

UNIVERSIDAD NACIONAL DEL CALLAO
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
UNIDAD DE INVESTIGACIÓN



INFORME FINAL DE INVESTIGACIÓN

**“APLICACIÓN DEL MÉTODO BARRERA LOGARÍTMICA
PARAMÉTRICA A UN PROGRAMA CONVEXO”**

AUTOR: EDINSON RAUL MONTORO ALEGRE

(PERIODO DE EJECUCIÓN: DEL 01.05.2020 AL 30.04.2021)

(RESOLUCIÓN DE APROBACIÓN N° 280-2020-R)

Callao – 2021

PERU

A handwritten signature in black ink, appearing to read "Edinson Raul Montoro Alegre", is placed over a light blue rectangular background.

Quai d'Or

DEDICATORIA

A mis hijos

A handwritten signature in black ink, appearing to read "Rui Mouta". The signature is written in a cursive style with a large, stylized initial 'R'.

AGRADECIMIENTO

A todas las personas que, de alguna manera, estuvieron motivándome en todo momento para la ejecución y culminación del proyecto.

A handwritten signature in black ink, appearing to read "Arui Quinto", is positioned in the lower right area of the page. The signature is written in a cursive style.

INDICE

Dedicatoria	iii
Agradecimiento	iv
Índice	5
Resumen	7
Abstract	8
Introducción	9
Capítulo I	11
Planteamiento del Problema	11
1.1. Descripción de la realidad problemática	11
1.2. Formulación del Problema	12
1.3. Objetivos	13
1.4. Limitantes de la Investigación	13
Capítulo II	14
Marco Teórico	14
2.1. Antecedentes	14
2.1.1. Internacionales	14
2.1.2. Nacionales	14
2.2. Marco	15
2.2.1. Teórico	15
2.2.2. Conceptual	17
2.3. Definiciones de Términos Básicos	17
Capítulo III	21
Hipótesis y Variables	21
3.1. Hipótesis	21
3.1.1. Hipótesis General	21
3.1.2. Hipótesis Específicas	21
3.2. Definición conceptual de variables	21



3.3. Operacionalización de variables	21
Capítulo IV	22
Diseño Metodológico	22
4.1. Tipo y diseño de la investigación	22
4.2. Método de Investigación	22
4.3. Población y muestra	22
4.4 Técnicas e instrumentos para recolección de la Información	22
4.5. Análisis y procesamiento de datos	22
4.6. Análisis de convergencia	66
Capítulo V	69
Resultados	69
Capítulo VI	70
Discusión de Resultados	70
6.1. Contrastación de Hipótesis	70
6.2. Responsabilidad ética	70
Conclusiones	72
Recomendaciones	73
Referencias Bibliográficas	74



RESUMEN

En este trabajo, se describe una implementación natural del clásico Método de la Función Barrera Logarítmica para un Programa Convexo Diferenciable. Para esto se asume que las funciones objetivo y las funciones restricción satisfacen la condición de Lipschitz con constante de Lipschitz $M > 0$.

En el método propuesto, la búsqueda lineal se hace a lo largo de las direcciones newton con respecto a la función barrera logarítmica estrictamente convexa cuando estamos demasiado lejos de la trayectoria central, y cuando estamos demasiado cerca de la trayectoria central con respecto a una métrica utilizada, solo reducimos el parámetro de barrera. Para dicho método se prueba que, el número de iteraciones que se necesitan para que el algoritmo converja a un ε -solución óptima es de $O\left((1+M^2)\sqrt{n}|\log \varepsilon|\right)$ o $O\left((1+M^2)n|\log \varepsilon|\right)$.

Palabras Claves: Programación Convexa, Método de Función Barrera, Método Punto-Interior



ABSTRACT

In this work, a natural implementation of the classic Logarithmic Barrier Function Method for a Differentiable Convex Program is described. For this, it is assumed that the objective functions and the restriction functions satisfy the Lipschitz condition with Lipschitz constant $M > 0$.

In the proposed method, the linear search is done along the Newton directions with respect to the strictly convex logarithmic barrier function when we are too far from the central path, and when we are too close to the central path with respect to a metric used, we only lower the barrier parameter. For this method it is proved that the number of iterations needed for the algorithm to converge to an ε -optimal solution is $O((1+M^2)\sqrt{n}|\log \varepsilon|)$ or $O((1+M^2)n|\log \varepsilon|)$.

Keywords: Convex Programming, Barrier Function Method, Point-Interior Method



INTRODUCCION

Los métodos clásicos de penalidad para resolver problemas de optimización restringida

$$\min\{f_0(x) / f_i(x) \leq 0 \ i = 1,2,3, \dots, n\}$$

Pueden ser divididos en dos clases: métodos de punto interior y exterior. Los métodos de punto interior operan en el interior de la región factible mientras que los métodos de punto exterior intentan obtener la solución desde la parte externa de la región factible.

Ambas clases pueden ser divididas en dos subclases. Los así llamados métodos paramétricos debido a que tienen un parámetro de control en la función de penalidad para controlar la convergencia hacia la solución. Los métodos no paramétricos no operan explícitamente con un parámetro de control, pero la convergencia hacia el mínimo es controlada por una sucesión de niveles de truncamiento de la región factible, convergiendo a z^* el valor mínimo desconocido del problema.

Métodos de punto interior paramétrico están basados en funciones penalidad de la forma

$$f_0(x) - \mu \sum_{i=1}^n \Psi(-f_i(x))$$

Donde μ denota un parámetro de control positivo y Ψ es una función de una variable η , tal que $\Psi(0+) = -\infty$. Debido a la singularidad que presenta en $\eta=0$, dichas funciones son llamadas Funciones Barrera.

Muchas elecciones para Ψ han sido propuestas en la literatura, por ejemplo, la función inversa $-\eta^{-1}$, la inversa de la función cuadrática $-\eta^{-2}$, pero la función $\log \eta$ es la que ha atraído la mayor atención en la literatura.

Métodos de punto interior no paramétrico están basados en funciones barrera de la forma

$$-q\Psi(z - f_0(x)) - \sum_{i=1}^n \Psi(-f_i(x))$$



Donde q es un factor de peso positivo. Su nombre puede ser algo engañoso, debido a que el nivel de truncamiento z puede ser visto como un parámetro de control. Sin embargo, ellos son llamados no paramétricos debido a que en esos métodos el parámetro de control no es usado explícitamente, pero este es ajustado automáticamente

$$z^k = z^{k-1} - \theta(z^{k-1} - f_0(x^k))$$

Donde $0 < \theta < 1$ es un factor de relajación, para asegurar que

$$z^{k-1} > z^k > z^*$$

Nuevamente aquí la función $\log(\eta)$ capta la atención de muchos.

En este trabajo de investigación se pretende utilizar una variante de la función logaritmo para a través del método de la función barrera aplicarla y obtener la solución de un programa convexo.



CAPITULO I

PLANTEAMIENTO DEL PROBLEMA

1.1 DESCRIPCIÓN DE LA REALIDAD PROBLEMÁTICA

Los métodos clásicos de penalidad para resolver problemas de optimización restringida

$$\min\{f_0(x) / f_i(x) \leq 0 \ i = 1,2,3, \dots, n \}$$

Pueden ser divididos en dos clases: métodos de punto interior y exterior. Los métodos de punto interior operan en el interior de la región factible mientras que los métodos de punto exterior intentan obtener la solución desde la parte externa de la región factible.

Ambas clases pueden ser divididas en dos subclases. Los así llamados métodos paramétricos debido a que tienen un parámetro de control en la función de penalidad para controlar la convergencia hacia la solución. Los métodos no paramétricos no operan explícitamente con un parámetro de control, pero la convergencia hacia el mínimo es controlada por una sucesión de niveles de truncamiento de la región factible, convergiendo a z^* el valor mínimo desconocido del problema.



Métodos de punto interior paramétrico están basados en funciones penalidad de la forma

$$f_0(x) - \mu \sum_{i=1}^n \Psi(-f_i(x))$$

Donde μ denota un parámetro de control positivo y Ψ es una función de una variable η , tal que $\Psi(0^+) = -\infty$. Debido a la singularidad que presenta en $\eta=0$, dichas funciones son llamadas Funciones Barrera.

Muchas elecciones para Ψ han sido propuestas en la literatura, por ejemplo, la función inversa $-\eta^{-1}$, la inversa de la función cuadrática $-\eta^{-2}$, pero la función $\log\eta$ es la que ha atraído la mayor atención en la literatura.

Métodos de punto interior no paramétrico están basados en funciones barrera de la forma

$$-q\Psi(z - f_0(x)) - \sum_{i=1}^n \Psi(-f_i(x))$$

Donde q es un factor de peso positivo. Su nombre puede ser algo engañoso, debido a que el nivel de truncamiento z puede ser visto como un parámetro de control. Sin embargo, ellos son llamados no paramétricos debido a que en esos métodos el parámetro de control no es usado explícitamente, pero este es ajustado automáticamente

$$z^k = z^{k-1} - \theta(z^{k-1} - f_0(x^k))$$

Donde $0 < \theta < 1$ es un factor de relajación, para asegurar que

$$z^{k-1} > z^k > z^*$$

Nuevamente aquí la función $\log(\eta)$ capta la atención de muchos.

1.2 FORMULACIÓN DEL PROBLEMA

La idea de usar la función logaritmo como herramienta para resolver problemas de optimización datan del año 1955 cuando Frish (1955a) lo usa para problemas de programación lineal. Posteriormente el mismo Frish (1955b) utiliza una función al que llamo Función Potencial Logarítmica para resolver un problema convexo, pero solo uso el gradiente de la función para mantenerse alejado de la frontera de la región factible. Años después Faicco y McCormick (1968) sistematiza las ideas de funciones barrera para resolver problemas de programación no lineal. Posteriormente el trabajo de Karmarkar (1984) marca un hito al publicar su algoritmo que resuelve un problema de programación lineal en tiempo polinomial. Luego Gill, Murray, Saunders, Tomlin y Wright (1986) logran deducir una relación entre el método e barrera logarítmica paramétrica y el método de Karmarkar pero para el caso de problemas de programación lineal. Vemos que tratan con más énfasis el caso de programación lineal, no habiendo tratado a profundidad, POR LO TANTO se puede definir el

Problema general: De qué manera un Programa convexo podría ser resuelto usando alguna variante de la función logaritmo?

Problema específico: De que forma el Método de Barrera Logarítmica puede solucionar el Programa Convexo?



1.3 OBJETIVOS

1.3.1. OBJETIVO GENERAL

Aplicar una variante de la función logaritmo a programación convexa.

1.3.2. OBJETIVO ESPECIFICO

Aplicar el Método de Barrera Logarítmico para resolver un Programa Convexo.

1.4 LIMITANTES DE LA INVESTIGACIÓN

Teóricos: La investigación, es del tipo teórica y se tiene como limitante o enmarcada dentro de la teoría de optimización paramétrica.

Temporales: No aplica

Espaciales: No aplica



CAPITULO II

MARCO TEÓRICO

Este capítulo está dedicado a una revisión de algunos conceptos matemáticos en el análisis de Programación No-Lineal.

2.1 Antecedentes

2.1.1 Internacionales

A nivel internacional tenemos los siguientes datos. Gill et al.(1986) fueron los primeros en derivar una relación entre el método de Karmarkar (1984) y el método de barrera logarítmica paramétrica para programación lineal. Su equipo también implemento el método de barrera logarítmica obteniendo resultados importantes. Sin embargo, ellos no probaron la polinomialidad como lo hizo Karmarkar (1984). Gonzaga (1989) fue el primero en probar la polinomialidad para una versión especial del método barrera logarítmica para programación lineal. En este método el parámetro de barrera es reducida por un factor $1-0.005/n$ lo cual significa que solo pasos muy pequeños son tomados. Lo que lo hacía en la práctica un método excesivamente lento.

Después, Roos y Vial (1990) y Gonzaga (1989) desarrollaron independientemente una versión más natural y practica de método de barrera logarítmica para programación lineal. En este método el parámetro de control se puede reducir por un factor arbitrario entre 0 y 1. Esto significa que los iterados no necesitan estar cerca de la trayectoria central y por tanto pasos grandes pueden darse. Ellos también probaron que este método necesitaba $O(nL)$ iteraciones.

2.1.2 Nacionales

La teoría de Métodos de Punto interior en nuestro país ha sido un área nueva de interés, en particular en la universidad Nacional Mayor de San Marcos, el profesor Edgar Quijano Urbano trabajo un algoritmo Primal-Dual basado en el método de barrera logarítmica aplicado a un problema de programación lineal (ver Edgar, 2019)



2.2 Marco

2.2.1 Teórico

Un modo de resolver problemas de optimización con restricciones en forma de desigualdad

$$(P) \begin{cases} \text{Min } f(x) \\ \text{s. a} \\ g_1(x) \leq 0 \\ g_2(x) \leq 0 \\ \vdots \\ g_m(x) \leq 0 \end{cases}$$

Es mediante la aproximación a este problema a través de una familia de problemas de minimización irrestricta

$$(P') \{ \text{Min } F(x) \text{ para } x \in R^n$$

Donde la función objetivo $F(x)$ del problema irrestricto es construido a partir de la función objetivo $f(x)$ y de las restricciones del problema restringido de tal modo que:

- $F(x)$ incluye un término de penalidad el cual incrementa el valor de $F(x)$ mientras que la restricción $g_i(x) \leq 0$ es violada. Es decir, violaciones grandes resultan en incrementos grandes.
- El mínimo x_F^* del problema (P') este próxima a la región factible y x_F^* está próxima al mínimo del problema (P) .

Usando esta aproximación, se espera que, como el tamaño del término penalidad en $F(x)$ aumenta, el mínimo x_F^* de $F(x)$ se aproximará al mínimo del problema (P) .

Por otro lado, un Programa Convexo es un problema de optimización de la forma

$$(P) \begin{cases} \text{Min } f(x) \\ \text{s. a} \\ g_1(x) \leq 0 \\ g_2(x) \leq 0 \\ \vdots \\ g_i(x) \leq 0 \end{cases}$$

En la cual todas las funciones involucradas son funciones convexas y la clave para resolver dicho problema es la aplicación de las famosas condiciones o teorema de Karush-Khunn-Tucker. Este resultado asocia un programa convexo con un sistema de ecuaciones algebraicas y desigualdades que con frecuencia pueden ser usadas para desarrollar procedimientos efectivos para calcular mínimos, y también pueden ser usados para obtener

información adicional acerca de la sensibilidad del valor mínimo del programa convexo a cambios en las restricciones.

A continuación presentamos parte del material bibliográfico utilizado en el marco teórico.

FRISCH, R. (1955a). **The Logarithmic Potential Method for Solving Linear Programming Problems.** *Memorandum, Institute of Economics, Oslo, Norway.*

FRISCH, R. (1955b). **The Logarithmic Potential Method of Convex Programming.** *Memorandum, Institute of Economics, Oslo Norway.*

PARISOT, G. R. (1961). **Resolution Numérique Approchee du Problème de Programmation Linéaire par Application de la Programmation Logarithmique.** *Revue Francaise de Recherche Operationelle*, 20, 227-259.

FIACCO, A. V., and McCormick, G. P. (1968). **Nonlinear Programming: Sequential Unconstrained Minimization Techniques,** *Wiley, New York, New York.*

KARMAKAR, N. (1984). **A New Polynomial-Time Algorithm for Linear Programming.** *Combinatorica*, 4, 373-395.

Gill, P. E., Murray, W., Saunders, M., A., Tomlin, J., A., and Wright, M., H. (1986). **On Projected Newton Barrier Methods for Linear Programming and an Equivalence to Karmarkar's Projective Method.** *Mathematical Programming*, 36, 183-209.

Gonzaga, C, C. (1989). **An Algorithm for Solving Linear Programming Problems in $O(nL)$ Operations,** *Progress in Mathematical Programming: Interior Point Methods, Edited by N., Megiddo , Springer Verlag, New York, 1-28.*

Roos, C., and Vial, J., P. (1990). **Long Steps with the Logarithmic Penalty Barrier Function in Linear Programming, Economic Decision Making: Games Economics and Optimization.** *Edited by J. Gabszewiecz, J. F., Richard, and L. Wolsey, Elsevier Science Publisher, Amsterdam, Holland, 433-441.*

Gonzaga C, C. (1991). **Large-Steps Path-Following Methods for Linear Programming: Barrier Function Method,** *SIAM Journal on Optimization*, 1, 33-41.

Renegar, J. (1988). **A Polynomial-Time Algorithm Based on Newton's Methods for Linear Programming,** *Mathematical Programming*, 40, 59-93.



2.2.2. Conceptual

La teoría de programación no lineal con restricciones y la teoría de programación paramétrica permite encontrar las condiciones necesarias para poder aplicar el método de barrera logarítmica a la programación convexa, lo cual permite el planteamiento de un algoritmo para resolver el problema en estudio.

2.3 Definiciones de términos básicos

Definición (Conjuntos convexos)

Un conjunto C es llamado convexo si para cualquier par de puntos x, y en el conjunto, el segmento que los une pertenece al conjunto C .

Definición (Semiespacio)

Si $x^* \in R^n$ y si $\alpha \in R$, entonces se define el semiespacio cerrado

$$F^+ = \{x \in R^n : x^* \cdot x \geq \alpha\}$$

$$F^- = \{x \in R^n : x^* \cdot x \leq \alpha\}$$

Análogamente Semiespacio abierto

$$G^+ = \{x \in R^n : x^* \cdot x > \alpha\}$$

$$G^- = \{x \in R^n : x^* \cdot x < \alpha\}$$

Definición (Combinación Convexa)

Si x^1, x^2, \dots, x^k son vectores en R^n , entonces la Combinación Convexa de x^1, x^2, \dots, x^k es cualquier vector de la forma $\lambda_1 x^1 + \lambda_2 x^2 + \dots + \lambda_k x^k$, donde los λ_i son números reales positivos tal que $\sum_{i=1}^k \lambda_i = 1$.

Definición (Capsula Convexa)

Sea D cualquier conjunto en R^n , la intersección de todos los conjuntos convexos que contienen a D es llamado la Capsula Convexa de D .

Definición (Función Convexa)

Sea $f(x)$ una función real valuada definida sobre un conjunto convexo C en R^n . Entonces

- La función $f(x)$ es convexa sobre C si

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

- La función $f(x)$ es estrictamente Convexa sobre C si



$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Definición (Problema Primal)

La formulación primal del problema de programación convexa suave:

$$(CP) \quad \text{mín} \{ f_0(y) : y \in F \}$$

Donde F denota la región factible, que es dada por

$$F = \{ y \in \mathbb{R}^n : f_i(y) \leq 0 ; 1 \leq i \leq n \};$$

las funciones $f_i(y)$, $0 \leq i \leq n$ son funciones convexas continuas con derivadas de primer y segundo orden en F , se denomina Problema Primal.

Definición (Condición de Suavidad)

Se dice que la **condición de suavidad** es satisfecha por f si su matriz hessiana satisface la *Condición de Lipschitz*.

Hipótesis

Supondremos que el interior de la región factible F denotado por F^0 es no vacío y acotado, sin pérdida de generalidad nosotros asumimos que $f_0(y)$ es lineal, esto quiere decir $f_0 = -b^T y$ y con $b \neq 0$. Si esto se puede introducir una variable adicional y_{m+1} , una restricción adicional $f_0 - y_{m+1} \leq 0$ y minimice y_{m+1} consecuentemente bajo ciertas consideraciones el problema a tratar será:

$$(CP) \quad \text{mín} \{ -b^T y : y \in F \}.$$

La formulación de *Wolfe(Ref.26)* del problema dual asociado con este problema primal es:



$$\begin{aligned} \min \quad & -b^T y - \sum x_i f_i(y) \\ (D) \quad & \text{s.a. } \sum_{i=1}^n x_i \nabla f_i(y) = b \\ & x_i \geq 0 \end{aligned}$$

Observación

Note que no existe simetría entre el problema primal y dual desde el punto de vista de programación lineal, porque el problema dual (D) contiene ambas variables x e y . Además, el problema dual no es necesariamente convexo.

Sin embargo, un resultado bien conocido (*Ref.26*) es que si y es una solución factible del problema dual (CP) y (\bar{y}, x) es una solución factible del problema dual (D), entonces se cumple

$$b^T y \leq b^T \bar{y} - \sum_{i=1}^n x_i f_i(\bar{y}).$$

Debido a la suposición que F es no vacío, se puede concluir que (D) tiene una solución mínima y los valores extremos son iguales.

Definición (Función Barrera Logarítmica)

Se define la función de barrera logarítmica asociada al problema (CP) como

$$\phi(y, \mu) = -b^T y / \mu - \sum_{i=1}^n \log[-f_i(y)],$$

donde μ es el parámetro de barrera. Las derivadas de primer y segundo orden de $\phi(y, \mu)$ están dadas por:

$$g(y, \mu) = \nabla \phi(y, \mu) = -b / \mu + \sum_{i=1}^n \nabla f_i(y) / [-f_i(y)] \quad 1(a)$$

$$H(y, \mu) = \nabla^2 \phi(y, \mu) = \sum_{i=1}^n \left[\nabla^2 f_i(y) / [-f_i(y)] + \nabla f_i(y) \nabla^T f_i(y) / f_i(y)^2 \right] \quad 1(b)$$



Si no hay confusión es posible escribir por brevedad g y H en lugar de $g(y, \mu)$ y $H(y, \mu)$.

Definición (Aproximación Cuadrática)

Se define la aproximación cuadrática $q_y(y, \mu)$ para $\phi(y, \mu)$ en el punto y , dado por:

$$q_y(y, \mu) = \phi(\bar{y}, \mu) + g^T(y, \bar{y}) + (1/2)(y - \bar{y})^T H(y - \bar{y}).$$

Nosotros usaremos la H -norma denotada por $\|\cdot\|_H$ para medir la distancia de puntos especialmente próximos o cercanos a la trayectoria central.

Definición (H-norma)

La definición de esta norma es dada por:

$$\|z\|_H = \sqrt{(z^T H z)}$$

Debido a que H es definida positiva, $\|\cdot\|_H$ satisface las condiciones para ser norma introducida esta notación,

Definición (Condición de Lipschitz)

Se define la condición de Lipschitz de la siguiente manera:

$$\exists M > 0 / \forall v \in R^n, \forall y, y + h \in F', \forall 1 \leq i \leq n$$

$$\left| v^T (\nabla^2 f_i(y+h) - \nabla^2 f_i(y)) v \right| \leq M \|h\|_H v^T \nabla^2 f_i(y) v. \quad (2)$$

Una condición similar es usada por Jarre (Ref.22). En general la condición puede ser difícil comprobación para un problema dado.



CAPITULO III

HIPOTESIS Y VARIABLES

3.1 HIPÓTESIS

Hipótesis general

Si las funciones involucradas en el Programa Convexo satisfacen condiciones de suavidad entonces se le puede aplicar el método de barrera logarítmica paramétrica.

Hipótesis específica

1. Si las hessianas de las funciones involucradas en el Programa Convexo satisfacen la condición de Lipschitz entonces se le aplica el método de barrera logarítmica paramétrica.
2. Si se aplica el método de barrera logarítmica paramétrico al Programa Convexo entonces se genera un esquema iterativo de solución para dicho problema.

3.2. Definición conceptual de las variables

Las variables identificadas en la hipótesis general, se pueden definir conceptualmente como:

3.2.1 VARIABLE DEPENDIENTE: La función Barrera Logarítmica

3.2.2 VARIABLE INDEPENDIENTE: Hessiana de la función

3.3. OPERACIONALIZACIÓN DE LAS VARIABLES.

VARIABLE INDEPENDIENTE	DEFINICIÓN	DIMENSIONES	INDICADORES	INDICES	METODO	TECNICA
Hessianas de las funciones	Segunda Derivada de la función	La teoría de programación no lineal	La condición de Lipschitz	Conjunto factible	Demostrativo Deductivo e inductivo	Construtivo
VARIABLE DEPENDIENTE	DEFINICIÓN	DIMENSIÓN	INDICADORES	INDICES	METODO	TECNICA
Función Barrera Logarítmica	Es la función barrera a utilizar	La teoría de Programación paramétrica	La dirección newton	El parametro de la función barrera	Demostrativo Deductivo e inductivo	Construtivo

CAPITULO IV

DISEÑO METODOLÓGICO

4.1. Tipo y diseño de la investigación

El tipo de investigación es básica, según Alva Lucía Marín Villada (2008), “También llamada investigación pura, teórica o dogmática”. Se caracteriza porque parte de un marco teórico y permanece en él; la finalidad radica en formular nuevas teorías o modificar las existentes, incrementando los conocimientos científicos o filosóficos, pero sin contrastarlos con ningún aspecto práctico.

Es un estudio básico porque mediante el cual, se buscará aportar conocimientos que permitan mejorar algunos detalles del marco teórico. El diseño es no experimental y tiene un enfoque cualitativo

4.2. Método de Investigación

Teniendo en consideración lo planteado y/o descrito en el proyecto, el método utilizado en el desarrollo es demostrativo e inductivo-deductivo

4.3. Población y muestra

Población: No aplica.

Muestra: No aplica

4.4. Técnicas e instrumentos de Recolección de la Información

La recolección de la información para resolver el problema del presente proyecto se realizara mediante búsqueda bibliográfica, búsqueda por internet, etc. asimismo, a través de la utilización de técnicas de análisis deductiva nos conducirán progresivamente a la resolución del problema planteado.

4.5. Análisis y procesamiento de datos. Los análisis de la información sobre el problema convexo determinaran las necesidades de conseguir información para su respectivo análisis y su incorporación en el estudio para la solución del problema convexo a través de la aplicación de la función barrera logarítmica.

A continuación presentamos, algunos comentarios, conceptos, proposiciones entre otros que nos permita establecer y obtener nuestro resultado.



4.5.1. MÉTODO DE LA FUNCIÓN DE BARRERA

Similarmente a las funciones de penalidad, las funciones de barrera son también usadas para transformar un problema restringido en un problema no restringido o en una sucesión de problemas no restringidos.

Estas funciones trabajan como una barrera que impide que la solución se salga de la región factible. Si la solución óptima ocurre en la frontera de la región factible el proceso se manifiesta desde el interior a la frontera de la región factible. Los problemas primal y de barrera se tratan abajo.

4.5.2 PROBLEMA PRIMAL

Consideremos el siguiente problema de optimización

$$\text{minimizar } f(x)$$

$$\text{s.a. } g(x) \leq 0$$

$$x \in X,$$

donde $g(x)$ es un vector función cuyos componentes son g_1, \dots, g_m . Aquí f, g_1, \dots, g_m son funciones continuas en E_n y X es un conjunto no vacío en E_n . Notar que cualesquiera restricciones de iguales, si en caso se presentan, son acomodadas con el conjunto X .

Alternativamente, en el caso de restricciones de igualdad lineal, podemos eliminarlas haciendo al respecto que algunas variables están en función de otras, para así reducir la dimensión del problema. La razón por la cual se hace esto necesario es que los métodos de la función de barrera requieren que el conjunto $\{x: g(x) < 0\}$ sea no vacío, lo cual podría obviamente no ser posible si las restricciones de igualdad $h(x) = 0$ fueran acomodadas con el conjunto de desigualdades como $h(x) \leq 0$ y $h(x) \geq 0$.

4.5.3 PROBLEMA DE BARRERA

minimizar $\theta(\mu)$

s.a. $\mu \geq 0$

donde $\theta(\mu) = \inf \{f(x) + \mu B(x); g(x) < 0 \in X\}$. Aquí B es una función de barrera que es no negativa y continua sobre la región $\{x: g(x) < 0\}$ y se aproxima a ∞ cuando la frontera de la región $\{x: g(x) < 0\}$ se aproxima desde el interior más específicamente, la función de barrera B es definida por

$$B(x) = \sum_{i=1}^m \phi[g_i(x)]$$

donde ϕ es una función de una variable que es continua sobre $\{y: y < 0\}$ y satisface

$$\phi(y) \geq 0, \text{ si } y < 0 \quad \wedge \quad \lim_{y \rightarrow 0^-} \phi(y) = \infty$$

Así, una función típica de barrera es de la forma

$$B(x) = \sum_{i=1}^m \frac{-1}{g_i(x)} \quad \vee \quad B(x) = \sum_{i=1}^m \ln(\min\{1, -g_i(x)\})$$

Notar que la segunda función de barrera es no diferenciable debido al término

$$\min\{1, -g_i(x)\}$$

Ahora, para ϕ es esencial solamente en una vecindad de $y=0$, se puede mostrar que la siguiente función de barrera, conocida como Función Logarítmica de Barrera de Frisch

$$B(x) = - \sum_{i=1}^m \ln[g_i(x)].$$

Nos referiremos a la función $f(x) + \mu B(x)$ como una Función Auxiliar. Idealmente nos gustaría que la función B tomara valor cero en la región $\{x: g(x) < 0\}$ y valor ∞ en la frontera.

Esto nos garantizaría que no caería en la región $\{x: g(x) \leq 0\}$ siempre que el problema de minimización inicie en un punto interior. Sin embargo, esta discontinuidad ϕ posee serias dificultades para algunos procesos computacionales. Por tanto, esta construcción ideal de B es reemplazada por el más realista requerimiento, o sea que B es no negativa y continua sobre la región $\{x: g(x) < 0\}$ y que se aproxima al infinito así como que la frontera es aproximada desde el interior. Notar que μB aproxima la función de barrera ideal descrita arriba cuando μ se aproxima a cero. Dado $\mu > 0$ y evaluando $\theta(\mu) = \inf \{f(x) + \mu B(x); g(x) < 0, x \in X\}$ no parece tan simple resolver el problema original debido a la presencia de la restricción $g(x) < 0$. Sin embargo, como resultado de la estructura de B , así comenzamos la optimización desde un punto en la región $S = \{x: g(x) < 0\} \cap X$ e ignoramos la restricción $g(x) < 0$, se tendrá un punto optimal en S . Esto resulta del hecho de que como aproximamos la frontera de $\{x: g(x) \leq 0\}$ de S , B se aproxima al infinito, lo cual impedirá que se caiga en el conjunto S . Esto se discutirá detalladamente más adelante.

4.5.4. MÉTODO DE NEWTON



Sea la función $F: R^n \rightarrow R^n$, se busca un cero de la función, es decir, busca un punto $\bar{\lambda} \in R^n$ tal que $F(\bar{\lambda}) = 0$, se realiza a través de un proceso iterativo, en el cual para cada iteración se parte de un punto dado $\lambda \in R^n$ y se recorre una aproximación de la dirección $\Delta \lambda$ tal que $F(\lambda + \Delta \lambda) = 0$. Esta aproximación se calcula a partir de un modelo lineal para la función F , dado por el truncamiento de su expansión en serie de Taylor en torno de λ ,

$$F(\lambda + \Delta \lambda) \approx F(\lambda) + J(\lambda) \Delta \lambda$$

Donde J es el jacobiano de F . El cálculo de la dirección $\Delta \lambda$ a ser tomada para cada iteración es realizado a través de la solución del sistema lineal $J(\lambda)\Delta \lambda = -F(\lambda)$.

Escribiremos el sistema como $F(\lambda) = 0$ donde $\lambda = (x, u, s)$ y

$$F(\lambda) = F(x, u, s) = \begin{bmatrix} A^T u + 8 - c \\ Ax - b \\ XSe - \mu e \end{bmatrix}$$

En este caso el jacobiano de F es dado por:

$$F(\lambda) = J(x, u, s) = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}$$

y la dirección de Newton $\Delta \lambda := (\Delta x, \Delta u, \Delta s)$; que calculamos a partir de un punto $(x, u, s) \in L^0$, dada por la solución del siguiente sistema lineal

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu e - XSe \end{bmatrix}$$

la dirección de Newton para la k -ésima iteración, entonces la nueva solución será:

$$\begin{aligned} x^{k+1} &= x^k + \beta_p^k \Delta x^k \\ u^{k+1} &= u^k + \beta_D^k \Delta u^k \\ s^{k+1} &= s^k + \beta_D^k \Delta s^k \end{aligned}$$



donde β_p^k y β_D^k son longitudes de paso para las variables primal y dual, respectivamente. La condición no negatividad para x^{k+1} y s^{k+1} hacen posible la elección de β_p^k y β_D^k .

Para preservar la no negatividad diremos

$$\beta_p^k := \text{mínimo} \left\{ 1, \alpha \text{ mínimo}_{\{i/(\Delta x^k)_i < 0\}} \left(-\frac{x_i^k}{(\Delta x^k)_i} \right) \right\},$$

$$\beta_D^k := \text{mínimo} \left\{ 1, \alpha \text{ mínimo}_{\{i/(\Delta s^k)_i < 0\}} \left(-\frac{s_i^k}{(\Delta s^k)_i} \right) \right\}$$

donde $0 < \alpha < 1$, además $(\Delta x^k)_i$ y $(\Delta s^k)_i$ es la i -ésima componente de Δx^k y Δs^k respectivamente; la razón por la que buscamos un α menor que 1 es evitar aproximarnos a la frontera del conjunto factible primal y del dual, respectivamente.

Finalmente notamos que las restricciones de igualdad del problema primal y dual se satisfacen, en efecto ahora del sistema se desprende que:

$$A\Delta x = 0$$

$$A^T \Delta u + \Delta s = 0$$

para la k -ésima iteración se tiene

$$A\Delta x^k = 0$$

$$A^T \Delta u^k + \Delta s^k = 0$$

Entonces, la nueva solución es factible.

Ahora regresando del sistema vamos a obtener de su tercera ecuación la siguiente expresión para Δs .

$$\Delta s = X^{-1}(\mu e - XSe - S\Delta x)$$



sustituyendo, en la primera ecuación de, tenemos

$$A^T \Delta u + X^{-1} S \Delta x = S - \mu X^{-1} e$$

Resolviendo para Δx , obtenemos

$$\Delta x = X S^{-1} (S - \mu X^{-1} e - A^T \Delta u)$$

Poniendo un $S = -A^T u + c$

$$\Delta x = X S^{-1} (-A^T u + c - \mu X^{-1} e - A^T \Delta u)$$

Reemplazando en la segunda ecuación:

$$-A X S^{-1} A^T \Delta u = -A X S^{-1} (-A^T u + c - \mu X^{-1} e)$$

resolviendo para Δu y sustituyendo, tenemos

$$A^T \left((A X S^{-1} A^T)^{-1} A X S^{-1} (c - A^T u - \mu X^{-1} e) \right) + X^{-1} S \Delta x = S - \mu X^{-1} e$$

luego al despejar

$$X^{-1} S \Delta x = -\mu X^{-1} e + S - A^T \left((A X S^{-1} A^T)^{-1} A X S^{-1} (c - A^T u - \mu X^{-1} e) \right)$$

Sea: $D^2 = X S^{-1}$ y realizando otra vez $S = c - A^T u$, entonces:

$$\begin{aligned} \Delta x &= (-D^2 - D^2 A^T (A D^2 A^T)^{-1} A D^2) (c - A^T u - \mu X^{-1} e) \\ &= (-D^2 - D^2 A^T (A D^2 A^T)^{-1} A D^2) c \\ &= -\mu \left(-D^2 - D^2 A^T (A D^2 - A^T)^{-1} A D^2 \right)^{X^{-1}} e \end{aligned}$$

si $\mu = 0$, entonces:

$$\Delta x = \left(-D^2 - D^2 A^T (AD^2 A^T)^{-1} AD^2 \right) c$$

donde Δx es la solución del sistema.

Si $\mu = 0$, el sistema de ecuaciones en conjunto con las restricciones de no negatividad para x y s representan las condiciones de optimalidad para los problemas primal y dual. La dirección de Newton calculada para $\mu = 0$, apunta hacia el óptimo de estos problemas. En general, se puede caminar poco a poco a lo largo de esta dirección sin que una de las componentes de x ó s sean negativos y consecuentemente, el progreso obtenido a lo largo de este camino es pequeño.

Para $\mu > 0$, entonces será un punto perteneciente a la trayectoria central. La dirección de Newton apunta hacia la trayectoria central, hacia el interior del octante no negativo.

Entonces la dirección calculada para algún $\mu > 0$, se obtiene desviando hacia el interior del octante no negativo, la dirección que apunta hacia el óptimo del problema. El objetivo de este desvío es mantenerse alejado de la frontera del octante de manera que se permita un mayor paso a lo largo de la dirección calculada, sin salirse de la región factible del problema.



4.5.5. PROPIEDADES DE LA TRAYECTORIA CENTRAL

Se puede probar que $\phi(y, \mu)$ es estrictamente convexa sobre este dominio F ver (Ref. 22), además toma infinitos valores sobre la frontera del conjunto factible. Esto permite concluir que para μ fijo, esta función alcanzará el mínimo valor en este dominio en un único punto, el cual es denotado por $y(\mu)$.

Las condiciones necesarias y suficientes Karush-Kuhn-Tucker para este mínimo son:

$$f_i(y) \leq 0, 1 \leq i \leq n \quad (3)$$

$$\sum_{i=1}^n x_i \nabla f_i(y) = b, x \geq 0 \quad (4)$$

$$-f_i(y)x_i = \mu - 1 \leq i \leq n \quad (5)$$

La función de barrera logarítmica no paramétrica:

$$q \log(b^T y - z) - \sum_{i=1}^n \log[-f_i(y)]$$

es relativamente próximo a la contraparte paramétrica ϕ . Lootsma (Ref.5) muestra en un contexto más general, que el punto de minimización para la función barrera logarítmica no paramétrica con nivel z de truncación, llamado el centro $y(z)$, minimiza la función logarítmica paramétrica para:

$$\mu = \frac{q}{[b^T y(z) - z]}$$

Así llamaremos al punto minimizador $y(\mu)$ "centro de (CP)" con respecto a μ .

La trayectoria central del problema es definido como el conjunto de centros $y(\mu)$, donde μ recorre de ∞ a 0. Note que $y(\mu)$ es primal factible y también $x(\mu)$ es dual factible, además las diferencias de valores de las funciones, llamada "dualidad Gap", que satisface la relación:

$$\begin{aligned} b^T y(\mu) - \sum_{i=1}^n x_i(\mu) \cdot f_i(y(\mu)) - b^T y(\mu) \\ = - \sum_{i=1}^n x_i(\mu) \cdot f_i(y(\mu)) = n\mu \end{aligned} \quad (6)$$

De aquí, si $\mu \rightarrow 0$, entonces $y(\mu)$ y $x(\mu)$ converge para soluciones óptimas de los problemas (CP) y (D), esto significa que la trayectoria central termina en la solución óptima del problema. De la literatura especializada sabemos que $b^T y(\mu)$ es monótonamente creciente si μ decrece. Para programación lineal es bien sabido que la función objetivo del dual a lo largo de la trayectoria central es monótona decreciente

si decrece μ . Sin embargo para programación convexa suave no podemos encontrar una demostración, ni aun para la programación cuadrática convexa.

Un primer resultado es el lema siguiente el cual nos dice que la función objetivo dual

$$b^T y(\mu) - \sum_{i=1}^n x_i(\mu) \cdot f_i(y(\mu))$$

es monótona decreciente si μ decrece.

Lema 4.5.1.-La función objetivo $b^T y(\mu)$ del problema primal (CP) es estrictamente monótona decreciente, y la función objetivo del problema dual (D),

$$b^T y(\mu) - \sum_{i=1}^n x_i(\mu) \cdot f_i(y(\mu)),$$

es monótonamente decreciente si μ decrece.

Prueba.-Nosotros tenemos esto $x(\mu)$ y $y(\mu)$ que satisface las condiciones de $K - K - K$ (3) - (5), tomando las derivadas con respecto a μ de la ecuación (4) y (5), nosotros obtendremos:

$$\left(\frac{d}{d\mu}\right)\left(\sum_{i=1}^n x_i \Delta f_i(y)\right) = \left(\frac{d}{d\mu}\right)b$$

$$\left(\frac{d}{d\mu}\right)(-f_i(y)x_i) = \left(\frac{d}{d\mu}\right)\mu$$

de la primera expresión se tiene:

$$x_1'(\mu) \nabla f_1(y(\mu)) + x_1(\mu) \left(\frac{d}{d\mu}\right) \nabla f_1(y(\mu)) + \dots + x_n'(\mu) \nabla f_n(y(\mu)) + x_n(\mu) \left(\frac{d}{d\mu}\right) \nabla f_n(y(\mu)) = 0$$

esto es igual



$$\sum_{i=1}^n x_i' \nabla f_i (y(\mu)) + \sum_{i=1}^n x_i H_i y' = 0 \quad (7)$$

de la segunda expresión se tiene

$$\begin{aligned} & -x_i' f_i (y) - x_i \left(\frac{d}{d\mu} \right) f_i (y(\mu)) \\ & -x_i' f_i (y) - x_i \nabla f_i (y)^T y' = 1, \quad i=1, \dots, n \end{aligned} \quad (8)$$

donde el primero denota la derivada con respecto a μ y H_i denota la matriz hessiana de $f_i (y)$. El jacobiano de este sistema de ecuaciones (7) es claramente no singular para $\mu > 0$; por lo tanto como una consecuencia del teorema de la función implícita. Nosotros podemos concluir que x' y y' existen para $\mu > 0$. Multiplicando (8) por x_i , tenemos:

$$\begin{aligned} x_i (-x_i' f_i (y) - x_i \nabla f_i (y)^T y') &= x_i \\ (-x_i' f_i (y) x_i - x_i^2 \nabla f_i (y)^T y') &= x_i \end{aligned}$$

De (5) tenemos que:

$$\mu x_i' - x_i^2 \nabla f_i (y)^T y' = x_i.$$

Multiplicando ésta ecuación por $\nabla f_i (y)$, sumando sobre i y usando (7) y (4), tenemos:

$$\nabla f_i (y) (\mu x_i' - x_i^2 \nabla f_i (y)^T y') = \nabla f_i (y) (x_i)$$

$$\mu x_i' \nabla f_i (y) - x_i^2 \nabla f_i (y) \nabla f_i (y)^T y' = x_i \nabla f_i (y)$$

sumando ambos lados desde 1 hasta n :



$$\sum_{i=1}^n [\mu x_i' \nabla f_i(y) - x_i^2 \nabla f_i(y) \nabla f_i(y)^T y'] = \sum_{i=1}^n [x_i \nabla f_i(y)]$$

$$\mu \sum_{i=1}^n x_i' \nabla f_i(y) - \sum_{i=1}^n x_i^2 \nabla f_i(y) \nabla f_i(y)^T y' = b$$

de (4) y (7) tenemos:

$$-\mu \sum_{i=1}^n x_i H_i y' - \sum_{i=1}^n x_i^2 \nabla f_i(y)^T y' = b .$$

Ahora, tomamos el producto interno con y' , obteniéndose

$$\begin{aligned} (y')^T \left[-\mu \sum_{i=1}^n x_i H_i y' - \sum_{i=1}^n x_i^2 \nabla f_i(y) \nabla f_i(y)^T y' \right] &= (y')^T b = b^T y' \\ b^T y' &= -\mu \sum_{i=1}^n x_i (y')^T H_i y' - \sum_{i=1}^n x_i^2 (\nabla f_i(y)^T y')^2 \\ &= -\mu \sum_{i=1}^n x_i (y')^T H_i y' - \sum_{i=1}^n x_i^2 (\nabla f_i(y)^T y')^2 \\ &= -\mu^2 \left[\sum_{i=1}^n (y')^T H_i y' [-f_i(y)] - \sum_{i=1}^n (\nabla f_i(y)^T y')^2 / [-f_i(y)]^2 \right] \\ &= -\mu^2 (y')^T \left[\sum_{i=1}^{nT} H_i y' [-f_i(y)] - \sum_{i=1}^n \nabla f_i(y) \nabla f_i(y)^T / [-f_i(y)]^2 \right]^T y' \end{aligned}$$

por lo tanto

$$b^T y = -\mu^2 (y')^T H y' < 0,$$

donde $h = h$. La última desigualdad se sigue porque H es definida positiva y $y \neq 0$. Esto prueba la primera parte del lema. Par probar la segunda parte bastara probar que:

$$\left(\frac{d}{d\mu} \right) \left(b^T y - \sum_{i=1}^n x_i f_i(y) \right)$$



es no negativa. Nosotros tenemos

$$\left(\frac{d}{d\mu}\right)\left(b^T y - \sum_{i=1}^n x_i f_i(y)\right)$$

que es igual

$$\begin{aligned} &= \left(\frac{d}{d\mu}\right)(b^T y + n\mu) \\ &= \left(\frac{d}{d\mu}\right)(b^T y) + \left(\frac{d}{d\mu}\right)(n\mu) = b^T y' + n \\ &= -\mu \sum_{i=1}^n x_i (y')^T H_i y' - \sum_{i=1}^n x_i^2 (\nabla f_i(y)^T y')^2 + n \end{aligned} \tag{10}$$

donde la última desigualdad sigue de (9). Elevando al cuadrado ambos lados de (8) y sumando cada i , tenemos

$$\left[-x_i' f_i(y) - x_i \nabla f_i(y)^T y'\right]^2 = 1^2 = 1$$

$$\left(-x_i' f_i(y)\right)^2 + (x_i)^2 \left(\nabla f_i(y)^T y'\right)^2 + 2x_i' f_i(y) x_i \nabla f_i(y)^T y' = 1$$

sumando ambos lados de $i=1$ hasta n

$$\sum_{i=1}^n \left(x_i' f_i(y)\right)^2 + \sum_{i=1}^n x_i^2 \left(\nabla f_i(y)^T y'\right)^2 + 2 \sum_{i=1}^n x_i' f_i(y) x_i \nabla f_i(y)^T y' = n \tag{11}$$

Usando (5) y (7), nosotros tenemos que

$$\begin{aligned} \sum_{i=1}^n x_i' f_i(y(\mu)) &= -\sum_{i=1}^n x_i H_i y' \\ (f_i(y) x_i) \left[\sum_{i=1}^n x_i' f_i(y(\mu)) \right] &= (-\mu) \left[-\sum_{i=1}^n x_i H_i y' \right] \\ \sum_{i=1}^n x_i' f_i(y) x_i \nabla f_i(y) &= \sum_{i=1}^n x_i H_i y' \\ (y')^T \left[\sum_{i=1}^n x_i' f_i(y) x_i \nabla f_i(y) \right] &= (y')^T \left[\mu \sum_{i=1}^n x_i H_i y' \right] \\ \sum_{i=1}^n x_i' f_i(y) x_i \nabla f_i(y) y' &= \mu \sum_{i=1}^n x_i (y')^T H_i y' \end{aligned} \tag{12}$$



Sustituyendo (11) y (12) en (10),

$$\begin{aligned}
 b^T y' + n &= -\mu \sum_{i=1}^n x_i (y')^T H_i y' - \sum_{i=1}^n x_i^2 (\nabla f_i (y)^T y')^2 + n \\
 &= -\mu \sum_{i=1}^n x_i (y')^T H_i y' - \sum_{i=1}^n x_i^2 (\nabla f_i (y)^T y')^2 + \sum_{i=1}^n (x_i f_i (y))^2 \\
 &= 2 \sum_{i=1}^n x_i f_i (y) x_i \nabla f_i (y)^T y'
 \end{aligned}$$

$$b^T y' + n = \sum_{i=1}^n (x_i f_i (y))^2 + \mu \sum_{i=1}^n x_i (y')^T H_i y' \geq 0 \quad (13)$$

esto prueba la segunda parte del lema.

Note que (9) y (13) juntos muestran que

$$0 < b^T y(\bar{\mu}) - b^T y(\mu) \leq n(\mu - \bar{\mu}) \quad (14)$$

4.5.6. ALGORITMO

En nuestro algoritmo, no necesitamos estar cerca de la trayectoria central, si estamos lejos de la trayectoria central, hacemos una búsqueda lineal a lo largo de la dirección de Newton con respecto a la función $\phi(y, \mu)$. La dirección de Newton asociado a la función $\phi(y, \mu)$ en y es dado por:

$$p(y, \mu) = -H(y, \mu)^{-1} g(y, \mu) = -H^{-1} g.$$

Si no hay confusión es posible escribir por brevedad p en vez de $p(y, \mu)$, estos procesos son repetidos hasta estar suficientemente cerca a la trayectoria central. Es decir nosotros paramos la búsqueda lineal.

Si $\|p\|_H \leq \tau$ donde τ es una cierta tolerancia.



Un criterio similar de proximidad es usado también por Jarre (Ref.22). En nuestro algoritmo usaremos:

$$\tau = 1/\lceil 8(1+2M)^2 \rceil,$$

que debe ser apropiado.

Note que $\|p\|_H = 0$ si y solamente si $y = y(\mu)$, si el criterio de proximidad es satisfecho, entonces reduciremos el parámetro de barrera μ como sigue: $\bar{\mu} = (1-\theta)\mu$ para algún $0 < \theta < 1$ y todo el proceso es repetido nuevamente hasta que el criterio de parada sea satisfecho.

4.5.7 Algoritmo de Barrera Logarítmica

A continuación presentamos el algoritmo para el método de la función barrera logarítmica.

ALGORITMO

Paso 0. θ es el factor de disminución, $0 < \theta < 1$; $\tau = 1/8(1+2M)^2$ es la tolerancia de proximidad; ε es un parámetro exacto o preciso; y^0 es un punto dado en la región factible y $\mu^0 \leq 1/\varepsilon$, tal que $\|p(y^0, \mu^0)\| \leq \tau$. Conjunto $y = y^0$; $\mu = \mu^0$.

Paso 1. Si $\mu \leq 1/\varepsilon$, entonces paramos.

Paso 2. Si $\|p\|_H \leq \tau$, entonces iremos al paso 6.

Paso 3. Conjunto $\bar{\alpha} = \arg \min_{\alpha > 0} \{\phi(y + \alpha p, \mu) : y + \alpha p \in F\}$.

Paso 4. Conjunto $y = \bar{\alpha} p$.

Paso 5. Ir al paso 2.

Paso 6. Conjunto $\mu = (1-\theta)\mu$.

Paso 7. Ir al paso 1.



Para escoger el punto inicial que satisface la entrada sugerida del algoritmo se refiere al lector revisar Jarre (Ref. 22) además Mehotra y Sun (Ref. 23).

4.5.8 Lemas Importantes

En la sección 4.6 probaremos resultados de complejidad sobre el algoritmo de barrera logarítmica. En esta sección trataremos con algunos lemas los cuales serán necesarios para obtener una cota superior para el número total de iteraciones tanto exterior como interior.

Los lemas son estructurados y terminados como sigue: Lema 4.5.1 da una cota superior para el error en la aproximación cuadrática, si las funciones $f_i(y)$ son lineales o cuadráticas, Lema 4.5.2 establece lo mismo que el Lema 4.5.1, pero ahora en general para funciones convexas restringidos. El lema 4.5.3 establece que si el nuevo punto siguiendo la dirección newton esta cerca entonces obtenemos puntos cerca de la trayectoria central. El Lema 4.5.4 establece que si se hace una buena búsqueda lineal a lo largo de la dirección de Newton entonces está garantizado un decrecimiento en el valor función barrera de una iteración centrada aproximadamente y el centro exacto. El lema 4.5.5 mejora las cotas dadas por el lema 4.5.4 y el Lema 4.5.6 establece que, dada una relación entre el valor objetivo en el centro exacto y una iteración centrada aproximadamente nos brinda una cota para la dualidad.

Lema 4.5.1. Si toda función $f_i(y)$ es lineal o cuadrática con matriz hessiana semidefinida positiva y si $y \in F'$ y $\|d\|_H < 1$, entonces $y+d \in F'$ y

$$|\phi(y+d, \mu) - q_y(y+d, \mu)| < \frac{\|d\|_H^3}{3(1-\|d\|_H)}.$$

Prueba. Se expande $\phi(y+d, \mu)$ es una serie de Taylor cerca a y .

$$\phi(y+d, \mu) = \phi(y+\mu) + q^T d + \frac{1}{2} d^T H d + O(d^3)$$

esto es:

$$\phi(y+d, \mu) = q_y(y+d, \mu) + \sum_{i=3}^{\infty} t_i. \quad (15)$$

Notar que ϕ solo tomará valores finitos en F' . Si se puede probar que $\sum_{i=3}^{\infty} t_i$ converge para un d , tal que $\|d\|_H < 1$, entonces $y+d \in F'$. Puede ser probado que



$$|t_i| \leq \left(\frac{1}{i}\right) \|d\|_H^i, \quad (16)$$

La prueba de esta desigualdad es del todo técnico y por consiguiente es omitido. De (16) y usando que $\|d\|_H$, se deriva que:

$$\sum_{i=3}^{\infty} |t_i| \leq \sum_{i=3}^{\infty} \|d\|_H^i / i = \frac{\|d\|_H^3}{3} + \frac{\|d\|_H^4}{4} + \frac{\|d\|_H^5}{5} + \dots$$

de otro lado, tenemos que:

$$\frac{\|d\|_H^3}{3} \leq \frac{\|d\|_H^3}{3} \rightarrow 3 \geq 3$$

$$\frac{\|d\|_H^3}{3} \leq \frac{\|d\|_H^4}{3} \rightarrow 3 \geq 4$$

$$\frac{\|d\|_H^3}{3} \leq \frac{\|d\|_H^5}{3} \rightarrow 3 \geq 5$$

$$\sum_{i=3}^{\infty} |t_i| \leq \frac{\|d\|_H^3}{3} + \frac{\|d\|_H^4}{4} + \frac{\|d\|_H^5}{5} + \dots$$

$$= \frac{\|d\|_H^3}{3} [1 + \|d\|_H + \|d\|_H^2 + \|d\|_H^3 + \dots]$$

$$\sum_{i=3}^{\infty} |t_i| \leq \sum_{i=3}^{\infty} \|d\|_H^i \leq \frac{\|d\|_H^3}{3(1 - \|d\|_H)}$$

Ahora aplicaremos la serie geométrica convergente

$$\frac{1}{1 - \|d\|_H} = 1 + \|d\|_H + \|d\|_H^2 + \|d\|_H^3 + \dots \text{ si } \|d\|_H < 1$$

luego

$$\sum_{i=3}^{\infty} \|d\|_H^i / i \leq \frac{\|d\|_H^3}{3} [1 + \|d\|_H + \|d\|_H^2 + \|d\|_H^3 + \dots] \leq \frac{\|d\|_H^3}{3}$$

por lo tanto



$$\sum_{i=3}^{\infty} |t_i| \leq \sum_{i=3}^{\infty} \|d\|_H^i / i \leq \frac{\|d\|_H^3}{3} \left(\frac{1}{1 - \|d\|_H} \right)$$

sustituyendo esto en (15) tenemos

$$|\phi(y+d, \mu) - q_y(y+d, \mu)| = \left| \sum_{i=3}^{\infty} t_i \right| \leq \sum_{i=3}^{\infty} |t_i| \leq \frac{\|d\|_H^3}{3} \left(\frac{1}{1 - \|d\|_H} \right)$$

Así, el lema queda probado.

Lema 4.5.2.-Si las funciones $f_i(y)$ satisfacen la condición relativa de Lipschitz (2) con la constante de Lipschitz $M > 0$ y si $y \in F'$ y

entonces $y+d \in F'$ y

$$|\phi(y+d, \mu) - q_y(y+d, \mu)| < \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M).$$

Prueba.-Usando los lemas previos, se analiza el cambio adicional en $\phi(y, u)$ causado por el cambio en la segunda derivada $\nabla^2 f_i(y)$ para el caso cuando $M > 0$. Sea $g_i(y+d)$ la aproximación cuadrática para $f_i(y+d)$ en y , quiere decir que acontece si f es una función con derivadas hasta de segundo orden continuas vecinas a y , entonces podemos escribir:

$$\begin{aligned} f_i(y+d) &= f_i(y) + f_i'(y)d + \frac{1}{2} f_i''(y)d^2 \\ &= f_i(y) + \nabla f_i(y)d + \frac{1}{2} d^T \nabla^2 f_i(y)d \\ &= g_i(y+d) \end{aligned}$$

Quiere decir que:

$$g_i(y+d) = f_i(y) + \nabla f_i(y)d + \frac{1}{2} d^T \nabla^2 f_i(y)d.$$



Definimos r_i como:

$$r_i = \left[f_i(y+d) - g_i(y+d) \right] / g_i(y+d), \text{ (error relativo)} \quad (17)$$

Se tiene:

$$f_i(y+d) = g_i(y+d)(1+r_i).$$

Ahora, usando el teorema de Taylor de segundo orden, para funciones de varias variables, mencionado en el apéndice donde:

$$f(\bar{x}+x) = f(\bar{x}) + \text{grad } f(\bar{x})x + \frac{1}{2}xH(\bar{x}+\mu x)x^t$$

donde $0 < \mu < 1$, lo cual se puede escribir como:

$$f(\bar{x}+x) = f(\bar{x}) + \text{grad } f(\bar{x})x + \frac{1}{2}xH(\bar{x})x^t + r(x),$$

en el que el residuo tiene la propiedad:

$$\lim_{x \rightarrow \infty} \frac{r(x)}{\|x\|^2} = 0$$

en nuestro Lema tenemos:

$$f_i(y+d) - g_i(y+d) = r(x)$$

donde:

$$r(x) = \frac{1}{2}xH(y+\mu d)x^t - \frac{1}{2}xH(y)x^t$$

$$r(x) = \frac{1}{2}(d^T(\nabla^2 f_i(y+\mu d) - \nabla^2 f_i(y))d).$$

Entonces se tiene que:

$$f_i(y+d) - g_i(y+d) = \frac{1}{2}d^T(\nabla^2 f_i(y+\mu d) - \nabla^2 f_i(y))d.$$

Aplicando el valor absoluto a ambos miembros de la última desigualdad se tiene:

$$\begin{aligned} |f_i(y+d) - g_i(y+d)| &= \left| \frac{1}{2}d^T(\nabla^2 f_i(y+\mu d) - \nabla^2 f_i(y))d \right| & (*) \\ &= \frac{1}{2} \left| d^T(\nabla^2 f_i(y+\mu d) - \nabla^2 f_i(y))d \right| \end{aligned}$$

Tenemos que la condición de Lipchitz viene dada por:

$$\exists M > 0, \text{ s.t.}, \forall y, y+h \in F', 1 \leq i \leq n$$



$$\begin{aligned}
\left| v^T (\nabla^2 f_i (y+h) - \nabla^2 f_i (y)) v \right| &\leq M \| \mu h \|_H d^T \nabla^2 f_i (y).d ; 0 < \mu < 1 & (I) \\
&= \frac{1}{2} M \mu \| h \|_H d^T \nabla^2 f_i (y).d ; 0 < \mu < 1 \\
&\leq \frac{1}{2} M \| d \|_H d^T \nabla^2 f_i (y).d
\end{aligned}$$

Usando el hecho que g_i es convexa y que $g_i (y + d / \|d\|_H) \leq 0$.

Como una consecuencia del Lema 5.1, donde $g_i (y + d)$ son cuadráticas, se puede dar una cota superior para $g_i (y + d)$, veamos:

$$g_i (y + d) = g_i \left((1 - \|d\|_H) y + \|d\|_H (y + d / \|d\|_H) \right), \text{ como } g_i \text{ es convexa.}$$

Entonces si \forall par de puntos $y: y + d / \|d\|_H$, con $0 < \|d\|_H < 1$, se cumple que:

$$g_i \left((1 - \|d\|_H) y + \|d\|_H (y + d / \|d\|_H) \right) \leq (1 - \|d\|_H) g_i (y) + \|d\|_H g_i (y + d / \|d\|_H)$$

como:

$$\begin{aligned}
g_i (y + d / \|d\|_H) &\leq 0 \\
\|d\|_H g_i (y + d / \|d\|_H) &\leq 0
\end{aligned}$$

entonces,

$$\begin{aligned}
g_i \left((1 - \|d\|_H) y + \|d\|_H (y + d / \|d\|_H) \right) &\leq (1 - \|d\|_H) g_i (y) + \|d\|_H g_i (y + d / \|d\|_H) \\
&\leq (1 - \|d\|_H) g_i (y)
\end{aligned}$$

por lo tanto:

$$\begin{aligned}
g_i (y + d) &\leq (1 - \|d\|_H) g_i (y) \\
\frac{1}{g_i (y + d)} &\geq \frac{1}{(1 - \|d\|_H) g_i (y)}
\end{aligned}$$



$$\frac{1}{-g_i(y+d)} \leq \frac{1}{-(1-\|d\|_H)g_i(y)} ; -g_i(y+d) \geq 0$$

$$\frac{1}{|g_i(y+d)|} \leq \frac{1}{-(1-\|d\|_H)g_i(y)} \quad (II)$$

Sustituyendo (I) y (II) en (17), se obtiene:

$$\left| \frac{f_i(y+d) - g_i(y+d)}{g_i(y+d)} \right| \leq \frac{1}{2} \frac{M \|d\|_H d^T \nabla^2 f_i(y) d}{-(1-\|d\|_H)g_i(y)} \quad (18)$$

Por otro lado:

$$g_i(y+d) = f_i(y) + \nabla f_i(y)^T d + \frac{1}{2} d^T \nabla^2 f_i(y) d$$

Evaluando cuando $d=0$; se tiene:

$$g_i(y+0) = f_i(y) + \nabla f_i(y)^T (0) + \frac{1}{2} (0)^T \nabla^2 f_i(y) (0)$$

$$g_i(y) = f_i(y)$$

Reemplazando en (III), se tiene:

$$|r_i| = \left| \frac{f_i(y+d) - g_i(y+d)}{g_i(y+d)} \right| \leq \frac{1}{2} \frac{M \|d\|_H d^T \nabla^2 f_i(y) d}{-(1-\|d\|_H)f_i(y)} \quad (IV)$$

sabemos que:

$$\frac{d^T \nabla^2 f_i(y) d}{|-f_i(y)|} \geq 0 \rightarrow \frac{d^T \nabla^2 f_i(y) d}{|-f_i(y)|} \leq \sum_{i=1}^n \frac{d^T \nabla^2 f_i(y) d}{|-f_i(y)|} \quad (\alpha)$$

como:

$$\sum_{i=1}^n \frac{d^T \nabla f_i(y)^T d}{[f_i(y)]^2} \geq 0 \quad (\beta)$$

Entonces sumamos ambas desigualdades, obteniéndose:



$$\begin{aligned}
\frac{d^T \nabla^2 f_i(y) \cdot d}{|-f_i(y)|} &\leq \sum_{i=1}^n \left[\frac{d^T \nabla^2 f_i(y) \cdot d}{|-f_i(y)|} + \frac{d^T \nabla f_i(y) \nabla f_i(y)^T \cdot d}{[f_i(y)]^2} \right] \\
&= d^T \left[\sum_{i=1}^n \left[\frac{\nabla^2 f_i(y)}{|-f_i(y)|} \right] + \frac{\nabla f_i(y) \nabla f_i(y)^T}{[f_i(y)]^2} \right] \cdot d \\
&= d^T \cdot H \cdot d \\
&= \|d\|_H^2
\end{aligned}$$

entonces:

$$\frac{d^T \nabla^2 f_i(y) \cdot d}{|f_i(y)|} \leq \|d\|_H^2 \tag{\gamma}$$

multiplicando (α) por $\frac{1}{2} \frac{M \|d\|_H}{(1 - \|d\|_H)}$,

$$\frac{1}{2} \frac{M \|d\|_H d^T \nabla^2 f_i(y) \cdot d}{(1 - \|d\|_H) [-f_i(y)]} \leq \frac{1}{2} \frac{M \|d\|_H^3}{(1 - \|d\|_H)}$$

Reemplazando la última desigualdad en (IV) tenemos:

$$\begin{aligned}
[r_i] &\leq \frac{1}{2} \frac{M \|d\|_H d^T \nabla^2 f_i(y) \cdot d}{-(1 - \|d\|_H) f_i(y)} \\
&\leq \frac{1}{2} \frac{M \|d\|_H^3}{(1 - \|d\|_H)}
\end{aligned}$$

Por otro lado tenemos que:

$$\|d\|_H < \min \left\{ \frac{1}{2}; \frac{1}{2M^{1/3}} \right\}$$

Entonces:

$$\|d\|_H < 1/2 \wedge \|d\|_H < \frac{1}{2M^{1/3}}$$

Luego:



$$\begin{aligned}
-\|d\|_H &> -\frac{1}{2} \\
2(1-\|d\|_H) &> 1 \\
\frac{1}{2(1-\|d\|_H)} &< 1
\end{aligned}$$

multiplicando por $M \|d\|_H^3$ en la última desigualdad, se tiene:

$$\frac{M \|d\|_H^3}{2(1-\|d\|_H)} < M \|d\|_H^3 \quad (\theta)$$

Por otro lado tenemos que:

$$\begin{aligned}
\|d\|_H &< \frac{1}{2M^{1/3}} \\
\|d\|_H^3 M &< \frac{1}{8}
\end{aligned}$$

Reemplazando esto último en (θ) :

$$\begin{aligned}
\frac{M \|d\|_H^3}{2(1-\|d\|_H)} &< M \|d\|_H^3 \\
&< \frac{1}{8}
\end{aligned}$$

entonces:

$$\frac{M \|d\|_H^3}{2(1-\|d\|_H)} < \frac{1}{8}$$

Llegándose establecer una cota para $|r_i|$:

$$|r_i| < \frac{1}{8}$$

Consecuentemente:

$$\begin{aligned}
|\phi(y+d, \mu) - q_y(y+d, \mu)| &= \left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n \log(-f_i(y+d)) - q_y(y+d, \mu) \right| \\
&= \left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n \log(-g_i(y+d)(1+r_i)) - q_y(y+d, \mu) \right| \\
&= \left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n (\log(-g_i(y+d)) + \log(1+r_i)) - q_y(y+d, \mu) \right| \\
&= \left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n \log(-g_i(y+d)) - \sum_{i=1}^n \log(1+r_i) - q_y(y+d, \mu) \right|
\end{aligned}$$

Por la desigualdad triangular, se tiene:

$$\leq \left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n \log(-g_i(y+d)) - q_y(y+d, \mu) \right| + \left| \sum_{i=1}^n \log(1+r_i) \right| \quad (\rho)$$

Como $g_i(y+d)$ es cuadrática la primera parte puede ser estimada por el Lema 5.1, por lo tanto:

$$\left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n \log(-g_i(y+d)) - q_y(y+d, \mu) \right| = |\phi^*(y+d, \mu) - q_y(y+d, \mu)|$$

entonces se tiene:

$$|\phi^*(y+d, \mu) - q_y(y+d, \mu)| \leq \|d\|_H^3 / 3(1 - \|d\|_H)$$

Reemplazando éste último en (ρ) , tenemos:

$$\begin{aligned}
\left| \frac{-b^T(y+d)}{\mu} - \sum_{i=1}^n \log(-g_i(y+d)) - q_y(y+d, \mu) \right| &= \left| \sum_{i=1}^n \log(1+r_i) \right| \quad (V) \\
&\leq \|d\|_H^3 / 3(1 - \|d\|_H) + \left| \sum_{i=1}^n \log(1+r_i) \right|
\end{aligned}$$

Porque $g_i(y+d)$ es cuadrática, la primera parte puede ser estimada por el lema 5.1, usando el hecho que $|r_i| \leq (1/8)$, lo cual implica que $\log(1+r_i) \leq \frac{9}{8}|r_i|$ y (18), la segunda parte puede ser estimada como sigue:

$$\sum_{i=1}^n \log(1+r_i) \leq \sum_{i=1}^n \left(\frac{9}{8} \right) |r_i|$$

y de (18) se tiene que:

$$|r_i| < \frac{1}{2} \frac{\|d\|_H M d^T \nabla^2 f_i(y) d}{(1 - \|d\|_H)(-f_i(y))}$$

Aplicando sumatoria $\sum_{i=1}^n$ y multiplicando por $\frac{9}{8}$ ambas desigualdades:

$$\begin{aligned} \frac{9}{8} \sum_{i=1}^n |r_i| &\leq \frac{9}{8} \sum_{i=1}^n \left(\frac{1}{2} \frac{\|d\|_H M d^T \nabla^2 f_i(y) d}{(1 - \|d\|_H)(-f_i(y))} \right) \\ &= \frac{9}{8} \left(\frac{1}{2} \frac{\|d\|_H M}{(1 - \|d\|_H)} \right) \sum_{i=1}^n \left[\frac{d^T \nabla^2 f_i(y) d}{-f_i(y)} \right] \end{aligned}$$

Entonces:

$$\left| \sum_{i=1}^n \log(1 + r_i) \right| \leq \frac{9}{8} \left(\frac{1}{2} \frac{\|d\|_H M}{(1 - \|d\|_H)} \right) \sum_{i=1}^n \left[\frac{d^T \nabla^2 f_i(y) d}{-f_i(y)} \right] \quad (\psi)$$

Por otro lado sabemos que:

$$\sum_{i=1}^n \frac{d^T \nabla^2 f_i(y) d}{(-f_i(y))} \geq 0 \text{ y } \sum_{i=1}^n \left[\frac{\nabla f_i(y)^T d}{f_i(y)} \right]^2 = \sum_{i=1}^n \frac{d^T \nabla f_i(y) \cdot \nabla f_i(y)^T d}{f_i(y)^2}$$

entonces se tiene que:

$$\begin{aligned} \sum_{i=1}^n \frac{d^T \nabla^2 f_i(y) d}{(-f_i(y))} &\leq \sum_{i=1}^n \frac{d^T \nabla^2 f_i(y) d}{-f_i(y)} + \sum_{i=1}^n \frac{d^T \nabla f_i(y) \cdot \nabla f_i(y)^T d}{f_i(y)^2} \\ &= \sum_{i=1}^n \left[\frac{d^T \nabla^2 f_i(y) d}{-f_i(y)} + \frac{d^T \nabla f_i(y) \cdot \nabla f_i(y)^T d}{f_i(y)^2} \right] \\ &= d^T \sum_{i=1}^n \left[\frac{\nabla^2 f_i(y)}{-f_i(y)} + \frac{\nabla f_i(y) \cdot \nabla f_i(y)^T}{f_i(y)^2} \right] d \\ &= d^T H d \\ &= \|d\|_H^2 \end{aligned}$$

entonces:

$$\sum_{i=1}^n \frac{d^T \nabla^2 f_i(y) d}{-f_i(y)} \leq \|d\|_H^2.$$

multiplicando en éste último por $\frac{9}{8} \frac{\|d\|_H M}{2(1 - \|d\|_H)} > 0$.



$$\begin{aligned} \frac{9}{8} \left[\frac{\|d\|_H \cdot M}{2(1-\|d\|_H)} \right] \sum_{i=1}^n \frac{d^T \nabla^2 f_i(y) \cdot d}{-f_i(y)} &\leq \frac{9}{8} \frac{\|d\|_H^3 \cdot M}{2(1-\|d\|_H)} \\ &= \frac{9}{16} \frac{\|d\|_H^3 \cdot M}{(1-\|d\|_H)} \\ &< 2M \cdot \frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} \end{aligned} \quad (\Delta)$$

por lo tanto (ψ) y de (Δ) se tiene:

$$\left| \sum_{i=1}^n \log(1+r_i) \right| < 2M \cdot \frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} \quad (\text{VI})$$

de (V) y (VI) se tiene una cota para ϕ :

$$\begin{aligned} \left| \phi(y+d, \mu) - q_y(y+d, \mu) \right| &< \frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} + \left| \sum_{i=1}^n \log(1+r_i) \right| \\ &< \frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} + 2M \cdot \frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} \\ &= (1+2M) \frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} \end{aligned}$$

por lo tanto se obtiene que:

$$\left| \phi(y+d, \mu) - q_y(y+d, \mu) \right| \leq \left[\frac{\|d\|_H^3 \cdot M}{3(1-\|d\|_H)} \right] (1+2M)$$

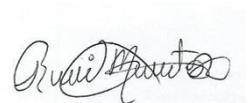
Lema 4.5.3.- Si $\|p\|_H \leq \frac{1}{8(1+2M)}$

Entonces:

$$\|y - y(\mu)\|_H \leq \frac{5}{2} \|p\|_H.$$

Prueba.- Sea h arbitrario, tal que: $\|h\|_H = \frac{3}{2} \|p\|_H$.

Consideremos los valores sobre el elipsoide:



$$\left\{ y + p + h : \|h\|_H = \frac{3}{2} \|p\|_H \right\}$$

Se tiene:

$$\begin{aligned} \|h + p\|_H &\leq \|h\|_H + \|p\|_H = \frac{3}{2} \|p\|_H + \|p\|_H \\ &= \frac{5}{2} \|p\|_H \leq \frac{5}{2} \cdot \frac{1}{8} \|p\|_H \leq \frac{5}{2} \cdot \frac{1}{8} \cdot \frac{1}{(1+2M)} \\ &= \frac{5}{16} \cdot \frac{1}{(1+2M)} = 0,3125 \cdot \frac{1}{(1+2M)} < 0, \hat{3} \frac{1}{(1+2M)} \\ &= \frac{5}{2} \|p\|_H < \frac{1}{3} \frac{1}{(1+2M)} \end{aligned} \quad (20)$$

Con la ayuda del Lema 5.2 y usando el hecho de que:

$$y + p = \arg \min_x q_y(x, \mu),$$

Obtenemos:

$$\begin{aligned} |\phi(y + p, \mu) - q_y(y + p, \mu)| &\leq \frac{\|p + h, \mu\|_H^3 (1 + 2M)}{3(1 - \|p + h, \mu\|_H)} \\ \phi(y + p + h, \mu) &> q_y(y + p + h, \mu) - \frac{\|p + h\|_H^3 (1 + 2M)}{3(1 - \|p + h\|_H)} \\ &= \phi(y, \mu) + g^T(p + h) + \frac{1}{2}(p + h)^T H(p + h) - \frac{\|p + h\|_H^3 (1 + 2M)}{3(1 - \|p + h\|_H)} \\ &= \phi(y, \mu) + g^T p + g^T h + \frac{1}{2}(p^T + h^T)H(p + h) - \frac{\|p + h\|_H^3 (1 + 2M)}{3(1 - \|p + h\|_H)} \\ &= \phi(y, \mu) + g^T p + g^T h + \frac{1}{2}\{p^T(H(p + h)) + h^T(H(p + h))\} - \frac{\|p + h\|_H^3 (1 + 2M)}{3(1 - \|p + h\|_H)} \\ &= \phi(y, \mu) + g^T p + g^T h + \frac{1}{2}\{p^T H p + p^T H h + h^T H p + h^T H h\} - \frac{\|p + h\|_H^3 (1 + 2M)}{3(1 - \|p + h\|_H)} \end{aligned}$$

$$\begin{aligned}
&= \phi(y, \mu) + g^T p + g^T h + \frac{1}{2} \{ p^T H p - 2g^T + h^T H h \} - \frac{\|p+h\|_H^3 (1+2M)}{3(1-\|p+h\|_H)} \\
&= \phi(y, \mu) + g^T p + \frac{1}{2} p^T H p + \frac{1}{2} h^T H h - \frac{\|p+h\|_H^3 (1+2M)}{3(1-\|p+h\|_H)} \quad (\alpha)
\end{aligned}$$

Por otro, sabemos que:

$$\|h+p\|_H < \frac{1}{3(1+2M)} < \frac{1}{3}, M > 0$$

$$\|h+p\|_H < \frac{1}{3}$$

$$1 - \|h+p\|_H > 1 - \frac{1}{3}$$

$$\frac{1}{1 - \|h+p\|_H} < \frac{3}{2}$$

multiplicando en la última desigualdad el valor de $\frac{\|h+p\|_H^3 (1+2M)}{3}$, tenemos:

$$\begin{aligned}
\frac{\|h+p\|_H^3 (1+2M)}{3(1-\|h+p\|_H)} &< \frac{3 \times 2 \|h+p\|_H^3 (1+2M)}{2} \\
&= \frac{1}{2} \|h+p\|_H (1+2M)
\end{aligned}$$

entonces,

$$-\frac{\|h+p\|_H^3 (1+2M)}{3(1-\|h+p\|_H)} > -\frac{1}{2} \|h+p\|_H^3 (1+2M)$$

reemplazando en (α) éste último, se tendrá:

$$\phi(y+p+h, \mu) \geq q_y(y+p, \mu) + \frac{1}{2} \|h\|_H^2 - \frac{1}{2} \|h+p\|_H^3 (1+2M) \quad (\beta)$$

se tiene que:

$$\|h+p\|_H \leq \frac{5}{2} \|p\|_H$$

$$\|h+p\|_H^3 \leq \frac{125}{8} \|p\|_H^3$$



$$\begin{aligned}\frac{1}{2}\|h+p\|_H^3(1+2M) &\leq \frac{125}{16}\|p\|_H^3(1+2M); M > 0 \\ -\frac{1}{2}\|h+p\|_H^3(1+2M) &\geq -\frac{125}{16}\|p\|_H^3(1+2M); M > 0\end{aligned}$$

Reemplazando en (β) lo último se da lo siguiente:

$$\phi(y+p+h, \mu) \geq q_y(y+p, \mu) + \frac{1}{2}\|h\|_H^2 - \frac{125}{16}\|p\|_H^3(1+2M)$$

se sabe que $\|h\|_H = \frac{3}{2}\|p\|_H$.

$$\begin{aligned}\phi(y+p+h, \mu) &\geq q_y(y+p, \mu) + \frac{1}{2}\left(\frac{9}{4}\|p\|_H^2\right) - \frac{125}{16}\|p\|_H^3(1+2M) \\ &= q_y(y+p, \mu) + \frac{9}{8}\|p\|_H^2 - \frac{125}{16}\|p\|_H^3(1+2M)\end{aligned}\quad (\gamma)$$

Por hipótesis se tiene:

$$\|p\|_H < \frac{1}{8(1+2M)}$$

$$8(1+2M)\|p\|_H < 1$$

$$9(1+2M)\|p\|_H^3 < \frac{9}{8}\|p\|_H^2, \|p\|_H > 0$$

Reemplazando en (β) se tiene:

$$\begin{aligned}\phi(y+p+h, \mu) &\geq q_y(y+p, \mu) + 9\|p\|_H^3(1+2M) - \frac{125}{16}\|p\|_H^3(1+2M) \\ &= q_y(y+p, \mu) + \frac{19}{16}\|p\|_H^3(1+2M) \\ &= q_y(y+p, \mu) + (1,81175)\|p\|_H^3(1+2M) \\ &> q_y(y+p, \mu) + \|p\|_H^3(1+2M)\end{aligned}$$

entonces:

$$\phi(y+p+h, \mu) \geq q_y(y+p, \mu) + \|p\|_H^3(1+2M)$$

usando el Lema 5.2, una vez más, tenemos que:

$$\phi(y+p, \mu) < q_y(y+p, \mu) + \frac{\|p\|_H^3(1+2M)}{3(1-\|p\|_H)}\quad (\theta)$$

tenemos por hipótesis que:



$$\frac{5}{2}\|p\|_H < \frac{1}{3(1+2M)}$$

$$\|p\|_H < \frac{2}{15(1+2M)} < \frac{2}{15}; M > 0$$

entonces:

$$\|p\|_H < \frac{2}{15}$$

$$3(1-\|p\|_H) > \frac{13}{5}$$

$$\frac{1}{3(1-\|p\|_H)} < \frac{5}{13} = 0,3846$$

$$\frac{1}{3(1-\|p\|_H)} < 0,3846 < 0.5$$

entonces:

$$\frac{1}{3(1-\|p\|_H)} < \frac{1}{2}$$

$$\frac{\|p\|_H^3(1+2M)}{3(1-\|p\|_H)} < \frac{1}{2}\|p\|_H^3(1+2M); \|p\|_H^3(1+2M) > 0$$

Reemplazo éste último en (θ) :

$$\begin{aligned} \phi(y+p, \mu) &< q_y(y+p, \mu) + \frac{\|p\|_H^3(1+2M)}{3(1-\|p\|_H)} \\ &< q_y(y+p, \mu) + \frac{1}{2}\|p\|_H^3(1+2M) \\ &< q_y(y+p, \mu) + \|p\|_H^3(1+2M) \end{aligned}$$

entonces:

$$\begin{aligned} \phi(y+p, \mu) &< q_y(y+p, \mu) + \|p\|_H^3(1+2M) \\ &< \phi(y+p+h, \mu) \end{aligned}$$

por lo tanto:

$$\phi(y+p, \mu) < \phi(y+p+h, \mu)$$



así en el centro del elipsoide, el valor potencial es menor que el valor sobre ésta frontera. Asimismo por la convexidad estricta de ϕ , el mínimo de ϕ está en el interior del elipsoide, lo cual significa que:

$$\|y - y(\mu)\|_H \leq \|p\|_H \leq \|p + h\|_H$$

Ahora usando (20) se tiene lo siguiente:

$$\|y - y(\mu)\|_H \leq \|p + h\|_H \leq \frac{5}{2} \|p\|_H$$

$$\|y - y(\mu)\|_H \leq \frac{5}{2} \|p\|_H.$$

Lema 4.5.4.-Si $\|p\|_H \geq \frac{1}{8(1+2M)}$, entonces el decrecimiento $\Delta\phi$ en la función potencial después de una búsqueda de línea a lo largo de la dirección de Newton p , satisface:

$$\Delta\phi \geq \frac{1}{140(1+2M)^2}.$$

Prueba.-Sea λ una longitud de paso, tal que:

$$\|\lambda p\|_H \leq \min\left\{\frac{1}{2}, \frac{1}{2M^{1/3}}\right\}$$

entonces como una consecuencia de Lema 5.2 se tiene:

$$\begin{aligned} |\phi(y + \lambda p, \mu) - q_y(y + \lambda p, \mu)| &< \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M) \\ - \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M) &< \phi(y + \lambda p, \mu) - q_y(y + \lambda p, \mu) < \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M) \end{aligned}$$

$$\phi(y + \lambda p, \mu) \leq q_y(y + \lambda p, \mu) + \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M)$$

ahora de la definición de q_y obtenemos:

$$\phi(y + \lambda p, \mu) \leq \phi(y, \mu) + g^T \lambda p + \frac{1}{2} (\lambda p)^T H \lambda p + \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M)$$

$$\phi(y, \mu) - \phi(y + \lambda p, \mu) \geq -\lambda g^T p - \frac{1}{2} \lambda^2 p^T H p - \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M) \quad (\varphi)$$



por otro lado:

$$\begin{aligned}
 p &= -H^{-1}g \\
 Hp &= -g \\
 p^T hp &= -p^T g \\
 p^T hp &= -g^T p \\
 \|p\|_H^2 &= -g^T p
 \end{aligned}$$

reemplazando la última igualdad en (φ) :

$$\phi(y, \mu) - \phi(y + \lambda p, \mu) \geq -\lambda \|p\|_H^2 - \frac{1}{2} \lambda^2 \|p\|_H^2 - \left[\frac{\|\lambda p\|_H^3}{3(1 - \|\lambda p\|_H)} \right] (1 + 2M)$$

Reemplazando λ por el valor $\frac{1}{9(1+2M)\|p\|_H}$, lo cual satisface (21):

$$\begin{aligned}
 \phi(y, \mu) - \phi(y + \lambda p, \mu) &\geq \frac{\|p\|_H^2}{9(1+2M)\|p\|_H} - \frac{\|p\|_H^2}{2 \times 81(1+2M)^2 \|p\|_H^2} \\
 &\quad - \left[\frac{\frac{\|p\|_H^3}{9^3(1+2M)^3 \|p\|_H^3}}{3 \left(1 - \frac{\|p\|_H}{9(1+2M)\|p\|_H} \right)} \right] (1+2M) \\
 &= \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} - \left[\frac{1}{\frac{9^3(1+2M)}{3(18M+8)}} \right] (1+2M) \\
 &= \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} - \frac{9(1+2M)^2}{9^3 \times 3(1+2M)^3(18M+8)} \\
 &= \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} - \frac{1}{243(1+2M)(18M+8)} \quad (\beta)
 \end{aligned}$$

Por otro lado, tenemos:



$$\begin{aligned} \|p\|_H &\geq \frac{1}{8(1+2M)} \\ \frac{\|p\|_H}{9(1+2M)} &\geq \frac{1}{72(1+2M)^2}, \quad 9(1+2M) > 0 \\ \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} &\geq \frac{1}{72(1+2M)^2} - \frac{1}{162(1+2M)^2} \\ &= \frac{90}{11664(1+2M)^2} \\ &= \frac{0.007716}{(1+2M)^2} \\ \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} &\geq \frac{0.007716}{(1+2M)^2} \quad (I) \end{aligned}$$

de la expresión (β) debemos encontrar una cota inferior para $-\frac{1}{243(1+2M)(18M+8)}$

se tiene que:

$$-\frac{1}{243(1+2M)(18M+8)} = -\frac{1}{243[9(1+2M)^2 - (2M+1)]}$$

por otro lado tenemos que:

$$8(2M+1)^2 \leq 9(1+2M)^2 - (2M+1), \quad M > 0$$

$$\frac{1}{9(1+2M)^2 - (2M+1)} \leq \frac{1}{8(2M+1)^2}$$

$$\frac{-1}{(1+2M)(18M+8)} \geq \frac{-1}{8(2M+1)^2}$$

$$\frac{-1}{243(1+2M)(18M+8)} \geq \frac{-1}{243 \times 8(2M+1)^2}$$

$$\frac{-1}{243(1+2M)(18M+8)} \geq \frac{-1}{1944(2M+1)^2}$$

$$= \frac{-0.0005144}{(2M+1)^2}$$

entonces:

$$\frac{-1}{243(1+2M)(18M+8)} > \frac{-0.007716}{(1+2M)^2} \quad (II)$$



De (I) y (II), tenemos el sistema de desigualdades:

$$\left. \begin{aligned} \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} &\geq \frac{0.007716}{(1+2M)^2} \\ \frac{1}{243(1+2M)(18M+8)} &\geq \frac{-0.0005144}{(1+2M)} \end{aligned} \right\} (+)$$

sumando ambas desigualdades tenemos:

$$\begin{aligned} \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} - \frac{1}{243(1+2)(18M+8)} &\geq \frac{1}{(2M+1)^2} \{0.007716 - 0.0005144\} \\ \frac{1}{(2M+1)^2} \{0.007716 - 0.0005144\} &= \frac{0.0072016}{(2M+1)^2} \\ &\geq \frac{0.0071428}{(2M+1)^2} \\ &= \frac{1}{140(2M+1)^2} \end{aligned}$$

entonces:

$$\begin{aligned} \frac{\|p\|_H}{9(1+2M)} - \frac{1}{162(1+2M)^2} - \frac{1}{243(1+2)(18M+8)} &\geq \frac{1}{140(2M+1)^2} \\ \Delta\phi = \phi(y, \mu) - \phi(y + \lambda p, \mu) &\geq \frac{1}{140(2M+1)^2} \\ \Delta\phi &\geq \frac{1}{140(2M+1)^2} \end{aligned}$$

Lema 4.5.5.-Si $\|p\|_H \leq \frac{1}{8}(1+2M)$, entonces:

$$\phi(y, \mu) - \phi(y(\mu), \mu) \leq 4\|p\|_H^2.$$

Prueba.-Sea d definida como $y(\mu) - y$, entonces se tiene:

$$\begin{aligned} d &:= y(\mu) - y, \\ y(\mu) &:= d + y, \end{aligned}$$

entonces, evaluamos $\phi(y, \mu)$ en el plano $y(\mu)$, obtenemos:



$$\phi(y(\mu), \mu) = \phi(y + d, \mu).$$

Ahora usando el Lema 5.2, se tiene:

$$\begin{aligned} |\phi(y + d, \mu) - q_y(y + d, \mu)| &< \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M) \\ - \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M) &< \phi(y + d, \mu) - q_y(y + d, \mu) < \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M) \end{aligned}$$

$$\begin{aligned} \phi(y(\mu), \mu) &\geq q_y(y + d, \mu) - \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M) \\ &= \phi(y, \mu) + g^T d + \frac{1}{2} d^T H d - \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M) \quad (\delta) \end{aligned}$$

Usando la dirección de Newton $p(y, \mu)$ asociado con $\phi(y, \mu)$:

$$\begin{aligned} p &= -H^{-1}g \\ Hp &= -g \\ (Hp)^T &= -g^T \\ p^T H &= -g^T \\ p^T H d &= -g^T d \\ p^T H d &= -g^T d \quad (\eta) \end{aligned}$$

Reemplazando (η) en (δ) se tiene:

$$\phi(y(\mu), \mu) \geq \phi(y, \mu) - p^T H d + \frac{1}{2} d^T H d - \left[\frac{\|d\|_H^3}{3(1 - \|d\|_H)} \right] (1 + 2M) \quad (\rho)$$

Usando la desigualdad de Cauchy – Schwarz.

Sean p y d dos vectores en \mathbb{R}^n por las desigualdades se tiene:

$$(p \cdot d)^2 \leq (p \cdot p) \cdot (d \cdot d)$$



Ahora multiplicamos por H^2 , donde H es una matriz hessiana:

$$\begin{aligned}
H^2(p.d)^2 &\leq H^2(p.p).(d.d) \\
(H.p.d)^2 &\leq (p.p)^T (H^2)^T (d.d) \\
&= p^T . p^T H^2 (d.d) \\
&= p^T . p^T H.H (d.d) \\
&= p^T . H . p . d^T H . d \\
&= (p^T . H . p) . (d^T H . d) \\
(H.p.d)^2 &\leq (p^T . H . p) . (d^T H . d) \\
(p^T . H . d)^2 &\leq (p^T . H . p) . (d^T H . d) \\
\sqrt{(p^T . H . d)^2} &\leq \sqrt{(p^T . H . p) . (d^T H . d)} \\
p^T . H . d &\leq \sqrt{p^T . H . p} . \sqrt{d^T H . d}
\end{aligned}$$

$$|p^T . H . d| \leq \|p\|_H \|d\|_H \quad (\theta)$$

Reemplazando (θ) en (ρ) , se tiene:

$$\begin{aligned}
p^T H d &\geq \phi(y, \mu) - \phi(y(\mu), \mu) + \frac{1}{2} d^T H d - \left[\frac{\|d\|_H^3}{3(1-\|d\|_H)} \right] (1+2M) \\
\|d\|_H \|d\|_H &\geq \phi(y, \mu) - \phi(y(\mu), \mu) + \frac{1}{2} d^T H d - \left[\frac{\|d\|_H^3}{3(1-\|d\|_H)} \right] (1+2M) \\
\phi(y, \mu) - \phi(y(\mu), \mu) &\leq \|p\|_H \|d\|_H - \frac{1}{2} d^T H d + \left[\frac{\|d\|_H^3}{3(1-\|d\|_H)} \right] (1+2M) \\
&= \|p\|_H \|d\|_H - \frac{1}{2} \|d\|_H^2 + \left[\frac{\|d\|_H^3}{3(1-\|d\|_H)} \right] (1+2M) \\
&\leq \|p\|_H \|d\|_H + \left[\frac{\|d\|_H^3}{3(1-\|d\|_H)} \right] (1+2M) \quad (\varepsilon)
\end{aligned}$$

como:



$$\|p\|_H \leq \frac{1}{8(2M+1)} < \frac{1}{8}; M > 0$$

$$\Rightarrow \|p\|_H \leq \frac{1}{8}$$

$$\Rightarrow \frac{5}{2} \|p\|_H \leq \frac{5}{16}$$

$$\|d\|_H \leq \frac{5}{2} \|p\|_H \leq \frac{5}{16}$$

$$\|d\|_H \leq \frac{5}{16}$$

$$\frac{1}{1-\|d\|_H} \leq \frac{1}{\left(1-\frac{5}{16}\right)}$$

$$\frac{1}{24(1-\|d\|_H)\|p\|_H} \leq \frac{1}{24\left(1+\frac{5}{16}\right)\|p\|_H} \quad (I)$$

de (I) y (II) tenemos:

$$\|d\|_H^3 \leq \frac{125}{8} \|p\|_H^3 \quad (i)$$

$$\frac{1}{24(1-\|d\|_H)\|p\|_H} \leq \frac{1}{24\left(1-\frac{5}{16}\right)\|p\|_H} \quad (ii)$$

(i) × (ii):

$$\frac{\|d\|_H^3}{24(1-\|d\|_H)\|p\|_H} \leq \frac{\frac{125}{8} \|p\|_H^3}{24\left(1-\frac{5}{16}\right)} \|p\|_H^2$$

reemplazando esto último en (□). Por otro lado tenemos:

$$\|p\|_H \leq \frac{1}{8(1+2M)}$$

$$(1+2M) \leq \frac{1}{8\|p\|_H}$$



multiplicando por $\frac{\|d\|_H^3}{3(1-\|d\|_H)} \geq 0$ a ambos miembros de la última desigualdad, tal que

$$\frac{\|d\|_H^3}{3(1-\|d\|_H)} \geq 0.$$

$$\frac{\|d\|_H^3 (1+2M)}{3(1-\|d\|_H)} \leq \frac{\|d\|_H^3}{24(1-\|d\|_H)\|p\|_H} \quad (\psi)$$

reemplazando (ψ) en (ε) , tenemos:

$$\phi(y, \mu) - \phi(y(\mu), \mu) \leq \|p\|_H \|d\|_H + \left[\frac{\|d\|_H^2}{24(1-\|d\|_H)\|p\|_H} \right] \quad (\nabla)$$

Por otro lado, usando el Lema 5.3 se obtiene:

$$\begin{aligned} \|y - y(\mu)\|_H &= \|d\|_H \leq \frac{5}{2} \|p\|_H \\ &= \|d\|_H \|p\|_H \leq \frac{5}{2} \|p\|_H^2 \end{aligned}$$

reemplazando éste último en (∇) :

$$\phi(y, \mu) - \phi(y(\mu), \mu) \leq \frac{5}{2} \|p\|_H + \left[\frac{\|d\|_H^3}{24(1-\|d\|_H)\|p\|_H} \right] \quad (\square)$$

del Lema 5.3 se obtiene:

$$\begin{aligned} \|d\|_H &\leq \frac{5}{12} \|p\|_H \\ \|d\|_H^3 &\leq \frac{125}{8} \|p\|_H^3 \end{aligned} \quad (I)$$



$$\begin{aligned}
\phi(y, \mu) - \phi(y(\mu), \mu) &\leq \|p\|_H + \left[\frac{125}{8 \times 24 \left(1 - \frac{5}{16}\right)} \right] \|p\|_H^2 \\
&= \frac{5}{2} \|p\|_H + \frac{125}{132} \|p\|_H^2 \\
&= \frac{455}{132} \|p\|_H^2
\end{aligned}$$

entonces:

$$\begin{aligned}
\phi(y, \mu) - \phi(y(\mu), \mu) &\leq (3.4469696) \|p\|_H^2 < 4 \|p\|_H^2 \\
\phi(y, \mu) - \phi(y(\mu), \mu) &\leq 4 \|p\|_H^2
\end{aligned}$$

Lema 4.5.6.-Si $\|p\|_H \leq \frac{1}{8(1+2M)}$, entonces:

$$|b^T y - b^T y(\mu)| \leq \frac{1}{2} \mu \sqrt{n}$$

Prueba.-Evaluando $g(y, \mu)$ en $y(\mu)$ se tiene:

$$\begin{aligned}
g(y(\mu), \mu) &= \nabla \phi(y(\mu), \mu) \\
&= \frac{-b}{\mu} + \sum_{i=1}^n \frac{\nabla f_i(y(\mu))}{[-f_i(y(\mu))]} .
\end{aligned}$$

En el punto $y(\mu)$ al plano que pasa por $(y(\mu))$ y tiene como normal $\nabla \phi(y(\mu), \mu)$, así el plano tangente a F' en $y(\mu)$ tiene al ecuación:

$$\begin{aligned}
[y - y(\mu)] \cdot \nabla \phi(y(\mu), \mu) &= 0 \\
[y - y(\mu)]^T \left\{ \frac{-b}{\mu} + \sum_{i=1}^n \frac{\nabla f_i(y(\mu))}{[-f_i(y(\mu))]} \right\} &= 0 \\
-[y - y(\mu)]^T \frac{b}{\mu} + [Y - Y(\mu)]^T \sum_{i=1}^n \frac{\nabla f_i(y(\mu))}{[-f_i(y(\mu))]} &= 0 \\
\frac{-b^T [y - y(\mu)]}{\mu} + \sum_{i=1}^n \frac{\nabla f_i(y(\mu))^T [Y - Y(\mu)]}{[-f_i(y(\mu))]} &= 0 \tag{23}
\end{aligned}$$

Como las funciones $f_i(y)$ son convexas.



Sea F^0 un conjunto convexo en R^n y $f_i : F^0 \rightarrow R$, una función diferenciable sobre F^0 , entonces f_i es convexa sobre F^0 , si y solo si:

Por otro lado tenemos que:

$$\|p\|_H < \frac{1}{8(1+2M)} < \frac{1}{8}; M > 0$$

entonces:

$$\|p\|_H < \frac{1}{8}$$

$$\|p\|_H^2 < \frac{1}{64}$$

$$4\|p\|_H^2 < \frac{4}{64} = \frac{1}{16}$$

$$4\|p\|_H^2 < \frac{1}{16}$$

Reemplazando éste último en (ξ) , se tiene:

$$-\frac{b^T (y - y(\mu))}{\mu} - \left\{ \sum_{i=1}^n \log \left(\frac{-f_i (y)}{-f_i (y(\mu))} \right) \right\} \leq \frac{1}{16} \quad (25)$$

definimos:

$$v_i = \frac{[f_i (y)]}{[-f_i (y(\mu))]} ; w = \frac{b^T (y - y(\mu))}{\mu}$$

de (24) tenemos que:

$$\begin{aligned} & \frac{-b^T (y(\mu))}{\mu} + \sum_{i=1}^n \frac{f_i (y) - f_i (y(\mu))}{[-f_i (y(\mu))]} \geq 0 \\ & = \frac{-b^T (y - y(\mu))}{\mu} + \sum_{i=1}^n \left[\frac{f_i (y)}{[-f_i (y(\mu))]} + 1 \right] \geq 0 \\ & = w - \sum_{i=1}^n v_i + n \geq 0 \end{aligned}$$

$$w \geq \sum_{i=1}^n v_i - n \quad (I)$$

$$\begin{aligned} f_i (y) & \geq f_i (y(\mu)) + \nabla f_i (y(\mu))^T (y - y(\mu)) \\ \nabla f_i (y(\mu))^T (y - y(\mu)) & \leq |f_i (y) - f_i (y(\mu))|, \forall y, y(\mu) \in F^0. \end{aligned}$$



Ahora dividiendo por $[-f_i (y(\mu))]$ en la última desigualdad, tenemos:

$$\frac{\nabla f_i(y(\mu))^T (y - y(\mu))}{[-f_i(y(\mu))]} \leq \frac{f_i(y) - f_i(y(\mu))}{[-f_i(y(\mu))]}$$

Aplicando la sumatoria a ambos miembros:

$$\sum_{i=1}^n \frac{\nabla f_i(y(\mu))^T (y - y(\mu))}{[-f_i(y(\mu))]} \leq \sum_{i=1}^n \frac{f_i(y) - f_i(y(\mu))}{[-f_i(y(\mu))]}$$

sumando a ambos miembros de éste último $\frac{-b^T (y - y(\mu))}{\mu}$, tenemos:

$$\frac{-b^T (y - y(\mu))}{\mu} + \sum_{i=1}^n \frac{\nabla f_i(y(\mu))^T (y - y(\mu))}{[-f_i(y(\mu))]} \leq \frac{-b^T (y - y(\mu))}{\mu} + \sum_{i=1}^n \frac{f_i(y) - f_i(y(\mu))}{[-f_i(y(\mu))]}$$

sustituyendo ésta desigualdad en (23), obtenemos:

$$0 \leq \frac{-b^T (y - y(\mu))}{\mu} + \sum_{i=1}^n \frac{f_i(y) - f_i(y(\mu))}{[-f_i(y(\mu))]} \quad (24)$$

ahora del Lema 4.5.5, se tiene:

$$\begin{aligned} \phi(y, \mu) - \phi(y(\mu), \mu) &\leq -\frac{b^T y}{\mu} - \sum_{i=1}^n \log(-f_i(y)) - \left\{ -\frac{b^T y}{\mu} - \sum_{i=1}^n \log(-f_i(y(\mu))) \right\} \leq 4 \|p\|_H^2 \\ &= -\frac{b^T (y - y(\mu))}{\mu} - \left\{ \sum_{i=1}^n \log(-f_i(y(\mu))) - \log(-f_i(y(\mu))) \right\} \leq 4 \|p\|_H^2 \\ &= -\frac{b^T (y - y(\mu))}{\mu} - \left\{ \sum_{i=1}^n \log \left(\frac{-f_i(y)}{-f_i(y(\mu))} \right) \right\} \leq 4 \|p\|_H^2 \end{aligned} \quad (\xi)$$

De (25) de igual modo se tiene:

$$\begin{aligned} -\frac{b^T (y - y(\mu))}{\mu} - \sum_{i=1}^n \log \left(\frac{-f_i(y)}{-f_i(y(\mu))} \right) &\leq \frac{1}{16} \\ &= w - \sum_{i=1}^n \log v_i \leq \frac{1}{16} \quad (II) \\ w &\leq \frac{1}{16} + \sum_{i=1}^n \log v_i \end{aligned}$$

De (I) y (II) se tiene que:

$$\sum_{i=1}^n v_i - n \leq w \leq \frac{1}{16} + \sum_{i=1}^n \log v_i$$



Dividiendo entre n , se tiene:

$$\frac{\sum_{i=1}^n v_i}{n} - 1 \leq \frac{w}{n} \leq \frac{1}{16n} + \frac{\sum_{i=1}^n \log v_i}{n}, \text{ donde: } \bar{V} = \frac{\sum_{i=1}^n v_i}{n} \quad (\beta)$$

$$\bar{V} - 1 \leq \frac{w}{n} + \frac{1}{16n} + \frac{\sum_{i=1}^n \log v_i}{n}$$

Ahora usando la desigualdad de la Media Aritmética – Geométrica:

$$\left(\prod_{i=1}^n v_i \right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^n v_i = \bar{V}$$

$$(V_1 \times V_2 \times \dots \times V_n)^{1/n} \leq \bar{V}$$

$$\log(V_1 \times V_2 \times \dots \times V_n)^{1/n} \leq \log \bar{V}$$

$$\frac{1}{n} \log(V_1 \times V_2 \times \dots \times V_n)^{1/n} \leq \log \bar{V}$$

$$\log(V_1 \times V_2 \times \dots \times V_n)^{1/n} \leq n \log \bar{V}$$

$$\sum_{i=1}^n \log V_i \leq n \log \bar{V}$$

$$\frac{1}{n} \sum_{i=1}^n \log V_i \leq \log \bar{V}$$

Reemplazando en (β) , se tiene:

$$\bar{V} - 1 \leq \frac{w}{n} \leq \frac{1}{16n} \leq \frac{1}{n} \sum_{i=1}^n \log V_i \leq \frac{1}{16n} + \log \bar{V}$$

entonces:

$$\bar{V} - 1 \leq \frac{w}{n} \leq \frac{1}{16n} + \log \bar{V} \quad (27)$$

como:

$$\log \bar{V} \leq \bar{V}; \quad \forall \bar{V} \geq 1$$

$\log \bar{V}$ es creciente para $\bar{V} \geq 1$.

Sea $h(\bar{V}) = \log \bar{V} - \bar{V} \leq 0; \quad \forall \bar{V} \geq 1$.

Derivando con respecto a \bar{V} :



$$\frac{dh(\bar{V})}{d\bar{V}} = \frac{1}{\bar{V}} - 1; \forall \bar{V} \geq 1$$

entonces:

$$\frac{dh(\bar{V})}{d\bar{V}} \leq 0.$$

Por lo tanto $h(\bar{V})$ es decreciente para $\bar{V} \geq 1$, entonces usando la expansión, para

$$\bar{V} \geq 1 + \frac{1}{2\sqrt{n}}.$$

Si: $\bar{V} \geq 1 + \frac{1}{2\sqrt{n}} \Rightarrow h(\bar{V}) \leq h\left(1 + \frac{1}{2\sqrt{n}}\right)$; $h(\bar{V})$ es decreciente.

Luego:

$$\begin{aligned} \frac{1}{16n} + \log \bar{V} - \bar{V} + 1 &\leq \frac{1}{16n} + \log\left(1 + \frac{1}{2\sqrt{n}}\right) - \left(1 + \frac{1}{2\sqrt{n}}\right) + 1 \\ &= \frac{1}{16n} + \log\left(1 + \frac{1}{2\sqrt{n}}\right) - \frac{1}{2\sqrt{n}} \\ &\leq \frac{1}{16n} + \log(1) + \log'(1) \cdot \frac{1}{2\sqrt{n}} + \log''(1) \cdot \frac{1}{2!} \frac{1}{4n} + \frac{\log'''(1)}{3!} \frac{2}{8n^{3/2}} - \frac{1}{2\sqrt{n}} \\ &= \frac{1}{16n} + \frac{1}{2\sqrt{n}} - \frac{1}{8n} + \frac{1}{24n^{2/3}} - \frac{1}{2\sqrt{n}} \\ &= \frac{1}{16n} - \frac{1}{8n} + \frac{1}{24^{3/2}} \end{aligned}$$

por lo tanto para todo \bar{V} satisfaciendo (27) se tiene:

$$\bar{V} < 1 + \frac{1}{2\sqrt{n}} \quad (I)$$

De la misma manera podemos probar que: $\bar{V} > 1 - \frac{1}{2\sqrt{n}}$, tenemos:

$$\bar{V} \leq 1 - \frac{1}{2\sqrt{n}}, \text{ para } 0 < \bar{V} < 1,$$

Entonces

$$\log \bar{V} < \bar{V}, \forall \bar{V} \in (0, 1).$$



Sea, $h(\bar{V}) = \log \bar{V} - \bar{V}$, $\forall \bar{V} \in \langle 0, 1 \rangle$.

Luego,

$$\frac{dh(\bar{V})}{d\bar{V}} = \frac{1}{\bar{V}} - 1 > 0.$$

Entonces: $h(\bar{V})$ es creciente.

$$h(\bar{V}) \leq h\left(1 - \frac{1}{2\sqrt{n}}\right)$$

$$\log \bar{V} - \bar{V} \leq \log\left(1 - \frac{1}{2\sqrt{n}}\right) - \left(1 - \frac{1}{2\sqrt{n}}\right)$$

$$\begin{aligned} \frac{1}{16n} + \log \bar{V} - \bar{V} + 1 &\leq \frac{1}{16n} + \log\left(1 - \frac{1}{2\sqrt{n}}\right) - \left(1 - \frac{1}{2\sqrt{n}}\right) + 1 \\ &= \frac{1}{16n} + \log\left(1 - \frac{1}{2\sqrt{n}}\right) + \frac{1}{2\sqrt{n}} \\ &\leq \frac{1}{16n} + \log(1) + \log'(1)\left(\frac{-1}{2\sqrt{n}}\right) + \frac{\log''(1)}{2!}\left(\frac{-1}{2\sqrt{n}}\right)^2 \\ &= \frac{\log'''(1)}{3!}\left(\frac{-1}{2\sqrt{n}}\right)^3 + \frac{1}{2\sqrt{n}} \\ &= \frac{1}{16n} - \frac{1}{2\sqrt{n}} - \frac{1}{24^{3/2}} + \frac{1}{2\sqrt{n}} \\ &= \frac{1}{16n} - \frac{1}{8n} - \frac{1}{24^{3/2}} \end{aligned}$$

Por lo tanto. Para todo \bar{V} satisfaciendo (27) se tiene:

$$\bar{V} > 1 - \frac{1}{2\sqrt{n}}; \forall \bar{V} \in \langle 0, 1 \rangle \quad (II)$$

De (I) y (II) tenemos:

si \bar{V} satisface (27):

$$\begin{aligned} 1 - \frac{1}{2\sqrt{n}} &< \bar{V} < 1 + \frac{1}{2\sqrt{n}} \\ |\bar{V} - 1| &< \frac{1}{2\sqrt{n}} \end{aligned}$$



por otro lado para n suficientemente grande:

$$\begin{aligned}\bar{V} - 1 &\approx \frac{w}{n} \\ \therefore \left| \frac{w}{n} \right| &\leq \frac{1}{2\sqrt{n}} \\ \left| \frac{-b^T(y - y(\mu))}{\mu n} \right| &\leq \frac{1}{2\sqrt{n}} \\ |b^T(y - y(\mu))| &\leq \frac{1}{2\sqrt{n}}; \forall \mu, n > 0 \\ |b^T(y - y(\mu))| &\leq \frac{1}{2} \mu \sqrt{n}\end{aligned}$$

4.6. ANÁLISIS DE CONVERGENCIA

En esta sección vamos a dar cotas para el número total de iteraciones externas e internas.

Teorema 6.1

Después de un mínimo de k iteraciones externas, con $k = O\left(\left(\frac{1}{\theta}\right) \log\left(\frac{n}{\varepsilon}\right)\right)$, el algoritmo termina con una ε -solución optimal para (CP).

Demostración

Se puede derivar una cota superior para la dualidad – Gap después de k iteraciones externas:

$$z^* - b^T y^k \leq z^* - b^T y(\mu^k) + |b^T y(\mu^k) - b^T y^k| \leq \mu^k \left(n + \left(\frac{1}{2}\right) \sqrt{n} \right)$$

donde

$$\mu^k = (1 - \theta)^k \mu^0$$

debido a (6) y al lema 5.6

esto significa que: $z^* - b^T y^k \leq \varepsilon$

es válido si: $(1 - \theta)^k \mu^0 \left(n + \left(\frac{1}{2}\right) \sqrt{n} \right) \leq \varepsilon$



Tomando en cuenta el algoritmo, se obtiene

$$k \geq \left[-\log \varepsilon + \log \left(n + \left(\frac{1}{2} \right) \sqrt{n} \right) + \log \mu^\circ \right] / [-\log(1-\theta)]$$

y si asumimos que $\mu^\circ \leq -\log(1-\theta)$

y desde que

$$\theta \leq -\log(1-\theta)$$

se cumple que:

$$k \geq \frac{1}{\theta} (-\log \varepsilon + \log n + \log \mu^\circ)$$

$$\therefore k = O\left(\frac{1}{\theta} \log\left(\frac{n}{\varepsilon}\right)\right)$$

Ahora damos una cota superior para el número total de iteraciones internas durante una iteración externa arbitraria.

Teorema 6.2

El número total P de iteraciones internas durante una iteración externa arbitraria satisface

$$P\delta \leq \left(1 + \frac{\theta}{(1-\theta)}\right)(\theta n + \sqrt{n})$$

donde δ es el factor de disminución o decrecimiento garantizado en cada iteración interna, dado por el lema 4.5.4.

Demostración

Denotamos por $\bar{\mu}$ parámetro de barrera usado en una iteración externa arbitraria, mientras que la cota inferior en la iteración externa anterior la denotamos por μ , Entonces, se cumple

$$\bar{\mu} = (1-\theta)\mu$$

Durante la última iteración externa se dan los iterados y°, y^1, \dots, y^P , donde y° es el iterado al inicio de la iteración externa. Debido al lema 5.4 se tiene

$$\phi(y^P, \bar{\mu}) \leq \phi(y^\circ, \bar{\mu}) - P\delta \dots (1)$$

Sabemos también que:



$$\phi(y, \bar{\mu}) = \phi(y, \mu) - \frac{\theta b^t y}{\mu} \dots (2)$$

Si lo sustituimos en (1), se obtiene:

$$P\delta \leq \phi(y^\circ, \mu) - \phi(y^p, \mu) + \left(\frac{\theta}{\mu}\right)(b^t y^p - b^t y^\circ) \dots (3)$$

y como y° es el más centrado, por el lema 4.5.5, se cumple

$$\phi(y^\circ, \mu) - \phi(y_{(\mu)}, \mu) \leq 1$$

Por lo tanto:

$$\phi(y^\circ, z) - \phi(y^p, z) \leq 1 + \phi(y_{(\mu)}, P) - \phi(y^p, \mu) \leq 1 \dots (4)$$

luego, lo reemplazamos en (3), y obtenemos

$$P\delta \leq 1 + \left(\frac{\theta}{\mu}\right)(b^t y^p - b^t y^\circ) \dots (5)$$

Por otro lado, como y° está aproximadamente centrado con respecto a μ y y^p con respecto a $\bar{\mu}$, obtenemos, usando el lema 5.6:

$$|b^t y^\circ - b^t y(\mu)| \leq \frac{1}{2} \mu \sqrt{n} \dots (6)$$

$$|b^t y^p - b^t y(\bar{\mu})| \leq \frac{1}{2} \bar{\mu} \sqrt{n} \dots (7)$$

Además, usando (14) se obtiene

$$b^t y(\bar{\mu}) - b^t y(\mu) \leq n(\mu - \bar{\mu}) \dots (8)$$

y usando las desigualdades (6), (7) y (8), se obtiene

$$\begin{aligned} b^t y^p - b^t y^\circ &= b^t y^p - b^t y(\bar{\mu}) - (b^t y^\circ - b^t y(\mu)) + b^t y(\bar{\mu}) - b^t y(\mu) \\ &\leq \frac{1}{2} \mu \sqrt{n} + \frac{1}{2} \bar{\mu} \sqrt{n} + n(\mu - \bar{\mu}) \\ &\leq \mu(\theta n + \sqrt{n}) \dots (9) \end{aligned}$$

Luego, este es sustituido en (5), para obtener

$$P\delta \leq 1 + \left[\frac{\theta}{1-\theta}\right](\theta n + \sqrt{n}).$$



CAPITULO V

RESULTADOS

Del Teorema 6.1, sabemos que el número total de iteraciones externas es como mínimo $k = O\left(\frac{1}{\theta} \log\left(\frac{n}{\varepsilon}\right)\right)$.

Por lo tanto, el número total de iteraciones de todo el proceso es $\left(\frac{1}{\delta}\right)\left(\frac{1}{\theta} + \frac{(\theta n + \sqrt{n})}{(1-\theta)}\right) O\left(\log \frac{n}{\varepsilon}\right)$

Esto hace que:

- i) Si tomamos $\theta = \left(\frac{v}{\sqrt{n}}\right)$, donde v es independiente de n , ε y M , entonces el algoritmo posee un $O\left((1+M^2)\sqrt{n} \log \varepsilon\right)$, de cota inferior para la iteración. En este caso podemos retornar a la vecindad de la trayectoria central en $O(1+M^2)$ pasos, mientras es cota inferior debe ser actualizada en $(\sqrt{n}|\log \varepsilon|)$ pasos.
- ii) Si tomamos $0 < \theta < 1$, independiente de n , ε y M entonces el algoritmo posee $O\left((1+M^2)n|\log \varepsilon|\right)$ de cota inferior para la iteración. Ahora podemos retornar a la vecindad de la trayectoria central $O\left((1+M^2)n\right)$ usando búsqueda lineal, mientras que la cota inferior debe ser actualizada en $O(|\log \varepsilon|)$ de tiempo.



CAPITULO VI

DISCUSIÓN DE RESULTADOS

6.1 CONTRASTACIÓN DE LA HIPOTESIS

En el proyecto de investigación se formuló la hipótesis general que, si en un programa convexo se satisface las condiciones de suavidad entonces se puede aplicar la función barrera logarítmica paramétrica, lo que permitió formular las siguientes hipótesis específicas:

1. Si las hessianas de las funciones involucradas en el Programa Convexo satisfacen la condición de Lipschitz entonces se le aplica el método de barrera logarítmica paramétrica.
2. Si se aplica el método de barrera logarítmica paramétrico al Programa Convexo entonces se genera un esquema iterativo de solución para dicho problema.

Se pudo observar que haciendo uso de las condiciones de Karush-Khun-Tucker (KKT) al problema primal y al problema barrera generados se obtiene un sistema de ecuaciones no lineales que caracterizan a la solución primal-dual y que conforman la trayectoria central. Para resolver dicho sistema se generan las direcciones newton los cuales permiten generar una sucesión de puntos en forma iterativa que siguen a la trayectoria central y haciendo variar el parámetro μ hacia cero nos permite aproximarnos a la solución optimal del programa convexo. Asimismo, las hipótesis asumidas y el esquema iterativo planteado permite proponer un algoritmo eficiente para resolver el problema planteado.

6.2 RESPONSABILIDAD ÉTICA

Conforme al código de ética de investigación de la Universidad Nacional del Callao aprobada por consejo Universitario N° 210-2017-CU del 06 de julio del 2017, en nuestra investigación se verificó su cumplimiento con la normativa institucional que regulan su proceso, se procedió con el rigor científico para su validación, fiabilidad y credibilidad de los métodos y fuentes de consulta utilizados con responsabilidad y transparencia en todo momento.



Es decir, que para la ejecución del proyecto se ha cumplido a cabalidad en lo establecido en el Reglamento General de Investigación, el Reglamento de Propiedad Intelectual y el Reglamento de participación de los docentes en proyectos de investigación aprobados por la Universidad Nacional del Callao. No se han falsificado o inventado datos o resultados total o parcialmente, ni se han plagiado datos, resultados, tablas, cuadros de otros autores o investigadores. Se ha cumplido con citar las referencias o fuentes bibliográficas, datos, resultados e información general de otros autores o investigadores, respetando sus derechos de autoría y de propiedad intelectual.

A handwritten signature in black ink, appearing to read "Gruai M. Pantoja", is centered on the page. The signature is written in a cursive style with some loops and flourishes.

CONCLUSIONES

A partir de los resultados obtenidos se ha podido establecer las siguientes conclusiones:

1. Las condiciones de suavidad en un programa convexo deben ser satisfechas para poder aplicar la función barrera logarítmica paramétrica.
2. El sistema obtenido al aplicar las condiciones de Karush-Khun-Tucker (KKT) al problema primal y al problema barrera generados puede ser resuelto usando el Método de Newton.
3. A medida que se varía el parámetro μ la solución del sistema anterior generan las direcciones newton los cuales permitirán generar la sucesión de puntos en forma iterativa que convergerán a la solución óptima del programa convexo.
4. El esquema iterativo planteado en el ítem 3 permite proponer un algoritmo que aproxima la solución óptima del programa convexo.
5. El número de iteraciones que se necesitan para que el algoritmo converja a un ε - solución óptima es de $O\left((1+M^2)\sqrt{n}|\log \varepsilon|\right)$ o $O\left((1+M^2)n|\log \varepsilon|\right)$.



RECOMENDACIONES

Al finalizar el proyecto de investigación podemos recomendar a los interesados en el área lo siguiente:

1. Estudiar otras condiciones que el programa convexo pueda satisfacer para poder aplicar la función barrera logarítmica.
2. Estudiar la aplicación de otros tipos de función en lugar de la función barrera logarítmica.
3. Plantear el uso de otro tipo de norma en lugar de H-norma.
4. Fomentar la investigación en el área.



REFERENCIAS BIBLIOGRÁFICAS

1. FRISCH, R., The Logarithmic Potential Method for Solving Linear Programming Problems, Memorandum, Institute of Economics, Oslo, Norway, 1955.
2. FRISCH, R., The Logarithmic Potential Method of Convex Programming, Memorandum, Institute of Economics, Oslo, Norway, 1955.
3. PARISOT, G.R., FRISCH, R., Résolution Numérique Approchée du Problème de Programmation Linéaire par Application de la Programmation Logarithmique, Revue Française de Recherche Opérationnelle. Vol. 20, pp. 227 – 259. 1961.
4. LOOTSMA, F.A., LOGARITHMIC Programming: A Method of Solving Nonlinear Programming Problems, Philips Research Reports. Vol. 22, pp. 329 – 344, 1967.
5. LOOTSMA, F.A., Constrained Optimization via Penalty Functions, Philips Research Report. Vol. 23, pp. 408 – 423, 1968.
6. FIACCO, A.V., and McCormick, G. P., Nonlinear Programming: Sequential Unconstrained Minimization Techniques, Wiley, New York, New York, 1968.
7. POLAK, E., Computational Method in Optimization, Academic Press, New York, New York, 1971.
8. HUARD, P., Resolution of Mathematical Programming with Nonlinear Constraints by the Methods of Centres, Nonlinear Programming, Edited by J. Abadie, North-Holland Publishing Company, Amsterdam, Holland, pp. 207 – 219, 1967.
9. LOOTSMA, F. A., Boundary Properties of Penalty Functions for Constrained Minimization. PhD Thesis. Technical University Eindhoven, Eindhoven, Holland, 1970.
10. LOOTSMA, F. A., A Survey of Methods for Solving Constrained Minimization Problems via Unconstrained Minimization, Numerical Methods for Nonlinear Optimization. Edited by F. A. Lootsma. Academic Press, New York, New York, 1972.
11. KARMARKAR, N., A New Polynomial-Time Algorithm for Linear Programming Combinatorica. Vol. 4, pp. 373 – 395, 1984.

12. TODD, M. J., Recent Developments and New Directions in Linear Programming, Technical Report No. 827. School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, New York, 1988.
13. GILL, P. E., MURRAY, W., SAUNDERS, M. A., TOMLIN, J. A., and WRIGHT, M.H.. On Projected Newton Barrier Methods for Linear Programming and an Equivalence to Karmarkar's Projective Method, Mathematical Programming. Vol. 36, pp. 183 – 209, 1986.
14. GONZAGA, C. C., An Algorithm for Solving Linear Programming Problems in $O(n^L)$ Operations, Progress in Mathematical Programming: Interior Point Methods, Edited by N. Megiddo, Springer Verlag, New York, pp. 1 – 28, 1989.
15. ROOS. C., and VIAL, J. P., Long Steps with the Logarithmic Penalty Barrier Function in Linear Programming, Economic Decision Making: Games, Economics and Optimization, Edited by J. Gabszeweez, J. f. Richard, and L. Wolsey, Elsevier Science Publisher, Amsterdam, Holland, pp. 433 – 441, 1990.
16. GONZAGA C.C., Large-Steps Path-Following Methods for Linear Programming: Barrier Function Method, SIAM Journal on Optimazation. Vol. 1, 1991.
17. RENEGAR, J.. A Polynomial-Time Algorithm Based on Newton's Method for Linear Programming, Mathematical Programming, Vol. 40, pp. 59 – 93, 1988.
18. VAIDYA, P. M.. An Algorithm for Linear Programming Which Requires $O(((m+n)n^2+(m+n)^{1.5}n) L)$ Arithmetic Operations, Mathematical Programming. Vol. 47, pp. 175 – 201, 1990.
19. DEN HERTOOG, D., ROOS, C., and TERLAKY, T., A Potential Reduction Variant of Renegar's Short-Step Path-Following Method for Linear Programming, Linear Algebra and Its Applications, Vol. 152, pp. 43 – 68, 1991.
20. SONNEVEND, G., An Analytical Centre for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming, Lecture Notes in Control and Information Science. Vol. 84, pp. 866 – 876, 1985.
21. JARRE, F., On the Convergence of the Method of Analytic Centers when Applied to Convex Quadratic Programs, Mathematical Programming, Vol. 49, pp. 341 – 358, 1990/91.



22. JARRE, F.. The Method of Analytic Centers for Smooth convex Programs, PhD Thesis, Institut für Angewandte Mathematik and Statistik, Universität Würzburg, Würzburg. Germany, 1989.
23. MEHROTRA, S., and SUN, J., An Interior Point Algorithm for Solving Smooth Convex Programs Based on Newton's Method, Technical Report No. 88 – 08, Department of Industrial Engineering and Management Science, Northwestern University, Evanston, Illinois, 1988.
24. NESTEROV, Yu. E., and NEMIROVSKY, A. S., Self-Concordant Functions and Polynomial Time Methods in Convex Programming, Central Economical and Mathematical Institute, USSR Academy of Science, Moscow, USSR, 1989.
25. DEN HERTOOG, D., ROOS, C., and TERLAKY, T., A Potential Reduction Method for a Class of Smooth Convex Programming Problems, SIAM Journal on Optimization, Vol. 2, 1992.
26. WOLFE, P., A Duality Theorem for Nonlinear Programming, Quarterly of Applied Mathematics, Vol. 19, pp. 239 – 244, 1961.
27. MEGIDDO, N., Pathways to the Optimal set in Linear Programming, Proceedings of the Mathematical Programming symposium of Japan, pp. 1 – 36, 1986.

