin}_{x_1, x_2, x_3} \{ & -5x_1 - 8x_2 - x_3 \mid 4x_1 + 2x_2 \leq y_1, 2x_1 + 4x_2 + x_3 \leq \\ & y_2, x_1, x_2, x_3 \geq 0\}, \end{aligned}$$

y el problema de nivel inferior se presenta como sigue:

$$\begin{aligned} \varphi(y_1, y_2) = \min_{x_1, x_2, x_3} \{ & -5x_1 - 8x_2 - x_3 \mid 4x_1 + 2x_2 \leq y_1, 2x_1 + 4x_2 + x_3 \leq \\ & y_2, x_1, x_2, x_3 \geq 0\}. \end{aligned}$$

La solución óptima al problema es $(x_1^*, x_2^*, x_3^*; y_1^*, y_2^*) = (1/3, 1/3, 0; 2, 2)$, conociendo esto iniciamos a resolverlo con el algoritmo propuesto de la siguiente manera:

Paso 0. Escogemos los vértices $y^1 = (5/2, 0)$ y $y^2 = (0, 10)$ que pertenecen al casco convexo de las estrategias del líder en el nivel superior. Fijamos como valor de tolerancia $\epsilon = 0.1$. Ahora, calculamos $\varphi(y^1) = 0$ y $\varphi(y^2) = -10$, sea $z_1 = +\infty$, de esta manera podemos construir la primer aproximación de la función de valor objetivo como sigue:

$$\Phi(y) = -y_2.$$

El Problema de Aproximación Entero (APE) que se agrega a la lista de problemas se presenta como sigue:

$$\begin{aligned} \min_{x,y} \{ & 3x_1 + 2x_2 + 6x_3 + 2y_1 \mid 4y_1 + y_2 = 10, 4x_1 + 2x_2 \leq y_1 \\ & 2x_1 + 4x_2 + x_3 \leq y_2, -5x_1 - 8x_2 - x_3 \leq -y_2 \\ & x_1, x_2, x_3 \geq 0, y_1, y_2 \in Z_+ \} \end{aligned}$$

Paso 1. Seleccionamos (AIP) de la lista de problemas.

Paso 2. Resolvemos (AIP) obteniendo por solución $(x_1^{1R}, x_2^{1R}, x_3^{1R}; y_1^{1R}, y_2^{1R}) = (0, 1/4, 0; 2, 2)$ con $z^R = 15/4$. Ya que z^R es menor que $+\infty$, vamos al Paso 3.

Paso 3. Debido a que $(y_1^{1R}, y_2^{1R}) = (2, 2)$ es diferente a los vértices en V , añadimos $(y_1^{1R}, y_2^{1R}) = (2, 2)$ a V , aceptamos $z_2 = +\infty$, $t = 2$ y vamos al Paso 4.

Paso 4. Hacer una subdivisión en $(y_1^{1R}, y_2^{1R}) = (2, 2)$, resultando 2 nuevos problemas, los cuales uno corresponde a $\text{conv} \{y^2 = (0, 10), (y_1^{1R}, y_2^{1R}) = (2, 2)\}$ y el segundo correspondiente a $\text{conv} \{(y_1^{1R}, y_2^{1R}) = (2, 2), y^1 = (5/2, 0)\}$, añadimos estos 2 nuevos problemas a la lista de problemas, cada uno descrito como sigue: el primero descrito con la aproximación

$$\Phi_1(y) = -17y_2/24 - 70/24,$$

y el segundo con la aproximación

$$\Phi_2(y) = -13y_2/6.$$

Entonces estos problemas son presentados como sigue:

$$(P^1) \quad \min_{x_1, x_2, x_3, y_1, y_2} \{ 3x_1 + 2x_2 + 6x_3 + 2y_1 \mid 4y_1 + y_2 = 10, 4x_1 + 2x_2 \leq$$

$$\leq y_1, 2x_1 + 4x_2 + x_3 \leq y_2, -5x_1 - 8x_2 - x_3 \leq \Phi_1(y), x_1, x_2, x_3 \geq 0, y_1, y_2 \in Z_+\}.$$

quitando y^1 de V , y

$$(P^2) \quad \min_{x_1, x_2, x_3, y_1, y_2} \{3x_1 + 2x_2 + 6x_3 + 2y_1 \mid 4y_1 + y_2 = 10, 4x_1 + 2x_2 \leq$$

$$\leq y_1, 2x_1 + 4x_2 + x_3 \leq y_2, -5x_1 - 8x_2 - x_3 \leq \Phi_2(y), x_1, x_2, x_3 \geq 0, y_1, y_2 \in Z_+\}.$$

quitando y^2 de V . Ir al paso 1.

Paso 1. Seleccionamos (P^1) de la lista de problemas, y vamos al Paso 2.

Paso 2. Resolvemos (P^1) quedando como solución $(x_1^{1R}, x_2^{1R}, x_3^{1R}; y_1^{1R}, y_2^{1R}) = (1/3, 1/3, 0; 2, 2)$ con $z^R = 17/3$. y ya que z^R es menor que z_2 , entonces vamos al Paso 3.

Paso 3. Como $(y_1^{1R}, y_2^{1R}) = (2, 2)$ es igual a uno de los vértices en V , guardamos la solución $(x_1^{1R}, x_2^{1R}, x_3^{1R}; y_1^{1R}, y_2^{1R}) = (0, 1/4, 0; 2, 2)$, seleccionamos $z_3 = 17/3$, $t = 3$, y vamos al Paso 1.

Paso 1. Seleccionamos (P^2) de la lista de problemas, y vamos al Paso 2.

Paso 2. Resolviendo (P^2) obtenemos la solución $(x_1^{1R}, x_2^{1R}, x_3^{1R}; y_1^{1R}, y_2^{1R}) = (1/3, 1/3, 0; 2, 2)$ con $z^R = 17/3$. Y como z^R es igual a z_3 , entonces vamos al Paso 3.

Paso 3. Ya que $(y_1^{1R}, y_2^{1R}) = (2, 2)$ es igual a uno de los vértices en V , guardamos la solución $(x_1^{1R}, x_2^{1R}, x_3^{1R}; y_1^{1R}, y_2^{1R}) = (1/3, 1/3, 0; 2, 2)$, actualizamos $z_4 = 17/3$, $t = 4$, y vamos al Paso 1.

Paso 1. La lista de problemas está vacía, aquí termina el algoritmo.

Concluimos que la última solución guardada $(x_1, x_2, x_3; y_1, y_2) = (1/3, 1/3,$

$0; 2, 2)$ con $z = 17/3$ es la solución obtenida mediante éste algoritmo que coincide con la ya conocida solución óptima del problema.

Capítulo 5

Conclusión

En éste trabajo presentamos dos aplicaciones de la vida real en las cuales aplicamos los métodos de solución aquí presentados. El primero un caso de lubricación elasto-hidrodinámica el cual se formuló como un problema implícito de complementariedad no lineal. El segundo un problema de transporte de gas natural el cual fue resuelto utilizando programación binivel.

En Capítulo 2 se propuso un algoritmo de Newton inexacto para resolver el Problema Implícito de Complementariedad No Lineal. Primeramente bajo ciertas suposiciones acerca de la función f reducimos el problema implícito de complementariedad al problema estándar, mostrando la convergencia de los puntos generados por el método de Newton a la solución única del problema original. Además, se construyó una secuencia de puntos por el método de Newton inexacto y mostramos su convergencia a la solución del problema implícito de complementariedad no lineal. Alrededor de todo esto, se establecieron condiciones suficientes de existencia y unicidad de las soluciones

de los problemas lineales auxiliares, cuya tasa de convergencia también fue estimada.

En Capítulo 3 se aplicó un algoritmo de perturbación para resolver el problema implícito de complementariedad. En principio mostramos la existencia y unicidad de la solución para problemas perturbados. Después, considerando una secuencia de perturbaciones tendientes a cero se mostró que generando una secuencia de soluciones para cada perturbación, éstas tienden a la solución del problema original cuya tasa de convergencia también fue establecida. Finalmente se justificó un algoritmo para la solución al problema.

En Capítulo 4 se expuso el problema de programación binivel entero mixto con un parámetro en el lado derecho del nivel inferior. Se mostraron algunas características importantes del problema que ayudaron posteriormente a describir un algoritmo de aproximación el cual, basado en el método branch and bound nos proporciona la solución al problema.

Los Capítulos 2 y 3 funcionan como herramientas para resolver los problemas planteados en el 3er Capítulo, porque problemas de dos niveles pueden ser reformulados como problemas de un sólo nivel con restricciones de complementariedad. Entonces, métodos de solución de problemas de complementariedad serán necesarios para buscar las soluciones de los problemas modificados de sólo un nivel.

Capítulo 6

Trabajo futuro

Las ideas que emergen para los problemas implícitos de complementariedad sugieren atacar más variedad de problemas de éste tipo, con funciones que presenten distintas características a las presentadas en éste trabajo y de esta manera crear métodos de solución más generales.

Para los problemas de programación binivel se presenta la idea de resolver la función de penalización exacta utilizando cálculo subdiferencial, esto podría aplicarse dado que la función antes mencionada es no diferenciable. La convexificación es también una opción que podría implementarse pero la desventaja que se presenta es que teóricamente podría funcionar pero al llevarlo a la práctica surgirían problemas con algunas funciones difíciles de resolver.

Una de las ideas más ambiciosas es el incursionar a los problemas MPEC (por sus siglas en inglés Mathematical Programs with Equilibrium Constraints) que tratan problemas de optimización con restricciones de comple-

mentariedad. Dándole nuevos enfoques y aplicando técnicas usadas en éste trabajo.

Bibliografía

- [1] Ahmad K., Kazmi K. R. and Rehman N. Fixed-Point Technique for Implicit Complementarity Problem in Hilbert Lattice. Springer Netherlands. Vol. 93, No. 1, pp. 67-72, APRIL 1997.
- [2] Bard, J. (1983). An algorithm for solving the general bilevel programming problem, *Mathematics of Operations Research*, Vol. 8, No. 2, pp. 260 – 282.
- [3] Bensoussan, A., and J.-L. Lions. Nouvelle formulation de problemes de controle impulsional et applications. C.R.Acad.Sci. 1973. Paris. Ser. A-B **276**, 1189-1192.
- [4] Bensoussan, A., M. Gourset, and J.-L. Lions., (1973). Controle impulsional et inequations quasi-variationnelles stationaires. C.R.Acad.Sci. 1973. Paris. Ser. A-B **276**, 1279-1284.
- [5] Berman A., Plemmons R. L. Nonnegative matrices in the mathematical sciences. - N.Y. - San Francisco - London: Acad. Press, 1979.

- [6] Bulavsky, V.A., Isac G. and V.V. Kalashnikov. Application of topological degree theory to complementarity problems. - In: A. Migdalas, P.M. Pardalos, and P. Varbrand (eds.). *Multinivel Optimization: Algorithms and Applications*, Dordrecht-Boston-London: Kluwer Academic Publishers, 1998. -P. 333-358.
- [7] Bulavsky, V.A., G. Isac, and V.V. Kalashnikov. Application of topological degree theory to semidefinite complementarity problems. *Optimization*, 2001, **49**, 405-423.
- [8] Capuzzo-Dolcetta, J. and U. Mosco. Implicit complementarity problems and quasi-variational inequalities. -In: R.W. Cottle, F. Giannessi, and J.-L. Lions (eds.) *Variational Inequalities and Complementarity Problems: Theory and Applications*, New York-London: John Wiley and Sons, 1980.
- [9] Dempe, S. (1987). A simple algorithm for the linear bilevel programming problem, *Optimization*, Vol. 18, No. 3, pp. 373 – 385.
- [10] Dempe, S. (2002). *Foundations of Bilevel Programming*, Dordrecht/London/Boston: Kluwer Academic Publishers.
- [11] Dempe, S. and Kalashnikov, V. (2005). *Discrete Bilevel Programming with Linear Lower Level Problems*, Preprint, TU Bergakademie Freiberg.
- [12] Dempe, S., Kalashnikov, V. and Ríos-Mercado, R.Z. (2005). Discrete bilevel programming: Application to a natural gas cash-out problem.

- European Journal of Operational Research*, Vol. 166, No. 2, pp. 469 – 488.
- [13] Dempe, S. and Schreier, H. (2006). *Operations Research – Deterministische Modelle und Methoden*, Teubner Verlag, Wiesbaden.
- [14] Dempe, S. and Zemkoho, A.B. (2008). *A Bilevel Approach to Optimal Toll Setting in Capacitated Networks*, Preprint, TU Bergakademie Freiberg, 2008.
- [15] Gao, D.Y. (2004). Canonical duality theory and solutions to constrained nonconvex quadratic programming, *Journal of Global Optimization*, Vol. 29, pp. 377 – 399.
- [16] Gao, D.Y. (2007). Solutions and optimality criteria to box constrained nonconvex minimization problems, *Journal of Industry and Management Optimization*, Vol. 3, No. 2, pp. 293 – 304.
- [17] Grygarová, L. (1970). Qualitative Untersuchung des I. Optimierungsproblems in mehrparametrischer Programmierung, *Applications of Mathematics*, Vol. 15, No. 4, pp. 276 – 295.
- [18] Isac, G. A special variational inequality and the implicit complementarity problem. J. Fac. Sci. Univ. Tokyo, Sect.IA, Math., 1990, **37**, 109-127.
- [19] Isac, G. Complementarity Problems. Lecture Notes in Mathematics, 1992, Berlin-Heidelberg: Springer Verlag, Vol. 1528.

- [20] Isac G., *Leray-Schauder Type Alternatives, Complementarity Problems and Variational Inequalities*, Springer Publisher, New York, 2006.
- [21] Isac, G., V.A. Bulavsky, and V.V. Kalashnikov. Exceptional families, topological degree and complementarity problems. *J. Global Opt.*, 1997, **10**, 2007-225.
- [22] Isac G., Bulavsky V. A. and Kalashnikov V. V., *Complementarity, Equilibrium, Efficiency and Economics*, Kluwer Academic Publishers, Dordrecht/Boston/London, 2002.
- [23] Isac, G., and V.V. Kalashnikov. Exceptional families of elements, Leray Schauder alternatives, pseudomonotone operators and complementarity. *J. Optim. Theory Appl.*, 2001, **109**, 69-83.
- [24] Isac G. and Németh S. Z. Duality of Implicit Complementarity Problems by Using Inversions and Scalar Derivatives. *Journal of Optimization Theory and Applications*. Springer Netherlands. Volume 128, Number 3. 2006.
- [25] Kalashnikov, V.V. Complementarity Problem and Generalized Oligopoly Models, Habilitation Thesis, Central Economics and Mathematics Institute, Moskow, 1995 (in Russian).
- [26] Kalashnikov, V.V., and Isac G. Solvability of implicit complementarity problems. *Annals of Operations Research* **116**, 199-221, 2002.

- [27] Kalashnikov, V.V. and Kalashnikova N.I. Optimizatsia. 42(59). P. 66 - 83. 1988. (in Russia).
- [28] Karmandian S. The nonlinear complementarity problem with applications. Part I. J. Optimiz. Theory and Applic. -1969.- V.4, Number 2.- P. 87-98.
- [29] Kinderlehrer, D., and Stampacchia G. An introduction to Variational Inequalities and Their applications, New York: Academic Press, 1980.
- [30] Kočvara M., Outrata J. V. On optimization of systems governed by implicit complementarity problems. Numerical Functional Analysis and Optimization, Volume 15, Issue 7 and 8, 1994, pages 869 - 887.
- [31] Kostreva, Michael M. Elasto-Hidrodinamic Lubrication: A Non-Linear Complementarity Problem. International Journal for numerical methods in fluids, Vol. 4, 377-397. 1984.
- [32] Kyparisis J. and Chi-Ming Ip. Solution behavior for parametric implicit complementarity problems. Mathematical Programming. Springer Berlin / Heidelberg. Volume 56, Numbers 1-3.agosto 1992.
- [33] Lee, B.S., Firdosh Khan M., and Salahuddin. Vector F-implicit complementarity problems with corresponding variational inequality problems. Appl. Math. Lett., 20, 433-438. 2007.

- [34] McLinden, L. (1980) *The complementarity problem for maximal monotone operator*. -In: R. W. Cottle e. a. (Eds), Variational Inequalities and Complementarity Problems. New York: Academic Press, -P. 251 - 270.
- [35] Mosco, U. *Implicit Variational Problems and Quasi-Variational Inequalities*, Lecture Notes in Math., No. 543, Springer-Verlag, 1976.
- [36] Mosco, U. *On some nonlinear quasi-variational inequalities and implicit complementarity problems in stochastic control theory*. -In: R.W. Cottle, F. Giannessi and J.-L. Lions (eds.), *Variational Inequalities and Complementarity Problems, Theory and Applications*, New York: John Wiley and Sons, 1980.
- [37] Nikaido, H. *Convex Structures and Mathematical Economics*, New York: Acad. Press, 1968.
- [38] Ortega, J., and W. C. Rheinboldt. (1970). *Iterative Solution of Nonlinear Equations in Several Variables*, New York: Academic Press.
- [39] Pang, J.-S. *The implicit complementarity problem*. -In: O.L. Mangasarian, R.R. Meyer, and S.M. Robinson (eds.), *Nonlinear Programming IV*, New York-London: Academic Press, 1981.
- [40] Pang J. S. *On the convergence of a basic iterative method for the implicit complementarity problem*. *Journal of Optimization Theory and Applications*. Springer Netherlands. Volume 37, Number 2. P. 149-162. 1982.

- [41] Pang J.-S. Inexact Newton methods for complementarity problem. *Math. Programming.* 1986. V.36, N I. p. 54-71.
- [42] Pang J.-S., Chan D. Iterative methods for variational and complementarity problems. *Math. Programming.*- 1982. - V. 24, N. 3. - P. 284-313.
- [43] Rockafellar, R.T. (1970). *Convex Analysis*, Princeton, New Jersey: Princeton University Press.
- [44] Siddiqi A. H. and Ansari Q. H. On the nonlinear implicit complementarity problem. *Internat. J. Math. and Math. Sci.* VOL. 16 NO. 4 (1993) 783-790.
- [45] Smith, M.J. A descent algorithm for solving monotone variational inequalities and monotone complementarity problems. *J. Optim. Theory Appl.*, 1984, **44**, 485-496.
- [46] Stackelberg, H.v. (1934). *Marktform und Gleichgewicht*, Vienna, Austria: Julius Springer, 1934. English translation: *The Theory of the Market Economy*, Oxford: Oxford University Press, 1952.
- [47] Wen, U.P. and Yang, Y.H. (1990). Algorithms for solving the mixed integer two level linear programming problem, *Computers and Operations Research*, Vol. 17, No. 2, pp. 133 – 142.
- [48] Wendell, R.E. (1982). A preview of a tolerance approach to sensitivity analysis in linear programming, *Discrete Mathematics*, Vol. 38, pp. 121 — 124.

- [49] Yao Jen-Chih. Generalized Quasi-Variational Inequality and Implicit Complementarity Problems. 1998.
- [50] Ye, J.J. and Zhu, D.L. (1995). Optimality conditions for bilevel programming problems, *Optimization*, Vol. 33, pp. 9 – 27.
- [51] Yeol Je Cho; Jun Li; Huang Nan-Jing. Mathematical and Computer Modelling Volume 45, Issues 7-8, April 2007, pp. 1001-1009.

Publicaciones

Artículos en Revistas Científicas

- STEPHAN DEMPE, VYACHESLAV KALASHNIKOV, NATALIYA KALASHNYKOVA Y AARÓN ARÉVALO FRANCO. A new approach to solving bi-level programming problems with integer upper level variables. - ICIC Express Letters (An International Journal of Research and Surveys, Japón, (ISSN 1881-803X), 2008, Vol. 3, No. 4(B), septiembre de 2009, pp. 1281-1287.
- VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA Y AARÓN ARÉVALO FRANCO. Inexact Newton Method to Solve Nonlinear Implicit Complementarity Problems. Aceptado para publicación en International Journal of Innovative Computing, Information and Control, ICIC International, Japón, ISSN 1349-4198. - 9 pp.

Capítulos y Contribuciones en Libros

- VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Multi-Valued Implicit Complementarity Problems: Perturbation Algorithm. - En: Jeng-Shyang Pan et al. (Eds), Proceedings of the 2nd International Conference on Innovative Computing, Information and Control, Kumamoto, Japan, September 5 - 7, 2007. - 4 pp.

- VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Solving Nonlinear Implicit Complementarity Problems with Inexact Newton Methods. - En: Jeng-Shyang Pan et al. (Eds), Proceedings of the 3rd International Conference on Innovative Computing, Information and Control (ICICIC'2008), Dalian, R.P. China, 18 - 20 de junio del 2008. - 4 pp.
- VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Inexact Newton Algorithm to Solve Nonlinear Complementarity Problems. - En: J.-S. Pan et al. (editors), Proceedings of the 8th International Conference on Intelligent System Design and Applications (ISDA'2008), Kaohsiung, Taiwan, 26 - 28 de noviembre de 2008, pp. 67 - 71. ISBN 978-0-7695-3382-7.

Memorias de Congresos con Revisión

- V.V. KALASHNIKOV, N.I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Solution of Implicit Complementarity Problems with Parameters. - En: R. Clute (editor), Proceedings of the International Business and Economics Research Conference (2006 IBERC & TLC), Las Vegas, NV, EE.UU., 02 - 06 de octubre del 2006. - 6 p. ISSN 1539 - 8757.
- V.V. KALASHNIKOV, N.I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Implicit Complementarity Problems: Solution with a Perturbation Algorithm. - En: H. Cancela y M. Urquhart (editores),

Anales del 13o Congreso Latino-Iberoamericano de Investigación Operativa (CLAIO XIII), Montevideo, Uruguay, 27 - 30 de noviembre del 2006. - 6 p.

- VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Solution of Multi-Valued Implicit Complementarity Problems with Parameters. - En: R. Clute (editor), Proceedings of the International Applied Business Research Conference (ABRC'2007), Mazatlán, México, 26 - 29 de marzo del 2007. - 8 p.
- VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA Y AARÓN ARÉVALO FRANCO. Implicit Complementarity Problem: Perturbation Algorithm. Proceedings of the 7th Annual IBER & TLC Conference, Las Vegas, NV, USA, October 1 - 4, 2007. - 6 pp.
- STEPHAN DEMPE, VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. An Algorithm to Solve Mixed-Integer Bi-level Programming Problems. - En: R. Clute (editor), Proceedings of the International Applied Business Research Conference (ABR & TCL'2009), Oahu, Hawaii, EE.UU., 05 - 09 de enero del 2009. - 12 pp.
- STEPHAN DEMPE, VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Solving Bi-level Programming Problems with Integer Upper Level Variables. - En:

R. Clute (editor), Proceedings of the International Applied Business Research Conference (IABR & TLC'2009), San Antonio, TX, EE.UU., 23 - 26 de marzo del 2009. - 13 pp.

- STEPHAN DEMPE, VYACHESLAV V. KALASHNIKOV, NATALIYA I. KALASHNYKOVA, Y AARÓN ARÉVALO FRANCO. Bi-level Programming Problems with Integer Upper Level Variables: Description and Solution. - En: J. Watada (editor), Proceedings of the 6th International Symposium on Management Engineering (ISME'2009), Dalian, P.R. China, 05 - 07 de agosto del 2009. - 6 pp.
- STEPHAN DEMPE, VYACHESLAV KALASHNIKOV, NATALIYA KALASHNYKOVA Y AARÓN ARÉVALO FRANCO. A new approach to solving bi-level programming problems with integer upper level variables. - Proceedings of the Second International Symposium on Intelligent Informatics (ISII2009), QinHuangDao, China, 13-15 de septiembre, 2009, pp.-6.