UNIVERSIDAD NACIONAL DE INGENIERIA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE MATEMÁTICA

El Método de Newton (Amortiguado) para Desigualdades Variacionales

por

Hermes Yesser Pantoja Carhuavilca

Tesís para Optar el
Título Profesional de:
Licenciado en Matemática

Prof. William Carlos Echegaray Castillo
Asesor

Lima, Noviembre de 2001.

RESUMEN

Este Trabajo presenta un algoritmo de Newton (Amortiguado) para resolver problemas de desigualdad variacional basado sobre la formulación del problema como un sistema de ecuaciones usando la aplicación Minty. El propósito de éste método es asegurar la convergencia y una convergencia cuadrática local bajo la suposición de regularidad. Bajo la suposición de regularidad débil y algunas condiciones mild, el algoritmo modificado demuestra que siempre existe una dirección descendente y converge a la solución

CONTENIDO

1 INTRODUCCION

Existen numerosos métodos localmente convergente para resolver problemas de desigualdad variacional finito–dimensional y problemas de complementariedad no–lineal [**6**]. Los más conocidos son aquellos que están basados en aproximaciones a una ecuación generalizada, siendo el método de Newton el más potente , el cual tiene la propiedad de convergencia cuadrática local bajo ciertas condiciones. La convergencia de este método requiere que uno empiece cerca a la solución. Para superar este incoveniente, muchos métodos iterativos extendidos han sido desarrollados. Estos métodos no garantizan una convergencia cuadrática

A fin de obtener una convergencia de tipo global para desigualdades variacionales se formularon algoritmos que incorporan el método de Newton con una búsqueda lineal de alguna función de mérito. Marcotte y Dussault sugirieron una convergencia global para desigualdades variacionales monótonas. Sin embargo la dimensión de los subproblemas no es reducido en ninguna iteración (es decir: Cada subproblema tiene el mismo número de variables como el problema original). En los ultimos años estos algoritmos están basados en aproximaciones a ecuaciones no–suaves. Robinson fue el primero en estudiar una de las funciones no–suaves llamadas B–Diferenciables [**14**].

Este trabajo presenta un algoritmo modificado de Newton(amortiguado) para resolver problemas de desigualdad variacional basada sobre la formulación de un sistema de ecuaciones B–diferenciable a través del uso de la aplicación Minty. Este algoritmo asegura la convergencia y la convergencia cuadrática local bajo la suposición de regularidad. Esta suposición puede ser debilitado en muchos casos. Introducimos entonces el concepto de regularidad débil y algunas modificaciones del algoritmo basada sobre esta suposición. Así, el método modificado de Newton (amortiguado) puede ser aplicado a un más amplio rango de situaciones.

El pre ente trabajo fue desarrollado tomando como referencia el paper:

" A nonsmooth Newton method for variational inequalities, I: Theory ".

De lo autores: **Baichun Xiao y Patrick Harker.**

Antes de empezar, indicaremos algunas notaciones usadas en éste trabajo:

Si $x \in \mathbb{R}^n$, denotamos $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

Además, si $H : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una función vectorial F-diferenciable, la matriz Jacobiana $\nabla H(x)$ es de orden $n \times m$ y su i-ésima columna es igual al vector gradiente $\nabla H_i(x)$ donde $H_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ es la i-ésima función componente de H. En ésta notación, el Jacobiamo es la transpuesta de la forma usual. La norma que se usa en todo éste trabajo es la norma euclidiana.

2 FORMULACIÓN DE LA ECUACIÓN DEL PROBLEMA DE DESIGUALDAD VARIACIONAL

En é te Capítulo definimos los problemas de desigualdad variacional, complementariedad lineal y no lineal; establecemos sus relaciones. Además presentamos la definición de una función B–diferenciable y varias propiedades importantes que se derivan de dicha función, formulamos los problemas como un sistema de ecuaciones usando la aplicación Minty. Finalizamos éste capítulo introduciendo una función mérito que permitirá desarrollar un algoritmo convergente en la búsqueda de la solución al sistema de ecuaciones.

2.1 Las definiciones del problema y hechos básicos

Definición 2.1.1 (Problema de Desigualdad Variacional.-)
Sea X un subconjunto no vacío de $\mathbb{R}^n$ y $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación. El Problema de Desigualdad Variacional, denotado por VI(X,F), es encontrar un $x^* \in X$ tal que:

$$F(x^*)^T(x - x^*) \geq 0, \qquad \forall x \in X \tag{2.1}$$

Uno tipicamente asume que el conjunto X es cerrado y convexo.

Definición 2.1.2 (Problema de Complementariedad No Lineal.-)
Sea $F : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación. El problema de complementariedad No Lineal, denotado por NCP(F), es encontrar un vector $x^* \in \mathbb{R}^n_+$ tal que:

$$F(x^*)^T x^* = 0, \qquad F(x^*) \in \mathbb{R}^n_+ \tag{2.2}$$

Cuando F es una función afín de x, decimos $F(x) = q + Mx$, donde $q \in \mathbb{R}^n$ es un vector dado y $M \in \mathbb{R}^{n\times n}$ es una matriz dada entonces, el problema NCP(F) se reduce al Problema de Complementariedad Lineal, lo cual denotaremos por LCP(q,M).

Definición 2.1.3 (Problema de Complementariedad Lineal (6).-)
Sea $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación. El Problema de Complementariedad Lineal, denotado como LCP(q,M), es encontrar un vector $x^* \in \mathbb{R}^n_+$ tal que:

$$q + Mx^* \geq 0, \quad x^* > 0 \tag{2.3}$$

$$(q + Mx^*)^T x^* = 0 \tag{2.4}$$

donde: $q \in \mathbb{R}^n$ es un vector dado y $M \in \mathbb{R}^{n\times n}$ es una matriz dada.

Proposición 2.1.1 *Sea $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función tal que $f \in C^1(X)$.*
Supongamos que existe un $x_0 \in X$, donde X es un subconjunto convexo, que satisface: $f(x_0) = Min_{x\in X} f(x)$.
Entonces x_0 es solución de la **Desigualdad Variacional**
$x_0 \in X : F(x_0)^T(x - x_0) \geq 0, \quad \forall x \in X$
donde $F(x) = \nabla f(x)$ para cada $x \in X$.

Demostración: Sea $x \in X$. Debido a que X es convexo, entonces:

$$(1-t)x_0 + tx = x_0 + t(x - x_0) \in X, \quad \forall \quad t \in I$$

donde $I = [0, 1]$.
Sea la función $\Phi : I \longrightarrow \mathbb{R}$ definido por: $\Phi(t) = f(x_0 + t(x - x_0)), \qquad 0 \leq t \leq 1.$
Se cumple que Φ alcanza su mínimo en $t = 0$, en efecto:
$\Phi(0) = f(x_0) = \text{Min}_{x\in X} f(x) = \text{Min}_{t\in I}\Phi(t)$, entonces $\Phi(t) \geq \Phi(0) \quad \forall t \in I$
Luego, $0 \leq \dfrac{\Phi(t) - \Phi(0)}{t} \approx \Phi'(0) \quad \forall t \in (0, 1]$
Entonces: $0 \leq \Phi'(0) = \nabla f(x_0)^T(x - x_0) = F(x_0)^T(x - x_0)$ $\square$

Debido a que f es diferenciable en x_0, el limite existe y es finito:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

Entonces uno de los tres casos puede ocurrir:

1. Si $a < x_0 < b$
   $f'_-(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$
   $f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$
   Como $f'_-(x_0) = f'_+(x_0) = f'(x_0)$
   Tenemos: $f'(x_0) = 0 \Longrightarrow f'(x_0)(x - x_0) \geq 0$

2. Si $x_0 = a$
   Como a es punto mínimo de f entonces existe una vecindad $V_\delta(a)$ para algún $\delta > 0$ tal que f está definido en $I \cap V_\delta(a)$ y tal que:
   $f(a) \leq f(x) \quad \forall x \in I \cap V_\delta(a)$
   $\Longrightarrow \forall h \in <0, \delta> : \ f(a + h) \geq f(a), \quad \frac{f(a+h) - f(a)}{h} \geq 0$
   Y por lo tanto $f'_+(a) = \lim_{h \to 0^+} \frac{f(a+h) - f(a)}{h} \geq 0$
   Entonces: $f'(x_0)(x - x_0) \geq 0$

3. Si $x_0 = b$
   Como b es punto mínimo de f entonces una vecindad $V_\delta(b)$ para algún $\delta > 0$ tal que f está definido en $I \cap V_\delta(b)$ y tal que:
   $f(b) \leq f(x) \quad \forall x \in I \cup V_\delta(b)$
   $\Longrightarrow \forall h \in <-\delta, 0> : \ f(b + h) \geq f(b), \quad \frac{f(b+h) - f(b)}{h} \leq 0$
   Y por lo tanto $f'_-(b) = \lim_{h \to 0^-} \frac{f(b+h) - f(b)}{h} \leq 0 \Longrightarrow f'(x_0)(x - x_0) \geq 0$

De $(1), (2), (3)$ tenemos:

$f'(x_0)(x - x_0) \geq 0, \qquad \forall x \in I.$ [ ]

La relación entre el Problema de Desigualdad Variacional VI(X,F) y un Problema de Optimización fracasa si F no es una aplicación gradiente de la función $f : \mathbb{R}^n \to \mathbb{R}$. De acuerdo al Principio de Simétria (Ver anexo), la simétria de ∇F proporciona una propiedad clave bajo lo cual el Problema de Desigualdad Variacional VI(X,F) es relacionado al programa matemático diferencial (2.5).

2.2 Formulación del Problema

Daremos algunos resultados previos que nos permitiran llegar a formular el problema de Desigualdad Variacional VI(X,F).

Teorema 2.2.1 (2) *Sean las funciones:*
$g_i(x) : \mathbb{R}^n \longrightarrow \mathbb{R}$ $i = 1, \ldots, m$; $h_j(x) : \mathbb{R}^n \longrightarrow \mathbb{R}$ $j = 1, \ldots, p$; *y el conjunto:*
$S = \{x \in X \subset \mathbb{R}^n : \; g_i(x) \leq 0, \; \forall i = 1, \ldots, m \; ; \; h_j(x) = 0, \; \forall j = 1, \ldots, p\} \subset \mathbb{R}^n$, *y sea* $x^* \in X$. *Además, suponiendo que:* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *es diferenciable en* x^*. *Si* x^* *es una solución local del problema Min* $f(x)$ *sujeto a* $x \in S$. *Entonces* $F_0 \cap T = \emptyset$ *donde:*
$F_0 = \{d \in \mathbb{R}^n : \nabla f(x^*)^T d < 0\} \quad y$
$T = \{d \in \mathbb{R}^n : d = \lim_{k \to \infty} \lambda_k (x_k - x^*), \quad \lambda_k > 0, \quad x_k \in \mathrm{S} \quad \forall k \in \mathbb{N} \quad y \quad x_k \to x^*\}$

Demostración:
Sea $d \in T$, por la diferenciabilidad de f en x^*, existe una vecindad $V(x^*, \epsilon)$, para algún $\epsilon > 0 / \; f(x_k) - f(x^*) = \nabla f(x^*)^T (x_k - x^*) + \| x_k - x^* \| \alpha(x^*; x_k - x^*) \ldots (*)$ donde $\alpha(x^*; x_k - x^*) \to 0$ cuando $x_k \to x^*$. Debido a que x^* es un punto óptimo local, para k suficientemente grande tenemos:$f(x_k) \geq f(x^*)$, por lo tanto de (*)
$\nabla f(x^*)^T (x_k - x^*) + \| x_k - x^* \| \alpha(x^*; x_k - x^*) \geq 0$
Multiplicando por $\lambda_k > 0$ y tomando el limite cuando $k \to \infty$, la desigualdad anterior implica que $\nabla f(x^*)^T d \geq 0$. Por lo tanto, demostramos que $d \in T$ implica que $\nabla f(x^*)^T d \geq 0$ entonces $d \notin F_0$ por lo tanto : $F_0 \cap T = \emptyset$ $\square$

Consideremo el iguiente problema $(\mathbf{P})$

$$\begin{array}{rl} Min & f(x) \\ \text{Sujeto} & \text{a:} \\ x & \in S \end{array}$$

Por el teorema 2.2.1, una condición de optimalidad necesaria en un mínimo local x es $F_0 \cap T \quad \emptyset$.

Ahora imponiendo una **Calificación de Restricción** $T = G' \cap H_0$ (ver lema A-3.1.2 del Anexo), donde:

$G' = \{d \in \mathbb{R}^n \ : \nabla g_i(\bar{x})^T d \leq 0 \quad \text{para cada} \quad i \in I\}$

donde $I = \{i = 1, \ldots, m : g_i(\bar{x}) = 0\}$

$H_0 = \{d \in \mathbb{R}^n \ : \nabla h_i(\bar{x})^T d = 0 \quad \text{para cada} \quad i = 1, \ldots, p\}$.

Luego usando el Teorema de Farkas(Ver Anexo), obtenemos las condiciones KKT. El teorema 2.2.2 demuestra lo mencionado.

Teorema 2.2.2 (Condiciones de Karush-Kuhn-Tucker (2).-)
Sean las funciones: $f(x) : \mathbb{R}^n \longrightarrow \mathbb{R}$; $\quad g_i(x) : \mathbb{R}^n \longrightarrow \mathbb{R} \quad i = 1, \ldots, m$; $h_j(x) : \mathbb{R}^n \longrightarrow \mathbb{R} \quad j = 1, \ldots, p$; *además definamos el conjunto:*
$S = \{x \in X \subset \mathbb{R}^n : \ g_i(x) \leq 0, \ \forall i = 1, \ldots, m \ ; \ h_j(x) = 0, \ \forall j = 1, \ldots, p\} \subset \mathbb{R}^n$
Consideremos el problema $(\mathbf{P})$. *Sea* $\bar{x}$ *solución local del problema. Supongamos que las funciones* f, g_i , $i = 1, \ldots, m$, h_j , $j = 1, \ldots, p$ *son diferenciables en* $\bar{x}$. *Supongamos que se satisface la calificación de restricción* $T = G' \cap H_0$. *Entonces,* $\bar{x}$ *es un punto KKT, es decir, existen escalares* $y_i \geq 0$ *para cada* $i = 1, \ldots, m$ *y* v_j *para cada* $j = 1, \ldots, p$ *tal que:*

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^{m} y_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} v_j \nabla h_j(\bar{x}) &= 0 \\ y_i g_i(\bar{x}) = 0 \qquad & i = 1, \ldots, m \\ y_i \geq 0 \qquad & i = 1, \ldots, m \\ h_j(\bar{x}) = 0 \qquad & j = 1, \ldots, p \end{aligned}$$

Demostración:

Debido a que $\bar{x}$ es solución factible del problema, entonces por el teorema (2.2.1), obtenemos $F_0 \cap T = \emptyset$. Además, debido a que la calificación de restricción se satisface en $\bar{x}$ tenemos : $F_0 \cap\ G' \cap H_0 = \emptyset$, es decir, el sistema $\mathbf{A}d \leq \mathbf{0}$ y $\mathbf{c}^T d > 0$ no tiene solución, donde las filas de $\mathbf{A}$ están dados por $\nabla g_i(\bar{x})^T$ para $i \in I, \nabla h_i(\bar{x})^T$ y $-\nabla h_i(\bar{x})^T$ para $i = 1 \dots p$, y $c = -\nabla f(\bar{x})$. Por el Teorema de Farkas el sistema $\mathbf{A^T}y = \mathbf{c}$ y $y \geq 0$ tiene una solución. Esto implica que existe escalares no negativos y_i para cada $i \in I$ y α_j, β_j para cada $j = 1 \dots p$ tal que:

$$\nabla f(\bar{x}) + \sum_{i \in I} y_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} \alpha_j \nabla h_j(\bar{x}) - \sum_{j=1}^{p} \beta_j \nabla h_j(\bar{x}) = 0$$

Haciendo $v_j = \alpha_j - \beta_j$ para cada j tenemos:

$$\nabla f(\bar{x}) + \sum_{i \in I} y_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} v_j \nabla h_j(\bar{x}) = 0$$

Fijando $y_i = 0$ para $i \notin I(\bar{x})$
tenemos finalmente:

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^{m} y_i \nabla g_i(\bar{x}) + \sum_{j=1}^{p} v_j \nabla h_j(\bar{x}) &= 0 \\ y_i g_i(\bar{x}) = 0 & \qquad i = 1, \dots, m \\ y_i \geq 0 & \qquad i = 1, \dots, m \\ h_j(\bar{x}) = 0 & \qquad j = 1, \dots, p \end{aligned}$$

$\square$

De acuerdo a la Proposición 2.1.1 y el Teorema 2.2.2 y teniendo en cuenta que $F = \nabla f$, tenemos:

Para el Problema de De igualdad Variacional VI(X,F) donde:

$$X = \{x \in \mathbb{R}^n / g_i(x) \leq 0,\ \forall i = 1, \ldots, m;\ h_j(x) = 0,\ \forall j = 1, \ldots, p\}$$

F, g, h, son continuamente diferenciable, si $\mathbf{x}^*$ es solución del problema y si una calificación de restricción es satisfecha en $\mathbf{x}^*$ entonces existen vectores $\mathbf{y}^* \in \mathbb{R}^m$ y $\mathbf{v}^* \in \mathbb{R}^p$ tal que las siguientes condiciones de complementariedad mixtas son satisfecha en $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{v}^*)$

$$F(\mathbf{x}) + \nabla g(\mathbf{x})\mathbf{y} + \nabla h(\mathbf{x})\mathbf{v} = 0 \tag{2.6}$$

$$\mathbf{y} \geq 0, \qquad g(\mathbf{x}) \leq 0, \qquad g(\mathbf{x})^T\mathbf{y} = 0 \tag{2.7}$$

$$h(\mathbf{x}) = 0 \tag{2.8}$$

Para escribir $(2.6) - (2.8)$ como un sistema de ecuaciones, definimos:

$$u_i^+ = \max\{u_i, 0\}, \quad u_i^- = \min\{u_i, 0\}, \qquad i = 1 \ldots m;$$

$$\mathbf{u}^+ = (\mathbf{u_1^+}, \ldots, \mathbf{u_m^+})^T, \quad \mathbf{u}^- = (\mathbf{u_1^-}, \ldots, \mathbf{u_m^-})^T$$

Debido a que $\mathbf{y} \geq 0$, sea $\mathbf{u}^+ = \mathbf{y}$. Debido a que $g(\mathbf{x}) \leq 0$ y $g(\mathbf{x})^T\mathbf{y} = 0$, podemos escribir que $\mathbf{u}^- = g(\mathbf{x})$. Así podemos verificar que $(2.6) - (2.8)$ es equivalente al siguiente sistema de ecuaciones no lineales:

$$F(\mathbf{x}) + \nabla g(\mathbf{x})\mathbf{u}^+ + \nabla h(\mathbf{x})\mathbf{v} = 0 \tag{2.9}$$

$$g(\mathbf{x}) - \mathbf{u}^- = 0 \tag{2.10}$$

$$h(\mathbf{x}) = 0 \tag{2.11}$$

Proposición 2.2.1 *Si $x_0 \in \mathbb{R}_+^n$ tal que $F(x_0)^T.(y - x_0) \geq 0 \quad \forall y \in \mathbb{R}_+^n$ es solución del Problema de Desigualdad Variacional $VI(\mathbb{R}_+^n, F)$ entonces el punto $x_0 \in \mathbb{R}_+^n$ es una solución al Problema de Complementariedad No Lineal (NCP).*

Demostración:

Supongamos que $x_0 \in \mathbb{R}^n_+$ es solución al Problema de Desigualdad Variacional, entonces: $y = x_0 + e_i, \quad e_i = (0, \dots, 1, \dots, 0)$, (1 en el i-ésimo lugar), es un elemento de $\mathbb{R}^n_+$, por lo tanto:

$0 \leq F(x_0)^T(x_0 + e_i - x_0) = F(x_0)^T.e_i = F_i(x_0) \quad$ ó $\quad F(x_0) \in \mathbb{R}^n_+$

De aquí, debido a que $y = 0 \in \mathbb{R}^n_+$, tenemos:

$0 \leq F(x_0)^T(-x_0) \Longrightarrow F(x_0)^T x_0 \leq 0$

Además $x_0, F(x_0) \in \mathbb{R}^n_+$ implica que $F(x_0)^T x_0 \geq 0$

Por lo tanto: $F(x_0)^T x_0 = 0$. Es decir x_0 es solución de NCP(F). $\square$

Debido a que $g(x) \leq 0$ y $x \in \mathbb{R}^n_+$, fijamos $g(\mathbf{x}) = -\mathbf{x}$, Luego de acuerdo a la proposición (2.2.1) el Problema de Desigualdad Variacional VI($\mathbb{R}^n_+$,F) se convierte en el Problema de Complementariedad No Lineal NCP.

Teniendo en cuenta ahora que el conjunto $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\}$, el sistema de ecuaciones $(2.6) - (2.8)$ se convierten en:

$$F(\mathbf{x}) + \nabla g(\mathbf{x})\mathbf{u}^+ = 0 \tag{2.12}$$

$$g(\mathbf{x}) - \mathbf{u}^- = 0 \tag{2.13}$$

En el contexto del Problema de Complementariedad No-lineal donde $g(\mathbf{x}) = -\mathbf{x}$ el sistema de ecuaciones $(2.12) - (2.13)$ se convierten en:

$$F(\mathbf{x}) - \mathbf{u}^+ = 0 \tag{2.14}$$

$$\mathbf{x} + \mathbf{u}^- = 0 \tag{2.15}$$

Ahora, teniendo en cuenta que: $\mathbf{u}^+ = -(-\mathbf{u})^-, \qquad \mathbf{u}^- = -(-\mathbf{u})^+$, reemplazamos en (2.14) y (2.15) respectivamente:

$$F(\mathbf{x}) + (-\mathbf{u})^- = 0 \tag{2.16}$$

$$\mathbf{x} - (-\mathbf{u})^+ = 0 \tag{2.17}$$

Luego tenemos $\mathbf{x} = -\mathbf{u}$, en efecto:

Debido a que $g(\mathbf{x}) = -\mathbf{x} \leq \mathbf{0}$, implica que $\mathbf{x} \geq \mathbf{0}$. Luego de (2.15): $\mathbf{x} = -\mathbf{u}^{-} \geq \mathbf{0}$, implica que $\mathbf{u}^{-} \leq 0$ es decir $\mathbf{u}^{-} = \mathbf{u}$ con lo cual $\mathbf{x} = -\mathbf{u}$.

De (2.17), tenemos: $\mathbf{x} = (-\mathbf{u})^{+}$, luego reemplazando en (2.16) :

$F((-\mathbf{u})^{+}) + (-\mathbf{u})^{-} = 0$

Por lo tanto obtenemos la forma reducida:

$$F(\mathbf{x}^{+}) + \mathbf{x}^{-} = 0 \tag{2.18}$$

donde:

$\mathbf{x}^{+} = (x_1^{+}, \ldots, x_m^{+})^T, \quad \mathbf{x}^{-} = (x_1^{-}, \ldots, x_m^{-})^T, \quad$ y

$x_i^{+} = \max(x_i, 0), \quad x_i^{-} = \min(x_i, 0), \quad i = 1, \ldots, m$. Las construcciones de $\mathbf{x}^{+}$ y $\mathbf{x}^{-}$ son frecuentemente referidos como la aplicación Minty.

Sea $H : \mathbb{R}^{n+m+p} \longrightarrow \mathbb{R}^{n+m+p}$ la función definida por:

$$H(z) = \begin{pmatrix} F(x) + \nabla g(x)\mathbf{u}^{+} + \nabla h(x)v \\ -g(x) + \mathbf{u}^{-} \\ -h(x) \end{pmatrix} \tag{2.19}$$

donde $z = (x, u, v)^T \in \mathbb{R}^{n+m+p}$. Usando esta función, el problema para resolver VI(X,F) se convierte en el problema para resolver el siguiente sistema de ecuaciones:

$$H(z) = 0 \tag{2.20}$$

Para encontrar un cero de la función H , el método ordinario de Newton no puede ser aplicado ya que H no es una función continuamente diferenciable. El objetivo básico de éste trabajo es desarrollar un método generalizado de Newton para las funciones los cuales no son F-diferenciable pero que tienen B-derivadas en todas partes.

2.3 Las definiciones de derivadas

Definición 2.3.1 (La F-Derivada (15).-)

Sean X e Y dos espacios normados, se dice que $F : X \longrightarrow Y$ es una aplicación F-diferenciable o llamada también Fréchet diferenciable en $x_0 \in X$ si y sólo si, existe una aplicación $T \in L(X,Y)$, tal que:

$F(x_0 + h) - F(x_0) = T(h) + o(\|h\|)$

donde $L(X,Y) = \{T : X \to Y \ / \ T \text{ es lineal y continua}\}$.

Esto implica, en particular que:

$$\lim_{\|h\| \to 0} \frac{\|F(x_0 + h) - F(x_0) - T(h)\|}{\|h\|} = 0 \tag{2.21}$$

A la funcional T la llamaremos F-Derivada ó derivada de Fréchet de F en x_0 y la denotaremos por $F'(x_0) := T$.

Si $X = \mathbb{R}^n$, entonces la derivada de Fréchet coincide con el concepto de derivada del análisis clásico.

Definición 2.3.2 (Derivada Direccional.-)

Sea $S \subseteq \mathbb{R}^r$ un conjunto no vacío y $f : S \to \mathbb{R}$ una función continua. Sea $z \in S$ y $d \in \mathbb{R}^r$ un vector diferente de cero tal que $z + \lambda d \in S$ para $\lambda > 0$ suficientemente pequeño. La Derivada Direccional de f en z a lo largo del vector d, denotado por $f'(z,d)$ es dado por el siguiente limite si este existe:

$$f'(z,d) = \lim_{\lambda \to 0^+} \frac{f(z + \lambda d) - f(z)}{\lambda}.$$

La derivada direccional de una función vectorial H de un vector z a lo largo de d, es denotado por $\mathrm{H}'(z,d)$.

Un concepto cercanamente relacionado a la derivada direccional es la B-**Derivada**

Definición 2.3.3 (La B-derivada.-)
Una función $H : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ se dice que es B-diferenciable en un punto z si H es lipschitziana en una vecindad de z y además existe una función $BH(z) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, llamada la B-Derivada de H en z, la cual es una función homogénea positiva de grado 1 (es decir: $BH(z)(\lambda v) = \lambda BH(z)(v), \quad \forall v \in \mathbb{R}^n, \quad \forall \lambda \geq 0$) tal que:

$$\lim_{v \to 0} \frac{H(z+v) - H(z) - BH(z)(v)}{\|v\|} = 0$$

Si H es B-diferenciable en todos los puntos de un conjunto S, entonces se dice que H es B-diferenciable en S.

Para ilustrar esta definición, consideremos las siguientes funciones lipschitzianas:
$F : \mathbb{R} \longrightarrow \mathbb{R}$, continuamente diferenciable en $\mathbb{R}$
$u^+ : \mathbb{R} \longrightarrow \mathbb{R}_+$,
$u^- : \mathbb{R} \longrightarrow \mathbb{R}_-$,
donde:
$R_+ = \{x \in \mathbb{R} : x \geq 0\}$,
$R_- = \{x \in \mathbb{R} : x < 0\}$
Sea $H : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definido por:

$$H(z) = \begin{pmatrix} F(x) - u^+ \\ x + u^- \end{pmatrix}$$

donde $z = (x, u)^T \in \mathbb{R}^2$. Luego tenemos que la B-derivada de H en un punto z a lo largo de la dirección $v = (dx, du)^T$ está dado por:

$$BH(z)v = \begin{pmatrix} \nabla F(x)dx - du^+ \\ dx + du^- \end{pmatrix},$$

donde:

$$du^+ = \begin{cases} du & ,\text{Si} \quad u > 0 \\ max(du, 0) & ,\text{Si} \quad u = 0 \\ 0 & ,\text{Si} \quad u < 0 \end{cases} \quad \text{y} \quad du^- = \begin{cases} 0 & ,\text{Si} \quad u > 0 \\ min(du, 0) & ,\text{Si} \quad u = 0 \\ du & ,\text{Si} \quad u < 0 \end{cases}$$

dicha función es una función homogénea positiva de grado 1 tal que

$$\lim_{v \to 0} \frac{H(z+v) - H(z) - BH(z)(v)}{||v||} = 0$$

En efecto:

debido a que:

$$\lim_{dx \to 0} \frac{F(x+dx) - F(x) - \nabla F(x)(dx)}{||dx||} = 0$$

$$\lim_{du \to 0} \frac{(u+du)^+ - (u)^+ - du^+}{||du||} = 0$$

$$\lim_{du \to 0} \frac{(u+du)^- - (u)^- - du^-}{||du||} = 0$$

Proposición 2.3.1 *En un espacio euclidiano finito dimensional $\mathbb{R}^n$, H es B-diferenciable en z si y sólo si es direccionalmente diferenciable en z. En este caso, la B-derivada y la derivada direccional son idénticas.*

Demostración:

($\Rightarrow$:) Por ser H, una aplicación B-diferenciable en z y sin pérdida de generalidad se puede suponer $v = \lambda d$, donde $||d|| = 1$ fijo, entonces:

$$\lim_{\lambda \to 0} \frac{H(z+\lambda d) - H(z) - BH(z)\lambda d}{||\lambda d||} = 0$$

Y por lo tanto:

$$\lim_{\lambda \to 0} ||\frac{H(z+\lambda d) - H(z)}{\lambda} - BH(z)d|| = 0$$

Es decir:

$BH(z)(d) = H'(z,d)$. Por lo tanto H es direccionalmente diferenciable en z

($\Leftarrow$:) De manera análoga: H es B-diferenciable en z. $\square$

La relación entre la F-Derivada y la derivada direccional está dirigido por el siguiente lema:

Lema 2.3.1 *Si* $H : \mathbb{R}^r \longrightarrow \mathbb{R}^r$ *es* F*-diferenciable en* z*, entonces la derivada direccional existe en* z *y está dada por:*
$H'(z,d) = \nabla H(z).d,$
para cada dirección $d \in \mathbb{R}^r$.

Demostración:
Dado que H es F-diferenciable en z, entonces para cualquier vector d y para $\lambda > 0$ suficientemente pequeño tenemos:
$H(z+\lambda d) = H(z) + \lambda \nabla H(z).d + \lambda \|d\| \alpha(z;\lambda d)$
Dividiendo por λ:
$\dfrac{H(z+\lambda d) - H(z)}{\lambda} = \nabla H(z).d + \|d\|\alpha(z;\lambda d)$
Cuando $\lambda \to 0^+ \Rightarrow \alpha(z;\lambda d) \to 0$ Así:
$H'(z,d) = \nabla H(z).d$ [ ]

Observación: Del lema anterior, es claro que si $H'(z,d)$ no es un operador lineal sobre d, entonces H no es F-diferenciable.

Definición 2.3.4 *La B-derivada* $BH(z)$ *se dice fuerte si:*

$$\lim_{(v,w)\to(0,0)} \frac{H(z+v) - H(z+w) - BH(z)(v-w)}{\|v-w\|} = 0$$

Proposición 2.3.2 (11)
Sea $H : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *localmente lipchitziana en un vector* z*:*

(i) *Si* H *es* F*-Diferenciable en* z*, entonces también es* B*-diferenciable, además* $BH(z) = \nabla H(z)$*. Inversamente, si* H *es* B*-diferenciable en* z *y si la* B*-Derivada* $BH(z)v$ *es lineal en* v*, entonces* H *es* F*-Diferenciable en* z*.*

(ii) *Si* H *es* B*-Diferenciable en* z*, entonces la* B*-Derivada es única. Además,* $BH(z)$ *es lipchitziana con el mismo módulo de* H*.*

Demostración:

(i) ($\Longrightarrow$) Como H es F-Diferenciable en z, existe una aplicación lineal continua $\nabla H(z) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, llamada la F-Derivada de H en z lo cual satisface la siguiente condición:

$\forall \epsilon > 0, \exists\, \delta > 0$ tal que $\|H(z+v) - H(z) - \nabla H(z)v\| \leq \epsilon\|v\|$

Como $\nabla H(z)$ es una función homogenea positiva de grado 1, Es decir : $\nabla H(z)(\lambda v) = \lambda \nabla H(z)(v)$ para todo $v \in \mathbb{R}^n$ y todo $\lambda \geq 0$, entonces haciendo $\nabla H(z) = BH(z)$, tenemos que H es B-Diferenciable en z.

($\Longleftarrow$) Como H es B-Diferenciable en z, y la B-Derivada $BH(z)v$ es lineal en v, entonces sólo es suficiente demostrar que $BH(z)v$ es continua de v.

Debido a que H-es B-diferenciable en z y H es continua en z:

$\forall \epsilon > 0$ y $\epsilon_1 = \epsilon_2 = \dfrac{\epsilon}{2}$ $\quad \exists\, \delta_1 > 0,\ \delta_2 > 0$:

$\|v\| < \delta_1 \Longrightarrow \|H(z+v) - H(z)\| < \epsilon_1$

$\|v\| < \delta_2 \Longrightarrow \|H(z+v) - H(z) - BH(z)(v)\| < \epsilon_2\|v\|$

Sea $\delta = \min\{\delta_1, \delta_2, 1\}$ y $\|v\| < \delta$.

Entonces:

$$\begin{aligned}
\|BH(z)(v)\| &\leq \|BH(z)(v) - H(z+v) + H(z)\| + \|H(z+v) - H(z)\| \\
&< \frac{\epsilon}{2}\|v\| + \frac{\epsilon}{2} \\
&< \epsilon
\end{aligned}$$

Por lo tanto $BH(z)(.)$ es continua en v.

(ii) Supongamos por el absurdo, que existen $BH_1(z)$, $BH_2(z)$ dos B-Derivadas, entonces para todo vector $v \in \mathbb{R}^n$ se tiene que :

$$\|BH_1(z)v - BH_2(z)v\| \leq \|H(z+v) - H(z) - BH_1(z)v\| + \|H(z+v) - H(z) - BH_2(z)v\|$$

Luego:

$$\frac{\|BH_1(z)v - BH_2(z)v\|}{\|v\|} \to 0, \quad \text{si} \quad \|v\| \to 0 \ldots (\dagger)$$

Pero $BH_1(z) - BH_2(z)$ es un operador lineal y continuo, luego debe ser necesariamente nulo. Asumamos que existe un $v \in \mathbb{R}^n$, tal que

$$\|(BH_1(z) - BH_2(z))v\| = \alpha, \quad \alpha > 0$$

Si hacemos $v = \lambda u, \quad \|u\| = 1$, entonces:

$$\frac{\|BH_1(z)(\lambda u) - BH_2(z)(\lambda u)\|}{\lambda} = \alpha > 0, \quad \text{si} \quad \lambda \to 0^+ \ldots (\dagger\dagger)$$

Teniendo en cuenta ($\dagger$), afirmamos que ($\dagger\dagger$) es una contradicción, por lo tanto, $BH_1 = BH_2$.

$BH(z)$ es lipschitziana, en efecto:

Sea $x, y \in \mathbb{R}^n$

$$\begin{aligned} \|BH(z)(x-y)\| &\leq \|H(z+(x-y)) - H(z) - BH(z)(x-y)\| + \\ &\quad \|H(z+(x-y)) - H(z)\| \\ &\leq \epsilon\|x-y\| + \lambda\|x-y\| \qquad \forall\, \epsilon > 0 \end{aligned}$$

Entonces:

$$\|BH(z)(x-y)\| \leq \lambda\|x-y\|$$

Teorema 2.3.1 (11) *Sea $H : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ B-Diferenciable en una vecindad de z. Las siguientes proposiciones son equivalentes:*

(a) La B-Derivada $BH(z)$ es fuerte, es decir:

$$\lim_{(x,y)\to(z,z)} \frac{H(x) - H(y) - BH(z)(x-y)}{\|x-y\|} = 0$$

(b) La F-Derivada $\nabla H(z)$ es fuerte en z, es decir:

$$\lim_{(x,y)\to(z,z)} \frac{H(x) - H(y) - \nabla H(z)(x-y)}{\|x-y\|} = 0$$

(c) $BH(.)$ es continua en z, es decir, para cada $\epsilon > 0$, existe una vecindad N de z tal que $\forall\ x \in N$ y $\forall\ v \in \mathbb{R}^n$

$$\|BH(x)v - BH(z)v\| \leq \epsilon\|v\|$$

Demostración:

$(a) \Longrightarrow (b)$: $BH(z)(v)$ es lineal en v, en efecto:

Sea $v \in \mathbb{R}^n$ un vector arbitrario. De acuerdo a la proposición de (a), tenemos:

$$\lim_{t\to 0^+} \frac{H(z) - H(z+tv) - tBH(z)(-v)}{t} = 0$$

$$\lim_{t\to 0^+} \frac{H(z+tv) - H(z) - tBH(z)(v)}{t} = 0$$

Sumando, obtenemos:

$$BH(z)(-v) = -BH(z)(v)$$

De manera análoga, si $t \to 0^-$.

Ahora, sean v_1 y v_2 dos vectores diferentes de cero con $v_1 + v_2 \neq 0$, entonces de acuerdo a lo supuesto por (a), tenemos:

$$\lim_{t\to 0^+} \frac{H(z+tv_1) - H(z-tv_2) - tBH(z)(v_1+v_2)}{t} = 0$$

$$\lim_{t\to 0^+} \frac{H(z+tv_1) - H(z) - tBH(z)(v_1)}{t} = 0$$

$$\lim_{t\to 0^+} \frac{H(z) - H(z-tv_2) - tBH(z)(v_2)}{t} = 0$$

Sumando las dos ultimas ecuaciones y restando la primera, establecemos que:

$$BH(z)(v_1+v_2) = BH(z)v_1 + BH(z)v_2$$

Como $BH(z)(v)$ es lineal en v y además H es B-diferenciable en z entonces H es F-diferenciable, luego: $BH(z) = \nabla H(z)$.

Por lo tanto la F-derivada $\nabla H(z)$ es fuerte.

$(b) \Rightarrow (c)$: Recordemos un hecho sobre sucesiones dobles:

Si $\{a_{pq}\}$ es una sucesión doble con límite en a, además si el límite $\lim\limits_{p\to\infty} a_{pq}$ existe para cada q fijo, entonces:

$\lim\limits_{q\to\infty}(\lim\limits_{p\to\infty} a_{pq})$ existe y es igual a a. (Ver Anexo)

Por lo supuesto en (b):

H tiene una fuerte F-derivada en z, entonces $BH(z) = \nabla H(z)$. Para cada vector y cerca a z, definimos:

$$g(y) = \mathrm{Sup}_{\|v\|=1}\|(BH(z) - BH(y))v\|$$

De acuerdo a la propiedad lipschitziana de la B-derivada, (Proposición 2.3.2(ii)) dicho supremo es finito. Sea $\{y^k\}$ una sucesión arbitraria que converge a z y sea $\{v^k\}$ una sucesión correspondiente con cada v^k que es un vector que logra el máximo en $g(y^k)$.

Sea t_l cualquier sucesión de escalares positivos tal que $t_l \to 0^+$.

Luego, por la desigualdad triangular:

$$\begin{aligned}\|(BH(z) - BH(y^k))v^k\| \;\leq\; & \|\frac{H(y^k + t_lv^k) - H(y^k) - t_lBH(z)v^k}{t_l}\| + \\ & +\|\frac{H(y^k + t_lv^k) - H(y^k) - t_lBH(y^k)v^k}{t_l}\|\end{aligned}$$

Consideremos la sucesión doble $(H(y^k + t_lv^k) - H(y^k) - t_lBH(z)v^k)/t_l$. Debido a que la F-derivada $\nabla H(z)$ es fuerte:

$$\lim_{k\to\infty}(\lim_{t_l\to 0^+}\|\frac{H(y^k + t_lv^k) - H(y^k) - t_lBH(z)v^k}{t_l}\|) = 0$$

De manera análoga:

$$\lim_{t_l\to 0^+}(\lim_{k\to\infty}\|\frac{H(y^k + t_lv^k) - H(y^k) - t_lBH(y^k)v^k}{t_l}\|) = 0$$

por el hecho mencionado sobre sucesiones dobles, tenemos que:

$\lim\limits_{k\to\infty} g(y^k) = 0$ es decir $\lim\limits_{k\to\infty} \|(BH(z) - BH(y^k))v^k\| = 0$. Luego se establece la continuidad de $BH(.)$ en z.

$(c) \implies (a)$: Debido a que $BH(.)$ es continua en z, es decir: Para cada $\epsilon > 0$, existe una vecindad N de z tal que para todo $x \in N$ y todo $u \in \mathbb{R}^n$

$$\|BH(x)u - BH(z)u\| \leq \epsilon\|u\|$$

Luego:

$$\mathrm{Sup}\|BH(x)u - BH(z)u\| \leq \epsilon\|u\|$$

Por el teorema del valor medio(Ver Anexo):

$$\|H(x) - H(y) - BH(z)(x-y)\| \leq \mathrm{Sup}\|BH(x)(x-y) - BH(z)(x-y)\| \leq \epsilon\|x-y\|$$

Entonces $BH(z)$ es fuerte. $\square$

2.4 Resultados Importantes

Para obtener varias propiedades útiles de la aplicación H definido por (2.19), los siguientes resultados son necesarios.

Lema 2.4.1 *La aplicación Minty es continuamente Lipschitziana*

Demostración:

Sean las aplicaciones:

$x^+ : \mathbb{R}^n \to \mathbb{R}^n_+$

$x^- : \mathbb{R}^n \to \mathbb{R}^n_-$ donde:

$\mathbb{R}^n_+ = \{x = (x_1, \dots, x_n)^T \in \mathbb{R}^n : x_i \geq 0\}$

$\mathbb{R}^n_- = \{x = (x_1, \dots, x_n)^T \in \mathbb{R}^n : x_i < 0\}$

Es suficiente demostrar que:

$$|x_i^+ - y_i^+| \leq |x_i - y_i| \quad \text{y} \quad |x_i^- - y_i^-| \leq |x_i - y_i|$$

Es decir, x^+ y x^- son continuamente lipschitziana con una constante lipschitziana $l = 1$.

Veamos:

1 : Si $x_i \geq 0, \quad y_i \geq 0$ entonces $x_i^+ = x_i, \quad x_i^- = 0; \quad y_i^+ = y_i, \quad y_i^- = 0$

Con lo cual:

$|x_i^- - y_i^-| = 0 \leq |x_i - y_i|$

$|x_i^+ - y_i^+| = |x_i - y_i|$

2 : Si $x_i < 0, \quad y_i < 0$ entonces $x_i^+ = 0, \quad x_i^- = x_i; \quad y_i^+ = 0, \quad y_i^- = y_i$

Con lo cual:

$|x_i^+ - y_i^+| = 0 \leq |x_i - y_i|$

$|x_i^- - y_i^-| = |x_i - y_i|$

3 : Si $x_i \geq 0, \quad y_i < 0$ entonces $x_i^+ = x_i, \quad x_i^- = 0; \quad y_i^+ = 0, \quad y_i^- = y_i$

Con lo cual:

$|x_i^+ - y_i^+| = x_i < x_i - y_i = |x_i - y_i|$

$|x_i^- - y_i^-| = |0 - y_i| = -y_i \leq x_i - y_i = |x_i - y_i|$

4 : Si $x_i < 0, \quad y_i \geq 0$ entonces $x_i^+ = 0, \quad x_i^- = x_i; \quad y_i^+ = y_i, \quad y_i^- = 0$

Con lo cual:

$|x_i^+ - y_i^+| = |-y_i| = y_i < y_i - x_i = |y_i - x_i| = |x_i - y_i|$

$|x_i^- - y_i^-| = |x_i| = -x_i \leq y_i - x_i = |y_i - x_i| = |x_i - y_i|$

$\square$

Proposición 2.4.1 *Sean las funciones:*

$F : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad g : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \quad h : \mathbb{R}^n \longrightarrow \mathbb{R}^p,$

$\nabla g : \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}, \quad \nabla h : \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times p}$ *continuamente lipschitziana en* $\mathbb{R}^n$

entonces la función H definida en (2.19) *es localmente continuamente lipschitziana.*

Demostración:

Probaremos primero que la suma de funciones lipschitziana es también lipschitziana:

Sean $p, r : X \to Y$ donde $X, Y \subset \mathbb{R}^n$, son lipschitzianas sobre X entonces:

$$\|p(x) - p(y)\|_Y \leq L_1\|x - y\|_X \qquad \forall x, y \in X$$

$$\|r(x) - r(y)\|_Y \leq L_2\|x - y\|_X \qquad \forall x, y \in X$$

$$\begin{aligned} \|(p+r)(x) - (p+r)(y)\|_Y &= \|(p(x) - p(y)) + (r(x) - r(y))\|_Y \leq \\ &\leq \|p(x) - p(y)\|_Y + \|r(x) - r(y)\|_Y \leq \\ &\leq L_1\|x - y\|_X + L_2\|x - y\|_X \end{aligned}$$

Por lo tanto:

$$\|(p+r)(x) - (p+r)(y)\|_Y \leq L\|x - y\|_X$$

donde: $L = L_1 + L_2$

Después probaremos que el producto de funciones lipschitziana es localmente continuamente lipschitziana:

Decimos que una aplicación $f : X \to Y$ donde $X, Y \subset \mathbb{R}^n$, es localmente lipschitziana sobre X, si $\forall x, y \in X$ existe una vecindad $V_\delta(x_0)$ tal que:

$$\|f(x) - f(y)\|_Y \leq L_1\|x - y\|_X \qquad \forall x, y \in V_\delta(x_0) \subset X$$

Sean $p, r : X \to Y,$ donde $X, Y \subset \mathbb{R}^n$, localmente lipschitziana sobre X.

Entonces haciendo $h = p.r,$ $x_0 \in X$ fijo. Tenemos:

$$\begin{aligned} \|h(x) - h(y)\|_Y &= \|p(x).r(x) - p(y).r(y)\|_Y = \\ &= \|(p(x) - p(y)).r(x) - p(y).(r(y) - r(x))\|_Y \leq \\ &\leq \|p(x) - p(y)\|_Y\|r(x)\|_Y + \|p(y)\|_Y\|r(x) - r(y)\|_Y \ldots (*) \end{aligned}$$

Como p y r son localmente lipschitziana sobre X

$$\|p(x) - p(y)\|_Y \leq L_1\|x - y\|_X \qquad \forall x, y \in V_\delta(x_0) \subset X$$

$$\|r(x) - r(y)\|_Y \leq L_2\|x - y\|_X \qquad \forall x, y \in V_\delta(x_0) \subset X$$

$\Longrightarrow$ de $\quad (*)$:

$$\leq (L_1\|x - y\|_X)\|r(x)\|_Y + (L_2\|x - y\|_X)\|p(y)\|_Y \ldots (**)$$

Ahora, por la continuidad de p y r:

$\forall \epsilon_1 > 0, \quad \exists \delta_1 > 0 \quad / \quad x \in V_{\delta_1}(x_0) \quad$ y $\quad \|x - x_0\|_X < \delta_1 \Longrightarrow \|r(x) - r(x_0)\|_Y < \epsilon_1 \ldots (\dagger)$

$\forall \epsilon_2 > 0, \quad \exists \delta_2 > 0 \quad / \quad y \in V_{\delta_2}(x_0) \quad$ y $\quad \|y - x_0\|_X < \delta_2 \Longrightarrow \|p(y) - p(x_0)\|_Y < \epsilon_2 \ldots (\dagger\dagger)$

De $(\dagger)$ y $(\dagger\dagger)$:
$\|r(x)\|_Y < \|r(x_0)\|_Y + \epsilon_1$
$\|p(y)\|_Y < \|p(x_0)\|_Y + \epsilon_2$
Escogemos $\epsilon_1 = \epsilon_2 = \frac{1}{2}$. Luego:
$\|r(x)\|_Y < \|r(x_0)\|_Y + \frac{1}{2}$
$\|p(y)\|_Y < \|p(x_0)\|_Y + \frac{1}{2}$
$\Longrightarrow$ de $\quad (**)$:
$\leq L_1\|x - y\|_X(\|r(x_0)\|_Y + \frac{1}{2}) + L_2\|x - y\|_X(\|p(x_0)\|_Y + \frac{1}{2}) =$
$\{L_1(\|r(x_0)\|_Y + \frac{1}{2}) + L_2(\|p(x_0)\|_Y + \frac{1}{2})\}\|x - y\|_X = L\|x - y\|_X$
donde:
$L = L_1(\|r(x_0)\|_Y + \frac{1}{2}) + L_2(\|p(x_0)\|_Y + \frac{1}{2})$
Por lo tanto $h = p.r$ es localmente lipshitziana.

Como F, g, h, ∇g, ∇h son continuamente lipschitziana y de acuerdo al lema (2.4.1): u^+ y u^- son continuamente lipschitziana. Además, dado que el producto de funciones lipschitziana es localmente lipschitziana, también la suma de funciones lipschitziana es lipschitziana, entonces de acuerdo como está definido H en (2.19) decimos que H es localmente lipschitziana. $\square$

Teorema 2.4.1 *Sea $H(z)$ definida por* (2.19). *Entonces:*

(a) *H es B-diferenciable en todas partes, y la B-derivada de H en un vector $z = (x, u, v)^T$ a lo largo de la dirección $d = (dx, du, dv)^T$, denotado por $H'(z, d)$, ó $BH(z)d$ está dado por:*

$$\begin{pmatrix} [\nabla F(x) + \sum_{i=1}^{m} u_i^+ \nabla^2 g_i(x) + \sum_{j=1}^{p} v_j \nabla^2 h_j(x)]dx + \nabla g(x) du^+ + \nabla h(x) dv \\ -\nabla g(x)^T dx + du^- \\ -\nabla h(x)^T dx \end{pmatrix} \tag{2.22}$$

donde:

$$du_i^+ = \begin{cases} du_i & , Si \quad u_i > 0 \\ max(du_i, 0) & , Si \quad u_i = 0 \\ 0 & , Si \quad u_i < 0 \end{cases} \quad y \quad du_i^- = \begin{cases} 0 & , Si \quad u_i > 0 \\ min(du_i, 0) & , Si \quad u_i = 0 \\ du_i & , Si \quad u_i < 0 \end{cases}$$

(b) *Denotando:*

$\alpha(z) = \{i \,/\, u_i > 0\}$

$\beta(z) = \{i \,/\, u_i = 0\}$

$\gamma(z) = \{i \,/\, u_i < 0\}$

Entonces H es F-diferenciable si y sólo si el conjunto β es vacío. Además si el conjunto β es vacío entonces la B-derivada $BH(z)$ es fuerte.

Demostración:

(a) De acuerdo a como fue definido H por (2.19), tenemos:

$$H(z+\lambda d)-H(z)=\begin{pmatrix} [F(x+\lambda dx)-F(x)+\nabla g(x+\lambda dx)(u+\lambda du)^+-\nabla g(x)u^+ + \\ \qquad +\nabla h(x+\lambda dx)(v+\lambda dv)^+-\nabla h(x)dv] \\ -g(x+\lambda dx)+(u+\lambda du)^- + g(x)-u^- \\ -h(x+\lambda dx)+h(x) \end{pmatrix}$$

Además, dado que:

$$H'(z,d)=\lim_{\lambda\to 0^+}\frac{H(z+\lambda d)-H(z)}{\lambda}=$$

$$=\lim_{\lambda\to 0^+}\frac{1}{\lambda}\begin{pmatrix} [F(x+\lambda dx)-F(x)+\nabla g(x+\lambda dx)(u+\lambda du)^+-\nabla g(x)u^+ + \\ \qquad +\nabla h(x+\lambda dx)(v+\lambda dv)^+-\nabla h(x)dv] \\ -g(x+\lambda dx)+(u+\lambda du)^- + g(x)-u^- \\ -h(x+\lambda dx)+h(x) \end{pmatrix}$$

$$=\begin{pmatrix} \lim_{\lambda\to 0^+}\frac{1}{\lambda}[F(x+\lambda dx)-F(x)+\nabla g(x+\lambda dx)(u+\lambda du)^+-\nabla g(x)u^+ + \\ \qquad +\nabla h(x+\lambda dx)(v+\lambda dv)^+-\nabla h(x)dv] \\ \lim_{\lambda\to 0^+}\frac{1}{\lambda}[-g(x+\lambda dx)+(u+\lambda du)^- + g(x)-u^-] \\ \lim_{\lambda\to 0^+}\frac{1}{\lambda}[-h(x+\lambda dx)+h(x)] \end{pmatrix}\ldots(*)$$

Pero, sabiendo que:

$$(1)\quad \lim_{\lambda\to 0^+}\frac{[F(x+\lambda dx)-F(x)]}{\lambda}=\nabla F(x)dx$$

$$(2)\quad \lim_{\lambda\to 0^+}\frac{1}{\lambda}[\nabla g(x+\lambda dx)(u+\lambda du)^+-\nabla g(x)u^+]=$$

$$=\lim_{\lambda\to 0^+}\frac{1}{\lambda}[\nabla g(x+\lambda dx)(u+\lambda du)^+-\nabla g(x)(u+\lambda du)^+]+$$

$$+\lim_{\lambda\to 0^+}\frac{1}{\lambda}[\nabla g(x)(u+\lambda du)^+-\nabla g(x)u^+]=$$

$$=\lim_{\lambda\to 0^+}\frac{1}{\lambda}[\nabla g(x+\lambda dx)-\nabla g(x)]\lim_{\lambda\to 0^+}(u+\lambda du)^+ +$$

$$+\lim_{\lambda\to 0^+}\nabla g(x)\lim_{\lambda\to 0^+}\frac{(u+\lambda du)^+-u^+}{\lambda}=\sum_{i=1}^{m}u_i^+\nabla^2 g_i(x).dx+\nabla g(x).du^+$$

$$(3)\ \lim_{\lambda\to 0^+}\frac{1}{\lambda}[\nabla h(x+\lambda dx)(v+\lambda dv)-\nabla h(x)v]=\sum_{j=1}^{p}v_j\nabla^2 h_j(x).dx+\nabla h(x).dv$$

$$(4)\ \lim_{\lambda\to 0^+}-(\frac{g(x+\lambda dx)-g(x)}{\lambda})=-\nabla g(x)^T dx$$

$$(5)\ \lim_{\lambda\to 0^+}\frac{(u+\lambda du)^- - u^-}{\lambda}=\nabla u^- .du=\frac{du^-}{du}.du=du^-$$

$$(6)\ \lim_{\lambda\to 0^+}-(\frac{h(x+\lambda dx)-h(x)}{\lambda})=-\nabla h(x)^T dx$$

Reemplazando $(1)-(6)$ en $(*)$ tenemos:

$$H'(z,d)=\begin{pmatrix} [\nabla F(x)+\sum_{i=1}^{m}u_i^+\nabla^2 g_i(x)+\sum_{j=1}^{p}v_j\nabla^2 h_j(x)]dx+\nabla g(x)du^+ + \nabla h(x)dv \\ -\nabla g(x)^T dx+du^- \\ -\nabla h(x)^T dx \end{pmatrix}$$

Por la proposición (2.4.1), H es continuamente lipschitziana en z, por lo tanto H es B-diferenciable en todas partes.

(b) $(\Longrightarrow)$: Primero demostraremos, que si el conjunto β es vacío implica que H es F-Diferenciable.

Debido a que el conjunto β es vacío implica que $u\neq 0$, luego: Asumamos que $u>0$, entonces:

$du_i^+ = du_i, \qquad u_i^+ = u_i \ldots (\dagger)$

$du_i^- = 0, \qquad u_i^- = 0 \ldots (\dagger\dagger)$

Utilizando $(\dagger)$ y $(\dagger\dagger)$ en (2.19) tenemos:

$$H(z)=\begin{pmatrix} F(x)+\nabla g(x)u+\nabla h(x)v \\ -g(x) \\ -h(x) \end{pmatrix}$$

Además, (†) y (††) en (2.22) tenemos:

$$H'(z,d) = \begin{pmatrix} [\nabla F(x) + \sum_{i=1}^{m} u_i \nabla^2 g_i(x) + \sum_{j=1}^{p} v_j \nabla^2 h_j(x)]dx + \dots \\ \qquad\qquad + \nabla g(x)du + \nabla h(x)dv \\ -\nabla g(x)^T dx \\ -\nabla h(x)^T dx \end{pmatrix} =$$

$$= \begin{pmatrix} \nabla F(x) + \sum_{i=1}^{m} u_i \nabla^2 g_i(x) + \sum_{j=1}^{p} v_j \nabla^2 h_j(x) & \nabla g(x) & \nabla h(x) \\ -\nabla g(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix} \begin{pmatrix} dx \\ du \\ dv \end{pmatrix}$$

Por lo tanto: $H'(z,d) = \nabla H(z)d$.

De manera análoga, si $u < 0$: $H'(z,d) = \nabla H(z)d$.

Por lo tanto H es F-Diferenciable. Debido a que H es F-Diferenciable entonces $BH(z) = \nabla H(z)$. Además como f, g, y h son continuamente diferenciable y u^+ es continua entonces $\nabla H(z)$ es continua es decir $BH(z)$ es continua por lo tanto la B-Derivada es fuerte.

($\Longleftarrow$) Reciprocamente, demostraremos: Si H es F-diferenciable entonces $\beta = \emptyset$.

Supongamos que $\beta \neq \emptyset$ no es vacío, es decir: existe al menos un elemento $u_i = 0$.

Hacemos $u_1 = 0$ y $du_1 > 0$, $du = (du_1, du_2, \dots, du_m)^T$

Entonces uno puede verificar que la $(n+1)$ ésima componente de $H'(z,d)$ es $-\nabla g_1(x)^T dx$.

Prueba:

La $(n+1)$ ésima componente de $H'(z,d)$ es:

$= -\nabla g_1(x)^T dx + du_1^- = -\nabla g_1(x)^T dx \quad$ pues $\quad du_1 > 0$

Mientras la $(n+1)$ ésima componente de $H'(z,-d)$ es:

$$= -\lim_{\lambda \to 0^+} \frac{g_1(x - \lambda dx) - g_1(x)}{\lambda} + \lim_{\lambda \to 0^+} \frac{(u_1 - \lambda du_1)^- - u_1^-}{\lambda}$$

$$= \nabla g_1(x)^T dx + \lim_{\lambda \to 0^+} \frac{d(-\lambda u_1)^-}{\lambda}, \quad \text{pues} \quad u_1 = 0$$

$$= \nabla g_1(x)^T dx + \lim_{\lambda \to 0^+} \frac{\min(0, -\lambda du_1)}{\lambda}, \quad \text{pues} \quad -\lambda u_1 = 0$$

$$= \nabla g_1(x)^T dx + \lim_{\lambda \to 0^+} \frac{-\lambda du_1}{\lambda}, \quad \text{pues} \quad -\lambda du_1 < 0$$

$= \nabla g_1(x)^T dx - du_1$

Es decir: $H'(z,-d) \neq -H'(z,d) \Longrightarrow H'$ no es lineal en d. Luego H no es F-diferenciable en z.

Por lo tanto: Si $\beta \neq \emptyset$ entonces H no es F-diferenciable en z. $\square$

2.5 Introducción de una función mérito

Para desarrollar un algoritmo convergente en la búsqueda de la solución al problema (2.20), necesitamos introducir una función mérito $\theta : \mathbb{R}^r \longrightarrow \mathbb{R}$ definida por

$$\theta(z) = \frac{1}{2} H(z)^T H(z) = \frac{1}{2} \|H(z)\|^2. \tag{2.23}$$

Aquí, $r = n + m + p$.

Claramente, encontrar un cero de H es equivalente a resolver un punto mínimo global z^* de θ con $\theta(z^*) = 0$.

La función θ es generalmente no F-diferenciable. Sin embargo, siempre es B-diferenciable como lo demuestra los siguientes resultados:

Teorema 2.5.1 *Sea $\theta : \mathbb{R}^r \longrightarrow \mathbb{R}$ definido por (2.23). Entonces*

(a) *θ es B-diferenciable en todas partes, y la B-derivada de θ en un vector $z = (x, u, v)^T$ a lo largo de la dirección $d = (dx, du, dv)^T$ está dado por: $\theta'(z, d) = H(z)^T H'(z, d)$ donde $H'(z, d)$ está dado por (2.22).*

(b) *Si $\beta(z) = \emptyset$ entonces θ es fuertemente F-diferenciable en z. Además, si z^* es solución de H entonces θ es fuertemente F-diferenciable en z^* y $\nabla\theta(z^*) = 0$*

Demostración

(*a*) Primero probaremos que $\theta(z)$ es continuamente lipschitziana:

Como H es localmente continuamente lipschitziana entonces existe una constante $L > 0$ tal que

$$\|H(z_1) - H(z_2)\| \leq L\|z_1 - z_2\| \quad \forall z_1, z_2 \in V_\delta(z_0) \subset \mathbb{R}^{n+m+p}$$

Por continuidad de H:

$$\forall \epsilon > 0, \quad \exists \delta_1 > 0 \quad / z \in V_{\delta_1}(z_0) \quad \text{y} \quad \|z - z_0\| \leq \delta_1 \Longrightarrow \|H(z) - H(z_0)\| < \epsilon$$

$\Longrightarrow \|H(z)\mid < \mid H(z_0)\| + \epsilon \quad \forall z \in V_{\delta_1}(z_0) \ldots (*)$

Escogemos $\epsilon = 1$, luego: $\Rightarrow \|H(z)\| < \|H(z_0)\| + \frac{1}{2}$ Además:

$$\begin{aligned}
\|\theta(z_1) - \theta(z_2)\| &= \tfrac{1}{2}\|(H(z_1) - H(z_2))^T(H(z_1) + H(z_2))\| \\
&\leq \tfrac{1}{2}\|(H(z_1) - H(z_2))\|\|H(z_1) + H(z_2)\| \\
&\leq \tfrac{1}{2}L\|z_1 - z_2\|\|H(z_1) + H(z_2)\| \\
&\leq \tfrac{1}{2}L\|z_1 - z_2\|\|H(z_1)\| + \tfrac{1}{2}L\,|z_1 - z_2\|\|H(z_2)\| \leq \ldots (**)
\end{aligned}$$

Utilizando $(*)$ en $(**)$ tenemos:

$\leq \frac{1}{2}L\|z_1 - z_2\|(\|H(z_0)\| + 1) + \frac{1}{2}L\|z_1 - z_2\|(\|H(z_0)\| + 1)$

$= L\|z_1 - z_2\|(\|H(z_0)\| + 1) = L\|z_1 - z_2\|$

donde: $L = L(\|H(z_0)\| + 1)$

Por lo tanto, θ es localmente lipschitziana $\Rightarrow$ θ es B-diferenciable en todas partes.

Después, probaremos: $\theta'(z, d) = H(z)^T H'(z, d)$

Veamos:

$$
\begin{aligned}
\theta'(z,d) &= \lim_{\lambda\to 0^+}\frac{\theta(z+\lambda d)-\theta(z)}{\lambda}\\
&= \lim_{\lambda\to 0^+}\frac{1}{2\lambda}[H(z+\lambda d)^T H(z+\lambda d)-H(z)^T H(z)]\\
&= \lim_{\lambda\to 0^+}\frac{1}{2\lambda}H(z+\lambda d)^T[H(z+\lambda d)-H(z)]+\\
&\quad+\lim_{\lambda\to 0^+}\frac{1}{2\lambda}[H(z+\lambda d)-H(z)]^T H(z)\\
&= \tfrac{1}{2}H(z)^T H'(z,d)+\tfrac{1}{2}H'(z,d)^T H(z)\\
&= \tfrac{1}{2}H(z)^T H'(z,d)+\tfrac{1}{2}H(z)^T H'(z,d)
\end{aligned}
$$

Por lo tanto:

$\theta'(z,d) = H(z)^T H'(z,d)$

(b) Como $\beta = \emptyset$, entonces H es F-diferenciable en z y debido a que H es continua en z implica que H es continuamente diferenciable en z (es decir : ∇H es continua en z) y de acuerdo a como θ está definida, entonces θ es continuamente diferenciable en z, es decir:

$\nabla\theta(z)$ es continua en z y como $\theta(z)$ es B-Diferenciable entonces θ es fuertemente F-Diferenciable.

Para probar la segunda parte:

Como: $\theta(z) = \dfrac{1}{2}H(z)^T H(z) = \dfrac{1}{2}\|H(z)\|^2 \Longrightarrow \nabla\theta(z) = H(z)^T\nabla H(z)$

Debido a que z^* es solución de H, tenemos $H(z^*) = 0$, por lo tanto: $\nabla\theta(z^*) = 0$.

Para probar que $\theta(z)$ es fuertemente F-diferenciable en z^*, sabiendo que $\nabla\theta(z^*) = 0$, sólo es suficiente demostrar que:

$$\lim_{(y,w)\to(z^*,z^*)}\frac{|\theta(y)-\theta(w)|}{\|y-w\|} = 0$$

Debido a que:

$$
\begin{aligned}
|\theta(y)-\theta(w)| &= \tfrac{1}{2}|H(y)^T H(y)-H(w)^T H(w)|\\
&= \tfrac{1}{2}|[H(y)-H(w)]^T[H(y)+H(w)]|
\end{aligned}
$$

Debido a la propiedad Lipschitziana de H, la expresión de arriba es acotada por:

$$\frac{1}{2}c\|y - w\|\|H(y) + H(w)\|,$$

donde $c > 0$ es la constante lipschitziana de H.

Con lo cual tenemos:

$$\lim_{(y,w)\to(z^*,z^*)} \frac{|\theta(y) - \theta(w)|}{\|y - w\|} \leq \lim_{(y,w)\to(z^*,z^*)} \frac{1}{2}c\|H(y) + H(w)\| = 0$$

Pues:

$H(z^*) = 0$. Por lo tanto:

$$\lim_{(y,w)\to(z^*,z^*)} \frac{|\theta(y) - \theta(w)|}{\|y - w\|} = 0$$

Se nota que en un cero de H, el conjunto β puede no ser vacío, ó que H puede no ser F-diferenciable. En otras palabras, el ser $\beta(z^*) = \emptyset$ no es condición necesaria para la F-diferenciabilidad de θ en z^*.

Llamamos el conjunto indice $\beta(z)$ como *el conjunto degenerado* y los indices de $\beta(z)$ serán los indices degenerados. Llamaremos un vector z un vector no-degenerado de H si $\beta(z) = \emptyset$.

Claramente, H y θ son F-diferenciable en todos los vectores no-degenerados. A través de éste trabajo, no asumimos que $\beta(z)$ es vacío en todas partes, pero que la medida del conjunto degenerado es cero: Es decir, esperamos encontrar raramente puntos degenerados durante el curso del algoritmo propuesto. Así H y θ tendrán fuerte B-derivada casi en todas partes.

3 EL ALGORITMO DE NEWTON AMORTIGUADO

El algoritmo de Newton(amortiguado) fue usado por Harker y Pang para resolver problemas de complementariedad y por Harker y Xiao para resolver problemas de complementariedad no lineal cuyos experimentos numéricos sugieren que el método de Newton(amortiguado) es generalmente más eficiente y robusto que el método tradicional de Newton . Esta aproximación toma muy pocas iteraciones que permiten converger en donde el método tradicional falla.
Describiremos el método de Newton(amortiguado) como sigue:

Paso 0 : Sea $z^0 \in \mathbb{R}^{n+m+p}$ un vector arbitrario inicial, y dados μ y σ escalares donde $\mu \in <0,1>$ y $\sigma \in <0,\frac{1}{2}>$, $\epsilon > 0$ pequeño
Paso $k(k=1,2,\ldots)$: Si $\|H(z^k)\| \leq \epsilon$ parar, en otro caso, generar z^{k+1} realizando los siguientes dos pasos:

1.– Resolver la ecuación de Newton:

$$H(z^k) + H'(z^k, d^k) = 0 \tag{3.1}$$

para la dirección $d^k \in \mathbb{R}^{n+m+p}$

2.– Sea $\lambda_k = \mu^{m_k}$ donde:

$$m_k = \text{Min}\{\{m \in \mathbb{Z}^+ \cup \{0\} \;/\; \theta(z^k) - \theta(z^k + \mu^m d^k) \geq 2\sigma\mu^m\theta(z^k)\}\} \tag{3.2}$$

Fijar $z^{k+1} = z^k + \lambda_k d^k$, e ir al paso $k+1$.

Para analizar este algoritmo, definimo

$$
\begin{aligned}
u_\alpha &= \{u_i \,/\, i \in \alpha(z)\},\\
u_\beta &= \{u_i \,/\, i \in \beta(z)\},\\
u_\gamma &= \{u_i \,/\, i \in \gamma(z)\},\\
g_\alpha(x) &= \{g_i(x) \,/\, i \in \alpha(z)\},\\
g_\beta(x) &= \{g_i(x) \,/\, i \in \beta(z)\},\\
g_\gamma(x) &= \{g_i(x) \,/\, i \in \gamma(z)\},\\
H_F(z) &= F(x) + \nabla g_\alpha(x) u_\alpha + \nabla h(x) v,\\
H_\alpha(z) &= -g_\alpha(x),\\
H_\beta(z) &= -g_\beta(x),\\
H_\gamma(z) &= -g_\gamma(x) + u_\gamma,\\
H_h(z) &= -h(x).
\end{aligned}
$$

Usando esta notación, la ecuación (3.1) puede ser escrito como:

$$
\begin{aligned}
H_F(z) + [\nabla F(x) + \sum_{i \in \alpha(z)} u_i \nabla^2 g_i(x) + \sum_{j=1}^{p} v_j \nabla^2 h_j(x)]dx +\\
+\nabla g_\alpha(x) du_\alpha + \nabla g_\beta(x) \max(0, du_\beta) + \nabla h(x) dv = 0 \qquad (3.3)
\end{aligned}
$$

$$H_\alpha(z) - \nabla g_\alpha(x)^T dx = 0 \qquad (3.4)$$

$$H_\beta(z) - \nabla g_\beta(x)^T dx + \min(0, du_\beta) = 0 \qquad (3.5)$$

$$H_\gamma(z) - \nabla g_\gamma(x)^T dx + du_\gamma = 0 \qquad (3.6)$$

$$H_h(z) - \nabla h(x)^T dx = 0 \qquad (3.7)$$

En el resto de este trabajo, discutiremos la existencia de una solución al sistema (3.1), la propiedad descendente de la búsqueda lineal y el comportamiento de la convergencia del algoritmo.

3.1 Factibilidad del subproblema

El establecimiento de la existencia de una solución del problema (3.1) es el primer objetivo así como la convergencia del algoritmo en mención. Aunque la condicione para la existencia son más débil que la regularidad, el uso de la uposición de regularidad(se verá mas adelante) permitirá a uno establecer algunos resultados importantes de convergencia.

Antes definiremos la noción de regularidad lo cual usaremos frecuentemente en éste trabajo y expresaremos sus propiedades, pero nec sitamos revisar algunos resultados de la teoria de matrices. Sea M una matriz $m \times n$, $I \subseteq \{1, \ldots, m\}$ y $J \subseteq \{1, \ldots, n\}$ subconjuntos de indices. Denotemos como M_{IJ} a la submatriz de M cuyas filas y columnas e tán indexados por I y J respectivamente. Para una matriz cuadrada M, M_{II} denota una submatriz principal de M indexados por I

Definición 3.1.1 *Dado una submatriz principal no-singular M_{II} de M, el Complemento Schur de M_{II} en M denotado por (M/M_{II}) es la matriz:*

$$(M/M_{II}) = M_{\overline{I}\overline{I}} - M_{\overline{I}I}(M_{II})^{-1}M_{I\overline{I}}$$

donde $\overline{I}$ es el complemento de I en $\{1, \ldots, m\}$

Proposición 3.1.1 *Sea M una matriz cuadrada real dada en forma particionada*

$$M = \begin{pmatrix} M_{II} & M_{I\overline{I}} \\ M_{\overline{I}I} & M_{\overline{I}\overline{I}} \end{pmatrix},$$

entonces:

(a) $Det[(M/M_{II})] = \dfrac{Det(M)}{Det(M_{II})}$, *siempre que M_{II} es no-singular.*

(b) *El Complemento Schur* (M_{KK}/M_{II}) *es una submatriz principal del Complemento Schur* (M/M_{II}), $\quad \forall I \subseteq K;\ K \subseteq \{1, \ldots, m\}$ *donde* M_{II} *es una submatriz principal no-singular de* M

(c) *El Complemento Schur de* M_{KK} *en* M *es igual al Complemento Schur de* (M_{KK}/M_{II}) *en* (M/M_{II}) *es decir:*

$$(M/M_{KK}) = \Big((M/M_{II})\Big/(M_{KK}/M_{II})\Big); \quad I \subseteq K \subseteq \{1, \ldots, m\}$$

donde M_{KK} *y* M_{II} *son submatrices principales no-singulares de* M.

Demostración:

(a) Como M_{II} es no-singular entonces: $\begin{pmatrix} M_{II}^{-1} & 0 \\ 0 & I \end{pmatrix}$ existe.

$$\Longrightarrow \begin{pmatrix} M_{II}^{-1} & 0 \\ 0 & I \end{pmatrix}.M = \begin{pmatrix} M_{II}^{-1} & 0 \\ 0 & I \end{pmatrix}.\begin{pmatrix} M_{II} & M_{I\bar{I}} \\ M_{\bar{I}I} & M_{\bar{I}\bar{I}} \end{pmatrix},$$

$$\Longrightarrow \begin{pmatrix} M_{II}^{-1} & 0 \\ 0 & I \end{pmatrix}.M = \begin{pmatrix} I & M_{II}^{-1}M_{I\bar{I}} \\ M_{I\bar{I}} & M_{\bar{I}\bar{I}} \end{pmatrix}$$

$$\Longrightarrow \begin{pmatrix} M_{II} & 0 \\ -M_{\bar{I}I} & I \end{pmatrix}.\begin{pmatrix} M_{II}^{-1} & 0 \\ 0 & I \end{pmatrix}.M =$$

$$= \begin{pmatrix} M_{II} & 0 \\ -M_{\bar{I}I} & I \end{pmatrix}.\begin{pmatrix} I & M_{II}^{-1}M_{I\bar{I}} \\ M_{I\bar{I}} & M_{\bar{I}\bar{I}} \end{pmatrix}$$

$$\Longrightarrow \begin{pmatrix} I & 0 \\ -M_{\bar{I}I}M_{II}^{-1} & I \end{pmatrix}.M = \begin{pmatrix} M_{II} & M_{I\bar{I}} \\ 0 & M_{\bar{I}\bar{I}} - M_{I\bar{I}}M_{II}^{-1}M_{I\bar{I}} \end{pmatrix}$$

Debido a que Det(AB)=Det(A).Det(B), donde A y B son matrices cuadradas. Entonces:

$$
\begin{aligned}
Det\left[\begin{pmatrix} I & 0 \\ -M_{\bar{I}I}M_{II}^{-1} & I \end{pmatrix}.M\right] &= Det\left[\begin{pmatrix} I & 0 \\ -M_{\bar{I}I}M_{II}^{-1} & I \end{pmatrix}\right].Det[M] \\
&= Det[M]
\end{aligned}
$$

y

$$
Det\left[\begin{pmatrix} M_{II} & M_{I\bar{I}} \\ 0 & M_{\bar{I}\bar{I}} - M_{I\bar{I}}M_{II}^{-1}M_{I\bar{I}} \end{pmatrix}\right] = Det[M_{II}].Det[M_{\bar{I}\bar{I}} - M_{I\bar{I}}M_{II}^{-1}M_{I\bar{I}}]
$$

$$
\Longrightarrow Det[(M/M_{II})] = \frac{Det(M)}{Det(M_{II})} \quad \square
$$

(b) Como K está dado tenemos:

$$
M = \begin{pmatrix} M_{I,I} & M_{I,K-I} & M_{I,\overline{K}} \\ M_{K-I,I} & M_{K-I,K-I} & M_{K-I,\overline{K}} \\ M_{\overline{K},I} & M_{\overline{K},K-I} & M_{\overline{K},\overline{K}} \end{pmatrix} = \begin{pmatrix} M_{KK} & M_{K\overline{K}} \\ M_{\overline{K}K} & M_{\overline{K}\overline{K}} \end{pmatrix}
$$

Por definición:

$(M_{KK}/M_{II}) = M_{K-I,K-I} - M_{K-I,I}(M_{II})^{-1}M_{I,K-I}$

y

$$
\begin{aligned}
(M/M_{II}) &= \begin{pmatrix} M_{K-I,K-I} & M_{K-I,\overline{K}} \\ M_{\overline{K},K-I} & M_{\overline{K},\overline{K}} \end{pmatrix} + \\
&\quad - \begin{pmatrix} M_{K-I,I} \\ M_{\overline{K},I} \end{pmatrix} (M_{II})^{-1} \begin{pmatrix} M_{I,K-I} & M_{I,\overline{K}} \end{pmatrix} \\
&= \begin{pmatrix} M_{K-I,K-I} - M_{K-I,I}(M_{II})^{-1}M_{I,K-I} & M_{K-I,\overline{K}} - M_{K-I,I}(M_{II})^{-1}M_{I,\overline{K}} \\ M_{\overline{K},K-I} - M_{\overline{K},I}(M_{II})^{-1}M_{I,K-I} & M_{\overline{K},\overline{K}} - M_{\overline{K},I}(M_{II})^{-1}M_{I,\overline{K}} \end{pmatrix}
\end{aligned}
$$

$\ldots(\alpha)$

Se observa claramente que:

(M_{KK}/M_{II}) es una submatriz principal de (M/M_{II}). $\forall I \subseteq K$ $[\ ]$

(c) Suponiendo que Q es la matriz inversa de M , donde:

$$Q = \begin{pmatrix} Q_{II} & Q_{I\bar{I}} \\ Q_{\bar{I}I} & Q_{\bar{I}\bar{I}} \end{pmatrix},$$

Si Q es la inversa de M, de $M.Q = I_{mm}$:

$$M_{II}.Q_{II} + M_{I\bar{I}}.Q_{\bar{I}I} = I_{II}$$

$$M_{II}.Q_{I\bar{I}} + M_{I\bar{I}}.Q_{\bar{I}\bar{I}} = 0_{I\bar{I}}$$

también, de $Q.M = I_{mm}$:

$$Q_{\bar{I}I}.M_{II} + Q_{\bar{I}\bar{I}}.M_{\bar{I}I} = 0_{\bar{I}I}$$

$$Q_{\bar{I}I}.M_{I\bar{I}} + Q_{\bar{I}\bar{I}}.M_{\bar{I}\bar{I}} = I_{\bar{I}\bar{I}}$$

Suponiendo que M_{II} es no-singular, despejando para $Q_{II}, Q_{I\bar{I}}, Q_{\bar{I}I}, Q_{\bar{I}\bar{I}}$ se obtiene:
$Q_{II} = M_{II}^{-1} + (M_{II}^{-1}.M_{I\bar{I}})D^{-1}(M_{\bar{I}I}.M_{II}^{-1})$
$Q_{I\bar{I}} = -(M_{II}^{-1}.M_{I\bar{I}}).D^{-1}$
$Q_{\bar{I}I} = -D^{-1}.(M_{\bar{I}I}.M_{II}^{-1})$
$Q_{\bar{I}\bar{I}} = D^{-1}$
donde:

$$D = M_{\bar{I}\bar{I}} - M_{\bar{I}I}M_{II}^{-1}M_{I\bar{I}}$$

Es decir, la matriz inversa de M está dada por:

$$M^{-1} = \begin{pmatrix} M_{II}^{-1} + (M_{II}^{-1}.M_{I\bar{I}})D^{-1}(M_{\bar{I}I}.M_{II}^{-1}) & -(M_{II}^{-1}.M_{I\bar{I}}).D^{-1} \\ -D^{-1}.(M_{\bar{I}I}.M_{II}^{-1}) & D^{-1} \end{pmatrix} \ldots (\dagger)$$

Como K está dado tenemos:

$$M = \begin{pmatrix} M_{I,I} & M_{I,K-I} & M_{I,\overline{K}} \\ M_{K-I,I} & M_{K-I,K-I} & M_{K-I,\overline{K}} \\ M_{\overline{K},I} & M_{\overline{K},K-I} & M_{\overline{K},\overline{K}} \end{pmatrix} = \begin{pmatrix} M_{KK} & M_{K\overline{K}} \\ M_{\overline{K}K} & M_{\overline{K}\overline{K}} \end{pmatrix}$$

$$\begin{aligned}(M/M_{KK}) &= M_{\overline{K}\overline{K}} + \\ &- \begin{pmatrix} M_{\overline{K},I} & M_{\overline{K},K-I} \end{pmatrix} . \begin{pmatrix} M_{II} & M_{I,K-I} \\ M_{K-I,I} & M_{K-I,K-I} \end{pmatrix}^{-1} . \begin{pmatrix} M_{I,\overline{K}} \\ M_{K-I,\overline{K}} \end{pmatrix}\end{aligned}$$

En virtud de (†), la matriz inversa de $\begin{pmatrix} M_{II} & M_{I,K-I} \\ M_{K-I,I} & M_{K-I,K-I} \end{pmatrix}^{-1}$ es:

$$= \begin{pmatrix} M_{II}^{-1} + (M_{II}^{-1}.M_{I,K-I})D^{-1}(M_{K-I,I}.M_{II}^{-1}) & -M_{II}^{-1}.M_{I,K-I}.D^{-1} \\ -D^{-1}.M_{K-I,I}.M_{II}^{-1} & D^{-1} \end{pmatrix}$$

donde: $D = M_{K-I,K-I} - M_{K-I,I}.(M_{II})^{-1}.M_{I,K-I}$

Reemplazando valores tenemos:

$$\begin{aligned}(M/M_{KK}) &= M_{\overline{K}\overline{K}} - \{M_{\overline{K}I}M_{II}^{-1}M_{I\overline{K}} + M_{\overline{K}I}M_{II}^{-1}M_{I,K-I}D^{-1}M_{K-I,I}M_{II}^{-1}M_{I\overline{K}} + \\ &-M_{\overline{K},K-I}D^{-1}M_{K-I,I}M_{II}^{-1}M_{I\overline{K}} - M_{\overline{K}I}M_{II}^{-1}M_{I,K-I}D^{-1}M_{K-I,\overline{K}} + \\ &+M_{\overline{K},K-I}D^{-1}M_{K-I,\overline{K}}\}\end{aligned}$$

Agrupando convenientemente:

$$\begin{aligned}(M/M_{KK}) &= M_{\overline{K}\overline{K}} - M_{\overline{K}I}M_{II}^{-1}M_{I\overline{K}} + \\ &-\left(M_{\overline{K},K-I} - M_{\overline{K}I}M_{II}^{-1}M_{I,K-I}\right)D^{-1}\left(M_{K-I,\overline{K}} - M_{K-I,I}M_{II}^{-1}M_{I\overline{K}}\right)\end{aligned}$$

De acuerdo a (α):

$$= \left((M/M_{II})/(M_{KK}/M_{II})\right)$$

$$\Longrightarrow (M/M_{KK}) = \left((M/M_{II})/(M_{KK}/M_{II})\right) \quad \square$$

Proposición 3.1.2 *Si M_{II} es no singular entonces M es no singular si, y solo si el Complemento Schur,(M/M_{II}), es también no singular.*

Demostración:

$(\Longrightarrow)$:

Supongamos que M es no-singular, es decir, $Det(M) \neq 0$, y dado que $Det(M_{II}) \neq 0$ pues M_{II} es no-singular entonces en virtud a la proposición $(3.1.1-(a))$:

$$Det[(M/M_{II})] = \frac{Det[M]}{Det[M_{II}]} \neq 0$$

Por lo tanto:

(M/M_{II}) es no-singular.

$(\Longleftarrow)$:

De manera similar:

En virtud a la proposición $(3.1.1-(a))$:

$$Det[M] = Det[(M/M_{II})].Det[M_{II}] \neq 0$$

Pues: $Det[(M/M_{II})]$ y $Det[M_{II}]$ son diferentes de cero, ya que (M/M_{II}) y M_{II} son matrices no-singulares.

Por lo tanto M es no-singular. $\square$

Definición 3.1.2 *Una matriz cuadrada real es llamada una P-Matriz si todos sus principales menores son positivos* [**9**].

Proposición 3.1.3 *Si M es una P-matriz entonces todas las submatrices principales de M y sus Complementos Schur son P-Matrices.*

Demostración:

Supongamo que M es una matriz cuadrada de orden m.

Probaremos primero que las submatrices principales de M son P-Matrices:

Sea M_{KK} una submatriz principal de M, $K \subseteq \{1, \ldots, m\}$.
Si $I \subset K \subseteq \{1, \ldots, m\}$ entonces de la forma como I está dado, M_{II} es una submatriz principal de M_{KK} y por lo tanto también es una submatriz principal de M.
Como M e una P-Matriz:

$$Det[M_{II}] > 0 \qquad \forall I \subseteq \{1, \ldots, m\}$$

Y como $I \subseteq K \subseteq \{1, \ldots, m\}$

$$Det[M_{II}] > 0 \qquad \forall I \subset K$$

Por lo tanto:

M_{KK} es una P-Matriz.

Luego probaremos que los Complementos Schur de M son P-Matrices.

Sea (M/M_{II}), Complemento Schur de M. Si $I \subseteq K \subseteq \{1, \ldots, m\}$, Para I arbitrario, luego de acuerdo a la proposición $(3.1.1-(b))$ tenemos (M_{KK}/M_{II}) es una submatriz principal de (M/M_{II}). Como M es una P-Matriz entonces $Det[M_{II}] > 0$ y $Det[M_{KK}] > 0$ pues M_{II}, M_{KK} son submatrices principales. Ahora, en virtud de la proposición $(3.1.1-(a))$:

$$Det[(M_{KK}/M_{II})] = \frac{Det[M_{KK}]}{Det[M_{II}]} > 0$$

Como I es arbitrario entonces todos los principales menores de (M/M_{II}) son positivos, por lo tanto:

(M/M_{II}) es una P-Matriz. [ ]

De la ecuación (3.6) se obtiene:

$$du_\gamma = -H_\gamma(z) + \nabla g_\gamma(x)^T dx.$$

Denotando

$\Delta(z) = \nabla F(x) + \sum_{i\in\alpha(z)} u_i \nabla^2 g_i(x) + \sum_{j=1}^{p} v_j \nabla^2 h_j(x),$

$du_\beta^+ = \max(0, du_\beta),$

$q_\beta(z) = H_\beta(z),$

$y = \begin{pmatrix} dx \\ du_\alpha \\ dv \end{pmatrix},$

$q_y(z) = \begin{pmatrix} H_F(z) \\ H_\alpha(z) \\ H_h(z) \end{pmatrix},$

$B(z) = \begin{pmatrix} \nabla g_\beta(x) \\ 0 \\ 0 \end{pmatrix},$ ....(Θ1)

$L(z) = \begin{pmatrix} \Delta(z) & \nabla g_\alpha(x) & \nabla h(x) \\ -\nabla g_\alpha(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix},$ ....(Θ2)

$P(z) = \begin{pmatrix} L(z) & B(z) \\ -B(z)^T & 0 \end{pmatrix},$

Entonces de (3.3) − (3.5) y (3.7) tenemos:

$$L(z)y + B(z)du_\beta^+ + q_y(z) = 0, \tag{3. }$$

$$[-B(z)^T y + q_\beta(z)]^T du_\beta^+ = 0, \tag{3.9}$$

$$-B(z)^T y + q_\beta(z) \geq 0, \tag{3.10}$$

$$du_\beta^+ \geq 0. \tag{3.11}$$

De la ecuacione (3.) − (3.11) definimos un ***problema de complementariedad lineal mixto***.

Si $L(z)$ es una matriz no singular, se puede resolver y en términos de du_β^+, de (3.8):

$$y = -L^{-1}(z)B(z)du_\beta^+ - L(z)^{-1}q_y(z)$$

Luego sustituyendo y en (3.9) − (3.10):

$$\left.\begin{aligned}
[B(z)^T L^{-1}(z)B(z)du_\beta^+ + B(z)^T L^{-1}(z)q_y(z) + q_\beta(z)]du_\beta^+ = 0 \\
B(z)^T[L^{-1}(z)B(z)du_\beta^+ + L^{-1}(z)q_y(z)] + q_\beta(z) \geq 0 \\
du_\beta^+ \geq 0
\end{aligned}\right\}\cdots(*)$$

El sistema resultante se convierte entonces en un ***problema de complementariedad lineal*** en du_β.

Proposición 3.1.4 *Considerando el siguiente problema de complementariedad lineal LCP(q,M): Dado $M \in \mathbb{R}^{n\times n}$, $q \in \mathbb{R}^n$. Encontrar $z \in \mathbb{R}^n$ sujeto a:*

$w - Mz = q, \quad w \in \mathbb{R}^n$

$w, z \geq 0,$

$w^T z = 0.$

Decimos que el problema LCP(q,M) tiene solución única, si M es una P-Matriz.

Demostración:

Supongamos que el problema LCP(q,M) no tiene solución única, es decir:

sean z_1 y z_2 solución del problema LCP(q,M). Entonces:

$w_1 - Mz_1 = q \qquad w_2 - Mz_2 = q$

$w_1, z_1 \geq 0 \qquad w_2, z_2 \geq 0$

$w_1^T.z_1 = 0 \qquad w_2^T.z_2 = 0$

Luego:

$w_1 - Mz_1 = w_2 - Mz_2 \Longrightarrow w_1 - w_2 = M(z_1 - z_2)$

$(w_1 - w_2)(z_1 - z_2)^T = M\|z_1 - z_2\|^2 \Longrightarrow (w_1 - w_2)(z_1^T - z_2^T) = M\|z_1 - z_2\|^2$

$w_1.z_1^T - w_1.z_2^T - w_2.z_1^T + w_2.z_2^T = M\|z_1 - z_2\|^2 \Longrightarrow -(w_1.z_2^T + w_2.z^T) = M\|z_1 - z_2\|^2$

Por consiguiente : M es una matriz negativa, en particular el elemento $M_{11} < 0$, lo cual es una contradicción pues M es una P-Matriz $(Det[M_{II}] > 0, \quad \forall I \subseteq \{1, .., n\})$.

Por lo tanto:

Si M es una P-Matriz entonces el problema LCP(q,M) tiene solución única. $\square$

La matriz en el problema de complementariedad lineal (*) se convierte entonces en $B(z)^T L(z)^{-1} B(z)$, lo cual es el Complemento Schur de $L(z)$ en $P(z)$. Haciendo:

$M = B(z)^T L^{-1}(z) B(z)$

$q = B(z)^T L^{-1}(z) q_y(z) + q_\beta(z)$

$z = du_\beta^+$

$w = B(z)^T L^{-1}(z) B(z) du_\beta^+ + B(z)^T L^{-1}(z) q_y(z) + q_\beta(z)$

y en virtud de la proposición (3.1.4), si $B(z)^T L(z)^{-1} B(z)$ es una P-Matriz, entonces el problema de complementariedad lineal $(*)$ tiene una solución única.

Una condición suficiente de $(3.8) - (3.11)$ para tener una solución única es que:

- La matriz $L(z)$ es no-singular
- El Complemento Schur de $L(z)$ en $P(z)$, $B(z)^T L(z)^{-1} B(z)$ es una P-Matriz.

3.2 Regularidad y Regularidad débil

Definición 3.2.1 (Regularidad) .- *Sea* $z = (x, u, v)^T$ *un vector arbitrario en* $\mathbb{R}^{n+m+p}$. *Entonces* z *es llamado un vector regular para la función* H *definido en* (2.19) *si:*

(*a*) $L(z) = \begin{pmatrix} \Delta(z) & \nabla g_\alpha(x) & \nabla h(x) \\ -\nabla g_\alpha(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix}$
es no-singular.

(*b*) *El Complemento Schur de* $L(z)$, *dado por* $B(z)^T L(z)^{-1} B(z)$ *es una P-Matriz, donde* $L(z)$ *y* $B(z)$ *están definidos en* (Θ2) *y* (Θ1) *respectivamente.*

Luego, si z es un vector regular entonces el sistema de ecuaciones $H(z) + H'(z, d) = 0$ tendrá solución única, con lo cual se obtendrá el vector dirección d.
La condición (a) de regularidad es similar a la condición de no-singularidad en *La Programación no-lineal* cuando aplicamos el método de Newton a ecuaciones F-diferenciables. Sin embargo, ya que H es solamente B-diferenciable, tenemos que imponer la condición (b) para obtener una dirección de Newton. En otras palabras, la condición (b), es adicionado para tener la suficiente fuerza en resolver ca os no-diferenciables. En lugar de poner todos los casos no-diferenciables en uno. Aquí "distinguiremos" las situaciones y encontraremos la condiciones bajo lo cual una dirección de Newton puede aún ser conseguido incluso i la F-derivada general no existe. Claramente, cuando el conjunto β en z es vacío, la regularidad es equivalente a la usual condición de no-singularidad.

Para resolver $y - (dx, du_\alpha, dv)^T$, se tiene que invertir la matriz de orden $(n+|\alpha|+p) \times (n+|\alpha|+p)$ el cual usualmente tiene menor dim nsión qu el Jacobiano de la función Lagrangiano. Es decir, $(n+m+p) \times (n+m+p)$. Si, además, la matriz $\Delta(z)$ es no-singular, resolver el problema para y es d scomponerlo en dos partes:

E decir, olucionar dx n (3.3) n t´rmino d du_α, du_β^+ y dv, ntonc s olucionar para du_α y dv n t´rmino d du_β^+ d (3.4) y (3.7). n éste caso, l sfu rzo d cálculo pu d ad má s r reducido por qu la dimensión d matrices a ser inv rtidas se convi rt n n $n \times n$ y $(|\alpha| + p) \times (|\alpha| + p)$ r pectivament .

Lema 3.2.1 *Sea M una matriz cuadrada real dada en forma particionada*

$$M = \begin{pmatrix} M_{II} & M_{I\bar{I}} \\ M_{\bar{I}I} & M_{\bar{I}\bar{I}} \end{pmatrix}$$

Supongamos que M_{II} es no-singular y que el Complemento Schur (M/M_{II}) es una P-Matriz. Si $I \subseteq K \subseteq \{1, \ldots, m\}$, entonces M_{KK} es no-singular y el Complemento Schur (M/M_{KK}) es una P-Matriz. En particular, M es no-singular.

Demostración:

Como K está dado, tenemos M_{II} es una submatriz principal no-singular de M_{KK}. En virtud a la proposición $(3.1.1 - (b))$, tenemos que el Complem nto Schur: (M_{KK}/M_{II}) es una submatriz principal del Complemento Schur (M/M_{II}).
De acuerdo a la proposición (3.1.3) y abiendo que (M/M_{II}) una P-Matriz ten mos entonc s que (M_{KK}/M_{II}) s una P-Matriz pu (M_{KK}/M_{II}) es una submatriz principal de (M/M_{II}). Además el compl mento Schur de (M_{KK}/M_{II}) n (M/M_{II}) denotado como:

$$(M/M_{II})/(M_{KK}/M_{II})), \qquad \text{s una P-Matriz} \qquad \ldots(*)$$

Como (M_{KK}/M_{II}) es una P-Matriz s d cir todas su principal s m nores son positivos entonces n particular (M_{KK}/M_{II}) es no-singular.
Si M_{II} es no-singular y l compl mento Schur (M_{KK}/M_{II}) s no-singular ntonc s por la proposición (3.1.2), M_{KK} es n singular.
n virtud a la proposición $(3.1.1 - (c))$: $(M/M_{KK}) = (M/M_{II})\ (M_{KK}/M_{II})$
Lu go, d acu rdo a (*) t n mos nton s qu : (M/M_{KK}) s una P-Matriz. $\square$

Proposición 3.2.1 *De acuerdo a como fueron definido $P(z)$, $L(z)$, y $B(z)$. Si $P(z)$ y $B(z)^T L(z)^{-1} B(z)$ son funciones continuas en z, entonces existe una vecindad N de z tal que para cada $z' \quad (x', u', v')^T \in N$*

$$(i)\ \ L'(z') = \begin{pmatrix} \Delta(z') & \nabla g_\alpha(x') & \nabla h(z') \\ -\nabla g_\alpha(x')^T & 0 & 0 \\ -\nabla h(x')^T & 0 & 0 \end{pmatrix} \text{ es no-singular.}$$

(ii) El Complemento Schur $B'(z')^T L'(z')^{-1} B'(z')$ es una P-matriz.

donde L' y B' estan definidas de manera análoga a L y B, y son evaluados en z' con el conjunto α definido en z.

Demostración

(i).- Como $P(.)$ es una función continua en z entonces la función matriz:

$$L : \mathbb{R}^{n+m+p} \to \mathbb{R}^{(n+|\alpha|+p)(n+|\alpha|+p)} \quad \text{es una función continua en z.}$$

Por lo tanto su función determinante:

$$Det[L(.)] : \mathbb{R}^{(n+|\alpha|+p)(n+|\alpha|+p)} \to \mathbb{R} \quad \text{también será una función continua en z.}$$

debido a que z es un vector regular entonces $L(z)$ es no-singular, $(Det[L(z)] \neq 0)$.
Supongamos que $Det[L(z)] > 0$. Por la continuidad de $Det[L(z)]$:
Para cada $\epsilon > 0, \quad \exists \delta > 0 \quad / \quad Det[L(z)] - \epsilon < Det[L'(z')] < Det[L(z)] + \epsilon$ siempre que $z' \in N(z, \delta)$.
Elegimos el δ que corresponde a $\epsilon = \dfrac{Det[L(z)]}{2}$ (Como $Det[L(z)] > 0 \to \epsilon > 0$) de manera que

$$\frac{Det[L(z)]}{2} < Det[L'(z')] < \frac{3Det[L(z)]}{2}$$

Siempre que $z' \in N(z, \delta)$.
Por lo tanto: $Det[L'(z')] > 0$ siempre que $z' \in N(z, \delta)$.

Supongamos que $Det[L(z)] < 0$. De manera análoga :$Det[L'(z')] < 0$ siempre que $z' \in N(z, \delta)$.

Por lo tanto: $L'(z')$ es no-singular siempre que $z' \in N(z, \delta)$.

(ii).- El Complemento Schur de $L(z)$ en $P(z)$ es:

$$(P(z)/L(z)) = 0_{|\beta|\times|\beta|} - (-B(z)^T.L(z)^{-1}.B(z)) = B(z)^T.L(z)^{-1}.B(z)$$

debido a que las matrices $B(z)$ y $L(z)$ son funciones continuas entonces la función matriz:

$$\begin{array}{rccl} T: & \mathbb{R}^{n+m+p} & \longrightarrow & \mathbb{R}^{|\beta|\times|\beta|} \\ & z & \longrightarrow & T(z) = (P(z)/L(z)) \end{array}$$

es una función continua en z.

Por lo tanto su función determinante $Det[T]$ es continua en z, luego:

$$Det[T_{kk}(z)] : \mathbb{R}^{|\beta|\times|\beta|} \to \mathbb{R}, \qquad \forall k = 1, \ldots, |\beta|$$

es también continua en z.

Donde T_{kk} es una submatriz principal de $(P(z)/L(z))$.

debido a que z es un vector regular, $(P(z)/L(z))$ es una P-Matriz entonces todos sus principales menores son positivos, es decir: $Det[T_{kk}(z)] > 0 \qquad \forall k = 1, \ldots, |\beta|$

Por la continuidad de $Det[T_{kk}(z)]$ en z:

$$\text{Para cada} \quad \epsilon > 0, \quad \exists\delta > 0 / \quad Det[T_{kk}(z)] - \epsilon < Det[T'_{kk}(z')] < Det[T_{kk}(z)] + \epsilon$$

Siempre que $z' \in N(z, \delta)$

Elegimos el δ que corresponde a $\epsilon = \dfrac{Det[T_{kk}(z)]}{2}$ (Como $Det[T_{kk}(z)] > 0 \to \epsilon > 0$)

de manera que

$$\frac{Det[T_{kk}(z)]}{2} < Det[T'_{kk}(z')] < \frac{3Det[T_{kk}(z)]}{2}$$

Siempre que $z' \in N(z, \delta)$.

Entonces : $Det[T'_{kk}(z')] > 0 \quad \forall k = 1, \ldots, |\beta|$.

Por lo tanto:

$(P'(z')/L'(z')) = B'(z')^T L'(z')^{-1} B'(z')$ es una P-matriz siempre que $z' \in N(z, \delta)$. $\square$

La siguiente proposición fue establecido por **Pang** para su formulación de ecuaciones no-lineales.

Proposición 3.2.2 *Sea $z = (x, u, v)^T$ un vector regular para la función H definida por (2.19). Entonces existe una vecindad N de z tal que cada vector en N es también regular. Además el vector d lo cual resuelve*

$$H(z^k) + H'(z^k, d) = 0,$$

para cada $z^k \in N$ es acotado: es decir.,$\|d\| \leq \xi; < \infty$ para cada norma vectorial $\|.\|$ en $\mathbb{R}^n$.

Demostración

Supongamos que z es un vector regular de la función H. Como F, g, y h son continuamente diferenciable entonces de acuerdo a como está definido $P(z)$, $L(z)$, y $B(z)$, tenemos que $P(z)$ y $B(z)^T L(z)^{-1} B(z)$ son funciones continuas en z, entonces de acuerdo a la proposición (3.2.1), existe una vecindad N de z tal que para cada $z' = (x', u', v')^T \in N$, la matriz :

$$L'(z') = \begin{pmatrix} \Delta(z') & \nabla g_\alpha(x') & \nabla h(z') \\ -\nabla g_\alpha(x')^T & 0 & 0 \\ -\nabla h(x')^T & 0 & 0 \end{pmatrix} \text{ es no-singular.}$$

y el Complemento Schur $B'(z')^T L'(z')^{-1} B'(z')$ es una P-matriz.

Donde L' y B' estan definidas de manera análoga a L y B, y son evaluados en z' con el conjunto α definido en z. Uno puede restringir N mucho más tal que para cada $z' \in N$:

$$\alpha(z) \subseteq \alpha(z'), \quad \gamma(z) \subseteq \gamma(z').$$

Con ideremos la matriz:

$$P(z') = \begin{pmatrix} L(z') & B(z') \\ -B(z')^T & 0 \end{pmatrix}$$

Tenemos:

$dim(L(z')) = n + |\alpha(z')| + p, \quad dim(L'(z')) = n + |\alpha(z)| + p$

dado que:

$$|\alpha(z)| \leq |\alpha(z')| \Longrightarrow dim(L'(z')) \leq dim(L(z'))$$

Es decir: $L'(z')$ es una submatriz principal de $L(z')$

Consideremos la matriz:

$$P'(z') = \begin{pmatrix} L'(z') & B'(z') \\ -B'(z')^T & 0 \end{pmatrix}$$

tenemos:

$dim(P(z')) = n + |\alpha(z')| + p + |\beta(z')| = n + m + p - |\gamma(z')|$

$dim(P'(z')) = n + |\alpha(z)| + p + |\beta(z)| = n + m + p - |\gamma(z)|$

debido a:

$$|\gamma(z)| \leq |\gamma(z')| \Longrightarrow dim(P(z')) \leq dim(P'(z'))$$

Es decir: $P(z')$ es una submatriz principal de $P'(z')$.

Luego, de acuerdo a la proposición (3.2.1):

$L'(z')$ es no-singular, y

$(P'(z')/L'(z'))$ es una P-Matriz.

Además, siendo $L'(z')$ una submatriz principal de $L(z')$, aplicamos entonces el lema (3.2.1), por consiguiente:

- $L(z')$ es no-singular.
- $(P'(z')/L(z'))$ es una P-Matriz.

Como $P(z')$ e una submatriz principal de $P'(z')$ entonces en virtud a la proposición $(3.1.1-b)$ tenemos:
$(P(z')/L(z'))$ es una submatriz principal de $(P'(z')/L(z'))$.
Luego por la proposición (3.1.3) tenemos:
Si $(P'(z')/L(z'))$ es una P-Matriz entonces $(P(z')/L(z'))$ es una P-Matriz, luego:

- $L(z')$ es no-singular.
- $(P(z')/L(z'))$ es una P-Matriz.

Concluimos por lo tanto que: z' es un vector regular.

Para probar que la solución es acotada en una vecindad de un vector regular, sea la única solución del sistema $(3.8)-(3.11)$ en $\tilde{z} \in N$:
$\begin{pmatrix} \tilde{y} \\ d\tilde{u}_\beta^+ \end{pmatrix}$, con
$d\tilde{u}_\beta^+ = \begin{pmatrix} d\tilde{u}_\beta^{++} \\ 0 \end{pmatrix}, \qquad d\tilde{u}_\beta^{++} > 0.$
donde $d\tilde{u}_\beta^{++} > 0$ representa aquellos componentes de $d\tilde{u}_\beta^+$ los cuales son estrictamente positivos, y $\mathbf{0}$ representa al vector cero.
De $(3.8)-(3.9)$ tenemos:
$L(\tilde{z})\tilde{y} + B(\tilde{z})d\tilde{u}_\beta^+ + q_y(\tilde{z}) = 0$
$[-B(\tilde{z})^T\tilde{y} + q_\beta(\tilde{z})]^T d\tilde{u}_\beta^+ = 0$
Luego:

$$L(\tilde{z})\tilde{y} + B(\tilde{z}) \begin{pmatrix} d\tilde{u}_\beta^{++} \\ 0 \end{pmatrix} + q_y(\tilde{z}) = L(\tilde{z})\tilde{y} + B_+(\tilde{z})d\tilde{u}_\beta^{++} + q_y(\tilde{z}) = 0, \tag{3.12}$$

$$[-B(\tilde{z})^T\tilde{y} + q_\beta(\tilde{z})]^T \begin{pmatrix} d\tilde{u}_\beta^{++} \\ 0 \end{pmatrix} = -B_+(\tilde{z})^T\tilde{y} + q_\beta^{++}(\tilde{z}) = 0, \tag{3.13}$$

donde $B_+(\tilde{z})$ y q_β^{++} son submatriz y subvector de $B(\tilde{z})$ y q_β respectivamente.
Dado que z es un vector regular entonces $P(z)$ es invertible:

Veamo : Si z es un vector regular entonces $L(z)$ es no-singular y el Complemento Schur de $L(z)$ en $P(z)$ e una P-Matriz, aplicando el Lema (3.2.1), $P(z)$ es no ingular entonces $P(z)$ es invertible.

Como P es continua en z además sabiendo que P^{-1} existe, entonces P^{-1} será continua en z, es decir:

$\forall \epsilon > 0, \quad \exists \delta > 0 \quad / \quad \forall z' \in N(z, \delta) \Longrightarrow \|P^{-1}(z) - P^{-1}(z')\| < \epsilon$

donde $\|.\|$ es una *norma matricial.*

Restringiendo la vecindad N, debe existir una cota superior uniforme $c > 0$ (depende solamente del vector z) tal que:

$\forall z' \in N(z, \delta) \Longrightarrow \|P^{-1}(z')\| < \|P^{-1}(z)\| + \epsilon = c$

Así, $P^{-1}(z')$ y la inversa de cualquier submatriz principal de $P(z')$ es uniformemente acotada en norma por c.

Es fácil ver que la matriz

$$P_+(\tilde{z}) = \begin{pmatrix} L(\tilde{z}) & B_+(\tilde{z}) \\ -B_+(\tilde{z})^T & 0 \end{pmatrix}$$

es una submatriz principal de $P(\tilde{z})$. Por lo tanto:

$$\begin{pmatrix} \tilde{y} \\ d\tilde{u}_\beta^{++} \end{pmatrix} = \begin{pmatrix} L(\tilde{z}) & B_+(\tilde{z}) \\ -B_+(\tilde{z})^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} -q_y(\tilde{z}) \\ -q_\beta^{++}(\tilde{z}) \end{pmatrix}$$

está acotada, y la cota depende solamente de z. [ ]

La condición de regularidad es bastante fuerte y no es satisfecha en muchos casos. Así, la aplicación del método de solución de arriba es muy limitado. Para hacer el algoritmo más aplicable, introduciremos el concepto de Regularidad Débil.

Definición 3.2.2 (Regularidad débil) *Sea $z = (x, u, v)^T$ un vector arbitrario en $\mathbb{R}^{n+m+p}$. El vector z es llamado un vector regular débilmente para la función H definido en (2.19) si existe una vecindad N de z tal que para cada $z' = (x', u', v')^T \in N$, la matriz*

$$L(z') = \begin{pmatrix} \Delta(z') & \nabla g_{\alpha'}(x') & \nabla h(z') \\ -\nabla g_{\alpha'}(x')^T & 0 & 0 \\ -\nabla h(x')^T & 0 & 0 \end{pmatrix}$$

es no-singular, donde $\alpha' = \alpha(z')$ y $L(z')$ está definido de manera análoga a $L(z)$.

Obviamente, si un vector es regular, esto es también regular débilmente. El siguiente ejemplo demuestra sin embargo, que lo inverso no es generalmente cierto.

Ejemplo 3.1 *Consideremos el Problema de Complementariedad No-lineal (NCP) donde*

$$F(x) = \begin{pmatrix} x_1^2 + x_1 + 2x_2 - 2 \\ 2x_1^2 - 2x_2 + x_2^2 - 2 \end{pmatrix} \tag{3.14}$$

De acuerdo a (2.14) – (2.15), el problema NCP puede ser formulado como el siguiente sistema de ecuaciones:

$$H(z) = \begin{pmatrix} F(x) - u^+ \\ x + u^- \end{pmatrix} = 0$$

donde $H : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, y $z = (x, u)^T \in \mathbb{R}^{2n}$. Luego:

$H_1(z) = x_1^2 + x_1 + 2x_2 - 2 - u_1^+ = 0$

$H_2(z) = 2x_1^2 - 2x_2 + x_2^2 - 2 - u_2^+ = 0$

$H_3(z) = x_1 + u_1^- = 0$

$H_4(z) = x_2 + u_2^- = 0$

Sea $z = (x, u)^T = (-\frac{1}{2}, 0, 1, 0)^T$ entonces:

$\alpha(z) = \{1\}$,

$\beta(z) = \{2\}$,

$\gamma(z) = \emptyset$.

En el problema NCP, tenemos:

$\Delta(z) = \nabla F(x)$,

$$B(z) = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix},$$

$$L(z) = \begin{pmatrix} \nabla F(x) & -I_\alpha^T \\ I_\alpha & 0 \end{pmatrix}.$$

$$P(z) = \begin{pmatrix} \nabla F(x) & -I_\alpha^T & -I_\beta^T \\ I_\alpha & 0 & 0 \\ I_\beta & 0 & 0 \end{pmatrix}$$

donde: I_α de orden $|\alpha| \times n$ es una submatriz de I, cuyas filas son indexadas por el conjunto α. I_β definido de manera análoga.

Evaluando, tenemos:

$$\nabla F(x) = \begin{pmatrix} 2x_1 + 1 & 2 \\ 4x_1 & 2x_2 - 2 \end{pmatrix} \Longrightarrow \nabla F(x)|_{(-\frac{1}{2},0)} = \begin{pmatrix} 0 & 2 \\ -2 & -2 \end{pmatrix},$$

$$I_\alpha = \begin{pmatrix} 1 & 0 \end{pmatrix},$$
$$I_\beta = \begin{pmatrix} 0 & 1 \end{pmatrix},$$

luego:

$$L(z) = \begin{pmatrix} 0 & 2 & -1 \\ -2 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix} \Longrightarrow L^{-1}(z) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{2} & -1 \\ -1 & -1 & -2 \end{pmatrix},$$ se observa que $L(z)$ es no-singular

Además:

$$P(z) = \begin{pmatrix} 0 & 2 & -1 & 0 \\ -2 & -2 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\Longrightarrow B(z)^T L(z)^{-1} B(z) = \begin{pmatrix} 0 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 0 & -\frac{1}{2} & -1 \\ -1 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = -\frac{1}{2} < 0$$

Entonce el complemento Schur de $L(z)$ en $P(z)$, dado por $B(z)^T L(z)^{-1} B(z)$ no e una P-Matriz. Por lo tanto z no es un vector regular, pues no satisface la condición (b) de la definición de regularidad (3.2.1). Sin embargo z es un vector regular débilmente, en efecto:

Sólo ba ta tomar una Vecindad N de z tal que $\forall z' \in N$, $\alpha(z) \subset \alpha(z')$, pero teniendo en cuenta que: $|\alpha(z)| = 1 \Rightarrow |\alpha(z')| \leq 2$, debido a que $u = (u_\alpha, u_\beta, u_\gamma)^T$, $u \in \mathbb{R}^2$. Luego $\beta(z') = \emptyset$, $\gamma(z') = \emptyset$. Entonces $L(z') = P'(z')$, luego por la continuidad de P, $P'(z')$ es no-singular, es decir $L(z')$ es no-singular. Por lo tanto z es un vector regular débilmente.

De este modo, la regularidad débil es de hecho más débil que la regularidad. $\square$

La proposición (3.2.1) describe la continuidad de la noción de regularidad. En otras palabras, si un vector z es regular, entonces en una pequeña vecindad de z es también regular. Del mismo modo, la regularidad débil posee la misma propiedad de continuidad.

Proposición 3.2.3 *Sea $z = (x, u, v)^T$ un vector regular débilmente para la función H. Entonces existe una vecindad N de z tal que cada vector en N es también regular débilmente.*

Demostración:

Debido a que z es un vector regular débilmente entonces existe una vecindad N de z tal que $\forall z' \in N$, $L(z')$ es no-singular

Probaremos que z' es un vector regular débilmente:

Elegimos una vecindad N' de z' de tal forma que $N'(z') \subset N(z)$, luego como cualquier vector $\tilde{z} \in N(z), L(\tilde{z})$ es no-singular, entonces en particular: $L(\tilde{z})$ también será no-singular $\forall \tilde{z} \in N'(z')$. Es decir existe una vecindad N' de z' tal que $\forall \tilde{z} \in N'$ se tiene $L(\tilde{z})$ es no-singular.

Por lo tanto z' es un vector regular débilmente. $\square$

Corolario 3.2.1 *Si $L(z)$ es una matriz no-singular y si $\beta(z) = \emptyset$ entonces z es un vector regular débilmente.*

Demostración:

Debido a que $\beta(z)$ $\emptyset$, existe una vecindad N de z tal que $\forall z' \in N$, $\beta(z') = \emptyset$. Luego, i re tringimos más la vecindad tal que $\alpha(z) = \alpha(z')$ $\Rightarrow L(z') = L'(z')$. Debido a que $L'(z')$ e no-singular entonces $L(z')$ es no-singular. Por lo tanto z es un vector regular débilmente . $\square$

Corolario 3.2.2 *Sea $L(z)$ una matriz no singular y $\beta_0(z)$ cualquier subconjunto de $\beta(z)$. Si la matriz*

$$\bar{L}(z) = \begin{pmatrix} \Delta(z) & \nabla g_{\alpha\cup\beta_0}(x) & \nabla h(x) \\ -\nabla g_{\alpha\cup\beta_0}(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix}$$

es no-singular $\forall \beta_0(z) \subseteq \beta(z)$, entonces z es un vector regular débilmente de H.

Demostración:

Debido a la continuidad de $L(z)$, y teniendo en cuenta que $L(z)$ es no-singular, entonces existe una vecindad N tal que $\forall z' \in N$, $L'(z')$ es no-singular. Ahora restringiendo la vecindad N tal que $\forall z' \in N$.

$\alpha(z) \subseteq \alpha(z')$

$\gamma(x) \subseteq \gamma(z')$

Por otra parte, si hacemos $\alpha(z') = \alpha(z) \cup \beta_0(z)$, $\beta_0(z) \subseteq \beta(z)$, es decir: $L(z')$ es no-singular. Por lo tanto z es un vector regular débilmente de H. $\square$

La regularidad débil desempeña un papel importante en el algoritmo modificado de Newton(amortiguado) que será pres ntado en el siguiente capítulo.

4 EL MÉTODO MODIFICADO DE NEWTON AMORTIGUADO

Este Capítulo pre enta una modificación del método de Newton (amortiguado), que es aplicable cuando la condición de regularidad no es satisfecha o cuando el sistema de ecuaciones $H(z^k)+H'(z^k,d^k) = 0$ no tenga solución, se presenta además algunos resultados que se necesitaran para la convergencia del algoritmo.

Ejemplo 4.1 *Consideremos el siguiente problema:*

$$Min \quad f(x) = 4x_1 - x_1^2 - x_2^2 + 2x_2$$

Sujeto a

$$g(x) = x_1^2 - 1 \leq 0$$
$$h(x) = x_1 - 2x_2 + 1 = 0$$

Este Problema es equivalente al siguiente sistema de ecuaciones:

$$\begin{aligned} H_1(z) &= 4 - 2x_1 + 2x_1u^+ + v \\ H_2(z) &= 2 - 2x_2 - 2v \\ H_3(z) &= -x_1^2 + 1 + u^- \\ H_4(z) &= -x_1 + 2x_2 - 1 \end{aligned}$$

Sea $z = (x, u, v)^T = (1, 1, 0, 0)^T$, donde $\alpha(z) = \emptyset; \beta(z) = 1; \gamma(z) = \emptyset$.
Luego para encontrar una dirección descente, resolveremos el siguiente Problema de Complementariedad Mixto Lineal(PCM) Mixto :

$$-2dx_1 + dv + 2du^+ + 2 = 0$$
$$-2dx_2 - 2dv = 0$$
$$-dx_1 + 2dx_2 = 0$$
$$(-2dx_1)du^+ \geq 0$$
$$-2dx_1 \geq 0, \quad du^+ > 0$$

E te PCL Mixto no tiene solución. Usaremos entonces el método modificado de Newton.

4.1 Propiedad Descendente

Demo traremos primero que el algoritmo de Newton (Amortiguado) descrito en la sección anterior genera una dirección descendente para la función norma $\theta(z)$.

Proposición 4.1.1 *Sea la dirección buscada d^k obtenida al resolver el sistema de ecuaciones:*

$$H(z^k) + H'(z^k, d^k) = 0. \tag{4.1}$$

Entonces

$$\theta'(z^k, d^k) = -2\theta(z^k) \leq 0. \tag{4.2}$$

En particular. $\theta'(z^k, d^k) < 0$ si $H(z^k) \neq 0$.

Demostración:

Por definición:

$\theta(z^k) = \frac{1}{2}H(z^k)^T H(z^k) = \frac{1}{2}\|H(z^k)\|^2 \geq 0$ y

$\theta'(z^k, d^k) = H(z^k)^T H'(z^k, d^k)$

Como d^k es obtenida de (4.1), tenemos:

$$H(z^k) + H'(z^k, d^k) = 0 \Longrightarrow H'(z^k, d^k) = -H(z^k)$$

Por lo tanto:

$\theta'(z^k, d^k) = H(z^k)^T.(-H(z^k)) = -H(z^k)^T.H(z^k) = -2\theta(z^k) \leq 0,$

y es solamente igual a cero cuando $H(z^k) = 0$.

En particular:

Si $H(z^k) \neq 0 \Longrightarrow \theta(z^k) > 0$, Luego $\theta'(z^k, d^k) = -2\theta(z^k) < 0$. $\square$

El siguiente corolario demue tra que una búsqueda lineal de Armijo puede ser realizado sucesivamente a través de un número finito de pasos.

Corolario 4.1.1 *Supongamos que z^k no es solución de $H(z) = 0$. Entonces existe un escalar $\tau_k \in (0, 1]$ tal que para todo $t \in < 0, \tau_k >$,*

$$\theta(z^k) - \theta(z^k + td^k) \geq 2\sigma t\theta(z^k)$$

donde $\sigma \in < 0, \frac{1}{2} >$

Demostración:

Para $t \in [0, 1]$, sea

$s(t) = \theta(z^k) - \theta(z^k + td^k) - 2\sigma t\theta(z^k)$

Entonces, $s(0) = \theta(z^k) - \theta(z^k) - 0 = 0$

Luego:

$$\begin{aligned} s'(0^+) = \lim_{t\to 0^+} \frac{s(t)}{t} &= \lim_{t\to 0^+} \frac{\theta(z^k) - \theta(z^k + td^k) - 2\sigma t\theta(z^k)}{t} \\ &= \lim_{t\to 0^+} \frac{\theta(z^k) - \theta(z^k + td^k)}{t} - 2\sigma\theta(z^k) \\ &= -\theta'(z^k, d^k) - 2\sigma\theta(z^k) \end{aligned}$$

Como:

$$\theta'(z^k, d^k) = -2\theta(z^k) \Longrightarrow s'(0^+) = 2\theta(z^k) - 2\sigma\theta(z^k) = 2(1 - \sigma)\theta(z^k)$$

Debido a que: $\sigma < \dfrac{1}{2} \Longrightarrow s'(0^+) > 0$. Por lo tanto, s es una función creciente.

Para $t \geq 0$ tenemos entonces $s(t) \geq 0$ luego:

$$s(t) = \theta(z^k) - \theta(z^k + td^k) - 2\sigma t\theta(z^k) \geq 0 \Longrightarrow \theta(z^k) - \theta(z^k + td^k) \geq 2\sigma t\theta(z^k) \quad \square$$

La interpretación del corolario (4.1.1) nos indica: si z^k no es un cero de la función H, entonces moviendose a lo largo de la dirección d^k nos permitirá decrecer lo "suficiente" a la función norma $\theta(z)$.

El algoritmo de Newton(amortiguado) genera una dirección descendente para la función norma $\theta(z)$. Las direcciones generadas por el algoritmo de Newton (amortiguado) no son de ningún modo las únicas direcciones descendentes para la función norma $\theta(z)$. Obviamente, la ventaja de la dirección de Newton es que ellos generan una razón cuadrática de convergencia.

4.2 Método modificado de Newton

Para modificar el algoritmo de Newton(amortiguado) y superar el problema mencionado anteriormente, necesitamos los siguientes resultados.

Definición 4.2.1 *Una función $F(x_1, x_2, \ldots, x_n)$ se llama homogénea de grado p si para todos los valores del parámetro λ y una cierta constante p se tiene la identidad* $F(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) = \lambda^p F(x_1, x_2, \ldots, x_n)$

Proposición 4.2.1 *Sea $d \in \mathbb{R}^{n+m+p}$. Si la derivada direccional de H en z a lo largo de la dirección d es homogénea de grado uno en d. (Se cumple $H'(z, -d) = -H'(z, d)$), entonces con tal de que $\theta'(z, d) \neq 0$, d ó $-d$ será una dirección descendente para la función norma $\theta(z)$.*

Demostración

Por definición tenemos: $\theta'(z, d) = H(z)^T H'(z, d)$

Si $\theta'(z, d) > 0$ (es decir, si d no es una dirección descendente), entonces debemos tener:

$\theta'(z, -d) = H(z)^T H'(z, -d) = -H(z)^T H'(z, d) = -\theta'(z, d) < 0 \Longrightarrow \theta'(z, -d) < 0$

Por lo tanto, $-d$ es una dirección descendente.

Si $\theta'(z, -d) > 0$, entonces debemos tener:

$\theta'(z, d) = H(z)^T H'(z, d) = -H(z)^T H'(z, -d) = -\theta'(z, -d) < 0 \Longrightarrow \theta'(z, d) < 0$

Por lo tanto, d es una dirección descendente. $\square$

El método modificado de Newton (amortiguado) es descrito como sigue:

1. Si el sistema (4.1) tiene una solución, realizaremos la búsqueda lineal a lo largo de la dirección d^k; en otro caso, fijamos $du_\beta = 0$, y resolver (3.3), (3.4), (3.7) y (3.6) para la dirección dx, du_α, dv y du_γ. Sea $d^k = (dx, du_\alpha, 0, du_\gamma, dv)^T$

2. Si d^k es una dirección descendente de θ en z, realizaremos una búsqueda lineal a lo largo de la dirección d^k para generar $z^{k+1} = z^k + \lambda_k d^k$; en otro caso, buscamos a lo largo de $-d^k$ para calcular $z^{k+1} = z^k - \lambda_k d^k$.

Cuando (4.1) tiene una solución, una dirección de Newton es obtenida. Esta dirección tiene toda la información de la derivada de H en z. Si el sistema no tiene solución, una dirección generada por la modificación anterior no es de Newton y contiene solamente información parcial de la derivada de H en z. Sin embargo, si el conjunto β es pequeño, lo cual usualmente se produce en la práctica, la dirección generada por el método modificado será una buena aproximación de la dirección de Newton.

Proposición 4.2.2 *Sea z un vector regular débil de H definido por* (2.19). *Si $\theta'(z, d) \neq 0$ donde d es la dirección generada por el algoritmo descrito anteriormente, entonces el método modificado de Newton(amortiguado) generará una dirección descendente para la función norma θ en z.*

Demostración

De acuerdo a la definición de regularidad débil, se asegura la existencia de $L^{-1}(z)$, por lo tanto dx, du_α, dv y du_γ puede ser resuelto para cualquier valor de du_β.
De acuerdo a la proposición 4.2.1, sólo es suficiente d mostrar que la derivada direccional generada por el algoritmo modificado de Newton es de grado uno en d.

Denotemos

$$d = (dx, du_\alpha, du_\beta, du_\gamma, dv)^T$$

Debido a que $\theta'(z,d) \neq 0$, z no es solución del sistema (4.1). Luego, fijamos $du_\beta = 0$, entonces:

$$H'(z,d) = \begin{pmatrix} \Delta(z)dx + \nabla g_\alpha(x)du_\alpha + \nabla h(x)dv \\ -\nabla g_\alpha(x)^T dx \\ -\nabla g_\beta(x)^T dx \\ -\nabla g_\gamma(x)^T dx + du_\gamma \\ -\nabla h(x)^T dx \end{pmatrix}$$

$$H'(z,-d) = -\begin{pmatrix} \Delta(z)dx + \nabla g_\alpha(x)du_\alpha + \nabla h(x)dv \\ -\nabla g_\alpha(x)^T dx \\ -\nabla g_\beta(x)^T dx \\ -\nabla g_\gamma(x)^T dx + du_\gamma \\ -\nabla h(x)^T dx \end{pmatrix}$$

Entonces:

$H'(z,-d) = -H'(z,-d)$

4.3 Regularidad débil

Se puede ver de la demostración anterior que bajo la supo ición de regularidad débil, el método modificado de Newton(amortiguado) solo fallará si $H(z) \neq 0$ y $\theta'(z,d) = 0$. Claramente, si $H(z) \neq 0$ y z es regular, la proposición (4.1.1) asegura una dirección descendente. Así, el método modificado de ewton nunca fallará en un punto regular. La siguiente proposición demuestra, cuando la regularidad es ausente, el método modificado de Newton(amortiguado) generará una dirección descendente si la regularidad débil y algunas condiciones mild son satisfechas.

Proposición 4.3.1 *Sea z un vector regular débilmente de H definido por* (2.19). *Si $H(z) \neq 0$, entonces:*
Siempre existe una dirección descendente d en z para la función norma θ a menos que se cumpla las **condiciones mild***:*

$$(i)\ H_\gamma(z) = 0$$

$$(ii)\ \nabla g_\alpha(x)^T H_F(z) = 0$$

$$(iii)\ \nabla h(x)^T H_F(z) = 0$$

$$(iv)\ H_F(z)^T \Delta(z) + g_\alpha(x)^T \nabla g_\alpha(x)^T + g_\beta(x)^T \nabla g_\beta(x)^T + h(x)^T \nabla h(x)^T = 0$$

$$(v)\ \nabla g_\beta(x)^T H_F(z) \geq 0$$

$$(vi)\ g_\beta(x) \geq 0$$

En particular, el método modificado de Newton(amortiguado) obtendrá una dirección descendente a menos que $(i) - (iv)$ sean satisfecha simultaneamente.

Demostración:

Supongamos que no existe una dirección descendente para la función norma en z, entonces para cualquier dirección $d \in \mathbb{R}^{n+m+p}$, tenemos $\theta'(z,d) \geq 0$. Por definición:
$\theta'(z,d) = H(z)^T H'(z,d)$
donde: $H(z)^T = \begin{pmatrix} H_F(z)^T & -g_\alpha(x)^T & -g_\beta(x)^T & H_\gamma(z)^T & -h(x)^T \end{pmatrix}$ y

$$H'(z,d) = \begin{pmatrix} \Delta(z)dx + \nabla g_\alpha(x)du_\alpha + \nabla g_\beta(x)\max(0, du_\beta) + \nabla h(x)dv \\ -\nabla g_\alpha(x)^T dx \\ -\nabla g_\beta(x)^T dx + \min(0, du_\beta) \\ -\nabla g_\gamma(x)^T dx + du_\gamma \\ -\nabla h(x)^T dx \end{pmatrix}$$

Luego:

$$\theta'(z,d) = H(z)^T H'(z,d)$$

$$
\begin{aligned}
& [H_F(z)^T\Delta(z) + g_\alpha(x)^T\nabla g_\alpha(x)^T + g_\beta(x)^T\nabla g_\beta(x)^T + \\
& - H_\gamma(z)^T\nabla g_\gamma(x)^T + h(x)^T\nabla h(x)^T]dx + \\
& + H_F(z)^T\nabla g_\alpha(x)du_\alpha + H_F(z)^T\nabla h(x)dv + H_\gamma(z)^T du_\gamma + \\
& + H_F(z)^T\nabla g_\beta(x)\max(0, du_\beta) - g_\beta(x)^T\min(0, du_\beta) \\
> \ & 0
\end{aligned} \tag{4.3}
$$

Fijando $dx = 0$, $du_\alpha = 0$, $du_\beta = 0$, $dv = 0$, y $du_\gamma \neq 0$, tenemos:
$H_\gamma(z)^T du_\gamma \geq 0 \qquad \forall du_\gamma \neq 0.$
Obviamente, este es solamente cierto cuando $H_\gamma(z) = 0$.
Por el mismo argumento:

- Fijando $dx = 0$, $dv = 0$, $du_\gamma = 0$, $du_\beta = 0$ y $du_\alpha \neq 0$, tenemos:
  $H_F(z)^T\nabla g_\alpha(x)du_\alpha \geq 0 \qquad \forall du_\alpha \neq 0.$
  este es solamente cierto cuando $H_F(z)^T\nabla g_\alpha(x) = 0$.

- Fijando $dx = 0$, $du_\gamma = 0$, $du_\beta = 0$, $du_\alpha = 0$ y $dv \neq 0$ tenemos:
  $H_F(z)^T\nabla h(x)dv \geq 0 \qquad \forall dv \neq 0.$
  este es solamente cierto cuando $H_F(z)^T\nabla h(x) = 0$.

- Fijando $du_\gamma = 0$, $du_\beta = 0$, $du_\alpha = 0$, $dv = 0$, $dx \neq 0$ tenemos:
  $[H_F(z)^T\Delta(z) + g_\alpha(x)^T\nabla g_\alpha(x)^T + g_\beta(x)^T\nabla g_\beta(x)^T + h(x)^T\nabla h(x)^T]dx = 0$
  $\forall dx \neq 0$
  este es solamente cierto cuando:

  $$H_F(z)^T\Delta(z) + g_\alpha(x)^T\nabla g_\alpha(x)^T + g_\beta(x)^T\nabla g_\beta(x)^T + h(x)^T\nabla h(x)^T = 0$$

Fijando $dx = 0$, $du_\alpha = 0$, $dv = 0$, $du_\gamma = 0$, y $du_\beta > 0$, tenemos:
$\nabla g_\beta(x)^T H_F(z) \geq 0$
Fijando $dx = 0$, $du_\alpha = 0$, $dv = 0$, $du_\gamma = 0$, y $du_\beta < 0$, tenemos:
$g_\beta(x) \geq 0.$

En particular: Si no se cumple alguna de las condiciones $(i) - (iv)$ entonces el método modificado de Newton amortiguado obtendrá una dirección descendente. En efecto:

De (4.3), fijamos $du_\beta = 0$ y elegimos apropiadamente valores para dx, du_α, dv y du_γ, es decir:

Sean $dx = 0$, $du_\alpha = 0$, $dv = 0$ y $du_\gamma > 0$. Además, si asumimos que la condición (i) no se cumple, es decir $(H_\gamma(z) < 0)$. Luego: $\theta'(z,d) < 0 = H_\gamma(z)^T du_\gamma < 0$ $\square$

Corolario 4.3.1 *Sea z un vector regular débilmente de H definido por* (2.19).
Si $H(z) \neq 0$, entonces:
Si $g_\beta(x) < 0$ implica que θ tiene una dirección descendente.

Demostración

Fijando $dx = 0$, $du_\alpha = 0$, $du_\gamma = 0$, $dv = 0$ y $du_\beta < 0$, tenemos de (4.3):

$$\begin{aligned}
\theta'(z,d) &= H_F(z)^T \nabla g_\beta(x) \max(0, du_\beta) - g_\beta(x)^T \min(0, du_\beta) \\
&= -g_\beta(x)^T \min(0, du_\beta) \\
&= -g_\beta(x)^T du_\beta < 0
\end{aligned}$$

Entonces:
$\theta'(z,d) < 0$. Es decir θ tiene una dirección descendente. $\square$

Semejante al caso regular, cuando un vector z, que no es solución, es regular débilmente, una dirección descendente generada por el algoritmo modificado de Newton (Amortiguado) es también acotado. Además, en una suficiente pequeña vecindad de tal vector, tenemos los siguientes resultados generales:

Proposición 4.3.2 *Sea z un vector regular débilmente de H. Si $\theta'(z,d) \neq 0$ donde d es la dirección generada por el algoritmo modificado de Newton (amortiguado), entonces existe una vecindad N de z tal que para cada vector $z' \in N$, la dirección generada por el algoritmo modificado de Newton (amortiguado) en z' es acotado.*

Demostración:

De acuerdo a la proposición (3.2.3), si z es un vector regular débilmente entonces existe una vecindad N de z tal que cada vector en N es también regular débilmente, luego si $z' = (x', u', v')^T \in N$ entonces z' es un vector regular débilmente.

Sea $z' = (x', u', v')^T \in N$ un vector y $d = (dx, du_\alpha, 0, du_\gamma)^T$ un vector dirección del algoritmo modificado de Newton (amortiguado). Sin perder la generalidad, asumamos que d es realmente una dirección descendente en z'. Debido a que z' es regular débilmente entonces $L(z')$ es no-singular por lo tanto, de la ecuacion (3.8), tenemos:

$$\begin{pmatrix} dx \\ du_\alpha \\ dv \end{pmatrix} = -L(z')^{-1}.q_y(z')$$

Además:

$du_\gamma = \nabla g_\gamma(x')^T dx - H_\gamma(z')$

donde: $L(z')$ y $q_y(z')$ son definidos como antes.

Como L es continua en z' además L^{-1} existe, entonces L^{-1} es continua en una vecindad $N(z, \delta)$ y por lo tanto $L(z')^{-1}$ es acotada uniformemente en norma por c, lo cual es solamente determinado por z.

du_γ también es acotado pues dx está acotado.

Por lo tanto d es acotado. [ ]

4.4 Resultados para la convergencia del algoritmo

De la discu ión previa, conocimos que mientras el sistema (4.1) tenga una solución, podemos encontrar una dirección descendente para la función norma. La búsqueda lineal entonces asegura que la sucesión $\{\theta(z^k)\}$ es estrictamente decreciente. De aquí, i el conjunto de nivel:

$$\{z \in \mathbb{R}^{n+m+p} : \|H(z)\| \leq \|H(z^0)\|\}$$

es acotado, la sucesión generada por el algoritmo tendrá al menos un punto de acumulación. Para establecer que el algoritmo converja, los siguientes resultados seran necesarios.

Lema 4.4.1 *Sea $V \subseteq \mathbb{R}^n$ y $G : V \to \mathbb{R}^n$ continuamente diferenciable. Sea G_i la i-ésima componente de G y ∇G es el jacobiano de G. Si para cada sucesión $\{z^k\}$, $\lim_{k\to\infty} \|z^k\| = \infty$,*

$$\lim_{k\to\infty} G_i(z^k) = \xi_i < \infty \qquad \forall i = 1, 2, \ldots, n,$$

y si

$$\nabla G^* = \lim_{k\to\infty} \nabla G(z^k)$$

existe, entonces ∇G^* es una matriz singular.

Demostración:

Como $\lim_{k\to\infty} \|z^k\| = \infty$ entonces $\{z^k\}$ es una sucesión no-acotada, entonces al menos uno de sus componentes no será acotada. Sin perder generalidad, asumimo que $\lim_{k\to\infty} z_1^k = \infty$. Sea ∇G^{1*} la primera columna de ∇G^*. Por la suposición, ∇G^* exi te. Por lo tanto podemos escribir:

$$\nabla G^{1*} = \lim_{k\to\infty} (\frac{\partial G_1(z^k)}{\partial z_1}, \frac{\partial G_2(z^k)}{\partial z_1}, \ldots, \frac{\partial G_n(z^k)}{\partial z_1})^T = (\eta_1, \eta_2, \ldots, \eta_n)^T$$

Demostraremos que ∇G^{1*} es un vector cero:

Como: $\lim\limits_{k\to\infty} G_i(z^k) = \xi_i < \infty \quad \forall i = 1,2,\ldots,n$ entonces dado $\epsilon > 0$, existe $K_1 \in \mathbb{N}$ tal que $\forall k > K_1$ y $\forall m > 0$ implica

$$|G_i(z^{k+m}) - G_i(z^k)| \leq \frac{\epsilon}{2} \Longrightarrow -\frac{\epsilon}{2} \leq G_i(z^{k+m}) - G_i(z^k) \leq \frac{\epsilon}{2}$$

donde $G_i : V \to \mathbb{R}$

Sea $z_i^{k+m} = z_i^k$ para $i = 2,3,\ldots,n$, y $z_1^{k+m} > z_1^k$.

Luego, dado que : $G \in C^1(V) \Longrightarrow G_i \in C^1(V) \qquad \forall i = 1,2,\ldots,n$

Además por el *Teorema del Valor Medio*:

Si G_i es diferenciable sobre un intervalo abierto que contiene el segmento rectilineo que va desde z_1^k hasta z_1^{k+m} entonces existe un $\lambda_i \in [0,1]$ tal que:

$$G_i(z^{k+m}) - G_i(z^k) = (z_1^{k+m} - z_1^k)\frac{\partial G_i}{\partial z_1}(z_i^*)\ldots(*)$$

donde $\quad z_i^* = \lambda_i z^{k+m} + (1-\lambda_i)z^k, \qquad 0 \leq \lambda_i \leq 1$

Luego de $(*)$ tenemos:

$$|G_i(z^{k+m}) - G_i(z^k)| = |(z_1^{k+m} - z_1^k)\frac{\partial G_i}{\partial z_1}(z_i^*)|$$

Pero como:

$$|G_i(z^{k+m}) - G_i(z^k)| \leq \frac{\epsilon}{2} \Longrightarrow |(z_1^{k+m} - z_1^k)\frac{\partial G_i}{\partial z_1}(z_i^*)| \leq \frac{\epsilon}{2}$$

De manera similar: como $\lim\limits_{k\to\infty} \dfrac{\partial G_i}{\partial z_1}(z^k)$ existe, entonces:

Para cada $\epsilon > 0$, existe $K_2 \in \mathbb{N}$ tal que $\forall k > K_2$:

$$|\frac{\partial G_i}{\partial z_1}(z^k) - \eta_i| \leq \frac{\epsilon}{2} \quad \Longrightarrow \quad |\eta_i - \frac{\partial G_i}{\partial z_1}(z^k)| \leq \frac{\epsilon}{2} \Longrightarrow |\eta_i| - |\frac{\partial G_i}{\partial z_1}(z^k)| \leq \frac{\epsilon}{2}$$
$$\Longrightarrow \quad |\eta_i| - \frac{\epsilon}{2} \leq |\frac{\partial G_i}{\partial z_1}(z^k)|$$

Sea $K = \max\{K_1, K_2\}$, entonces $\forall k > K$ y $m > 0$

$$(|\eta_i| - \frac{\epsilon}{2})|(z_1^{k+m} - z_1^k)| \leq |\frac{\partial G_i}{\partial z_1}(z^k)||(z_1^{k+m} - z_1^k)| \leq \frac{\epsilon}{2}$$

Luego:

$(|\eta_i| - \frac{\epsilon}{2})|(z_1^{k+m} - z_1^k)| \leq \frac{\epsilon}{2}, \quad \forall \epsilon > 0$

Probaremos que $|\eta_i| \leq \epsilon \quad \forall \epsilon > 0$

En efecto: Supongamos que existe un $\epsilon > 0$ tal que $|\eta_i| > \epsilon \Longrightarrow |\eta_i| - \frac{\epsilon}{2} > 0$

Escogiendo: $\epsilon = (|\eta_i| - \frac{\epsilon}{2})|(z_1^{k+m} - z_1^k)| > 0$

Tenemos:

$$(|\eta_i| - \frac{\epsilon}{2})|(z_1^{k+m} - z_1^k)| \leq \frac{(|\eta_i| - \frac{\epsilon}{2})|(z_1^{k+m} - z_1^k)|}{2}$$

Lo cual es una contradicción, por lo cual concluimos:

$|\eta_i| \leq \epsilon \quad \forall \epsilon > 0 \qquad \Longrightarrow \qquad \eta_i = 0 \quad \forall i$

Por lo tanto: ∇G^{1*} es un vector cero. Entonces ∇G^* es singular $\square$

Teorema 4.4.1 *Sea F continuamente diferenciable, $g_i \quad (i = 1, \ldots, m)$ y $\quad h_j$ $(j = 1, \ldots, p)$ son dos veces continuamente diferenciable, y H es la función definida por (2.19). Sea $z^0 \in \mathbb{R}^{n+m+p}$ y el conjunto de indices $\alpha \subset \{1, 2, \ldots, m\}$ arbitrario. Si para cualquier vector $y = (x, u_\alpha, v)^T \in \mathbb{R}^{n+m_\alpha+p}$ con $u_\alpha > 0$, la matriz:*

$$L(y) = \begin{pmatrix} \Delta(y) & \nabla g_\alpha(x) & \nabla h(x) \\ -\nabla g_\alpha(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix}$$

es no-singular, y para cualquier sucesión y^k con

$\lim_{k \to \infty} \|y^k\| = \infty$

el limite de la matriz función $L(y^k)$ existe y es no-singular, entonces el conjunto de nivel

$\{z \in \mathbb{R}^{n+m+p} : \|H(z)\| \leq \|H(z^0)\|\}$

es acotada.

Demostración:

Sea

$$M = \{z \in \mathbb{R}^{m+n+p} \ / \ \|H(z)\| \leq \|H(z^0)\|\}$$

Supongamos que M no es acotada. Entonces existe una sucesión $\{w^k\}(w^k \in M)$ que no es acotada, es decir:

$\lim\limits_{k\to\infty} \|w^k\| = \infty$

Y además

$\|H(w^k)\| \leq \|H(z^0)\|$

Denotemos:

$\alpha(z) = \{i/u_i > 0\}$

$\beta(z) = \{i/u_i = 0\}$

$\gamma(z) = \{i/u_i < 0\}$

Debido a que

$\lim\limits_{k\to\infty} \|w^k\| = \infty$

Implica que: $\lim\limits_{k\to\infty} \|y^k\| = \infty, \quad y^k = (x^k, u^k_\alpha, v^k)^T$

y/o

$\lim\limits_{k\to\infty} \|u^k_\gamma\| = \infty$

Consideremos dos situaciones:

(i.-) $\lim\limits_{k\to\infty} \|y^k\| = \infty$

Debido a que existe solamente un número finito de combinaciones de los conjuntos indices, podemos asumir que existe una subsucesión de $\{w^k\}$. Sea $\{\tilde{w}^k\}$ dicha subsucesión, en donde:

$\tilde{w}^1 = (\tilde{x}^1, \tilde{u}^1_{\alpha(\tilde{w}^1)}, \tilde{u}^1_{\beta(\tilde{w}^1)}, \tilde{u}^1_{\gamma(\tilde{w}^1)}, \tilde{v}^1)^T$

$\tilde{w}^2 = (\tilde{x}^2, \tilde{u}^2_{\alpha(\tilde{w}^2)}, \tilde{u}^2_{\beta(\tilde{w}^2)}, \tilde{u}^2_{\gamma(\tilde{w}^2)}, \tilde{v}^2)^T$

$\vdots$

$\tilde{w}^k = (\tilde{x}^k, \tilde{u}^k_{\alpha(\tilde{w}^k)}, \tilde{u}^k_{\beta(\tilde{w}^k)}, \tilde{u}^k_{\gamma(\tilde{w}^k)}, \tilde{v}^k)^T$

Tal que:

$\alpha(\tilde{w}^1) = \alpha(\tilde{w}^2) = \ldots = \alpha(\tilde{w}^k) = \ldots,$

$\beta(\tilde{w}^1) = \beta(\tilde{w}^2) = \ldots = \beta(\tilde{w}^k) = \ldots,$

$\gamma(\tilde{w}^1) = \gamma(\tilde{w}^2) = \ldots = \gamma(\tilde{w}^k) = \ldots$

Sea: $\tilde{w}^k = (\tilde{x}^k, \tilde{u}_\alpha^k, \tilde{u}_\beta^k, \tilde{u}_\gamma^k, \tilde{v}^k)^T = (\tilde{y}^k, \tilde{u}_\beta^k, \tilde{u}_\gamma^k)^T$

donde: $\tilde{y}^k = (\tilde{x}^k, \tilde{u}_\alpha^k, \tilde{v}^k)^T$

Además, debido a que

$\|H(\tilde{w}^k)\| \leq \|H(z^0)\|$,

y en virtud al teorema de Bolzano-Weierstrass, como $\{\|H(\tilde{w}^k)\|\}$ es una sucesión acotada, entonces existe una subsucesión convergente. Sin perder la generalidad, sea:

$\lim\limits_{k\to\infty} \|H(\tilde{w}^k)\| = \eta < \infty$.

Denotemos:

$$H_0(\tilde{y}^k) = \begin{pmatrix} H_F(\tilde{y}^k) \\ -g_\alpha(\tilde{x}^k) \\ -h(\tilde{x}^k) \end{pmatrix},$$

donde $\tilde{y}^k = (\tilde{x}^k, \tilde{u}_\alpha^k, \tilde{v}^k)^T$.

Como H_F, g_α, h son continuamente diferenciable entonces H_0 es continuamente diferenciable. Aplicando el lema 4.4.1 pues:

- $\lim\limits_{k\to\infty} \|\tilde{y}^k\| = \infty$
- H_0 es continuamente diferenciable
- $\lim\limits_{k\to\infty} \|H(\tilde{w}^k)\| < \infty \Longrightarrow \begin{cases} \lim\limits_{k\to\infty} H_F(\tilde{y}^k) < \infty \\ \lim\limits_{k\to\infty} g_\alpha(\tilde{x}^k) < \infty \\ \lim\limits_{k\to\infty} h(\tilde{x}^k) < \infty \end{cases}$
- $\lim\limits_{k\to\infty} \nabla H_0(\tilde{y}^k) = \lim\limits_{k\to\infty} L(\tilde{y}^k)$ existe.

Entonces:

$\lim\limits_{k\to\infty} \nabla H_0(\tilde{y}^k) = \lim\limits_{k\to\infty} L(\tilde{y}^k)$ es una matriz singular lo cual contradice la suposición que $\lim\limits_{k\to\infty} L(\tilde{y}^k)$ es no-singular.

(ii.-) Consideremos ahora:

$\lim\limits_{k\to\infty} \|y^k\| < \infty$ y $\lim\limits_{k\to\infty} \|u_\gamma^k\| = \infty$

Denotemos:

$H_\gamma(z) = -g_\gamma(x) + u_\gamma$

Debido a que

$$\|H_\gamma(w^k)\| \leq \|H(w^k)\| \leq \|H(z^0)\|$$

Tenemos:

$$\| - g_\gamma(x^k) + u_\gamma^k\| \leq \|H(z^0)\|$$

Por desigualdad triangular

$$\|u_\gamma^k\| \leq \| - g_\gamma(x^k) + u_\gamma^k\| + \|g_\gamma(x^k)\| \leq \|H(z^0)\| + \|g_\gamma(x^k)\|$$

Entonces:

$$\|u_\gamma^k\| \leq \|H(z^0)\| + \|g_\gamma(x^k)\|$$

Como

$\lim_{k\to\infty} \|y^k\| < \infty \Longrightarrow \lim_{k\to\infty} \|x^k\| < \infty$

Es decir $\{x^k\}$ es una sucesión acotada entonces posee una subsucesión $\{\tilde{x}^k\}$ convergente, sin perder generalidad asumamos que $\{x^k\}$ es convergente, debido a que g es continua entonces $g(\{x^k\})$ es convergente, luego $g(\{x^k\})$ es acotada. Entonces concluimos: $\|u_\gamma^k\| < \infty$. Lo cual es una contradicción $\square$

Teorema 4.4.2 *Supongamos que una sucesión $\{z^k\} = \{(x^k, u^k, v^k)^T\}$ generada por el algoritmo modificado de Newton(amortiguado) es acotado. Si*

1. *$z^* = (x^*, u^*, v^*)^T$ es un punto de acumulación de la sucesión.*

2. *z^* es un vector regular de la función H.*

entonces $H(z^*) = 0$.

Demostración:

Supongamos que $H(z^*) \neq 0$. Por la naturaleza del método modificado de Newton (amortiguado), $\theta(z^k)$ es estrictamente decreciente. Debido a que z^* es regular, el sistema de ecuaciones:

$$H(z^*) + H'(z^*, d^*) = 0$$

tiene una solución d^*, luego:

$$\theta'(z^*, d^*) = -2\theta(z^*) < 0 \Longrightarrow \theta(z^*) > 0 \ldots (\dagger)$$

Además: en virtud a la proposición 3.2.2, d^* es acotada es decir:

$$\lim_{k\to\infty} \|d^k\| = \|d^*\| < \infty,$$

entonces la sucesión de direcciones buscadas d^k es acotada. Sin embargo como z^* es un <u>punto</u> de acumulación de z^k, existe entonces una subsucesión $\{\tilde{z}^k\}$ tal que $\lim_{k\to\infty} \tilde{z}^k \quad z^*$. Sin perder la generalidad asumiremos que $\lim_{k\to\infty} z^k = z^*$.
Ahora, debido a que d^k es una dirección descendente de θ en z entonces:

$$z^{k+1} = z^k + \lambda^k d^k$$

Pero como la sucesión $\{\lambda^k\}$ es acotada, tenemos entonces que: $d^* = 0$. Sabiendo que $H(z^*) \neq 0$ y en virtud al corolario (4.1.1) tenemos que:

Existe un escalar $t \in (0, 1)$ tal que

$$\theta(z^*) - \theta(z^* + td^*) > 2\sigma t\theta(z^*)$$

cuando $d^* \quad 0$ tenemos : $2\sigma t\theta(z^*) < 0 \Longrightarrow \theta(z^*) < 0 \ldots (\dagger\dagger)$
De $(\dagger)$ y $(\dagger\dagger)$ nos lleva a una contradicción, en consecuencia $H(z^*) = 0$ $\square$
Se nota que la suposición de regularidad es usado para inferir que la sucesión de direcciones buscada d^k es acotada.

Teorema 4.4.3 *Supongamos que una sucesión* $\{z^k\} = \{(x^k, u^k, v^k)^T\}$ *generada por el algoritmo modificado de Newton (amortiguado) es acotado. Si*

1. $z^* = \{(x^*, u^*, v^*)^T\}$ *es un punto de acumulación de la sucesión.*
2. z^* *es un vector regular débilmente de la función* H.
3. *Se cumple las condiciones* $(i) - (iv)$ *de la proposición* (4.3.1) *y* $g_\beta(x^*) = 0$

entonces $H(z^*) = 0$

Demostración:

Si $g_\beta(x^*) = 0$, entonces de la condición (iv) en la proposición (4.3.1), tenemos:

$$H_F(z^*)^T \Delta(z^*) + g_\alpha(x^*)^T \nabla g_\alpha(x^*)^T + h(x^*)^T \nabla h(x^*)^T = 0$$

Esto, junto con $(ii) - (iii)$:
$\nabla g_\alpha(x^*)^T H_F(z^*) = 0$
$\nabla h(x^*)^T H_F(z^*) = 0$
demuestra que el sistema de ecuaciones:

$$w^T . \begin{pmatrix} \Delta(z^*) & \nabla g_\alpha(x^*) & \nabla h(x^*) \\ -\nabla g_\alpha(x^*)^T & 0 & 0 \\ -\nabla h(x^*)^T & 0 & 0 \end{pmatrix} = 0$$

tiene una solución

$$w = \begin{pmatrix} H_F(z^*) \\ -g_\alpha(x^*) \\ -h(x^*) \end{pmatrix}$$

Además, debido a que z^* es un vector regular débilmente entonces la matriz

$$\begin{pmatrix} \Delta(z^*) & \nabla g_\alpha(x^*) & \nabla h(x^*) \\ -\nabla g_\alpha(x^*)^T & 0 & 0 \\ -\nabla h(x^*)^T & 0 & 0 \end{pmatrix}$$

es no-singular, por lo tanto el vector w debe ser cero.

E to, junto con la condición (i) : $-g_\gamma(x^*) + u_\gamma = 0$, tenemos:

$$H(z^*) = \begin{pmatrix} H_F(z^*) \\ -g_\alpha(x^*) \\ -g_\beta(x^*) \\ -g_\gamma(x^*) + u_\gamma \\ -h(x^*) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Por lo tanto $H(z^*) = 0$. $\square$

Observación: Si un punto de acumulación z^* es un vector regular entonces se cumplirá la condición 3 del teorema 4.4.3, en efecto:

De acuerdo al teorema 4.4.2, $H(z^*) = 0$, tenemos entonces: $H_\gamma(z^*) = 0$ y

$$\begin{pmatrix} H_F(z^*) \\ -g_\alpha(x^*) \\ -h(x^*) \\ -g_\beta(x^*) \end{pmatrix}^T \cdot \begin{pmatrix} \Delta(z^*) & \nabla g_\alpha(x^*) & \nabla h(x^*) & \nabla g_\beta(x^*) \\ -\nabla g_\alpha(x^*)^T & 0 & 0 & 0 \\ -\nabla h(x^*)^T & 0 & 0 & 0 \\ -\nabla g_\beta(x^*)^T & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Es decir, se cumple la condición 3 del teorema 4.4.3.

Sin embargo, cuando z^* es un punto regular débilmente y la condición 3 es satisfecha en z^*, no implica que z^* sea un vector regular. Así, la si cancia del método modificado de Newton(amortiguado) es que, durante un paso intermedio, e requiere regularidad débil y algunas **condiciones mild** para asegurar una dirección descendente; en un punto de acumulación, se requiere condiciones más débil que regularidad para garantizar la consecución de la solución.

El iguiente ejemplo demuestra que un punto de acumulación puede satisfacer las condiciones 1 y 2 en el teorema 4.4.3 pero no implica que sea un vector regular.

Ejemplo 4.2 *Consideremos el problema de complementariedad no–lineal:*

$(x_1^2 + 2x_1 + 2x_2 - 3)x_1 = 0$

$(2x_1^2 - x_1 - 2x_2 + x_2^2 - 1)x_2 = 0$

donde:

$$F(x) = \begin{pmatrix} x_1^2 + 2x_1 + 2x_2 - 3 \\ 2x_1^2 - x_1 - 2x_2 + x_2^2 - 1 \end{pmatrix}$$

El problema NCP puede ser formulado como el siguiente sistema de ecuaciones:

$H_1(z) = x_1^2 + 2x_1 + 2x_2 - 3 - u_1^+ = 0,$

$H_2(z) = 2x_1^2 - x_1 - 2x_2 + x_2^2 - 1 - u_2^+ = 0,$

$H_3(z) = x_1 + u_1^- = 0,$

$H_4(z) = x_2 + u_2^- = 0$

El algoritmo modificado de Newton(amortiguado) converge a un punto de acumulación $z^* = (1, 0, -1, 0)^T$ con $\gamma(z^*) = 1, \beta(z^*) = 2$ y $\alpha(z^*) = \emptyset$

En el punto $z^* = (x^*, u^*)^T = (1, 0, -1, 0)^T$, tenemos:

$$\nabla F(x^*) = \begin{pmatrix} 4 & 2 \\ 3 & -2 \end{pmatrix}, \qquad I_\beta = (0, 1)$$

Luego:

$$L(z^*) = \begin{pmatrix} 4 & 2 \\ 3 & -2 \end{pmatrix} \implies L^{-1}(z^*) = \begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{14} & -\frac{2}{7} \end{pmatrix}$$

$$P(z^*) = \begin{pmatrix} 4 & 2 & 0 \\ 3 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Se observa que $L^{-1}(z^*)$ es no–singular, además:

$$\left(L(z^*)/P(z^*) \right) = \begin{pmatrix} 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{14} & -\frac{2}{7} \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -\frac{2}{7} < 0$$

Debido a que el Complemento Schur $L(z^*)$ no es una P-Matriz, entonces: z^* no es un punto regular.

Dado que en el NCP $g(\mathbf{x}) = -\mathbf{x}$ entonces : $g((x_1, x_2)) = (-x_1, -x_2)$
$\nabla g((x_1, x_2)) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
Luego: $\nabla g_\beta(x^*) = \nabla g_2((1,0)) = (0,-1)^T; g_\beta(x^*) = 0; H_\gamma(z^*) = -g_\gamma(x^*) + u_\gamma = 0.$
Pero se verificará fácilmente que z^* es un vector regular débilmente, en efecto: sólo basta tomar una vecindad N de z^* de tal forma que $\emptyset = \alpha(z^*) \subseteq \alpha(z') \quad \forall z' \in N$ tal que $L_{\alpha(z')}$ es no-singular. Además satisface $g_\beta(z^*) = 0$ y $H_\gamma(z^*) = 0$. Es decir z^* satisface la condición 3 del teorema 4.4.3. Así, z^* es solución del problema, sin embargo, z^* no es regular.

4.5 Convergencia Cuadrática

Para muchos algoritmos de programación cuadrática secuencial, *El Efecto de Maratos* e refiere a situaciones en lo cual (x^k, λ^k) puede ser arbitrariamente cerca a (x^*, λ^*) y el tamaño de paso puede fallar al reducir la función mérito. En otras palabras, la unidad del tamaño de paso y el resultado de la razón de la convergencia cuadrática se convierte en inalcanzable.

En ésta sección, demostramos que los algoritmos descritos en la sección previa pueden lograr una razón de convergencia cuadrática en una vecindad de un vector límite si ciertas condiciones son satisfechas. Demostraremos que *El Efecto de Maratos* puede ser eliminado del algoritmo descrito anteriormente. Para establecer el principal re ultado, necesitamos asumir que f, g_i $(i = 1, \ldots, m)$, h_j, $(j = 1, \ldots, p)$ son dos veces continuamente diferenciable.

Para probar el siguiente teorema, necesitamos primero establecer el igiente lema.

Lema 4.5.1 *Supongamos que una subsucesión $\{z^k\}$ generada por el algoritmo modificado de Newton (amortiguado) converge a un vector regular débilmente $z^* = (x^*, u^*, v^*)^T$ de la función H definida por* (2.19)*, entonces la sucesión $\{d^k/\|H(z^k)\|\}$ es acotado.*

Demostración

Con ideremos dos casos:

(a) Asumiremos que $d^k = (dx^k, du^k_\alpha, du^k_\beta, du^k_\gamma, dv^k)^T$ es una solución del sistema de ecuaciones:

$$H(z^k) + H'(z^k, d^k) = 0$$

De forma análoga a (3.12) y (3.13) obtenemos el siguiente sistema:

$$L(z^k)y^k + B_+(z^k).du^{k++}_\beta + q_y(z^k) = 0 \tag{4.4}$$

$$-B_+(z^k)^T y^k + q^{++}_\beta(z^k) = 0 \tag{4.5}$$

donde: $B_+(z^k)$ y q^{++}_β son submatriz y subvector de $B(z^k)$ y q_β respectivamente, Por lo tanto:

$$\begin{pmatrix} y^k \\ du^{k++}_\beta \end{pmatrix} = -\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} q_y(z^k) \\ q^{++}_\beta(z^k) \end{pmatrix}$$

donde:

$$y^k = \begin{pmatrix} dx^k \\ du^k_\alpha \\ dv^k \end{pmatrix}$$

y

$$du^{k++}_\beta = \{du^{k+}_\beta : du^{k+}_\beta > 0\}$$

debido a que:

$$\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix}^{-1}$$

existe para z^k suficientemente cerca a z^*, entonces es acotado en norma por una constante c lo cual depende solamente de z^*.

Así, aplicando la desigualdad de Minkowski, tenemos:

$$\|y^k\| \leq \left\| \begin{pmatrix} y^k \\ du_\beta^{k++} \end{pmatrix} \right\| < c. \left\| \begin{pmatrix} q_y(z^k) \\ q_\beta^{++}(z^k) \end{pmatrix} \right\|$$
$$< c.(\|q_y(z^k)\| + \|q_\beta^{++}(z^k)\|) \ldots (*)$$

como:

$$q_y(z) = \begin{pmatrix} H_F(z) \\ H_\alpha(z) \\ H_h(z) \end{pmatrix} \quad \text{y} \quad q_\beta^{++}(z) = H_\beta^{++}(z)$$

consecuentemente:

$$\frac{\|y^k\|}{\|H(z^k)\|} \leq c.\Big(\| \begin{pmatrix} H_F(z) \\ H_\alpha(z) \\ H_h(z) \end{pmatrix} \| / \|H(z^k)\| + \|H_\beta(z^k)^{++}\| / \|H(z^k)\|\Big)$$

notando que

$$\|H(z^k)\|^2 = \|H_F(z^k)\|^2 + \|H_\alpha(z^k)\| + \|H_\beta(z^k)\| + \|H_\gamma(z^k)\|^2 + \|H_h(z^k)\|^2$$

Así:

$$\{y^k / \|H(z^k)\|\} \text{ o } \left\{ \begin{pmatrix} dx^k \\ du_\alpha^k \\ dv^k \end{pmatrix} / \|H(z^k)\| \right\} \text{ es acotado.}$$

De forma similar de $(*)$

$$\|du_\beta^+ + \| \leq \| \begin{pmatrix} y^k \\ du_\beta^{k++} \end{pmatrix} \|$$

entonces $\{du_\beta^{k++} / \|H(z^k)\|\}$ es acotado.

Debido a que

$$du_\gamma^k = -H_\gamma(z^k) + \nabla g_\gamma(x^k)^T dx^k \ldots (\dagger)$$

es decir, du_γ^k es una función lineal de dx^k, y sabiendo que $\{dx^k\}$ es acotado, pues $\{dy^k\}$ es acotado, luego:

$$\frac{\|du_\gamma^k\|}{\|H(z^k)\|} = -\frac{H_\gamma(z^k)}{\|H(z^k)\|} + \nabla g_\gamma(x^k)^T \frac{\|dx^k\|}{\|H(z^k\|}$$

por lo tanto:

$\{du_\gamma^k / \|H(z^k)\|\}$ es acotado.

Entonces: $\{d^k / \|H(z^k)\|\}$ es acotado.

(*b*) Si el sistema de ecuaciones:

$$H(z^k) + H'(z^k, d^k) = 0$$

no tiene solución, entonces la dirección $d^k = (dx^k, du_\alpha^k, 0, du_\gamma^k, dv^k)^T$ es formado fijando $du_\beta = 0$. De (4.4) tenemos:

$L(z^k)y^k + q_y(z^k) = 0$

sabiendo que $L(z^k)$ es no-singular entonces:

$y^k = -L(z^k)^{-1}q_y(z^k)$

Luego

$$\frac{\|y^k\|}{\|H(z^k)\|} = \|\begin{pmatrix} dx^k \\ du_\alpha^k \\ dv^k \end{pmatrix}\| / \|H(z^k)\| = \frac{\|-L(z^k)^{-1}.q_y(z^k)\|}{\|H(z^k)\|} \leq c.\frac{\|q_y(z^k)\|}{\|H(z^k)\|}$$

claramente $\{y^k / \|H(z^k)\|\}$ es acotado por una constante lo cual depende solamente de z^*.

De la misma forma como en (†)

$\{du_\gamma^k / \|H(z^k)\|\}$ es acotada

Por lo tanto: $\{d^k / \|H(z^k)\|\}$ es acotada

Usando este resultado, ahora probaremos el teorema.

Teorema 4.5.1 *Supongamos que una subsucesión* $\{z^k\}$ *generada por el algoritmo modificado de Newton (amortiguado) converge a un vector regular* $z^* = (x^*, u^*, v^*)^T$ *de la función* H *definida por* (2.19), *entonces para* k *suficientemente grande,* $m_k = 0$. *En otras palabras, un tamaño de unidad de paso es eventualmente conseguido.*

Antes demostraremos la siguiente proposición:

Proposición 4.5.1 *Suponiendo que una subsucesión $\{z^k\}$ generada por el algoritmo modificado de Newton(amortiguado) converge a un vector regular $z^* = (x^*, u^*, v^*)^T$ de la función H tal que:*

$$\theta(z^k) - \theta(z^k + d^k) < 2\sigma\theta(z^k)$$

Entonces:

$\lim_{k\to\infty} d^k = 0, \quad y \quad \lim_{k\to\infty} w^k = z^*, \quad donde : w^k = z^k + d^k$

Demostración:

Debido a que:

$\lim_{k\to\infty} H(z^k) = H(z^*) = 0$

y en virtud al lema 4.5.1, la sucesión $\{\frac{d^k}{\|H(z^k)\|}\}$ es acotada, luego:

$$w^k - z^k = d^k \quad \Longrightarrow \quad \frac{\|w^k - z^k\|}{\|H(z^k\|} = \frac{\|d^k\|}{\|H(z^k\|} \leq M \quad \Longrightarrow 0 \leq \|w^k - z^k\| \leq M\|H(z^k)\|$$

por el teorema de Sandwich:

$$\lim_{k\to\infty} \|w^k - z^k\| = 0 \quad \Longrightarrow \quad \lim_{k\to\infty} \|d^k\| \quad \Longrightarrow \lim_{k\to\infty} d^k = 0$$

por lo tanto: $\lim_{k\to\infty} w^k = z^*$ $\square$

Demostración del Teorema:

Por la condición de Armijo, tenemos:

$$\theta(z^k) - \theta(z^k + \mu^{m_k} d^k) \geq 2\sigma\mu^{m_k}\theta(z^k)$$

Entonces, es suficiente demostrar que dado la subsucesión $\{\bar{z}^k\}$, la siguiente desigualdad se cumple:

$$\theta(z^k) - \theta(z^k + d^k) \geq 2\sigma\theta(z^k)$$

Para todo k uficientemente grande, es decir $m_k = 0$

Demo traremos por contradicción:

Suponiendo que la conclusión no es verdadera, entonces existe una subsucesión $\{z^k\}$ tal que:

$$\theta(z^k) - \theta(z^k + d^k) < 2\sigma\theta(z^k) \ldots (\dagger)$$

De acuerdo a la proposición anterior, tenemos:

$$\lim_{k\to\infty} d^k = 0, \quad \text{y} \quad \lim_{k\to\infty} w^k = z^*, \quad \text{donde} : w^k = z^k + d^k$$

Sea:

$$\theta(z^k) - \theta(w^k) = \frac{1}{2}\sum_{i=1}^{5} T_i$$

donde:

T_1, T_2, T_3, T_4 estan definidos de la siguiente manera:

$$\begin{aligned}
T_1 &= H_F(z^k)^T H_F(z^k) - H_F(w^k)^T H_F(w^k) \\
&= \Big(F(x^k) + \nabla g_{\alpha(z^*)}(x^k)u_{\alpha(z^*)} + \nabla h(x^k)v\Big)^T \Big(F(x^k) + \nabla g_{\alpha(z^*)}(x^k)u_{\alpha(z^*)} + \\
&\quad +\nabla h(x^k)v\Big) + \\
&\quad -\Big(F(x^k + dx) + \nabla g_{\alpha(z^*)}(x^k + dx)(u_{\alpha(z^*)} + du_{\alpha(z^*)}) + \nabla h(x^k + dx)(v + dv)\Big)^T \\
&\quad \Big(F(x^k + dx) + \nabla g_{\alpha(z^*)}(x^k + dx)(u_{\alpha(z^*)} + du_{\alpha(z^*)}) + \nabla h(x^k + dx)(v + dv)\Big) \\
T_2 &= H_{\alpha(z^*)}(z^k)^T H_{\alpha(z^*)}(z^k) - H_{\alpha(z^*)}(w^k)^T H_{\alpha(z^*)}(w^k) \\
&= g_{\alpha(z^*)}(x^k)^T g_{\alpha(z^*)}(x^k) - g_{\alpha(z^*)}(x^k + dx)^T g_{\alpha(z^*)}(x^k + dx) \\
T_3 &= H_{\beta(z^*)}(z^k)^T H_{\beta(z^*)}(z^k) - H_{\beta(z^*)}(w^k)^T H_{\beta(z^*)}(w^k) \\
&= g_{\beta(z^*)}(x^k)^T g_{\beta(z^*)}(x^k) - g_{\beta(z^*)}(x^k + dx)^T g_{\beta(z^*)}(x^k + dx) \\
T_4 &= H_{\gamma(z^*)}(z^k)^T H_{\gamma(z^*)}(z^k) - H_{\gamma(z^*)}(w^k)^T H_{\gamma(z^*)}(w^k) \\
&= \Big(-g_{\gamma(z^*)}(x^k) + u_{\gamma(z^*)}\Big)^T \Big(-g_{\gamma(z^*)}(x^k) + u_{\gamma(z^*)}\Big) + \\
&\quad -\Big(-g_{\gamma(z^*)}(x^k + dx) + u_{\gamma(z^*)} + du_{\gamma(z^*)}\Big)^T \Big(-g_{\gamma(z^*)}(x^k + dx) + u_{\gamma(z^*)} + du_{\gamma(z^*)}\Big) \\
T_5 &= H_h(z^k)^T H_h(z^k) - H_h(w^k)^T H_h(w^k) \\
&= h(x^k)^T h(x^k) - h(x^k + dx)^T h(x^k + dx)
\end{aligned}$$

Los conjuntos indices: α, β y γ están definidos en el punto z^*. Restringiendo la vecindad podemos asumir que para k suficientemente grande:

$\alpha(z^*) \subseteq \alpha(z^k)$, y $\gamma(z^*) \subseteq \gamma(z^k)$.

Debido a que la ucesión $\{d^k/\|H(z^k)\|\}$ es acotada, sin perder generalidad asumiremo que $\{d^k/\|H(z^k)\|\}$ converge al vector $\tilde{d}$ $(\tilde{x}, \tilde{u}, \tilde{v})^T$. Es decir:

Si $d^k = (dx^k, du^k, dv^k)^T \implies \dfrac{dx^k}{\|H(z^k)\|} \to \tilde{x}; \quad \dfrac{du^k}{\|H(z^k)\|} \to \tilde{u}; \quad \dfrac{dv^k}{\|H(z^k)\|} \to \tilde{v}$

Además, debido a que la sucesión normalizada $\{H(z^k)/\|H(z^k)\|\}$ es acotada, sin perder generalidad asumiremos que converge al vector $\tilde{H} = (\tilde{H}_F, \tilde{H}_\alpha, \tilde{H}_\beta, \tilde{H}_\gamma, \tilde{H}_h)^T$ es decir:

Si $H = (H_F, H_\alpha, H_\beta, H_\gamma, H_h)^T$. Entonces:

$\dfrac{H_F}{\|H(z^k)\|} \to \tilde{H}_F; \quad \dfrac{H_\alpha}{\|H(z^k)\|} \to \tilde{H}_\alpha; \quad \dfrac{H_\beta}{\|H(z^k)\|} \to \tilde{H}_\beta;$

$\dfrac{H_\gamma}{\|H(z^k)\|} \to \tilde{H}_\gamma; \quad \dfrac{H_h}{\|H(z^k)\|} \to \tilde{H}_h$

Debido a que z^k es un vector regular, para k suficientemente grande, tenemos:

$$H(z^k) + H'(z^k, d^k) = 0 \tag{4.6}$$

Así:

$$\frac{H(z^k)}{\|H(z^k)\|} + \frac{H'(z^k, d^k)}{\|H(z^k)\|} = 0 \tag{4.7}$$

Se cumple para todo k, luego para $k \to \infty$ obtenemos:

$$\tilde{H}_F(z^*) + \nabla H_F(z^*)^T.\tilde{d} = 0, \tag{4.8}$$

$$\tilde{H}_{\alpha(z^*)}(z^*) + \nabla H_{\alpha(z^*)}(z^*)^T.\tilde{x} = 0, \tag{4.9}$$

$$\tilde{H}_{\beta(z^*)}(z^*) + \nabla H_{\beta(z^*)}(z^*)^T.\tilde{x} = 0, \tag{4.10}$$

$$\tilde{H}_{\gamma(z^*)}(z^*) + \nabla H_{\gamma(z^*)}(z^*)^T. \begin{pmatrix} \tilde{x} \\ \tilde{u}_{\gamma(z^*)} \end{pmatrix} = 0, \tag{4.11}$$

$$\tilde{H}_h(z^*) + \nabla H_h(z^*)^T.\tilde{x} = 0. \tag{4.12}$$

Con_iderando el término T_1:

$$\begin{aligned} T_1 &= H_F(z^k)^T H_F(z^k) - H_F(w^k)^T H_F(w^k) \\ &= \|H_F(z^k)\|^2 - \|H_F(w^k)\|^2 \\ &= 2(\theta_F(z^k) - \theta_F(w^k)) \end{aligned}$$

Por el Teorema del Valor Medio:

$$T_1 = -2\nabla\theta_F(\tilde{z}^k).d^k = -2H_F(\tilde{z}^k)^T\nabla H_F(\tilde{z}^k)^T.d^k$$

donde: $\tilde{z}^k = tz^k + (1-t)w^k$, $0 < t < 1$

Como: $\{\tilde{z}^k\} \to z^*$

$$\begin{aligned} \lim_{k\to\infty} \frac{T_1}{\|H(z^k)\|^2} &= \lim_{k\to\infty} \frac{-2H_F(\tilde{z}^k)^T\nabla H_F(\tilde{z}^k)^T.d^k}{\|H(z^k)\|^2} \\ &= -2\tilde{H}_F(z^*)^T\nabla H_F(z^*)^T.\tilde{d} \qquad = 2\tilde{H}_F(z^*)^T\tilde{H}_F(z^*)\ldots(I) \end{aligned}$$

Similarmente se puede obtener:

$$\lim_{k\to\infty} \frac{T_5}{\|H(z^k)\|^2} = 2\tilde{H}_h(z^*)^T\tilde{H}_h(z^*)\ldots(II)$$

Debido a que:

$\alpha(z^*) \subseteq \alpha(z^k)$ y $\gamma(z^*) \subseteq \gamma(z^k)$

$H_{\alpha(z^k)}(z^k) = -g_{\alpha(z^k)}(x^k) \Longrightarrow H_{\alpha(z^*)}(z^k) = -g_{\alpha(z^*)}(x^k)$

De manera análoga:

$H_{\alpha(z^*)}(w^k) = -g_{\alpha(z^*)}(x^k + dx)$

$H_{\gamma(z^*)}(z^k) = -g_{\gamma(z^*)}(x^k) + du^k_\gamma$

$H_{\gamma(z^*)}(w^k) = -g_{\gamma(z^*)}(x^k + dx) + (u^k + du^k)_\gamma$

Luego, $H_{\alpha(z^*)}(z^k)$ y $H_{\alpha(z^*)}(w^k)$, $H_{\gamma(z^*)}(z^k)$ y $H_\gamma(z^*)(w^k)$ tiene la misma forma función respectivamente, entonces:

$$\begin{aligned} \lim_{k\to\infty} \frac{T_2}{\|H(z^k)\|^2} &= \lim_{k\to\infty} \frac{-2H_{\alpha(z^*)}(\tilde{z}^k)^T\nabla H_{\alpha(z^*)}(\tilde{z}^k)^T.dx^k}{\|H(z^k)\|^2} \\ &= -2\tilde{H}_{\alpha(z^*)}(z^*)^T\nabla H_{\alpha(z^*)}(z^*)^T.\tilde{x} = 2\tilde{H}_{\alpha(z^*)}(z^*)^T\tilde{H}_{\alpha(z^*)}(z^*)\ldots(III) \end{aligned}$$

De forma análoga:

$$\begin{aligned} \lim_{k\to\infty} \frac{T_4}{\|H(z^k)\|^2} &= \lim_{k\to\infty} \frac{-2H_{\gamma(z^*)}(\tilde{z}^k)^T\nabla H_{\gamma(z^*)}(\tilde{z}^k)^T.\begin{pmatrix} dx^k \\ du^k_\gamma \end{pmatrix}}{\|H(z^k)\|^2} \\ &= 2H_{\gamma(z^*)}(z^*)^T H_{\gamma(z^*)}(z^*)\ldots(IV) \end{aligned}$$

Consideremos T_3: observamos que para algún $i \in \beta(z^*)$, la función $H_{\beta(z^*)}(z^k)$ y $H_{\beta(z^*)}(w^k)$ pueden tener diferentes funciones:

Ejemplo: Sea $i \in \beta(z^k) \cap \gamma(w^k)$; $\quad (\beta(z^k) \subseteq \beta(z^*))$

Si $i \in \beta(z^k)$,

$H_i(z^k) = -g_i(x^k)$

Si $i \in \gamma(w^k)$:

$H_i(w^k) = -g_i(x^k+dx)+(u_i+du_i)^- = -g_i(x^k+dx)+du_i^- = -g_i(x^k+dx)+\min(0,du_i)$

Luego, $H_{\beta(z^*)}(z^k)$ y $H_{\beta(z^*)}(w^k)$ pueden tener diferentes funciones.

Notamos además que para cualquier $i \in \beta(z^*)$, si $i \in \alpha(z^k)$ y $i \in \alpha(w^k)$, o si $i \in \beta(z^k)$ y $i \in \beta(w^k)$, $H_i(z^k)$ y $H_i(w^k)$ tienen la misma función:

En efecto:

Si $i \in \beta(z^*) \cap \alpha(z^k) \cap \alpha(w^k)$

$H_i(z^k) = -g_i(x^k)$

$H_i(w^k) = -g_i(x^k + dx)$

Luego, $H_i(z^k)$ y $H_i(w^k)$ tienen la misma función.

De manera análoga si $i \in \beta(z^*) \cap \beta(z^k) \cap \beta(w^k)$.

Ahora, por el teorema del valor medio:

$$\lim_{k\to\infty} \frac{(H_i(z^k))^2 - (H_i(w^k))^2}{\|H(z^k)\|^2} = 2(\bar{H}_i)^2 \tag{4.13}$$

para todo $i \in \beta(z^*) \cap \bar{\gamma}(z^k) \cap \bar{\gamma}(w^k)$, donde $\bar{\gamma}(z^k) = \alpha(z^k) \cup \beta(z^k)$.

Similarmente, para $i \in \beta(z^*) \cap \gamma(z^k) \cap \gamma(w^k)$, por el teorema del valor medio y notando que: $H_i(z^k) = -g_i(x^k) + (u_i^k)^-$

Luego: $\lim\limits_{k\to\infty} H_i(z^k) = -g_i(x^*)$, también obtenemos (4.13)

Es imposible que $i \in \alpha(z^k) \in \gamma(w^k)$, pues:

$H_i(z^k) = -g_i(x^k)$

$H_i(w^k) = -g_i(x^k + dx) + (u_i^k + du_i^k)^-$

$H_i(z^k)$ y $H_i(w^k)$ tienen diferentes formas.

De mánera análoga, es imposible que $i \in \gamma(z^k) \cap \alpha(w^k)$.

Cuando k es suficientemente grande, el resto tenemos que considerar:

(i) $i \in \beta(z^k) \cap \gamma(w^k)$ ó

(ii) $i \in \gamma(z^k) \cap \beta(w^k)$

Si (i) es el caso, de la ecuación de Newton:

$$g_i(x^k) + \nabla g_i(x^k)^T dx - \min(0, du_i) = 0$$

Luego:

$$\begin{aligned} g_i(x^k + dx) &= g_i(x^k) + \nabla g_i(x^k)^T dx + \mathrm{O}(\|dx\|) \\ &= \min(0, du_i) + \mathrm{O}(\|dx\|) \\ &= \min(0, du_i) + \mathrm{O}(\|H(z^k)\|) \end{aligned}$$

Debido a que $i \in \gamma(w^k)$

$H_i(w^k) = -g_i(x^k + dx) + (u_i + du_i)^- = -g_i(x^k + dx) + du_i^-$

Sabiendo que $du_i^- = \min(0, du_i)$, tenemos:

$H_i(w^k) = -g_i(x^k + dx) + \min(0, du_i) = -\mathrm{O}(\|H(z^k)\|)$

por lo tanto:

$(H_i(z^k))^2 - (H_i(w^k))^2 = 2(H_i(z^k))^2 + \mathrm{O}(\|H(z^k)\|^2)$.

Luego se cumple (4.13). Similarmente para el caso (ii).

Entonces, tenemos:

$$\lim_{k\to\infty} \frac{T_3}{\|H(z^k)\|^2} = 2\tilde{H}_\beta^T \tilde{H}_\beta$$

Con todos los resultados obtenidos anteriormente (I, II, III, IV), y dividiendo en ambos lados de $(\dagger)$ por $\|H(z^k)\|^2$, obtenemos:

$$\frac{\theta(z^k) - \theta(z^k + d^k)}{\|H(z^k)\|^2} < 2\sigma \frac{\theta(z^k)}{\|H(z^k)\|^2}$$

Y tomando el límite, obtenemos:

$$\tilde{H}^T \tilde{H} \leq \sigma \tilde{H}^T \tilde{H}$$

como $\sigma < \dfrac{1}{2}$, la desigualdad implica que $\tilde{H} = 0$, lo cual no es cierto. $[\ \ ]$

De la demostración anterior, uno puede ver que la condición de regularidad es importante porque esto asegura que la dirección d^k es la solución de (3.1). Bajo la suposición de regularidad débil, (3.1) no puede tener una solución.

Así no se puede obtener:

$$\tilde{H} + H'(\tilde{z}, \tilde{d}) = 0$$

Sin embargo, si la sucesión $\{z^k\}$ generado por el algoritmo modificado de Newton (amortiguado) converge a un vector regular débilmente z^*, y si:

$$\lim_{k \to \infty} \left(\frac{H(z^k)}{\|H(z^k)\|} + \frac{H'(z^k, d^k)}{\|H(z^k)\|} \right) = 0 \tag{4.14}$$

Entonces por el mismo argumento, el tamaño de paso se obtendrá cuando k sea suficientemente grande. Por consiguiente, tenemos el siguiente corolario.

Corolario 4.5.1 *Supongamos que una subsucesión* $\{z^k\}$ *generado por el algoritmo modificado de Newton (amortiguado) converge a un vector regular débilmente* $z^* = (x^*, u^*, v^*)^T$ *de la función* H *definido por* (2.19). *Si* (4.14) *se cumple, entonces para* k *suficientemente grande,* $m_k = 0$. *En otras palabras, el tamaño de paso es eventualmente conseguido.*

Teorema 4.5.2 *Supongamos que una subsucesión* $\{z^k\}$ *generada por el algoritmo modificado de Newton (amortiguado) converge a un vector regular* $z^* = (x^*, u^*, v^*)^T$ *de la función* H *definida por* (2.19). *Entonces existe una constante* $c > 0$ *tal que para* k *suficientemente grande*

$$\|z^{k+1} - z^k\| \leq c\|z^k - z^*\|^2 \tag{4.15}$$

Demostración

Si una subsucesión $\{z^k\}$ generado por el algoritmo modificado de Newton (amortiguado) converge a un vector regular $z^* = (x^*, u^*, v^*)^T$ entonces de acuerdo al teorema (4.5.1):

$z^{k+1} = z^k + d^k$ para k suficientemente grande.

U ando la misma notaciones de las ecuaciones (3.12) y (3.13) tenemos:

$$\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} y^k \\ du_\beta^{k++} \end{pmatrix} + \begin{pmatrix} q_y(z^k) \\ q_\beta^{k++}(z^k) \end{pmatrix} = 0 \qquad (4.16)$$

donde

$$y^k = \begin{pmatrix} dx^k \\ du_\alpha^k \\ dv^k \end{pmatrix}$$

$du_\beta^{k++} > 0$,

y B_+ y $q_\beta^{++}(z^k)$ son la submatriz y el subvector de B y q_β respectivamente.

Se nota que:

$$\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} - x^* \\ u_\alpha^{k+1} - u_\alpha^* \\ v^{k+1} - v^* \\ u_\beta^{k+1++} - u_\beta^* \end{pmatrix} =$$

$$= \begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} x^k - x^* \\ u_\alpha^k - u_\alpha^* \\ v^k - v^* \\ u_\beta^{k++} - u_\beta^* \end{pmatrix} - \begin{pmatrix} q_y(z^k) \\ q_\beta^{++}(z^k) \end{pmatrix} \qquad (4.17)$$

donde $z^* = (x^*, u^*, v^*)^T$ es una solución del sistema de ecuaciones. Debido a que

$$q_y(z^k) = \begin{pmatrix} H_F(z^k) \\ -g_\alpha(x^k) \\ -h(x^k) \end{pmatrix}$$

y

$q_\beta^{++}(z^k) = \{-g_\beta(z^k) : du_\beta^k > 0\}$

Ahora como:

$$\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} x^k - x^* \\ u_\alpha^k - u_\alpha^* \\ v^k - v^* \\ u_\beta^{k++} - u_\beta^* \end{pmatrix} =$$

$$
= \begin{pmatrix} L(z^k)\begin{pmatrix} x^k - x^* \\ u_\alpha^k - u_\alpha^* \\ v^k - v^* \end{pmatrix} + B_+(z^k)(u_\beta^{k++} - u_\beta^*) \\ -B_+(z^k)^T \begin{pmatrix} x^k - x^* \\ u_\alpha^k - u_\alpha^* \\ v^k - v^* \end{pmatrix} \end{pmatrix} \tag{4.18}
$$

Luego, la ecuación (4.18) puede ser escrito como:

$$
\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} x^k - x^* \\ u_\alpha^k - u_\alpha^* \\ v^k - v^* \\ u_\beta^{k++} - u_\beta^* \end{pmatrix} =
$$

$$
= \begin{pmatrix} \Delta(z^k)(x^k - x^*) + \nabla g_\alpha(x^k)(u_\alpha^k - u_\alpha^*) + \nabla h(x^k)(v^k - v^*) + \nabla g_\beta^+(x^k)(u_\beta^{k++} - u_\beta^*) \\ -\nabla g_\alpha(x^k)^T(x^k - x^*) \\ -\nabla h(x^k)^T(x^k - x^*) \\ -\nabla g_\beta^+(x^k)^T(x^k - x^*) \end{pmatrix}
$$

Luego de la ecuación (4.17) tenemos:

$$
\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} - x^* \\ u_\alpha^{k+1} - u_\alpha^* \\ v^{k+1} - v^* \\ u_\beta^{k+1++} - u_\beta^* \end{pmatrix} =
$$

$$
= \begin{pmatrix} \Delta(z^k)(x^k - x^*) + \nabla g_\alpha(x^k)(u_\alpha^k - u_\alpha^*) + \nabla h(x^k)(v^k - v^*) + \nabla g_\beta^+(x^k)(u_\beta^{k++} - u_\beta^*) \\ -\nabla g_\alpha(x^k)^T(x^k - x^*) \\ -\nabla h(x^k)^T(x^k - x^*) \\ -\nabla g_\beta^+(x^k)^T(x^k - x^*) \end{pmatrix} +
$$

$$
-\begin{pmatrix} H_F(z^k) \\ -g_\alpha(x^k) \\ -h(x^k) \\ -g_\beta(x^k) \end{pmatrix}
$$

$$
= \begin{pmatrix} \nabla H_F(z^k)^T(z^k - z^*) \\ -\nabla g_\alpha(x^k)^T(x^k - x^*) \\ -\nabla h(x^k)^T(x^k - x^*) \\ -\nabla g_\beta^+(x^k)^T(x^k - x^*) \end{pmatrix} - \begin{pmatrix} H_F(z^k) \\ -g_\alpha(x^k) \\ -h(x^k) \\ -g_\beta(x^k) \end{pmatrix}
$$

Por lo tanto:

$$
\begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} - x^* \\ u_\alpha^{k+1} - u_\alpha^* \\ v^{k+1} - v^* \\ u_\beta^{k+1++} - u_\beta^* \end{pmatrix} = \begin{pmatrix} H_F(z^*) - H_F(z^k) + \nabla H_F(z^k)^T(z^k - z^*) \\ -g_\alpha(x^*) + g_\alpha(x^k) - \nabla g_\alpha(x^k)^T(x^k - x^*) \\ -h(x^*) + h(x^k) - \nabla h(x^k)^T(x^k - x^*) \\ -g_\beta(x^*) + g_\beta(x^k) - \nabla g_\beta(x^k)^T(x^k - x^*) \end{pmatrix} \tag{4.19}
$$

Bajo la suposición de las funciones F, g, h y la función lagrangiana H_F son dos veces continuamente diferenciable, el lado derecho de la ecuación (4.19) está acotado en norma por:

$$
c\|z^* - z^k\|^2
$$

donde el valor de c es dependiente de la segunda derivada de H_F, g y h en z^*. La matriz del lado izquierdo de la ecuación (4.19) es no singular, luego su inversa es acotada en norma por una constante $\bar{c}$ lo cual también es dependiente de z^*.

En efecto:

$$
\left\| \begin{pmatrix} x^{k+1} - x^* \\ u_\alpha^{k+1} - u_\alpha^* \\ v^{k+1} - v^* \\ u_\beta^{k+1++} - u_\beta^* \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} L(z^k) & B_+(z^k) \\ -B_+(z^k)^T & 0 \end{pmatrix}^{-1} \right\| .c\|z^k - z^*\|^2 \leq l\|z^k - z^*\|^2
$$

E decir xi te una constante c lo cual depende solamente de z^* tal que para cada k uficientemente grande:

$$\|z^{k+1} - z^*\| \leq c\|z^k - z^*\|^2 \quad \square$$

Cuando z^* es olamente regular débilmente, la convergencia cuadrática en la vecindad de la olución no es garantizado en general. La principal razón es que la dirección d^k no es necesariamente una solución al sistema (3.1).

4.6 Propiedades del método básico

En esta sección discutiremos el algoritmo de Newton(amortiguado) básico y el algoritmo modificado para el sistema de ecuaciones no-diferenciables en lo cual la desigualdad variacional puede ser formulado. A diferencia del enfoque de Pang, el cual usa el operador "Min" para convertir un problema de desigualdad variacional en un sistema de ecuaciones, obtenemos un sistema de ecuaciones a través del uso de la aplicación Minty. La diferencia básica entre las dos formulaciones es que la formulación Minty pone ambas restricciones de desigualdad y sus variables duales en el sistema de ecuaciones, mientras que la formulación del operador "Min" solamente considera uno de esos dos a la vez. En general es dificil decir cual formulación e mejor que el otro. El siguiente ejemplo demuestra que un punto regular bajo una formulación puede no ser regular bajo otra formulación y viceversa.

Ejemplo 4.3 *Consideremos el siguiente Problema de Complementariedad donde:*

$$F(x) = \begin{pmatrix} x_1^2 + x_1 + 2x_2 - 3 \\ 2x_1^2 - 2x_2 + x_2^2 - 2 \end{pmatrix}$$

Como demostraremos en la siguiente sección:

La formulación de la aplicación Minty lleva al siguiente sistema de ecuaciones:

$$H(x) = F(x^+) + x^- = \begin{pmatrix} H_1(x) \\ H_2(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Luego:

$$\begin{pmatrix} H_1(x) \\ H_2(x) \end{pmatrix} = \begin{pmatrix} (x_1^+)^2 + x_1^+ + 2x_2^+ - 3 \\ 2(x_1^+)^2 - 2x_2^+ + (x_2^+)^2 - 2 \end{pmatrix} + \begin{pmatrix} x_1^- \\ x_2^- \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} H_1(x) \\ H_2(x) \end{pmatrix} = \begin{pmatrix} (x_1^+)^2 + x_1^+ + 2x_2^+ - 3 + x_1^- \\ 2(x_1^+)^2 - 2x_2^+ + (x_2^+)^2 - 2 + x_2^- \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

entonces:

$H_1(x) = (x_1^+)^2 + x_1^+ + 2x_2^+ - 3 + x_1^- = 0$

$H_2(x) = 2(x_1^+)^2 - 2x_2^+ + (x_2^+)^2 - 2 + x_2^- = 0$

La formulación del operador "min" genera:

$$H(x) = \text{Min}(x, F(x)) = \begin{pmatrix} H_1(x) \\ H_2(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Luego:

$H_1(x) = \text{Min}(x_1, F_1(x)) = (x_1, x_1^2 + x_1 + 2x_2 - 3) = 0$

$H_2(x) = \text{Min}(x_2, F_2(x)) = (x_2, 2x_1^2 - 2x_2 + x_2^2 - 2) = 0$

- En el punto $x^{(1)} = (1, 1)^T$

Formulación Minty:

$\alpha(x^{(1)}) = \{1, 2\}$

$\beta(x^{(1)}) = \emptyset$

$\gamma(x^{(1)}) = \emptyset$

Y el jacobiano de F es:

$$\nabla F(x^{(1)}) = \begin{pmatrix} 2x_1^{(1)} + 1 & 2 \\ 4x_1^{(1)} & x_2^{(1)} - 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 4 & 0 \end{pmatrix}$$

Lo cual es no-singular. Por lo tanto $x^{(1)}$ es un punto regular bajo la formulación de la aplicación Minty.

Formulación del operador "Min":

$H_1(x^{(1)}) = \text{Min}\{1,1\} = 1 = F_1(x^{(1)}) = x^{(1)}$

$H_2(x^{(1)}) = \text{Min}\{1,-1\} = -1 = F_2(x^{(1)})$

Por definición tenemos:

$\alpha(x) = \{i : x > F(x)\}$

$\beta(x) = \{i : x = F(x)\}$

$\gamma(x) = \{i : x < F(x)\}$

Ahora, como:

$F_1(x^{(1)}) = x^{(1)} \Longrightarrow \beta(x^{(1)}) = \{1\}$

$F_2(x^{(1)}) < x^{(1)} \Longrightarrow \alpha(x^{(1)}) = \{2\}$

Luego: $\gamma(x^{(1)}) = \emptyset$ Entonces:

$$\nabla F_{\alpha\alpha}(x^{(1)}) = \frac{\partial F_2}{\partial x_2} = -2 + 2x_2 = 0$$

Por lo tanto, este punto no es regular.

- En el punto $x^{(2)} = (2,0)^T$

Formulación Minty:

$\alpha(x^{(2)}) = \{1\}$

$\beta(x^{(2)}) = \{2\}$

$\gamma(x^{(2)}) = \emptyset$

Y el jacobiano de F es:

$$\nabla F(x^{(2)}) = \begin{pmatrix} 2x_1^{(2)} + 1 & 2 \\ 4x_1^{(2)} & x_2^{(2)} - 2 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 8 & -2 \end{pmatrix}$$

Luego:

$F_{\alpha\alpha} = F_{11} = \dfrac{\partial F_1}{\partial x_1} = 2x_1^{(2)} + 1 = 5$

El Complemento Schur:

$F_{\beta\beta} - F_{\beta\alpha}.(F_{\alpha\alpha})^{-}1.F_{\alpha\beta}$

Debido a que:

$F_{\beta\beta} = -2, \quad F_{\beta\alpha} = 8 \quad F_{\alpha\beta} = 2$

Entonces:

$$F_{\beta\beta} - F_{\beta\alpha}.(F_{\alpha\alpha})^{-1}.F_{\alpha\beta} = -2 - 8.\frac{1}{5}.2 = -\frac{16}{5} < 0$$

Entonces el Complemento Schur no es una P-Matriz.

Por lo tanto, $x^{(2)}$ no es un punto regular.

Formulación del operador "Min"

$H_1(x) = \text{Min}\{x_1^{(2)}, (x_1^{(2)})^2 + x_1^{(2)} + 2x_2^{(2)} - 3\} = \text{Min}\{2, 3\} = 2$

$H_2(x) = \text{Min}\{x_2^{(2)}, 2(x_1^{(2)})^2 - 2x_2^{(2)} + (x_2^{(2)})^2 - 2\} = \text{Min}\{0, 6\} = 0$

Como:

$F_1(x) > x, \quad F_2(x) > x \quad \Longrightarrow \gamma(x^{(2)}) = \{1, 2\}$

El punto $x^{(2)} - (2, 0)^T$ es regular ya que todos los indices estan el conjunto γ.

5 ALGORITMO DE NEWTON AMORTIGUADO PARA PROBLEMAS DE COMPLEMENTARIEDAD

Hasta ahora hemos discutido el algoritmo de Newton (amortiguado) y us propiedades teóricas cuando aplicamos a desigualdades variacionales. Debido a que los Problemas de Complementariedad es un caso especial de los problemas de desigualdad variacional, el algoritmo debería tener algunas caracteristicas especiales y variaciones cuando aplicamos a éste caso especial. En este capítulo, exploraremos estas caracteristicas y variaciones en el contexto de los Problemas de Complementariedad no-lineal.

5.1 Método de Newton amortiguado para Problemas de Complementariedad

Siendo $g(x) = -x$, el Problema de Desigualdad Varicional VI(X,F) se convierte en el problema de Complementariedad NCP(F). En este caso el sistema de ecuaciones $H(z) = 0$ dado por $(2.9) - (2.11)$, se convierte en:

$$H(z) = \begin{pmatrix} F(x) - u^+ \\ x + u^- \end{pmatrix} = 0 \tag{5.1}$$

donde $H : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, y $z = (x, u)^T \in \mathbb{R}^{2n}$. La B-derivada de H en el punto z a lo largo de la dirección $d = (dx, du)^T$ está dado por:

$$H'(z, d) = \begin{pmatrix} \nabla F(x)dx - I.du^+ \\ I.dx + du^- \end{pmatrix} \tag{5.2}$$

En efecto:

$$\begin{aligned} H'(z,d) &= \lim_{\lambda\to 0^+} \frac{H(z+\lambda d)-H(z)}{\lambda} \\ &= \begin{pmatrix} [\lim\limits_{\lambda\to 0^+} \frac{F(x+\lambda dx)-F(x)}{\lambda}] - [\lim\limits_{\lambda\to 0^+} \frac{(u+\lambda du)^+ - u^+}{\lambda}] \\ (\lim\limits_{\lambda\to 0^+} \frac{(x+\lambda dx)-x}{\lambda}) + [\lim\limits_{\lambda\to 0^+} \frac{(u+\lambda du)^- - u^-}{\lambda}] \end{pmatrix} \end{aligned}$$

Luego:

$$H'(z,d) = \begin{pmatrix} \nabla F(x)dx - I.du^+ \\ I.dx + du^- \end{pmatrix}$$

donde I es una matriz identidad $n\times n$ y du^+ y du^- tienen el mismo significado como antes. Definimos:

$H_F(z) = F(x) - I_\alpha^T u_\alpha,$

$H_\alpha(z) = x_\alpha,$

$H_\beta(z) = x_\beta,$

$H_\gamma(z) = x_\gamma + u_\gamma,$

La ecuación de newton:

$$H(z^k) + H'(z^k, d^k) = 0 \tag{5.3}$$

puede ser escrito como:

$$H_F(z) + \nabla F(x)dx - I_\alpha^T du_\alpha - I_\beta^T \max(0, du_\beta) = 0, \tag{5.4}$$

$$H_\alpha(z) + I_\alpha dx = 0, \tag{5.5}$$

$$H_\beta(z) + I_\beta dx + \min(0, du_\beta) = 0, \tag{5.6}$$

$$H_\gamma(z) + I_\gamma dx + du_\gamma = 0, \tag{5.7}$$

donde I_α es una submatriz $|\alpha| \times n$ de I cuyas filas son indexados por el conjunto α, también I_β y I_γ son definidos de manera análoga.

Similarmente uno puede definir los conceptos de regularidad y regularidad débil.

Definición 5.1.1 *Sea $z = (x,u)^T$ un vector arbitrario en $\mathbb{R}^{2n}$. Entonces z es llamado un vector regular para la función H si*

$$(a)\quad L(z) = \begin{pmatrix} \nabla F(x) & -I_\alpha^T \\ I_\alpha & 0 \end{pmatrix} \qquad \text{es no-singular.}$$

(b) *El complemento Schur de $L(z)$ en la matriz*

$$P(z) = \begin{pmatrix} \nabla F(x) & -I_\alpha^T & -I_\beta^T \\ I_\alpha & 0 & 0 \\ I_\beta & 0 & 0 \end{pmatrix}$$

dado por $(I_\beta, 0).L(z)^{-1}.(I_\beta, 0)^T$, es una P-Matriz.

Definición 5.1.2 *Sea $z = (x,u)^T$ un vector arbitrario en $\mathbb{R}^{2n}$, entonces z es llamdo un vector regular débilmente para la función H si existe una vecindad N de z tal que para cada vector $z' = (x',u')^T \in N$, la matriz*

$$L(z') = \begin{pmatrix} \nabla F(x') & -I^T_{\alpha(z')} \\ I_{\alpha(z')} & 0 \end{pmatrix}$$

es no-singular.

El algoritmo modificado de Newton (amortiguado) descrito previamente generará una dirección descendente para la función norma en un vector z si z es un vector regular.

Proposición 5.1.1 *Sea z un vector regular débilmente de H definido por* (2.19). *Si $H(z) \neq 0$, entonces siempre existe una dirección descendente d en z para la función norma θ a menos que:*

(i) $L_1 = H_\gamma(z) = 0$;

(ii) $L_2 = I_\alpha H_F(z) = 0$;

(iii) $L_3 = H_F(z)^T \nabla F(z) + \sum_{i\in\alpha(z)} x_i + \sum_{i\in\beta(z)} x_i = 0;$

(iv) $L_4 = I_\beta H_F(z) \leq 0;$

(v) $L_5 = x_\beta \leq 0$

En particular, si $x_\beta > 0$, entonces siempre existirá una dirección descendente para la función norma.

Demostración:

Supongamos que no existe una dirección descendente para la función norma en z, entonces para cualquier dirección $d \in \mathbb{R}^{2n}$, tenemos $\theta'(z,d) \geq 0$. Por definición:
$\theta'(z,d) = H(z)^T H'(z,d)$
donde:

$$\begin{aligned} H(z)^T &= \begin{pmatrix} H_F(z)^T & x_\alpha^T & x_\beta^T & H_\gamma(z)^T \end{pmatrix} \quad \text{y} \\ H'(z,d) &= \begin{pmatrix} \nabla F(x)dx - I_\alpha^T du_\alpha - I_\beta^T \min(0, du_\beta) \\ I_\alpha dx \\ I_\beta dx + \min(0, du_\beta) \\ I_\gamma dx + du_\gamma \end{pmatrix} \end{aligned}$$

Luego:

$$\begin{aligned} \theta'(z,d) &= H(z)^T H'(z,d) \\ &= [H_F(z)^T \nabla F(x) + x_\alpha^T I_\alpha + x_\beta^T I_\beta + H_\gamma(z)^T I_\gamma] dx \\ &\quad - H_F(z)^T I_\alpha^T du_\alpha + H_\gamma(z)^T du_\gamma \\ &\quad - H_F(z)^T I_\beta^T \max(0, du_\beta) + x_\beta^T \min(0, du_\beta) \\ &\geq 0 \end{aligned}$$

Fijando $dx = 0$, $du_\alpha = 0$, $du_\beta = 0$, y $du_\gamma \neq 0$, tenemos:
$H_\gamma(z)^T du_\gamma \geq 0 \qquad \forall du_\gamma \neq 0.$
Obviamente, este es solamente cierto cuando $H_\gamma(z) = 0$.

Por el mismo argumento:

- Fijando $dx = 0$, $du_\gamma = 0$, $du_\beta = 0$ y $du_\alpha \neq 0$, tenemos:
  $H_F(z)^T I_\alpha^T du_\alpha \leq 0 \qquad \forall du_\alpha \neq 0.$
  este es solamente cierto cuando $H_F(z)^T I_\alpha^T = I_\alpha H_F(z) = 0.$

- Fijando $du_\gamma = 0$, $du_\beta = 0$, $du_\alpha = 0$, $dx \neq 0$ tenemos:
  $[H_F(z)^T \nabla F(x) + x_\alpha^T I_\alpha + x_\beta^T I_\beta + H_\gamma(z)^T I_\gamma] dx = 0, \quad \forall dx \neq 0$
  este es solamente cierto cuando:
  $$\begin{aligned} H_F(z)^T \nabla F(x) + x_\alpha^T I_\alpha + x_\beta^T I_\beta + H_\gamma(z)^T I_\gamma &= H_F(z)^T \nabla F(z) + \\ &+ \sum_{i \in \alpha(z)} x_i + \sum_{i \in \beta(z)} x_i \\ &= 0 \end{aligned}$$

Fijando $dx = 0$, $du_\alpha = 0$, $du_\gamma = 0$, y $du_\beta \neq 0$, es trivial ver que los coeficientes de $\max(0, du_\beta)$ tiene que ser mayor o igual a cero, y los coeficientes de $\min(0, du_\beta)$ tiene que ser menor o igual a cero, esto nos lleva a :

$H_F(z)^T I_\beta^T = I_\beta H_F(z) \leq 0$

$x_\beta \leq 0.$ $\square$

Si z es un vector regular débilmente y algunas de las condiciones $(i) - (v)$ no son satisfechas, el siguiente procedimiento es implementado.

Paso 1:

Para alguna componente l_{ij} de L_j $(j = 1 \ldots 5)$ que viole las condiciones de arriba, fijar d_{ij}, la dirección asociada con l_{ij}, a ser 1 o -1 de tal forma que $l_{ij}.d_{ij} < 0$

Paso 2:

Realizar una búsqueda lineal a lo largo de la dirección d obtenida en el paso 1.

Es fácil ver que el algoritmo de Newton(amortiguado) con ésta modificación generará una dirección descendente para la función norma. La convergencia del teorema y la razón cuadratica de convergencia en la vecindad de solución se mantendrá bajo la mismas condiciones.

5.2 Formulación alternativa para el Problema de Complementariedad No Lineal

Siguiendo el concepto de la aplicación Minty, y de acuerdo a (2.18), definamos la aplicación $H : \mathbb{R}^n \to \mathbb{R}^n$ por

$$H(x) = F(x^+) + x^- \tag{5.8}$$

donde:

$x_i^+ = \max(x_i, 0)$, $\quad x_i^- = \min(x_i, 0)$, $\quad x^+ = (x_1^+, x_2^+, \ldots, x_n^+)^T$ $\quad$ y

$x^- = (x_1^-, x_2^-, \ldots, x_n^-)^T$

En otras palabra, x^+ resuelve el NCP(2.2) y solo si resuelve el sistema de ecuaciones (5.8).

En efecto:

Si x^+ resuelve el NCP, es decir:

$$F(x^+)^T x^+ = 0, \qquad F(x^+) \in \mathbb{R}_+^n$$

Luego:

$$F(x^+)^T x^+ = -(x^-)^T x^+ = 0 \Longrightarrow (F(x^+) + x^-)^T x^+ = 0$$

Entonces: $H(x) - F(x^+) + x^- = 0$,

es decir x es solución del sistema de ecuaciones (5.8).

El otro sentido de manera análoga.

Debido a la estructura especial de los problemas de complementariedad uno tiene los siguientes resultados concerniente a la acotación del conjunto nivel si F es una función P-Uniforme.

Teorema 5.2.1 *Sea la aplicación $F : \mathbb{R}^n_+ \to \mathbb{R}^n$ continuamente diferenciable y una función P-uniforme; es decir, $\nabla F(x)$ es una P-Matriz para todo $x \in \mathbb{R}^n_+$. Sea $x^0 \in \mathbb{R}^n$ un punto arbitrario. Entonces el conjunto de nivel*

$$\{x \in \mathbb{R}^n : \|H(x)\| \leq \|H(x^0)\|\}$$

es acotado

Demostración:

Por contradicción, asumamos que el conjunto de nivel de H no es acotado; es decir, asumamos que existe una sucesión $\{x^k\}$ con $\lim_{k\to\infty} \|x^k\| = \infty$ tal que:

$$\|H(x^k)\| \leq \|H(x^0)\|$$

Definamos:

$$\alpha(x^k) = \{j : x^k_j > 0\}, \quad \beta(x^k) = \{j : x^k_j = 0\}, \quad \gamma(x^k) = \{j : x^k_j < 0\}$$

y

$$x^k_\alpha = \{x^k_j : j \in \alpha(x^k)\}, \quad x^k_\beta = \{x^k_j : j \in \beta(x^k)\}, \quad x^k_\gamma = \{x^k_j : j \in \gamma(x^k)\}$$

Debido a que $\|x^k\| \to \infty$, se debe tener $|x^k_j| \to \infty$ para $j \in \alpha(x^k)$ o $j \in \gamma(x^k)$. Si $x^k_j \to \infty$ para $j \in \alpha(x^k)$, entonces existe una subsucesión $\{\tilde{x}^l\}$ de $\{x^k\}$ tal que los conjuntos indices son constantes; es decir,

$$\begin{aligned}
\alpha(\tilde{x}^1) &= \alpha(\tilde{x}^2) = \ldots = \alpha(\tilde{x}^l) = \ldots,\\
\beta(\tilde{x}^1) &= \beta(\tilde{x}^2) = \ldots = \beta(\tilde{x}^l) = \ldots,\\
\gamma(\tilde{x}^1) &= \gamma(\tilde{x}^2) = \ldots = \gamma(\tilde{x}^l) = \ldots.
\end{aligned}$$

Debido a que $\|H(x^l)\| \leq \|H(x^0)\|$ para todo l, existe una subsucesión $\{\tilde{x}^q\}$ de $\{\tilde{x}^l\}$ tal que

$$\lim_{q\to\infty} \|H(\tilde{x}^q)\| = \eta < \infty$$

Luego para todo $\epsilon > 0$ existe un $L > 0$ tal que $\forall\ m > 0$

$$\|H(\tilde{x}^{L+m}) - H(\tilde{x}^L)\|^2 \leq \epsilon$$

Esto implica:

$$\begin{aligned}
\epsilon &\geq \|H(\tilde{x}^{L+m}) - H(\tilde{x}^L)\|^2 \\
&= \|F_\alpha((\tilde{x}^{L+m})^+) - F_\alpha((\tilde{x}^L)^+)\|^2 + \|F_\beta((\tilde{x}^{L+m})^+) - F_\beta((\tilde{x}^L)^+)\|^2 \\
&\quad + \|F_\gamma((\tilde{x}^{L+m})^+) - F_\gamma((\tilde{x}^L)^+) + (\tilde{x}^{L+m})^- - (\tilde{x}^L)^-\|^2 \\
&\geq \|F_\alpha((\tilde{x}^{L+m})^+) - F_\alpha((\tilde{x}^L)^+)\|^2
\end{aligned}$$

Debido a que:

$$\begin{aligned}
F_\alpha((\tilde{x}^{L+m})^+) - F_\alpha((\tilde{x}^L)^+) &= \nabla F_{\alpha\alpha}(\theta(\tilde{x}^{L+m})^+ + (1-\theta)(\tilde{x}^L)^+)((\tilde{x}^{L+m})^+ - (\tilde{x}^L)^+) \\
&\quad + \nabla F_{\alpha\beta}(\theta(\tilde{x}^{L+m})^+ + (1-\theta)(\tilde{x}^L)^+).0 \\
&\quad + \nabla F_{\alpha\gamma}(\theta(\tilde{x}^{L+m})^+ + (1-\theta)(\tilde{x}^L)^+).0
\end{aligned}$$

para algún $\theta \in [0, 1]$, por la asumida diferenciabilidad de F, se tiene:

$$\begin{aligned}
&\|F_\alpha((\tilde{x}^{L+m})^+) - F_\alpha((\tilde{x}^L)^+)\|^2 \\
&= \|\nabla F_{\alpha\alpha}(\theta(\tilde{x}^{L+m})^+ + (1-\theta)(\tilde{x}^L)^+)\|^2 \|((\tilde{x}^{L+m})^+ - (\tilde{x}^L)^+)\|^2 \leq \epsilon
\end{aligned}$$

Haciendo $m \rightarrow \infty$, se tiene $\|((\tilde{x}^{L+m})^+ - (\tilde{x}^L)^+)\|^2 \rightarrow \infty$, en efecto:
Debido a que $\|x^k\| \rightarrow \infty$, luego $\{x^k\}$ no es de cauchy.
Entonces $\|((\tilde{x}^{L+m})^+ - (\tilde{x}^L)^+)\|^2 \rightarrow \infty$
de este modo se debe tener:

$$\lim_{m \rightarrow \infty} \|\nabla F_{\alpha\alpha}(\theta(\tilde{x}^{L+m})^+ + (1-\theta)(\tilde{x}^L)^+)\|^2 = 0$$

Lo cual contradice la suposición de que F es una función P-uniforme.
Entonces concluimos:
$\|x_\alpha^k\| < \infty$.

Si $x_j^k \to -\infty$ para algún $j \in \gamma(x^k)$, y por el mismo argumento anterior, existe una sub ucesión $\{\tilde{x}^q\}$ de $\{x^k\}$ tal que los conjuntos indices son constantes para todo q y $\|H(\tilde{x}^q)\| \leq \|H(x^0)\|$. Entonces, para cualquier $\epsilon > 0$, existe un $L > 0$ tal que para todo $m > 0$,

$$\|H(\tilde{x}^{L+m}) - H(\tilde{x}^L)\|^2 \leq \epsilon$$

De este modo

$$\begin{aligned} \epsilon &\geq \|H(\tilde{x}^{L+m}) - H(\tilde{x}^L)\|^2 \\ &\geq \|F_\gamma((\tilde{x}^{L+m})^+) - F_\gamma((\tilde{x}^L)^+) + (\tilde{x}^{L+m})^- - (\tilde{x}^L)^-\|^2 \end{aligned} \tag{5.9}$$

Debido a que $(\tilde{x}^{L+m})^+$ y $(\tilde{x}^L)^+$ no son dependientes de los valores de $\tilde{x}_\gamma^{L+m}$ y $\tilde{x}_\gamma^L$, es decir, $F((\tilde{x}^{L+m})^+)$ y $F((\tilde{x}^L)^+)$ son finitos debido a que $\tilde{x}_\alpha$ y $\tilde{x}_\beta$ son acotados. Así:

$$\lim_{m\to\infty} \|F_\gamma((\tilde{x}^{L+m})^+) - F_\gamma((\tilde{x}^L)^+) + (\tilde{x}^{L+m})^- - (\tilde{x}^L)^-\|^2 = \infty,$$

lo cual contradice (5.9), entonces concluimos

$\|x_\gamma^k\| < \infty$.

Por los argumentos establecidos decimos que la sucesión $\{x^k\}$ debe ser acotado y de este modo, el conjunto de nivel de H es acotado. $\Box$

Proposición 5.2.1 *Sea H definido en (5.8) entonces:*

La Derivada de H en un vector x está dado por:

$$\frac{\partial H_i}{\partial x_j}.d_j = \begin{cases} F_{ij} & Si \quad j \in \alpha(x) \\ F_{ij} max(0, d_j) + I_{ij} min(0, d_j) & Si \quad j \in \beta(x) \\ I_{ij}.d_j & Si \quad j \in \gamma(x) \end{cases}$$

donde

$F_{ij} = \dfrac{\partial F_i}{\partial x_j}$,

$$I_{ij} = \begin{cases} 0 & Si \quad j \neq i, \\ 1 & Si \quad j = i, \end{cases}$$
$\alpha(x) = \{j : x_j > 0\}$,

$\beta(x) = \{j : x_j = 0\}$,

$\gamma(x) = \{j : x_j < 0\}$,

Demostración:

La i-ésima fila de la B-Derivada :

$$BH_i(x).d = \lim_{\lambda \to 0^+} \frac{F_i((x+\lambda d)^+) - F_i(x^+)}{\lambda} + \lim_{\lambda \to 0^+} \frac{(x_i + \lambda d_i)^- - x_i^+}{\lambda}$$

Debido a que:

- $\lim\limits_{\lambda \to 0^+} \dfrac{F_i((x+\lambda d)^+) - F_i(x^+)}{\lambda} = \nabla F_i(x^+).\nabla(x^+).d = \nabla F_i(x^+).d^+$
- $\lim\limits_{\lambda \to 0^+} \dfrac{(x_i + \lambda d_i)^- - x_i^+}{\lambda} = I_{ij}.d_j^-$

Por lo tanto:

$$BH_i(x).d = \begin{pmatrix} \frac{\partial F_i}{\partial x_1} & \frac{\partial F_i}{\partial x_2} & \cdots & \frac{\partial F_i}{\partial x_n} \end{pmatrix} . \begin{pmatrix} \frac{\partial x_1^+}{\partial x_1} & 0 & \cdots & 0 \\ 0 & \frac{\partial x_2^+}{\partial x_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \frac{\partial x_n^+}{\partial x_n} \end{pmatrix} . \begin{pmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{pmatrix} + I_{ij}.d_j^-$$

Ahora, tenemos:

- Si $j \in \alpha(x)$:

  $x_j > 0 \quad$ y $\quad dx_j^+ = dx_j = d_j, \quad dx_j^- = d_j^- = 0$

  entonces:

  $BH_i(x).d_j = \dfrac{\partial H_i}{\partial x_j}(x).d_j = \dfrac{\partial F_i}{\partial x_j}.d_j = F_{ij}.d_j$

- Si $j \in \beta(x)$:

  $x_j = 0 \quad \text{y} \quad dx_j^+ = d_j^+ = \max(0, d_j), \quad , dx_j^- = d_j^- = \min(0, d_j)$

  entonces:

  $$BH_i(x).d_j = \frac{\partial H_i}{\partial x_j}(x).d_j = \frac{\partial F_i}{\partial x_j}.d_j^+ + I_{ij}d_j^- = \frac{\partial F_i}{\partial x_j}\max(0, d_j) + I_{ij}\min(0, d_j)$$

- Si $j \in \gamma(x)$:

  $x_j < 0 \quad \text{y} \quad dx_j^+ = d_j^+ = 0, \quad , dx_j^- = dx_j = d_j$

  entonces:

  $$BH_i(x).d_j = \frac{\partial H_i}{\partial x_j}(x).d_j = I_{ij}d_j \qquad \square$$

Aplicando el algoritmo de Newton(amortiguado) a el sistema (5.8) uno necesita resolver el siguiente Problema de Complementariedad mixta lineal para generar la dirección d^k durante la iteración k:

$$F_\alpha((x^k)^+) + (x_\alpha^k)^- + F_{\alpha\alpha}d_\alpha + F_{\alpha\beta}\max(0, du_\beta) + I_{\alpha\beta}\min(0, d_\beta) + I_{\alpha\gamma}d_\gamma = 0 \ldots (\dagger)$$

$$F_\gamma((x^k)^+) + (x_\gamma^k)^- + F_{\gamma\alpha}d_\alpha + F_{\gamma\beta}\max(0, du_\beta) + I_{\gamma\beta}\min(0, d_\beta) + I_{\gamma\gamma}d_\gamma = 0 \ldots (\dagger\dagger)$$

$$F_\beta((x^k)^+) + (x_\beta^k)^- + F_{\beta\alpha}d_\alpha + F_{\beta\beta}\max(0, du_\beta) + I_{\beta\beta}\min(0, d_\beta) + I_{\beta\gamma}d_\gamma = 0 \ldots (\dagger\dagger\dagger)$$

Como $I_{\alpha\beta} = 0 = I_{\alpha\gamma}$ en ($\dagger$) tenemos:

$$F_\alpha((x^k)^+) + (x_\alpha^k)^- + F_{\alpha\alpha}d_\alpha + F_{\alpha\beta}\max(0, du_\beta) = 0 \tag{5.10}$$

Como $I_{\gamma\beta} = 0$ y $I_{\gamma\gamma} = 1$ en ($\dagger\dagger$) tenemos:

$$F_\gamma((x^k)^+) + (x_\gamma^k)^- + F_{\gamma\alpha}d_\alpha + F_{\gamma\beta}\max(0, du_\beta) + d_\gamma = 0 \tag{5.11}$$

Como $I_{\beta\beta} = 1$, $I_{\beta\gamma} = 0$ y $x_\beta^k = 0$ en ($\dagger\dagger\dagger$) tenemos:

$$F_\beta((x^k)^+) + F_{\beta\alpha}d_\alpha + F_{\beta\beta}\max(0, du_\beta) + \min(0, d_\beta) = 0 \tag{5.12}$$

donde: $F_{\alpha\alpha}$ es la matriz

$$\left(\frac{\partial F_i}{\partial x_j} \right) \qquad \forall i, j \in \alpha$$

y las otras matrices $F_{\alpha\beta}, F_{\gamma\alpha}, F_{\gamma\beta}$ y $F_{\beta\beta}$ son definidos de manera análoga.

Denotamos:

$H_\alpha = F_\alpha((x^k)^+) + (x_\alpha^k)^-$

$H_\gamma = F_\gamma((x^k)^+) + (x_\gamma^k)^-$

$H_\beta = F_\beta((x^k)^+)$

$d_\beta^+ = \max(0, d_\beta)$

$w_\beta = H_\beta + F_{\beta\alpha} d_\alpha + F_{\beta\beta} d_\beta^+$

Notamos que $w_\beta = -\min(0, d_\beta)$ y así de $(5.10) - (5.12)$ tenemos:

$$H_\alpha + F_{\alpha\alpha} d_\alpha + F_{\alpha\beta} d_\beta^+ = 0 \tag{5.13}$$

$$w_\beta = H_\beta + F_{\beta\alpha} d_\alpha + F_{\beta\beta} d_\beta^+ \geq 0 \tag{5.14}$$

$$d_\beta^+ \geq 0, \quad w_\beta^T . d_\beta^+ = 0 \tag{5.15}$$

$$d_\gamma = -H_\gamma - F_{\gamma\alpha} d_\alpha - F_{\gamma\beta} d_\beta^+ \tag{5.16}$$

Demostraremos que $w_\beta^T d_\beta^+ = 0$

- Si $d_\beta = 0$ es obvio que $w_\beta^T . d_\beta^+ = 0$
- Si $d_\beta \neq 0$ entonces
  - Si $d_\beta > 0$ tenemos:
    $w_\beta = -\min(0, d_\beta) = 0 \Longrightarrow w_\beta^T d_\beta^+ = 0$
  - Si $d_\beta < 0$ tenemos:
    $w_\beta = -d_\beta, \quad d_\beta^+ = 0 \Longrightarrow w_\beta^T d_\beta^+ = 0$

En todos los casos $w_\beta^T d_\beta^+ = 0$. $\square$

De (5.13), si $F_{\alpha\alpha}$ es no-singular, tenemos:

$d_\alpha = -F_{\alpha\alpha}^{-1}F_{\alpha\beta}.d_\beta^+ - F_{\alpha\alpha}^{-1}H_\alpha$

Reemplazando en (5.14):

$$\left.\begin{array}{rcl} w_\beta & = & H_\beta + F_{\beta\alpha}(-F_{\alpha\alpha}^{-1}F_{\alpha\beta}d_\beta^+ - F_{\alpha\alpha}^{-1}H_\alpha) + F_{\beta\beta}d_\beta^+ \\ & = & (F_{\beta\beta} - F_{\beta\alpha}F_{\alpha\alpha}^{-1}F_{\alpha\beta})d_\beta^+ + H_\beta - F_{\beta\alpha}F_{\alpha\alpha}^{-1}H_\alpha \ \geq \ 0 \\ & & d_\beta^+ \ \geq \ 0 \\ & & w_\beta^T d_\beta^+ \ = \ 0 \end{array}\right\} \cdots (*)$$

Entonces: $(*)$ define un Problema de Complementariedad Lineal.

Las condiciones de Regularidad en este caso es logicamente definido por:

(i) $F_{\alpha\alpha}$ es no-singular.

(ii) $F_{\beta\beta} - F_{\beta\alpha}(F_{\alpha\alpha})^{-1}F_{\alpha\beta}$, el Complemento Schur, es una P-Matriz.

Cuando las condiciones (i) y (ii) son satisfechas en x^k, x^k es llamada un vector regular de $H(x)$. Claramente, si x^k es un vector regular de H, el sistema $(5.10) - (5.12)$ tiene una solución única. Si F es una función P-Uniforme, entonces cualquier $x \in \mathbb{R}^n$ es regular. i.e, el sistema $(5.13) - (5.16)$ siempre tendrá una solución única cuando F es uniformemente P.

Si existe una vecindad N de x^k tal que para cada vector $x' \in N$, la matriz $F_{\alpha'\alpha'}$ es no singular, donde $\alpha' = \alpha(x')$, entonces llamamos x^k vector regular débilmente. El algoritmo modificado de Newton(amortiguado) obtendrá una dirección descendente para la función norma a meno que:

(i) $L_1 = H_\gamma(x) = 0;$

(ii) $L_2 = H_\alpha(x)^T F_{\alpha\alpha} + H_\beta(x)^T F_{\beta\alpha} + H_\gamma(x)^T F_{\gamma\alpha} = 0;$

(iii) $L_3 = H_\alpha(x)^T F_{\alpha\beta} + H_\beta(x)^T F_{\beta\beta} \geq 0;$

(iv) $L_4 = H_\beta(x) \leq 0.$

Si x es un vector regular y una de las componenetes de $F_\beta(x^+)$ es po itiva, entonces existe una dirección descendente para la función norma debido a la violación de (iv). En particular, si x es un vector regular débilmente y $F_\beta(x^+) = 0$, entonces x es solución de (5.8). En el caso donde x es un vector regular débilmente y algunas de las condiciones $(i) - (iv)$ no es satisfecha, el procedimiento para generar una dirección descendente es la misma como la que se describió anteriormente: es decir:

(i) Para alguna componente l_{ij} de $L_j(j \quad 1, i..4)$ el cual viola la condición, fijamos d_{ij}, la dirección asociada con l_{ij}, a ser 1 o -1 tal que $l_{ij} \times d_{ij} < 0$;

(ii) Realizar una búsqueda lineal a lo largo de la dirección obtenida en (i).

Se puede ver que aunque la clasificación de los conjuntos α, β y γ y la funcionH_α, H_β y H_γ son diferentes bajo dos formulaciones alternativas, la condición de regularidad y regularidad débil son muy similares.

En el capítulo anterior, señalamos que el algoritmo básico de Newton(amortiguado) puede fallar en cualquier paso intermedio si la condición de regularidad no es satisfecha, y el algoritmo modificado puede llevarse a cabo bajo condiciones más débiles. Los siguientes ejemplos muestran cómo trabaja el algoritmo modificado.

Ejemplo 5.1 *Consideremos el problema de complementariedad no-lineal donde:*

$$F(x) = \begin{pmatrix} -\frac{3}{2}x_1^2 + \frac{13}{2}x_1 + 2x_2 - 5 \\ 2x_1^2 - 2x_2 + x_2^2 - 2 \end{pmatrix}$$

La solución a este NCP(F) es $x^* = (1, 0)^T$ con $\alpha(x^*) = \{1\}$ y $\beta(x^*) = \{2\}$.

En efecto:

Este problema de complementariedad no-lineal NCP(F) es equivalente al siguiente istema de ecuaciones:

$$H(x) = F(x^+) + x^- = 0 \ldots (\dagger)$$

Luego:

$$\begin{aligned} H_1(x) &= F_1(x^+) + x_1^- &= -\tfrac{3}{2}(x_1^+)^2 + \tfrac{13}{2}(x_1^+) + 2(x_2^+) - 5 + x_1^- = 0 \\ H_2(x) &= F_2(x^+) + x_2^- &= 2(x_1^+)^2 - 2(x_2^+) + (x_2^+)^2 - 2 + x_2^- = 0 \end{aligned}$$

Sea el punto inicial $x^0 = (2,0)^T$, luego:

$H_1 = -\frac{3}{2}(2)^2 + \frac{13}{2}(2) - 5 = 2$

$H_2 = 2(2)^2 - 2 = 6$

Ahora, para encontrar una dirección descendente, resolvemos el sistema de ecuaciones (5.13) – (5.16). Pero como $\alpha(x^0) = \{1\}$, $\beta(x^0) = \{2\}$ y $\gamma(x^0) = \emptyset$, entonces:

$$\begin{aligned} H_1 + F_{11}d_1 + F_{12}d_2^+ &= 0 \\ H_2 + F_{21}d_1 + F_{22}d_2^+ &\geq 0, \qquad d_2^+ \geq 0 \\ (H_2 + F_{21}d_1 + F_{22}d_2^+)^T d_2^+ &= 0 \end{aligned}$$

Por definición:

$F_{11} = \dfrac{\partial F_1}{\partial x_1} = -3x_1 + \dfrac{13}{2}|_{(2,0)} = \dfrac{1}{2}$

$F_{12} = \dfrac{\partial F_1}{\partial x_2} = 2$

$F_{21} = \dfrac{\partial F_2}{\partial x_1} = 4x_1|_{(2,0)} = 8$

$F_{22} = \dfrac{\partial F_2}{\partial x_2} = 2x_2 - 2|_{(2,0)} = 2$

Reemplazando en el sistema de ecuaciones anterior, tenemos:

$$\begin{aligned} 2 + 0.5dx_1 + 2dx_2^+ &= 0 \ldots (I) \\ 6 + 8dx_1 - 2dx_2^+ &\geq 0, \qquad dx_2^+ \geq 0 \ldots (II) \\ (6 + 8dx_1 - 2dx_2^+)^T d_2^+ &= 0 \ldots (III) \end{aligned}$$

Este LCP no tiene una solución, en efecto:

De (I): $dx_1 = -4 - 4dx_2$

Reemplazando en (II), tenemos:

$6 + 8(-4 - 4dx_2) - 2dx_2^+ = -26 - 34dx_2^+ < 0$

Lo cual no puede ser, por lo tanto, el sistema no tiene solución.

Ahora, aplicamos **el algoritmo modificado de Newton(amortiguado)**

Fijamos: $dx_2^+ = 0$

Reemplazando en (I), tenemos $dx_1 = -4$, por lo tanto:

La dirección descendente será: $d = (-4, 0)^T$.

Luego el siguiente punto será:

$x^{(1)} = x^0 + \lambda d$

Para encontrar el valor del parametro λ realizamos una búsqueda lineal. Después de dicha búsqueda $\lambda = \dfrac{1}{2}$. Por lo tanto:

$x^{(1)} = (2, 0)^T + \dfrac{1}{2}(-4, 0)^T = (1, 0)^T$

Reemplazando $x^{(1)}$ en $(\dagger)$, obtenemos:

$$H(x^{(1)}) = 0$$

Entonces: $x^* = x^{(1)} = (1, 0)^T$ es solución al problema de complementariedad no lineal NCP(F).

Notemos que el vector x^* no es un vector regular, en efecto:

Debido a que $x^* = (1, 0)^T$, tenemos:

$\alpha(x^*) = \{1\}$

$\beta(x^*) = \{2\}$

$\gamma(x^*) = \emptyset$

Recordemos que, para que el vector x^* sea un vector regular debe satisfacer las dos condiciones:

(i) $F_{\alpha\alpha}$ es no-singular

(ii) $F_{\beta\beta} - F_{\beta\alpha}(F_{\alpha\alpha})^{-1}F_{\alpha\beta}$, el Complemento Schur, es una P-Matriz.

Como:

$$\nabla F(x) = \begin{pmatrix} -3x_1 + \frac{13}{2} & 2 \\ 4x_1 & 2x_2 - 2 \end{pmatrix}$$

Entonces:

- $F_{\alpha\alpha} = F_{11} = -3x_1 + \frac{13}{2}|_{(1,0)} = \dfrac{7}{2} \neq 0$, luego $F_{\alpha\alpha}$ es no-singular.
- sabiendo que: $F_{\alpha\alpha} = \frac{7}{2}$, $\quad F_{\alpha\beta} = 2$, $\quad F_{\beta\alpha} = 4$, $\quad F_{\beta\beta} = -2$
  $\Longrightarrow F_{\beta\beta} - F_{\beta\alpha}(F_{\alpha\alpha})^{-1}F_{\alpha\beta} = -2 - 4(\dfrac{7}{2})^{-1}2 = -\dfrac{30}{7} < 0.$
  Luego el Complemento Schur no es una P-Matriz.

Por lo tanto x^* no es un punto regular, pero es regular débilmente. En efecto: Sólo basta tomar una vecindad N de x^* de tal forma que $\alpha(x^*) \subseteq \alpha(x')$ tal que $L_{\alpha(x')}$ sea no singular.

El problema anterior puede también ser formulado como un sistema de ecuaciones a través del uso del operador "min".

Esta formulación nos lleva a:

En el punto inicial $x^0 = (2,0)^T$:

$H_1(x^0) = \text{Min}\{F_1(x^0), x_1^0\} = \text{Min}\{-\frac{3}{2}(x_1^+)^2 + \frac{13}{2}(x_1^+) + 2(x_2^+) - 5, x_1\} = \text{Min}\{2,2\} = 2$

$\Longrightarrow H_1(x^0) = F_1(x^0) = x_1^0 = 2 \ldots (*)$

$H_2(x^0) = \text{Min}\{F_2(x^0), x_2^0\} = \text{Min}\{2(x_1^+)^2 - 2(x_2^+) + (x_2^+)^2 - 2, x_2\} = \text{Min}\{6,0\} = 0$

$\Longrightarrow H_2(x^0) = 0 \Longrightarrow F_2(x^0) > x_2^0 \ldots (**)$

donde lo conjuntos indices: α, β y γ están denotados de la siguiente forma:

$\alpha(x^k) = \{i : x_i^k > F_i(x^k)\}$

$\beta(x^k) = \{i : x_i^k = F_i(x^k)\}$

$\gamma(x^k) = \{i : x_i^k < F_i(x^k)\}$

Según ésta denotación, de (*) y (**):

$\alpha(x^0) = \emptyset, \quad \beta(x^0) = \{1\}, \quad \gamma(x^0) = \{2\}$

Luego: tenemos el siguiente problema de complementariedad lineal LCP:

$H_1 + F_{11}dx_1^+ \geq 0, \quad dx_1^+ \geq 0$

$(H_1 + F_{11}dx_1^+)^T dx_1^+ = 0$

$dx_2^+ = -H_2 - F_{21}dx_1^+$

Reemplazando valores:

$2 + 0.5dx_1^+ \geq 0, \quad dx_1^+ \geq 0$

$(2 + 0.5dx_1^+)^T dx_1^+ = 0$

$dx_2^+ = -8dx_1^+$

La única solución de este sistema de ecuaciones es:

$dx_1^+ = 0$, debido a que $dx_2^+ = -8dx_1^+ = 0$, entonces allí no se puede encontrar una dirección descendente diferente de cero. Por lo tanto, aquí el operador "min" fallará. Ahora sea el punto inicial $x^0 = (0,0)^T$, de manera análoga se puede verificar que es un vector regular débilmente (no es regular).

El subproblema LCP en este punto se convierte en:

$H_\beta + F_{\beta\beta}d_\beta^+ \geq 0, \qquad d_\beta^+ \geq 0$

$(H_\beta + F_{\beta\beta}d_\beta^+)^T d_\beta^+ = 0$

Como $\beta(x) = \{1,2\}$ entonces:

$-5 + \frac{13}{2}dx_1^+ + 2dx_2^+ \geq 0, \quad dx_1^+ \geq 0, \quad (-5 + \frac{13}{2}dx_1^+ + 2dx_2^+)^T dx_1^+ = 0$

$-2 - 2dx_2^+ \geq 0, \quad dx_2^+ \geq 0, \quad (-2 - 2dx_2^+)^T dx_2^+ = 0$ Se observa, que el sistema de arriba no tiene solución. Además, fijando $dx_1 = dx_2 = 0$, tampoco ayuda a la situación.

Sin embargo, debido a que:

$$H_\beta(x^0)^T F_{\beta\beta} = (-5, -2) \begin{pmatrix} 6.5 & 2 \\ 0 & -2 \end{pmatrix} = (-32.5, -6)$$

Usando el algoritmo modificado descrito anteriormente tenemos:

Debido a que $-32.5 < 0$, entonces fijamos $d_1 = 1$

Debido a que $-6 < 0$, entonces fijamos $d_2 = 1$

De este modo obtenemos la dirección descendente $d = (1, 1)^T$. $\square$

6 CONCLUSIONES

Una de las dificultades que presenta el método tradicional de Newton e la diferenciabilidad. Cuando la función en consideración no es F–diferenciable en ciertos puntos, el método puede fallar. El algoritmo de Newton(Amortiguado) supera dichas dificultades aplicando el Método de Newton a las funciones B – diferenciables. La caracteristica de éste nuevo método se basa fundamentalmente en que identifica aquellas partes en la cual puede causar problemas la no diferenciabilidad de la función.

Bajo la suposición de regularidad, el algoritmo modificado de Newton (amortiguado) posee una convergencia cuadrática local. Bajo la suposición de regularidad débil y algunas **Condiciones Mild**, el algoritmo modificado de Newton crea siempre una dirección descendente y converge a la solución. Por lo tanto éste algoritmo de Newton Amortiguado es adecuado para muchas aplicaciones donde la condición de regularidad está ausente.

En el contexto de un programa no-lineal, el algoritmo es del tipo de programación secuencialmente cuadrática con dos característica distintas:

- No hace uso de la función de penalidad.

- Evita el Efecto de Maratos

Con éste trabajo desarrollado estamos en condiciones de realizar experimentos numéricos y comparar los resultados de estos experimentos con los métodos tradicionales como es el método de Newton y se puede mostrar que éste método es más eficiente y robusto que aquellos métodos tradicionales para resolver problemas de complementariedad no–lineal.

BIBLIOGRAFIA

[1] M. Avriel. *Nonlinear Programming Analysis and Methods.* Prentice-Hall. (1976).

[2] M.S. Bazaraa and C.M Shetty. *Nonlinear Programming: Theory and Algorithms.* John Wiley and Sons, 1979

[3] F. Facchinei. "A Simply Constrained optimization reformulation of KKT y tem arising from variational inequalities". T cnical Univer ity of Dre den. Institute of Numerical Mathematics. (1997).

[4] M. Fukushima. "Optimization Based Globally Convergent Methods for the Nonlinear Complementarity Problem".

[5] M. Fuku hima. "Equivalent differentiable optimization problem and de cent m thods for asymmetric variational inequality probl m ". Technical report 9007. Department of Applied Mathematic and Physics. Faculty of Engineering. Kyoto Univer ity (Kyoto. Japan. 19 9).

[6] P.T. Harker and J. Pang. "Finite-Dimen ional Variational Inequality and onlinear Complementarity Problems: Survey of Theory, Algorithms and Aplication ". *Mathematical Programming B-4* . (1990) 1-60.

[7] P.T. Harker and B. Xiao. " ewton' method for the nonlinear complementarity problem: A B-diff rentiable equation approach". *Mathematical programming 48* (1990) 339-357.

[8] D. Kinderlehrer. "An Introduction to variational In qualiti and th ir aplications." Academic Pr s. (19 0).

[9] Malcolm Keswell. " Notes on Second Order "Sufficient" Conditions for Characterising Extrema of Bivariate and Multivariate Functions".

[10] J.S. Pang. "Minimization of Locally Lipschitzian Functions". *Society for Industrial and Applied Mathematics* (1991) 57-82.

[11] J.S. Pang. "Newton's method for B-differentiable equations." *Mathematics of Operations Research 15* (1990) 311-341.

[12] J.S. Pang. " A B-differentiable equation based, globally, and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems". *Mathematical Programming 51* (1991) 101-131.

[13] S.M. Robinson. "False umerical Convergence in some Generalized Newton Methods". Department of Industrial Engineering, University of Wisconsin (Madison, WI, 1995).

[14] S.M. Robinson. "An implicit-function theorem for a class of non mooth functions." Manuscript. Department of Industrial Engineering, 1991.

[15] A. Shapiro. "On concepts of directional differentiability". Research Report 73/ 8(18). Department of Mathematics and Applied Mathematics, University of South Africa (Pretoria, South Africa. 1988)

[16] B. Xiao and P. Harker. " A nonsmooth Newton method for variational inequalities, I: Theory ". Department of System Engineering, University of Pennsylvania, Philadelphia. 1993