UNIVERSIDAD NACIONAL DEL CALLAO

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
UNIDAD DE INVESTIGACIÓN

INFORME FINAL DE INVESTIGACIÓN

"TRAYECTORIAS QUE CONDUCEN AL CONJUNTO OPTIMAL EN PROGRAMACIÓN LINEAL"

AUTOR: EDINSON RAUL MONTORO ALEGRE

LINEA DE INVESTIGACIÓN: MATEMÁTICA APLICADA

(PERIODO DE EJECUCIÓN: DEL 01.07.2021 AL 30.06.2022)

(RESOLUCIÓN DE APROBACIÓN Nº 436-2021-R)

Callao, 2022

PERU

IMFORMACIÓN BÁSICA

I.1 FACULTAD

CIENCIAS NATURALES Y MATEMÁTICA

I.2 UNIDAD DE INVESTIGACIÓN

Unidad de Investigación de la Facultad de Ciencias Naturales y Matemática.

I.3 TITULO DEL PROYECTO DE INVESTIGACIÓN

Trayectorias que Conducen al Conjunto Optimal en Programación Lineal

I.4 AUTOR

Edinson Raúl Montoro Alegre
Codigo Orcid: 0000-0002-8237-9469
DNI: 09627181

I.5 LUGAR DE EJECUCIÓN

Facultad de Ciencias Naturales y Matemáticas

I.6 TIPO

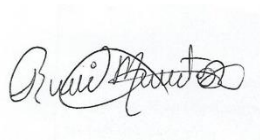
Investigación Básica. Investigación aplicada, cuantitativa y transversal.

I.7 TEMA OCDE

1.01.02 Matemática Aplicada

DEDICATORIA

A mis hijos



AGRADECIMIENTOS

A todas aquellas personas que luchan por un Perú mejor y no desfallecen

INDICE

Resumen

En este trabajo de investigación se presentan trayectorias continuas que conducen al conjunto de Soluciones Optimas del Problema de Programación Lineal. Dichas trayectorias son derivadas a partir de la función Barrera logarítmica ponderada. Asimismo, las ecuaciones que definen a las trayectorias son bilineales y poseen propiedades de simetría Primal-Dual Suaves.

Abstract

In this research work, continuous trajectories are presented that lead to the set of Optimal Solutions of the Linear Programming problem. Said trajectories are derived from the weighted logarithmic barrier function. Likewise, the equations that define the trajectories are bilinear and have Smooth Primal-Dual symmetry properties.

INTRODUCCION

Los algoritmos iterativos para Optimización No-Lineal usualmente se caracterizan por asignar un punto $x \in S \subset R^n$ (usualmente convexo) un “próximo punto” $x' = f(x) \epsilon\, S$. Es decir, dado un punto inicial x^0, el esquema iterativo genera una sucesión de puntos $\{x^k\}$, donde $x^{k+1} = f(x^k)$, que converge a la solución.

Es con frecuencia muy útil utilizar la idea de considerar la versión “infinitesimal” de algoritmos iterativos en el siguiente sentido: Dado el esquema iterarivo $x' = f(x)$, considerar la Ecuación Diferencial Ordinaria

$$\dot{x} = f(x) - x = x' - x \qquad (*)$$

Cuando esta ecuación posee única solución a través de x^0 , entonces éste determina una trayectoria (o curva) $x = x(t)$ tal que la tangente de la trayectoria en cualquier x es igual a la recta determinada por x y x'.

Si un algoritmo genera el punto x' “cerca” a x entonces la trayectoria puede ser una buena aproximación a la secuencia generada por el algoritmo. Esto será cierto al menos durante las etapas posteriores de la ejecución si la sucesión converge a un punto solución. Si el algoritmo en cambio da pasos “largos” durante las primeras etapas, entonces la ruta puede ser una mala aproximación.

En el presente trabajo de investigación se pretende hacer un estudio de las trayectorias solución o curvas que surgen de la idea de relacionar el algoritmo iterativo de Programación Lineal.

I.PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

1.1 Descripción de la realidad problemática

Algoritmos para Programación Lineal pueden con frecuencia ser interpretados como procedimientos que siguen a una trayectoria o curva. Esta interpretación ya fue aplicada: Al Método Simplex [4], al algoritmo de punto fijo de Scarf [19], al algoritmo de Lemke [13] para el problema de Complementariedad Lineal, a los métodos de Homotopía a las ecuaciones lineal por partes [5] y a la mayoría de métodos de Optimización No-Lineal. Este tema también es tratado en el libro de García y Zangwill [8]. Por otro lado, los algoritmos para Programación Lineal de Murty [17] y Mangasariam [15] también están basados en trayectorias naturales que conducen a las soluciones óptimas.

Las trayectorias correspondientes a algoritmos discretos para la optimización No-lineal se estudiaron en [6, 8, 10]. Por otro lado, la analogía con las ecuaciones diferenciales es bien conocida. Varios investigadores han trabajado sobre trayectorias para llegar a soluciones de Programación Lineal. Nazareth [18] interpreta el Algoritmo de Karmarkar [12] como un método de homotopía con reinicio. En [2] y [16] se estudian resultados acerca de la versión infinitesimal del algoritmo de Karmarkar y otros relacionados. Smale [20] muestra que la trayectoria generada por el Algoritmo Simplex auto-dual [4] puede ser aproximada por el método Newton-trayectoria para resolver un cierto sistema de ecuaciones no-lineales.

1.2 Formulación del problema

Por otro lado, como ya se ha mencionado es muy instructivo y didáctico el considerar la versión “infinitesimal” de algoritmos iterativos en el siguiente sentido: Dado el esquema iterativo $x' = f(x)$, considerar la ecuación diferencial ordinaria

$$x = f(\dot{x}) - x = x' - x$$

Cuando ésta ecuación posee única solución a través de x^0 entonces éste determina una trayectoria $x = x(t)$ tal que la tangente de la trayectoria en cualquier punto x será la línea recta determinada por los puntos x' y x. Si un punto genera el punto x' “cerca” a x entonces la trayectoria puede ser una buena aproximación a la sucesión generada por el algoritmo. Esto será cierto al menos durante las etapas posteriores de la ejecución si la sucesión converge a un punto solución. Si el algoritmo da pasos largos (o grandes) durante las primeras etapas, entonces la trayectoria puede ser una mala aproximación.

Estas ideas ya fueron consideradas por diversos investigadores desde el año 1968 como se puede ver en [6, 8, 10]. Asimismo las diferentes aplicaciones dadas por Nazareth [18], Frisch [7] y Greenberg [10] pero ninguno ha tratado o puesto énfasis en las trayectorias solución relacionadas al Problema de Programación Lineal. Por lo tanto, se puede definir

1.2.1 Problema General

¿De qué manera se puede describir a las trayectorias solución de un Problema de Programación Lineal?

1.2.2 Problema Específico:

¿Qué propiedades tienen las trayectorias solución de un Problema de Programación Lineal?

1.3 Objetivos

1.3.1 Objetivo General

Describir a las trayectorias solución de un Problema de Programación Lineal.

1.3.2 Objetivo Específico

Deducir las propiedades que tienen las trayectorias solución de un Problema de Programación Lineal.

1.4 Justificación

La comprensión, identificación y solución de problemas de Programación Lineal resultan ser muy importantes en las diferentes áreas como: Análisis de riesgos, administración de empresas, planeación urbana, contabilidad, ingeniería y banca.
En ese sentido, el desarrollo y ampliación de todo tipo de conocimiento nuevo en el área de la Programación Lineal repercutirá en mejorar las técnicas de aplicación en las diferentes áreas antes mencionadas. El presente trabajo de investigación pretende aportar y explorar nuevas ideas para obtener técnicas y/o métodos nuevos de solución.

1.5 Limitantes de la Investigación

a. **Limitante Teórico.**

El trabajo de investigación esta circunscrita en la teoría de Métodos de Punto interior, es decir **el limitante teórico** de nuestra investigación está dentro de las fronteras de la aplicación de los métodos de punto interior aplicados a Programación Lineal, funciones barreras y funciones convexas diferenciables. Aún quedan buena parte de problemas que no son considerados como: Programación No Lineal no diferenciable y Programación No Lineal Cuasi-Convexa y que están fuera de las fronteras antes mencionadas.

b. **Limitante Temporal:**

El tipo de investigación es básica, también llamada investigación Pura, teórica o dogmática por tal motivo no está sujeto al factor tiempo ni mucho menos pueda éste interferir como variable a la hora de realizar la investigación. El **limitante temporal** no se aplica a este tipo de investigación.

c. **Limitante Espacial**:

El tipo de investigación es básica, también llamada investigación Pura, teórica o dogmática por tal motivo no está sujeto al factor espacio o lugar ni mucho menos pueda éste interferir como variable a la hora de realizar la investigación. **El limitante espacial** no se aplica a este tipo de investigación

II.EL MARCO TEORICO

Esta sección está dedicada a una revisión de algunos conceptos matemáticos en Programación Lineal, Método de Función Barrera, Teoría de la Dualidad.

2.1 Antecedentes

2.1.1 Internacionales

A nivel internacional tenemos los siguientes datos: Fiacco (1968) fue el primero en derivar la teoría relacionada a los métodos de punto interior. Greenberg et al (1972) trabajo las trayectorias solución relacionadas a Problemas de Programación No Lineal y posteriormente García et al (1981) analizaron las trayectorias solución de optimización relacionadas a ecuaciones de Punto Fijo y Problema de Equilibrio.

Después Nazareth (1986) interpreta el algoritmo de Karmarkar [12] como un método de homotopía con reinicio, análogamente, Smale (1986) muestra que la trayectoria generada por el algoritmo Simplex Auto-Dual [4] puede ser aproximado por el Método de Newton-trayectoria resolviendo un cierto sistema de ecuaciones no lineales.

2.1.2 Nacionales

A nivel nacional tenemos los siguientes trabajos: Quijano (2019) presenta un método que sigue la trayectoria central para resolver un problema de programación lineal, el método le permite deducir un algoritmo primal-dual de pasos cortos utilizando una apropiada medida de proximidad. En cambio, Luna (2020) desarrolla y describe un método de punto interior primal-dual para resolver el problema de programación lineal el cual utiliza una matriz de escalamiento para deducir dos direcciones de descenso.

2.2 Marco

2.2.1 Teórico

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

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

Es mediante la aproximación a este problema a través de una familia de problemas de minimización irrestricta Fiacco (1968)

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

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´) esté 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

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

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-Khun-Tucker (Gill). 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.

En nuestro caso, vamos a considerar el siguiente Problema de Programación Lineal en forma Estándar

$$(P)\begin{cases} \text{maximizar}\ c^{t}\ x \\ \qquad \text{s.a} \\ Ax = b \\ x \geq 0 \end{cases}$$

Donde A $\epsilon \mathbb{R}^{mxn}$, $b \ \epsilon\ \mathbb{R}^{m}$, $c, x\ \epsilon\ \mathbb{R}^{n}$

La función barrera logarítmica, generalmente usado en optimización con restricciones no lineales, puede ser aplicada al problema de Programación Lineal Estándar, dando lugar al siguiente problema.

$$
(P_u)\begin{cases} \quad \text{maximizar}\; c^t\, x + u \displaystyle\sum_i \ln x_i \\ \text{s.a} \\ Ax = b \\ x > 0 \end{cases}
$$

Donde $u > 0$ suele ser pequeño. El enfoque de la función barrera es válido si existe un $x > 0$ sujeto a $Ax = b$

A partir de estos problemas y aplicando las condiciones de Karush-Khunn-Tucker se pretende obtener un sistema de ecuaciones no lineales que permita caracterizar a las trayectorias solución. Asímismo, se deberá imponer ciertas condiciones o hipótesis para asegurar la solución del sistema obtenido.

Se debe entender que las trayectorias solución son esenciales para el diseño y análisis de algoritmos para optimización y describir trayectorias para problemas de Programación Lineal en forma Estándar nos permitirá idear e imaginar nuevas formas de solución.

2.2.2 Conceptual

Esperamos encontrar las condiciones necesarias para caracterizar a las trayectorias solución a través del sistema no lineal que surge al aplicar las condiciones de Karush-Khunn-Tucker. Asimismo, esperamos deducir alguna técnica adecuada para resolver dicho sistema no lineal.

2.3 Definición de términos básicos

Comenzamos revisando las direcciones de búsqueda de los tres importantes algoritmos punto – interior. De acuerdo a Shanno y Bagchi [14], todos aquellos métodos pueden ser clasificados como métodos de función barrera porque sus direcciones de búsqueda pueden ser obtenidas de una aproximación Newton (linealización) de las condiciones Karush – Kuhn – Tucken(K-K-T) de un problema de programación lineal aumentado con un término **barrera logarítmica** en la función objetivo.

Definición 1. (Frish [7])

El problema primal aumentado es:

$$
(P_\mu)\begin{cases} \min \mathrm{c}^t \mathrm{x} - \mu \sum_{\mathrm{i}=1}^{n} L_n \mathrm{x}_i \\ s.a \quad Ak = b \\ k > 0 \end{cases}
$$

y el problema dual aumentado es:

$$
(D_\mu)\begin{cases} \max \mathrm{b}^t \mathrm{y} + \mu \sum_{\mathrm{i}=1}^{n} L_n (C_i - A_i^t y) \\ s.a \quad A^t y < c \end{cases}
$$

Es necesario revisar algunas teorías de optimización que nos serán de gran utilidad en el desarrollo de la investigación.

1. Método de Barrera (Gill [3])

Los métodos de barrera se pueden aplicar a problemas del tipo

$$
\begin{cases} \min \mathrm{f(x)} \\ \mathrm{s.a\ g(x)} \leq 0;\ \mathrm{x} \in \mathrm{D} \end{cases} \tag{1}
$$

Donde $\mathrm{D} \subseteq \mathbb{R}^n$ $\quad f: \mathbb{R}^n \to \mathbb{R}, g: \mathbb{R}^n \to \mathbb{R}^m$ $\ g = (g_1, g_2 \ldots g_m), g_i \in C^0\ (D)$

Si hacemos

$$
\Omega = \{\mathrm{x} \in \mathrm{D}\ /\ \mathrm{g(x)} \leq 0\} \tag{2}
$$

y asumimos que el interior del conjunto viable (que lo denotamos por Ω^0) es no vacío es decir

$$
\Omega^0 = \{\mathrm{x} \in \mathrm{D}\ /\ \mathrm{g(x)} < 0\} \neq \emptyset \tag{3}
$$

Así mismo supongamos que:

$$
\inf_{\mathrm{x} \in \Omega^0} \mathrm{f(x)} = \inf_{\mathrm{x} \in \Omega} \mathrm{f(x)} = \mathrm{v} > -\infty \tag{4}
$$

Con esta hipótesis podemos eliminar ciertos casos patológicos, donde la solución no puede ser aproximada por puntos del interior del conjunto viable.

Supongamos que (1) tiene minimizador global, entonces podemos transformar (1) en un problema irrestricto con función objetivo $f(x) + tB(x), t > 0$ donde B es la función barrera.

Definición 2. (Fiacco [6]**)**

Una función $B: \Omega^0 \rightarrow \mathbb{R}_+$ es llamada Barrera para el problema (1) si satisface las siguientes condiciones:

(i) $B(x)$ *está definida y es continua para todo* $x \in \Omega^0$

(ii) $B(x) \geq o$ *para todo* $x \in \Omega^0$

(iii) *Si* $\{x_k\} \subset \Omega^0, g(x_k) < 0 \ \forall_k$ *y* $\lim_{k\rightarrow +\infty} g_i(x_k) = 0$ *para algún* $i \in \{1, \ldots, m\}$

Entonces $\lim_{k\rightarrow +\infty} B(x_k) = +\infty$

Ejemplo:

$$B(x) = -\sum_{i=1}^{m} \log\bigl(-g_i(x)\bigr) ; x \in \Omega^0 \text{ (La barrera logaritmica)} \qquad (5)$$

$$B(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)} ; x \in \Omega^0 \text{ (La barrera inversa)}$$

La función dada en (5) puede asumir valores negativos y por tanto no cumple la condición (ii). Pero para el caso en que Ω sea limitado podremos que trabajar con otra función que si satisface (ii). En efecto:

Sea $M \in \mathbb{R}$ tal que $B(x) > M$ para todo $x \in \Omega^0$ y consideremos

$$\widetilde{B}(x) = -\sum_{i=1}^{m} \log\bigl(-g_i(x)\bigr) - M$$

Esta función satisface:

(i) *Obvio*

(ii) *Obvio*

(iii) *Sea* $i \in \{1, \dots, m\} \quad / \lim_{k \to +\infty} g_i(x_k) = 0$ *y* $j \in \{1, \dots, m\} - \{i\}$

con $\lim_{k \to +\infty} g_j\,(x_k) = b_j < 0$

$$\Longrightarrow \widetilde{B}(x) = -\log(-g_1(x)) \dots - \log(-g_i(x)) \dots - M$$

$$\lim_{k \to +\infty} \widetilde{B}(x) = -\log(-b_1) \dots + \infty .. - M = +\infty$$

Ahora, el problema en barrera asociado a $\widetilde{B}$ será

$$\begin{cases} \min f(x) + t\,\widetilde{B}\,(x) \\ \quad s.a\ x \in \Omega^0 \end{cases} \quad \text{para } t > 0$$

que coincide con

$$\begin{cases} \min f(x) + tB(x) - tM \\ \quad s.a\ x \in \Omega^0 \end{cases} \quad \text{para } t > 0$$

y que es equivalente a

$$\begin{cases} \min f(x) + tB(x) \\ \quad s.a\ x \in \Omega^0 \end{cases} \quad \text{para } t > 0$$

Así la función logarítmica puede ser usada como barrera sin ningún prejuicio

ALGORITMO:

Escoger $t_0 > 0$ y tomar $k := 0$

1. Calcular $x_k \equiv x(t_k)$ solución global de

$$\begin{cases} \min f(x) + t_k\,B(x) \\ s.a.\,x \in \Omega^0 \end{cases} \tag{6}$$

2. Escoger t_{k+1} tal que $0 < t_{k+1} < t_k$ y tomar $k := k + 1$ y volver al paso (1)

Definimos la siguiente función $Q(x, t) = f(x) + tB(x)$ y vamos a probar las propiedades fundamentales del algoritmo.
Los siguientes lemas y teoremas, pueden ser hallados en Megido [16]

Lema 1
Sea $\{x_k\}$ una sucesión generada por el algoritmo entonces

(a) $Q(x_{k+1}, t_{k+1}) \leq Q(x_k, t_k)$

(b) $B(x_k) \leq B(x_{k+1})$

(c) $f(x_{k+1}) \leq f(x_k)$

Demostración

(a) $Q(x_{k+1}, t_{k+1}) = f(x_{k+1}) + t_{k+1}B\,(x_{k+1})$
$\leq f(x_k) + t_{k+1}B(x_k)$ Por defin. x_{k+1}
$\leq f(x_k) + t_kB(x_k)$ por el paso (2) del algoritmo
$= Q(x_k, t_k)$

(b) $Q(x_{k+1}, t_{k+1}) = f(x_{k+1}) + t_{k+1}B(x_{k+1}) \leq f(x_k) + t_{k+1}B(x_k)$

y por otro lado

$$Q(x_k, t_k) = f(x_k) + t_kB(x_k) \leq f(x_{k+1}) + t_kB(x_{k+1})$$

Sumando ambas desigualdades

$$f(x_{k+1}) + f(x_k) + t_{k+1}B(x_{k+1}) + t_kB(x_k)$$
$$\leq f(x_{k+1}) + f(x_k) + t_{k+1}B(x_k) + t_kB(x_{k+1})$$

$$(t_{k+1} - t_k)B(t_{k+1}) \leq (t_{k+1} - t_k)B(x_K)$$

Pero t_{k+1}-$t_k \leq 0$, entonces

$$B(t_{k+1}) \geq B(x_k)$$

(c) $f(x_{k+1}) + t_{k+1}B(x_{k+1}) \leq f(x_k) + t_{k+1}B(x_k)$

$\leq f(x_k) + t_{k+1}B(x_{k+1})$ por (b)

$$f(x_{k+1}) \leq f(x_k)$$

TEOREMA 2

Sean $f: \mathbb{R}^n \to \mathbb{R}$ y $g: \mathbb{R}^n \to \mathbb{R}^m$ funciones continuas en $\mathbb{R}^n$. Supongamos que el conjunto Ω definido en (2) satisface (3) y (4). Entonces, todo punto de acumulación de la sucesión $\{x_k\}$ generada por el algoritmo donde x_k es una solución global del problema (6) con $t = t_k$ y $t_k \to 0$ $(k \to +\infty)$, es una solución global del problema (1).

Demostración

Sea $\bar{x}$ un punto de acumulación de la sucesión $\{x_k\}$ y sea $\left\{x_{k_j}\right\} \to \bar{x}(j \to +\infty)$.

Como$\{x_k\} \subset \Omega^0$ es decir $g(x_k) < 0 \; \forall \; k$ y g continua $\Longrightarrow g(\bar{x}) \leq 0$ luego $\bar{x} \in \Omega$

Ahora, debemos probar que

$f(\bar{x}) = \bar{v};$ $\bar{v}$ =min$\{f(x)/x \in \Omega\}$

Supongamos $f(\bar{x}) > \bar{v}$ definimos $\delta = (f(\bar{x}) - \bar{v}) \,/\, 2 > 0$ y por la definición de infimun

$$\exists \; x_{k_l} \in \Omega^0 \text{tal que } f\left(x_{k_l}\right) < \bar{v} + \delta$$

la sucesión $\{f(x_k)\}$ es no-creciente, por tanto, $\{f(x_{k_j})\}$ tambien es no-creciente y tenemos que

$$f\left(x_{k_j}\right) \geq \lim_{j-+\infty} f\left(x_{k_j}\right) = f(\bar{x}) = 2\delta + \bar{v}$$

$$f\left(x_{k_j}\right) - f\left(x_{k_l}\right) \geq \bar{v} + 2\delta - f(x_{k_l}) > \delta \qquad (7)$$

Entonces tenemos que para j suficientemente grande (j >>>1)

$$0 \geq Q\left(x_{k_j}, t_{k_j}\right) - Q(x_{k_l}, t_{k_j})$$

$$= f\left(x_{k_j}\right) + t_{k_j} B\left(x_{k_j}\right) - f\left(x_{k_l}\right) - t_{k_j} \mathrm{B}(x_{k_l})$$

y debido a (7)

$$> \delta + \underbrace{t_{k_j}}_{>0} \underbrace{\left(B\left(x_{k_j}\right) - B(x_{k_l})\right)}_{>0} > \delta$$

$$\geq \delta \text{ pues } B(x_{k_j}) \geq B(x_{k_0})$$

$$\geq \delta > 0$$

esto conduce a una contradicción, luego

como $t_k \longrightarrow 0(k \longrightarrow +\infty) \Longrightarrow t_{k_j} \longrightarrow 0(j \longrightarrow +\infty)$

$$\therefore f(\bar{x}) = \bar{v}$$

2. Programación Convexa: Teorema de Karush – Kuhn-Tucker

Las siguientes definiciones son extraídas de Fiacco [6] y Cotle [3]
Consideremos el siguiente programa:

$$(P)\begin{cases} \qquad\qquad min\, f(x), \\ s.a\ g_1(x) \leq 0, g_2(x) \leq 0, \ldots\ g_m(x) \leq 0 \\ donde\ x\ \in C \subset \mathbb{R}^n, \end{cases}$$

Definición 3

Se da las siguientes definiciones

- Si $x\ \in C$ tal que $g_i(x) \leq 0, \forall i = 1, \ldots m,$ entonces, x es llamado punto viable de (P).
- Definimos F como, $F = \{x \in \mathbb{R}^n / x\ es\ un\ punto\ viable\ de\ (P)\}$ es decir
  F es la Región viable para (P)
- (P) es consistente, si la región viable para (P) es no vacío
- (P) es súper consistente, si existe un punto viable x para (P) tal que $g_i(x) < 0, \forall i = 1, \ldots m$. A dicho punto x se le llama Punto de Slater.

- Si (P) es un programa consistente y si x^* es un punto viable para (P) talque $f(x^*) \leq f(x)$ para todo punto viable x para (P) $\Longrightarrow x^*$ es una solución para (P)

Dado un programa (P) denotamos el ínfimo de la función objetivo $f(x)$ en la región viable F para (P) por MP esto es:

$$MP = \inf_{x \in F}\{f(x)\}$$

Investiguemos la sensibilidad del valor de

$$MP = \inf_{x \in F}\{f(x)\}$$

Para cambios ligeros en las restricciones o sea definamos para cada $z \in \mathbb{R}^m$ un programa (P) como sigue.

Para esto consideremos el problema P(z) asociado al programa (P) el cual es llamado Programa Perturbado de (P).

$$P(z)\begin{cases} \min f(x) \\ s.a\ g_1(x) \leq z_1, \dots, g_m(x) \leq z_m \\ donde\ x \in C \subset \mathbb{R}^n\ y\ f(x), \dots, g_1(x), \dots, g_m(x) \\ son\ funciones\ convexas\ definidas\ en\ un\ conjunto\ convexo\ C \end{cases}$$

Y en forma análoga definamos $MP(Z)$

$$\text{MP(z)}\begin{cases} MP_z = \inf_{x \epsilon F_z}\{f(x)\} \\ \\ \text{donde} \\ F_{(z)} = \{x\ \epsilon \mathbb{R}^n / x \epsilon C\ g_i(x) \leq z_i \forall_i\} \end{cases}$$

A hora definimos la función $MP(.)$ como

$$z \to MP(z)$$

Donde el dominio de la función $MP(.)$ es

$$MP = \{z \in \mathbb{R}^m / F(z)\ es\, no\ vacío\}$$

Es fácil notar que $MP(0) = MP$ si z = 0

Teorema 4

Si (*P*) es un programa convexo y *P*(*z*) es la perturbación de (*P*) para $z \in \mathbb{R}^m$, entonces la función $MP(z)$ es convexo y su dominio es un subconjunto convexo de $\mathbb{R}^m$. Si (*P*) es super consistente, entonces 0 es un punto interior del dominio de MP.

Demostración

(i) Probaremos que $dom(MP)$ es convexo. En efecto dado $z^{(1)}, z^{(2)} \in dom\ MP$

Por demostrar que $\lambda\ z^1 + (1-\lambda)z^{(2)} \in dom\ (MP)\ \forall\ \lambda \in [0,1]$

Como $z^{(1)} \in dom\ (MP) \Rightarrow F(z^{(1)})$ es no vacio $\Rightarrow \exists x^{(1)}\ /\ g(x^{(1)}) \leq z^1$

$z^2 \in dom\ (MP) \Rightarrow F(z^2)$ es no vacío $\Rightarrow \exists x^{(2)}\ /\ g(x^{(2)}) \leq z^2$

Luego $g(\lambda\ x^1 + (1-\lambda)x^2) \leq \lambda\ g(x^1) + (1-\lambda)g(x^2) \leq \lambda\ z^1 + (1-\lambda)z^2 \forall\ \lambda \in [0,1]$

$$\Rightarrow F(\lambda\ z^1 + (1-\lambda)z^2) \neq \phi \forall\ \lambda \in [0,1]$$

$\Rightarrow \lambda\ z^1 + (1-\lambda)z^2 \in dom\ (MP)$

(ii) Probaremos que $MP(z)$ es convexo dados $z^1, z^2 \in dom\ (MP)$

Por demostrar que

$MP(\lambda\ z_1 + (1-\lambda)z_2) \leq \lambda\ MP(z^1) + (1-\lambda)MP(z^2) \quad \forall\ \lambda \in [0,1]$

Si $MP(\lambda\ z^1 + (1-\lambda)z^2)$ es finito;

$MP(\lambda\ z^1 + (1-\lambda)z^2) =inf\{f(x): x \in C, g(x) \leq \lambda\ z^1 + (1-\lambda)z^2\}$

= inf$\{f(x): x = \lambda\ x^1 + (1-\lambda)x^2$ donde $x^1, x^2 \in C, g(x) \leq \lambda\ z^1 + (1-\lambda)z^2\}$

$\leq$inf$\{f(x): x = \lambda\ x^1 + (1-\lambda)x^2 donde x^1, x^2 \in C\ y\ \lambda\ g(x^1) + (1-\lambda)g(x^2) \leq \lambda\ z^1 + (1-\lambda)z^2\}$

$\leq$inf$\{f(\lambda\ x^1 + (1-\lambda)x^2): x^1, x^2 \in C\ y g(x^1) \leq z^1; g(x^2) \leq z^2\}$

$\leq$inf$\{\lambda\ f(x^1) + (1-\lambda)f(x^2): x^1, x^2 \in C\ y g(x^1) \leq z^1; g(x^2) \leq z^2\}$

$= MP(z^1) + (1-\lambda)MP(z^2)$

Si $MP(\lambda\ z^1 + (1-\lambda)z^2) = -\infty$

$\Rightarrow MP(\lambda\ z^1 + (1-\lambda)z^2) \leq \lambda\ MP(z^1) + (1-\lambda)MP(z^2)$

Se cumple

$\therefore MP(z)$es convexo en su dominio

(iii) Probaremos que 0 es un punto interior del dominio de $MP(z)$. En efecto, como (*P*) es superconsistente

$\Longrightarrow$ existe $w \in \mathbb{R}^n$tal que $g_i(w) < 0$ para $i = 1, \dots, m$

Sea $r = \min\{-g_i(w): 1 \leq i \leq m\}$ luego

para algún $z \in B(0, r)$ con r > 0

$\Longrightarrow -r < z_i < r; \forall i = 1, \dots m$

$\Longrightarrow g_i(w) \leq -r < z_i; i = 1, \dots m$

$\Longrightarrow g_i(w) < z_i; i = 1, \dots m$

$\Longrightarrow g(w) < z$;

$\Longrightarrow z \in dom(MP)$

$$\therefore \exists r > 0 / B(0, r) \subset dom(MP()$$

Teorema 5

Sea $f(x)$ una función convexa definida en un conjunto convexo C con $C^0 \neq \emptyset\ en\ \mathbb{R}^n$. Si $x^{(0)}$ es un punto interior de $C \Longrightarrow \exists d$ en $\mathbb{R}^n$ tal que

$$f(x) \geq f\left(x^{(0)}\right) + d(x - x^0) \forall x \in C$$

Demostración:

ver [3]

Lema 6

Sea C es convexo con$C^0 \neq \emptyset$ en $\mathbb{R}^n$

Si $x \in C^0\ e\ y\ \in C$, entonces el segmento semi abierto que une$x$ a y denotado por

$$[x, y >= \{\lambda\, x + (1 - \lambda)y: 0 < \lambda \leq 1\}$$

Está totalmente conformado por puntos interiores de C

Demostración.

Ver [3]

OBSERVACIONES: Se derivadas del Lema 6

(i) Si z^1 y $z^2 \in$ dom (MP) y si MP(z^1) $= -\infty$

$entonces \;\; MP(z) = -\infty \qquad \forall z \in [z^1, z^2 >$

En efecto: sea

$[z^1, z^2) = \{\lambda\, z^1 + (1 - \lambda) z^2 : 0 < \lambda \leq 1\}$

entonces;

$MP(\lambda\, z^1 + (1 - \lambda) z^2) \leq \lambda\, MP(z^1) + (1 - \lambda) MP(z^2)$

(ii) Si $MP(z^0)$ es finito en punto interior del dominio de $MP(z)$, entonces $MP(z)$ es finito en todo su dominio.

En efecto:

Por contradicción sea $MP(z^1) = -\infty \quad z^1 \in dom(MP)$ y como $z^0 \in \; dom(MP)$

$$\Rightarrow \lambda\, z^1 + (1 - \lambda) z^0 \in \; dom(MP)$$

$\Rightarrow z^2 = \lambda\, z^1 + (1 - \lambda) z^0$ para λ^0

Luego

$$z^0 = \frac{1}{1 - \lambda_0} z^2 - \frac{\lambda}{1 - \lambda_0} z^1$$

$$z^0 = \frac{\lambda_0}{\lambda_0 - 1} z^1 + \frac{1}{1 - \lambda_0} z^2$$

Definamos $\beta_0 = \frac{\lambda_0}{\lambda_0 - 1}$y por la parte (i) tenemos

$z^0 = \beta_0 z^1 + (1 - \beta_0) z^2 0 < \beta < 1$

$M(z^0) = -\infty \Rightarrow \Leftarrow$

$\therefore MP(z)$ es finito en todo su domino.

Teorema 7

Si (P) es un programa convexo superconsistente tal que $MP = MP(0)$ es finito, entonces $MP(z)$ es finito en todo su domino y existe un vector $\lambda \in \mathbb{R}^m$ tal que $\lambda \geq 0$ y $MP(z) \geq MP(0) - \lambda\, z$.

Demostración

Por el teorema (5) con z = 0 punto interior del *dom*(MP) se tiene $\exists \lambda \in \mathbb{R}^m$tal que

$$MP(z) \geq MP(0) - \lambda\, z \quad \forall z \in dom(MP)$$

Probemos ahora que $\lambda \geq 0$. Por contradicción: Supongamos que $\lambda_i < 0$ para algún i.

Por otro lado, desde que $0 \in int(domMP)$, entonces $\exists r > 0$

$$B(0,r) \subset dom(MP)$$

Tomemos el $z^i = \left(0, \dots 0, \frac{r}{2}, 0 \dots 0\right)$ vemos que $z^i \in B(0,r) \longrightarrow z^i \in dom(MP)$

$$MP\left(z^i\right) \geq MP(0) - \lambda\, z^i = MP(0) - \frac{r}{2} \qquad (*)$$

Pero como $z^i \geq 0$

y como $F(0) \subset F(z^1, \dots, z^m)$ entonces

$MP\left(z^i\right) \leq MP(0) = MP$

luego de (*) tenemos

$MP \geq MP\left(z^i\right) \geq MP - \frac{r}{2}\lambda_i$

$$0 \geq -\frac{r}{2}\lambda_i > 0 \Rightarrow\Leftarrow$$

Observación 1:

Si (*P*) es un programa convexo para el cual MP es finito y para el cual existe un $\lambda \geq 0$

$$MP(z) \geq MP(0) - \lambda\, z \quad \forall z \in dom(MP)$$

Entonces λ es llamado <u>vector sensible</u> para (*P*) y el teorema (7) simplemente garantiza que los programas convexos superconsitentes siempre tiene vectores sensibles.

Se debe tener en cuenta que si $MP(z)$ es diferenciable en $z = 0 \Rightarrow$ el vector λ (en *teorema 7*) puede ser tomada como $\nabla MP(0)$ ya que $MP(z)$ es una función convexa.

Teorema 8

Supongamos que

$$(P)\begin{cases} \min f(x) \\ s.a\ g_1(x) \leq 0, \dots, g_m(x) \leq 0, \\ donde\ x\ \in C \end{cases}$$

es un programa convexo para el cual existe un vector sensible λ,

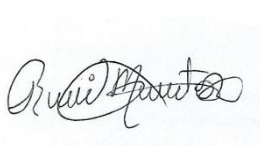
entonces

$$MP = \inf_{x \in C}\{f(x) + \sum_{i=1}^{m} \lambda_i\, g_i(x)\}$$

Demostración

Para cada z en el dominio de $MP(z)$, sabemos que existe λ (por hipotesis)

$$MP(z) \geq MP - \lambda\, z$$



En particular, si $x \in C$, entonces $z = \left(g_1(x), \dots, g_m^{(x)}\right) = g(x)$ está en el dominio de $MP(z)$.Asi

$$MP\big(g(x)\big) + \sum_{i=1}^{m} \lambda_i \, g_i(x) \geq MP$$

Luego, se sigue inmediatamente de la definición de $MP\big(g(x)\big)$ que

$$f(x) \geq MP(g(x))$$

consecuentemente;

$$f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x) \geq MP$$

para todo $x \in C$; Así mismo

$$\inf_{x \in C} \{f(x) + \sum_{i=1}^{m} \lambda_i \, g_i(x)\} \geq MP \qquad (i)$$

Por otro lado, debido a que $\lambda_i \geq 0 \quad \forall i$

$$\Rightarrow \inf \{f(x) + \sum_{\mathrm{i}=1}^{\mathrm{m}} \lambda_i \, g_i(x) : \mathrm{x} \in C\}$$

$$\leq \inf \{f(x) + \sum_{\mathrm{i}=1}^{\mathrm{m}} \lambda_i \, g_i(x) : x \in C, g(x) \leq 0\}$$

$$\leq \inf\{f(x) : x \in C, g(x) \leq 0\} = MP \qquad (ii)$$

De (i) y (ii)

$$\therefore MP = \inf_{x \in C} \{f(x) + \sum_{\mathrm{i}=1}^{\mathrm{m}} \lambda_i \, g_i(x)\}$$

Observación 2:

Si (*P*) es un programa convexo superconsistente tal que MP es finito entonces el *teorema 7* implica que existe un vector sensible λ para (*P*) y la tesis del *teorema 8* se cumple.

Definición 9

El Lagrangeano $L(x, \lambda)$ de un programa convexo

$$(P) \begin{cases} \min f(x) \\ s.a \; g_1(x) \leq 0, \dots, g_m(\mathrm{x}) \leq 0 \\ donde \; x \in C \end{cases}$$

es la función definida por:

$$L(x,\lambda) = f(x) + \sum_{\text{i}=1}^{\text{m}} \lambda_i \, g_i(x) \qquad \forall x \in C \quad \lambda \geq 0.$$

Note que si (*P*) es un programa convexo supersconsistente tal que MP es finito entonces existe el vector sensible λ y

$$MP = \inf_{x \in C} \{L(x,\lambda)\}$$

Teorema 10 (Karush – Kuhn – Tucker en Forma de punto silla)

Supongamos que (*P*) es un programa convexo superconsistente. Entonces $x^* \in C$ es una solución de (*P*) ⇔ existe en $\lambda^* \in \mathbb{R}^m$tal que:

(1) $\lambda^* \geq 0$,

(2) $L(x^*,\lambda) \leq L(x^*,\lambda^*) \leq L(x,\lambda^*) \quad \forall x \in C \; y \; \forall \lambda \geq 0$

(3) $\lambda_i^* \, g_i(x^*) = 0 \quad \forall i = 1,2, \dots, m$

Demostración

Supongamos que (*P*) es un programa convexo superconsistente. Entonces $x^* \in C$ es una solución de (*P*) ⇔ existe en $\lambda^* \in \mathbb{R}^m$tal que:

(1) $\lambda^* \geq 0$,

(2) $L(x^*,\lambda) \leq L(x^*,\lambda^*) \leq L(x,\lambda^*) \quad \forall x \in C \; y \; \forall \lambda \geq 0$

(3) $\lambda_i^* \, g_i(x^*) = 0 \quad \forall i = 1,2, \dots, m$

Demostración

Si x^* es solución de (*P*), entonces $x^* \in C, g(x^*) \leq 0 y \, f(x^*) = MP$ por (teorema 7), existe $\lambda^* \geq 0$ en $\mathbb{R}^m$

Consecuentemente por (teorema 8) tenemos

$$f(x^*) = \inf_{x \in C} \{L(x,\lambda^*)\} \tag{*}$$

También desde $\lambda^* \geq 0, g_i(x^*) \leq 0; \forall i = 1, \dots, m$ se sigue:

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*) \leq f(x^*)$$

Pero de (*)

$$f(x^*) \leq f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*)$$

entonces

$$f(x^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*) = L(x^*, \lambda^*)$$

y

$$\lambda_i^* \, g_i(x^*) = 0; \quad \forall i = 1, \dots, m$$

en conclusión

$$L(x^*, \lambda^*) \leq L(x, \lambda^*); \;\; \forall x \in C \;\; \lambda_i^* \, g_i(x^*) = 0, \;\; \forall i = 1, \dots, m.$$

Para probar la desigualdad de la izquierda de (2) note que para un $\lambda \geq 0$ en $\mathbb{R}^m$,

$$L(x^*, \lambda^*) - L(x^*, \lambda) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*) - f(x^*) - \sum_{i=1}^{m} \lambda_i \, g_i(x^*) = \; - \sum_{i=1}^{m} \lambda_i \, g_i(x^*) \geq 0$$

ya que $\lambda_i^* \, g_i(x^*) = 0; \lambda_i \geq 0; \; g_i(x^*) \leq 0 \; \forall i = 1, \dots, m$

entonces

$$L(x^*, \lambda^*) - L(x^*, \lambda) \geq 0; \quad \forall \; \lambda \geq 0 \; en \; \mathbb{R}^m$$

Esto completa la prueba de (1), (2) y (3) cuando x^* es solución de (*P*)

Probemos el recíproco:

Supongamos $x^* \in C$ y $\lambda^* \geq 0$ en $\mathbb{R}^m$ satisfaciendo (1, 2 y 3) y tomando en $\lambda^i \; \epsilon \; \mathbb{R}^n$ para algún $i, \; 1 \leq i \leq m$

$$\lambda_j^i = \begin{cases} \lambda_j^* & ; \quad si \quad j \neq i \\ \lambda_j^* + 1 & ; \quad si \quad j = 1 \end{cases}$$

Entonces $\lambda^i \geq 0$ y por (2) se tiene

$$0 \geq L\left(x^*, \lambda^{(i)}\right) - L(x^*, \lambda^*) = g_i(x^*)$$

Como i fue arbitrario, entonces es válido para todo i.

Luegox^* es un punto viable de (*P*)

Por otro lado, (2) implica que

$$f(x^*) = L(x^*, 0) \leq L(x^*, \lambda^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*)$$

pero $\lambda_i^* \geq 0, g_i(x^*) \leq 0 \quad \forall i = 1, \dots, m$así que

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*) \leq f(x^*)$$

concluimos que

$$f(x^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*) = L(x^*, \lambda^*)$$

y también que

$$\lambda_i^* \, g_i(x^*) = 0 \ \forall i = 1, \dots, m$$

Observación 3:

a) Si $A \subset B \Rightarrow \inf_{x \in B}\{h(x)\} \leq \inf_{x \in A}\{h(x)\}$

b) Si $h(x) \leq k(x) \forall x \in A \Rightarrow \inf_{x \in A} \{h(x)\} \leq \inf_{x \in A}\{k(x)\}$

Veamos que:

$$f(x^*) = L(x^*, \lambda^*) = \inf\{L(x, \lambda^*) : x \in C\}$$

$$\leq \inf\{L(x, \lambda^*) : x \in C, g(x) \leq 0\}$$

$$= \inf\left\{f(x) + \sum_{i=1}^{m} \lambda_i^* \, g_i(x^*) : x \in C, g_1(x) \leq 0, \dots, g_m(x) \leq 0\right\}$$

$$\leq \inf\{f(x) : x \in C, g_1(x) \leq 0, \dots, g_m(x) \leq 0\} = MP$$

además

$$MP \leq f(x^*)$$

por lo tanto

$$MP = f(x^*).$$

así que x^* es la solución de (*P*).

Asi un punto (x^*,λ^*) tal que $x \in C, \lambda^* \geq 0$ y satisface la desigualdad

$$L(x^*,\lambda) \leq L(x^*,\lambda^*) \leq L(x,\lambda^*) \ \forall x \ \in C, \lambda \geq 0$$

es llamado punto silla para el lagrangeano de (*P*) y también se cumple la condición (3)

$$\lambda_i^* \, g_i(x^*) = 0; i = 1,2, \dots, m$$

Esto muestra que si (x^*,λ^*) es un punto de silla de el langrangeano de un programa convexo (*P*); entonces:

(1) MP es finito y x^* es una solución de (*P*)

(2) Las condiciones de complementariedad

$$\lambda_i^* \, g_i(x^*) = 0; i = 1,2, \dots, m$$

es satisfecha.

Teorema 11 (de Karush – Kuhn – Tucker de forma gradiente)

Supongamos que (*P*) es un programa convexo superconsistente tal que la función objetivo $f(x)$ y las funciones restringidas $g_1(x), \dots, g_m(x)$ tienen primera derivada parcial continua en C para (*P*). Si x^* es viable para (*P*) y un punto interior de C, entonces x^* es una solución de (*P*) si y solo si existe un $\lambda^* \in \mathbb{R}^m$ tal que:

$(\mathbf{1})$ $\lambda_i^* \geq 0$ para $i = 1,2, \dots, m$

$(\mathbf{2})$ $\lambda_i^* \, g_i(x^*) = 0$ para $i = 1,2, \dots, m$;

$(\mathbf{3})$ $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \, \nabla g_i(x^*)$

Demostración

(⇒) Si x^* es solución de (*P*), entonces por (teorema 10) existe∃ λ^* en $\mathbb{R}^m$ tal que (1) y (2) se cumple

$$L(x^*,\lambda^*) \leq L(x,\lambda^*), \forall x \in C$$

así que x^* es un mínimo global para $h(x) = L(x,\lambda^*)$ en C desde que $x^* \in intC$ y $h \in C^1(C) \Rightarrow \nabla h(x^*) = 0$ esto satisface (3).

($\Leftarrow$) Supongamos que $x^* \in C$ y $\lambda^* \in \mathbb{R}^m$ satisfaciendo las condiciones (1),(2) y (3). Sea$x$ un punto viable para (*P*), entonces

$$\lambda^* g(x) \leq 0 \quad f(x) + \lambda^* g(x) \leq f(x) \tag{i}$$

por la convexidad de f tenemos

$$f(x) \geq f(x^*) + \nabla f(x^*)(x - x^*) \tag{ii}$$

y por convexidad de g tenemos

$$g(x) \geq g(x^*) + \nabla g(x^*)(x - x^*)$$

$$\lambda_i^* g_i(x) \geq \lambda_i^* g_i(x^*) + \lambda_i^* \nabla g_i(x^*)(x - x^*) \tag{iii}$$

Sumando (ii) y (iii) y usando (i)

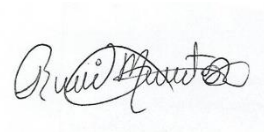
$$f(x) \geq f(x) + \sum \lambda_i^* g_i(x) \geq f(x^*) + \sum \lambda_i^* g_i(x^*) + \underbrace{[\nabla f(x^*) + \sum \lambda_i^* g_i(x^*)]}_{0} (x - x^*)$$

$$f(x) \geq f(x^*) + \sum \lambda_i^* g_i(x^*) \geq f(x^*)$$

Como x fue arbitrario

$$f(x) \geq f(x^*) \qquad \forall\, x \in F$$

Luego, x^*es un minimizador global



III. HIPOTESIS Y VARIABLES

3.1. Hipótesis

3.1.1 Hipótesis general

Dado el Problema de Programación Lineal

$$(P)\begin{cases} \text{maximizar } c^t\, x \\ \qquad\qquad\qquad \text{s.a} \\ \quad Ax = b \\ x \geq 0 \end{cases}$$

y una apropiada función barrera, entonces, se puede deducir el Sistema de Ecuaciones no lineales que definen a las trayectorias Solución del problema.

3.1.2 Hipótesis específica

Dado el Problema de Programación Lineal (P), su Problema de Barrera asociada

$$(P_u)\begin{cases} \quad \text{maximizar } c^t\, x + u \displaystyle\sum_i \ln x_i \\ \text{s.a} \\ Ax = b \\ x > 0 \end{cases}$$

Y el Sistema No Lineal de ecuaciones que definen a la trayectoria solución, entonces el conjunto de soluciones de dicho Sistema No lineal brinda las propiedades de las Trayectorias Solución.

3.2. Definición conceptual de las variables

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

VARIABLE DEPENDIENTE: Trayectorias Solución

VARIABLE INDEPENDIENTE: Sistema de ecuaciones no-lineales

3.3.OPERACIONALIZACIÓN DE LAS VARIABLES.

VARIABLE	DIMENSIONES	INDICADORES	INDICE	Métodos y Técnicas	TÉCNICA
Variable Independiente: Sistema de Ecuaciones No-lineales asociada al PPL	Espacio R^n	Comportamiento o variación de los parámetros del Problema de Programación Lineal	Problema Lineal	Análisis	Constructiva
Variable Dependiente: Trayectoria Solución	Espacio R^n	Comportamiento o variación del Conjunto Solución del Sistema de Ecuaciones No-lineal.	Problema Barrera	Análisis	Constructiva

IV.DISEÑO METODOLÓGICO

4.1. Tipo y diseño de la investigación

La investigación es del tipo 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, en incrementar los conocimientos científicos o filosóficos, pero sin contrastarlos con ningún aspecto práctico”

Es un estudio básico porque se buscará aportar conocimientos que permitan mejorar algunos detalles del marco teórico. El diseño es no experimental.

4.2. Método de Investigación

Los métodos utilizados en la investigación serán de:

Análisis y síntesis, pues se analizarán los elementos del problema en estudio para encontrar su relación entre sí. A su vez, la síntesis se producirá sobre la base de los resultados dados por el análisis. En este contexto Analizamos el conjunto de ecuaciones que definen a la trayectoria solución para establecer características de la misma en el espacio Rn.

Inductivo-Deductivo, pues utilizaremos la inducción como una forma de razonamiento por medio de la cual se pasa del conocimiento de casos particulares a un conocimiento más general. Asimismo, usaremos la deducción otra forma de razonamiento, mediante el cual pasamos de un conocimiento general a otro de menor nivel o particular. En el trabajo de investigación estamos constantemente utilizando este método a la hora de plantear y desarrollar los teoremas y lemas correspondientes.

4.3. Población y muestra

Por el tipo de investigación no utilizamos ninguna técnica de recolección de datos como la observación, experimentación o encuesta. Por lo tanto, no recolectamos datos de una muestra o población.

Población: No Aplica
Muestra: No Aplica

4.4. Lugar de Estudio.

Los ambientes de la Facultad de Ciencias Naturales y Matemática de la UNAC, precisando que gran parte del trabajo se desarrollará en el domicilio por la situación sanitaria que atraviesa el país.

4.5. Técnicas e instrumentos de Recolección de la Información

Por el tipo de investigación no utilizamos ninguna técnica de recolección de datos como la observación, experimentación o encuesta.

Por ser un trabajo netamente “matemático” (Teórico – abstracto); no se requiere procedimientos especiales para la recolección de la información. Lo que se realiza es una búsqueda y revisión bibliográfica: (libros de especialidad, páginas web, artículos, revistas especializadas, etc.), asimismo, a través de la utilización de técnicas de análisis – síntesis, inductivo-deductivo nos conducirán progresivamente a la resolución del problema planteado.

4.6. Análisis y procesamiento de datos.

Los análisis de la información sobre el problema de programación lineal determinaran las necesidades de conseguir información para su respectivo análisis y su incorporación en el estudio para la solución del problema barrera a través de la aplicación de la función barrera apropiada. Asimismo, se podra resolver el sistema de ecuaciones no lineales asosiada al problema en estudio.

4.7. Aspecto Etico de Investigación

De a cuerdo al Código de Ética de investigación de la UNAC en esta investigación se ha cumplido con la normativa de la Universidad. Es decir no se ha falsificado o inventado datos o resultados total o parcialmente, ni se ha plagiado datos, resultados , tablas cuadros de otros autores o investigadores

A continuación presentamos, observaciones, conceptos, proposiciones entre otros que nos permitirá establecer y obtener nuestros resultados

Trayectorias que conduce al Conjunto Optimal en Programación Lineal

1. Introducción

Los algoritmos iterativos para optimización no lineal usualmente asignan para algún punto $x \in S \subset \mathbb{R}^n$(en la mayoría de los casos convexo) un "próximo punto" $x' = f(x) \in S$.
Es decir, dado un punto inicial x°, el esquema iterativo genera una sucesión de puntos $\{x^k\}$, donde $x^{k+1} = f(x^k)$que converge a la solución.
Con frecuencia es instructivo considerar la versión "infinitesimal" de algoritmos iterativos en el siguiente sentido: Dado el esquema iterativo $x' = f(x)$se considera la ecuación diferencial.

$$\dot{x} = f(x) - x = x' - x$$

Cuando esta ecuación posee única solución a través de x° entonces éste determina una trayectoria (o curva) $x = x(t)$ tal que la tangente de la trayectoria en cualquier x es igual a la recta determinada por x y x'.
Si un algoritmo genera el punto x' "cerca" a x entonces la trayectoria puede considerarse como una buena aproximación a la secuencia generada por el algoritmo. Esto será cierto al menos durante las etapas posteriores de la ejecución del algoritmo si la secuencia generada converge a un punto de solución. En cambio, si el algoritmo da "pasos largos" durante las primeras etapas, entonces la trayectoria puede ser una mala aproximación.
En este trabajo se estudia las trayectorias solución relacionadas a la función barrera para el Problema de Programación Lineal y de acuerdo a la siguiente distribución. En la sección 2, describimos las trayectorias para problemas de programación lineal en forma estándar. En la sección 3, desarrollamos esencialmente la misma teoría dentro de un marco Primal-Dual o simétrico. En la sección 4, analizamos algunas propiedades de las tangentes de las trayectorias. En la sección 5, se trata algunas propiedades de diferenciabilidad y en la sección 6, se considera el comportamiento de las trayectorias cerca de los vértices.

2. Definiciones Previas

Usaremos aplicaciones de las funciones de barrera logarítmica, problemas con restricción Jacobiana, etc.

2.1 Las Condiciones de Karush Khunn Tucker (KKT

Consideremos el siguiente problema de Programación convexa.

$$(I)\left\{\begin{array}{c} \min f(x) \\ \text{S.A} \\ g_i(x) = 0 \ ; \ i = 1, \dots, n \\ x \in C \end{array}\right.$$

$donde\ f, g_i: C \to \mathbb{R}\ \ C\epsilon\mathbb{R}^n convexo\ \ f, g_i$ son todas convexas.
El problema (I) se le asocia la función Lagrangeana

$$L = (x, \lambda) = f(x) + \lambda_i g_i(x)\ de\ \mathbb{R}^n$$

Así mismo, se asume que el problema (I) es superconsistente, es decir

$$\exists\ x \in C\ /\ g_i(x) < 0 \qquad \forall i = \overline{1, \dots, n}$$

Además, los f, g_i poseen primeras derivadas parciales continuas, entonces $x^*\ \epsilon\ C$ es el mínimo si y solo si

$$i) \quad \exists\ \lambda_i^*\ \geq 0\ , \quad i = \overline{1, \dots, n}$$
$$ii) \quad \nabla_x\ L(x^*, \lambda^*) = 0$$
$$iii) \quad g_i\ (x^*) = 0 \qquad \forall i = \overline{1, n}$$
$$\lambda_i g_i(x^*) = 0 \qquad \forall i = \overline{1, n}$$

Método de Barrera

Se usan para transformar un problema de optimización con restricciones en otros sin ellas, donde fijan una "barrera" que evitará que los puntos generados se salgan de la región factible.

3.Problema de Barrera

Es el problema de optimización del tipo siguiente

$$\text{Minimizar}\ \ \theta(x, u) = f(x) + u\beta(x)$$
$$G(x) < 0$$

Los problemas con restricciones del tipo

$$g_i(x) = 0 \quad \forall i = \overline{1, \dots, n}$$

serán tratados tomando las restricciones

$$g_i(x) \leq (0 +\ \varepsilon) \wedge\ g_i(x) \geq (0 -\ \varepsilon)$$

donde:

$$G(x) = (g_1(x), g_2(x), \dots, g_n(x)),$$

es el conjunto de las restricciones del problema, $B(x)$ es la función de barrera y "u" el parámetro de barrera.

3.1 Las funciones de barrera se definen de la siguiente manera:

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

Donde $\emptyset\ (y)$ es una función continua y positiva para $y < 0$;

$$\lim_{y \to 0}(\emptyset\ (y)) = \infty$$

En la mayoría de los casos el punto óptimo se encuentra muy cercano a la frontera, tomando la función de barrera valores muy elevados. Aquí es donde aparece la utilidad del parámetro de barrera $(u > 0)$ que tendrá la función de reducir el valor de la barrera a medida que nos aproximemos a la frontera.

Función Barrera Logarítmica

Consideremos un problema de Programación Lineal en forma estándar.

$$(P)\begin{cases} \text{maximizar } c^t\,x \\ \qquad\qquad\text{s.a} \\ \quad Ax = b \\ x \geq 0 \end{cases}$$

Donde A $\epsilon\mathbb{R}^{mxn}$, $b\ \epsilon\ \mathbb{R}^m$, $c, x\ \epsilon\ \mathbb{R}^n$
La función barrera logarítmica, generalmente usado en optimización con restricciones no lineales, pero puede ser aplicado al problema de Programación Lineal, dando lugar al siguiente problema.

$$(P_u)\begin{cases} \quad \text{maximizar } c^t\,x + u\displaystyle\sum_i \ln x_i \\ \text{s.a} \\ Ax = b \\ x > 0 \end{cases}$$

Donde $u > 0$ suele ser pequeño. El enfoque de la función barrera es válido si existe un $x\ > 0$ sujeto a $Ax = b$
D_x es la matriz diagonal de orden $d \times d$ cuya diagonal son las componentes de x.

$$D_x = \begin{bmatrix} x_1 & & 0 \\ & x_2 & \\ 0 & & x_n \end{bmatrix}$$

Un vector $x\ > 0$ es una solución óptima para (P_u)si y solamente si existe un vector $y\ \epsilon\ \mathbb{R}^n$

$$(P_u)\begin{cases} \text{maximizar } c^t\,x + u\displaystyle\sum_i \ln x_i = f(x,u) \\ \qquad\text{s.a} \\ \qquad Ax = b \\ \qquad x > 0 \end{cases}$$

$$(P_u)\begin{cases} \qquad \min h(x,u) = -f(x,u) \\ \text{s.a} \\ Ax = b \\ x > 0 \end{cases}$$

Para $u > 0$, el problema (P_u)tiene única solución óptima $\chi(u)$, entonces (P_u) tiene una única solución óptima $\forall u > 0$ ya que su función objetivo es estrictamente cóncava.
Hemos estudiado para $-f\ (x,u)$ es convexa, para ello hacemos que es mínimo y utilizando el teorema de condición de K – K – T.
Consideremos el siguiente problema de Programación convexa.

$$(P_u)\begin{cases} \qquad \min h(x,u) = -f(x,u) \\ \text{s.a} \\ \quad g_i\,(x) = 0 \quad , \quad i = \overline{1,n} \\ x\ \epsilon\ C \end{cases}$$

Donde $f, g_i : C \longrightarrow \mathbb{R}$ / $C \subset \mathbb{R}^n$convexo y f , g_i son todos convexos.

Se le asocia la función Lagrangeana correspondiente que es una función de las variables del problema (*) y de un nuevo conjunto de variables, una para cada restricción llamadas multiplicadores de Lagrange.

$$L = (x\,,\lambda) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) \,/\, \lambda \,\epsilon \mathbb{R}^m$$

(P_u) es superconsistente, entonces

$$\exists\, x \,\epsilon\, C \;/\; g_i(x) < 0 \qquad \forall i = \overline{1,n}\;,$$

además, los f, g_i poseen primeras derivadas parciales continuas entonces $x^* \in C$ es mínimo (P_u)si solamente si:

$$\begin{aligned} &i) \quad \exists\, \lambda_i^* \geq 0 \\ &ii) \quad \nabla x\, L(x^*, \lambda^*) = 0 \\ &ii) \quad g_i x^* = 0 \quad \wedge \quad \lambda_i g_i(x^*) = 0 \quad \forall i = \overline{1,n} \end{aligned}$$

$$(P_u)\begin{cases} \min h(x,u) = \\ \text{s. a} \\ \qquad Ax = b \\ x > 0 \end{cases} \qquad \begin{cases} -c^t\, x - u \displaystyle\sum_i \ln xi \\ \text{s. a} \\ \qquad Ax = b \\ x > 0 \end{cases}$$

Donde $\varphi(x) = \sum_j \ln x_j$; es la suma de funciones estrictamente convexa.

Además, el parámetro $u > 0$ entonces $u\,\varphi(x)$ sigue siendo estrictamente convexa.

Además, $-C^t x$ es una función lineal entonces, es convexa.

Donde $h(x\,,u)$ es estrictamente convexa tal que $h > 0$

Por lo tanto, $x(u)$ es único mínimo global.

Si el (P_u)posee solución óptima, $x(u)$, entonces le podemos aplicar el teorema de las condiciones de K-K-T.

Veamos el Lagrangeana de (P_u)

$$L = (x, \lambda) = -C^t x - u\,\varphi(x) - (y,z)^t \begin{bmatrix} b & -Ax \\ & -x \end{bmatrix}$$

Como estamosminimizando

Funciones matriciales

$$L = (x, \lambda) = -C^t x - u\,\varphi(x) - y^t(\underbrace{b - Ax}_{\text{g1}}) - \underbrace{z^t(-x)}_{\text{g2}}$$

(P_u) deberá satisfacer:

$$\begin{aligned} &i) \quad \exists\, (y,z) \geq 0 \\ &ii) \quad \nabla x\, L(x,y,z) = 0 \;\leftrightarrow -\,C^t \;\;;\;\; u\left(\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n}\right) + y^t A + Z^t = 0 \end{aligned}$$

Aplicando transpuesta

$$\longrightarrow\; -c - u \begin{pmatrix} \frac{1}{x_1} & & & \\ & \frac{1}{x_2} & & \\ & & \ddots & \\ & & & \frac{1}{x_n} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + \; A^t y + z = 0$$

$$-c - u\, x^{-1} e + A^t y + z = 0 \qquad donde\ \chi = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$$

$iii) \quad Ax = b$

$$x > 0$$

$\rightarrow z^t x = 0 \qquad\qquad se\ concluye\ z = 0$

Por lo tanto

$i) \quad y \geq 0$

$$\left.\begin{aligned} &ii)\ -c - u\, x^{-1} e + A^t y = 0 \ \rightarrow\ -c - u\, x^{-1} e + A^t \\ &\qquad\qquad\qquad\qquad Ax = b \end{aligned}\right] \qquad (*)$$

Definimos el siguiente Sistema como “Sistema (0)” a partir de (*):

$$\longrightarrow (O) \ldots\ldots\ldots\ldots \begin{cases} u\, x^{-1} e + A^t y = -c \\ \qquad Ax = b \end{cases}$$

El sistema (0) tiene una única solución óptima en x, $\forall\, u > 0$
La matriz Jacobiana de (0) es el siguiente:

$$J = J_u(x, y) = \begin{pmatrix} -ux^{-2} & -A^t \\ A & 0 \end{pmatrix}$$

Supongamos que A es de rango completo m $(m \leq n)$

El rango de una matriz es el número máximo de columnas (filas respectivamente) que son sea linealmente independiente (l.i).
Sea $A \epsilon\, M_{mxn}$ $\rightarrow$ a) Si $m \neq n \wedge rang\ (A) = \min\{m, n\}$
diremos que la matriz Aes una matriz de rango completo
$\longrightarrow rg\ (A) = n$

En este caso, el valor de y se determina de forma única por el valor de x.
También se cumple en este caso que la matriz $A\, A^t$ es definida positiva y por lo tanto no singular.

$Si\ rg(A) = n$ diremos que A es una matriz no singular o regular
Si $rg(A) < n$ diremos que A es una matriz singular.

El sistema de ecuaciones lineales

$$J_u(x, y) = \begin{pmatrix} u \\ 0 \end{pmatrix}$$

Puede ser interpretado como un problema de “mínimo cuadrado” o un problema de proyección

Mínimos cuadrados es una técnica de análisis numérico enmarcada dentro de la optimización matemática en la que dados un conjunto de pares ordenados: variables independientes, variable dependiente y una familia de funciones, se intenta encontrar la función continua, dentro de dicha familia, que mejor se aproxime a los datos (un mejor ajuste).

Buscamos minimizarla la expresión.

<ins>Proposición 1</ins>:
El problema (P_u)es ilimitado para todo $u > 0$ o tiene una solución óptima para todo $u > 0$

Prueba
Dado el conjunto

$$L(t) = \{x \geq 0;\ c^t x = t\ , Ax = b\}$$

Y por hipótesis tiene interior no vacío$(\mathcal{F}_0), \forall t\ \epsilon I\ (abierto)$, entonces

Primero supongamos que la función $\emptyset(x) = \sum_j \ln x_j$ es no acotado en $L(t)$

$$\longrightarrow (P_u) \text{ es no acotado } \forall\, u > 0 \text{ (valores positivos)}$$

Supongamos ahora que $\emptyset(x)$ es acotado en $L(t), t \in I$ (abierto)

$$\longrightarrow \emptyset(x) = \sum_j \ln x_j \text{ es acotado inferiormente}$$

Luego, $g(x) \leq \emptyset(x)$

Donde $g(t)$ denota el máximo valor de $\emptyset(x)$ sobre $L(t)$

$$\left[\begin{array}{c} Si\ \emptyset(x) \text{ es estrictamente cóncava, entonces hay un único máximo} \\ x = x(t)\ \text{ ; para cada } t\epsilon I \end{array}\right]$$

<ins>Consideremos en primer lugar:</ins> $\infty\, \epsilon\, I$, es decir la función la función $C^t x$ es acotada inferiormente. Además, existe un rayo contenido en el interior de la región factible, entonces $C^t x \longrightarrow \infty$

Por lo tanto, como el dominio es poliédrico, el rayo de la función $\emptyset(x)$ es acotada inferiormente.

Por lo tanto (P_u) es no acotado $\forall\, u > 0$.

Como, el sistema (O) tiene solución para algún $u > 0$, entonces se determina un camino único y continuó $x = x(u)$.

Asimismo, cuando A posee rango completo, también determina un camino continuo

$$y = y(u)$$

Como u⟶0, vamos a suponer que

$$\lim_{u \to 0} x(u) = \bar{x}$$
$$\lim_{u \to 0} y(u) = \bar{y}$$

Existen y lo denotamos por $\bar{x}$ y $\bar{y}$

Se deduce que A$\bar{x} = b$,$\bar{x} \geq 0 \wedge A^t \, \bar{y} \geq C$

Por otra parte, para cada "j" $\bar{x}_j \geq 0$, $A_j^t \bar{y} \geq C_j$

Resulta que $\bar{x}$ y $\bar{y}$ son soluciones óptimas para (P)y (D)vamos a modificar la función objetivo de manera que se disponga de una solución inicial. Tenga en cuenta que el lugar de (P_u) podemos trabajar con un problema de la forma.

$$(P_u)\begin{cases} \text{maximizar } c^t x + u \sum_i w_i \ln x_i \\ \text{s. a} \\ \qquad Ax = b \\ x > 0 \end{cases} \longrightarrow (P_u)\begin{cases} \text{minimizar} -c^t x - u \sum_i w_i \ln x_i \\ \text{s. a} \\ \qquad Ax = b \\ x > 0 \end{cases}$$

Análogamente al caso anterior lo vamos a desarrollar

Veamos el Lagrangeano de $(P_u\,(w))$

$$L(x, \lambda) = -C^t x - u\, \varphi(x).w - (y)^t (b - Ax)$$

De $(P_u\,(w))$ deberá satisfacer:

$i)$ $\quad \exists\, y \geq 0$

$ii)$ $\quad \nabla x\, L(x, y) = 0 \leftrightarrow -c^t - u\left(\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n}\right) + y^t A = 0$

aplicando transpuesta

$$\longrightarrow -c - u \begin{pmatrix} \frac{1}{x_1} & & \\ & \ddots & \\ & & \frac{1}{x_n} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} + A^t y = 0$$

$$-c - u\, X^{-1} w + A^t y = 0 \qquad donde\ X = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$$

$iii)$ $\quad Ax = b$

$\qquad x > 0$

Por lo tanto

$i)$ $\quad y \geq 0$

$ii)$ $-c - u\, X^{-1} w + A^t y = 0$

Entonces:

$$(P_u\,(w))\begin{cases} u x^{-1} w + A^t y = -\mathrm{c} \\ \qquad Ax = b \end{cases}$$

Específicamente, $w = D_{x^{\circ}}(A^t\, y^0 - 0)$ ya que $D_x{}^{-1} = X^{-1}$

Por lo tanto, dado un par de puntos factibles en el interior serán soluciones del problema Primal y Dual debido a lo que anunciamos anteriormente. Entonces, para $u > 0$ el sistema (0) tiene única solución$(x(u), y(u))$.

Por otro lado, se desprende que, si multiplicamos la primera fila del sistema (0) por $(x(u))$, la segunda por $(y(u))$ Y luego sumamos

$$\begin{array}{l} u\,x^{-1}.x(u) - A^t y.x(u) = -c^t\,x(u) \\ \qquad\qquad Ax.y(t) = b.y(t) \\ \hline by(u) - c^t x(u) = \underbrace{u\,x^{-1}.x(u) - A^t y.x(u) + Ax.y(u)} \end{array}$$

$$by(u) - c^t x(u) = \eta u$$

El valor óptimo se encuentra entre $by(u)$ y $c^t x(u)$

Entonces $c^t x(u)$ tiende al valor óptimo a medida que $u \longrightarrow 0$

Sea $v(u) = C^t x(u)$ donde $x(u)$ es la solución óptima de (Pu), y sea $V(o)$ que denota al valor óptimo de (P).

Acabamos de probar que

$$V(u) \longrightarrow V(o)$$

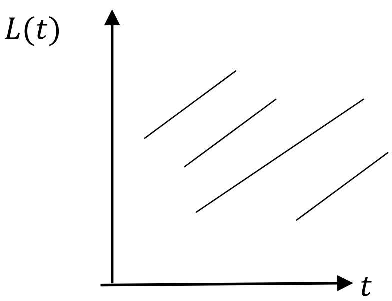
A medida que

$$u \longrightarrow 0$$

Obviamente $x(u)$ es también solución del siguiente problema:

$$(\bar{P}_u)\begin{cases} Minimizar - C^t x - uP(x) \\ s.a. \\ Ax = b \\ \qquad C^t x(u) = V(u) \\ x > 0 \end{cases}$$

Ahora si $L(t)$ es no acotada, existirá un rayo "acotado inferiormente" pues se aleja de la frontera, a lo largo de una de las variables que tiende al infinito mientras que los otros acotados están lejos de cero.



El conjunto $L(V(0))$ es acotado, entonces $L(V(0)) \leq g(t)$ (es una función cóncava t)

Esta concavidad implica que $g(t)$es acotada superiormente

$$t \longrightarrow V(0)$$

Sea N el conjunto de todos los índices j sujeto a $x_j = 0$ en cada solución óptima

Sea

$$\emptyset_N(x) = \sum_{j \in \mathbb{N}} \ln x_j$$

y

$$\emptyset_B(x) = \emptyset(x) - \emptyset_N(x)$$

Sea $\xi_j(u)$ la j – ésima componente del vector $x(u)$, es decir de la solución óptima en u.

Afirmamos que $x(0)$ es igual al límite de $x(u)$ a medida que u tiende a cero

$$\longrightarrow \lim_{u \to 0} x(u) = x(0)$$

Esta solución se caracteriza por el siguiente sistema

$$\frac{1}{x_i} - A_i^t y = -\lambda C_i \quad (j \notin \mathbb{N})$$

$$Ax = b$$

$$x_i = 0 \quad (j \in \mathbb{N})$$

$$C^t x = V(0)$$

Donde λ es un multiplicador correspondiente a la ecuación

$$C^t x = V(0)$$

y A_i^t es la j-esima fila de A^t., entonces

$$\lim_{u \to 0} Xk(u) = 0 \qquad para\ todo\ \ k = \overline{1,n}$$

Satisfase el sistema de ecuaciones e implica por lo tanto

$$X(u) \to X(0)$$

$$u \to 0$$

Proposición 2:

Si para algún $u > 0$ el sistema (0) tiene una solución $x > 0$ entonces para todo los $u > 0$ hay una solución $x(u)$ de modo que la trayectoria $x(u)$ es continua y el límite de $x(u)$ mientras u tiende a cero, existe y converge a una solución óptima al problema de programación lineal (P).

Podemos resolver el problema de programación lineal a partir de lo siguiente:

Si $u = 1$, hemos observando anteriormente que es solución óptima de $(P_u(w))$

Si $0 < u < 1$, entonces

$$\lim_{u \to 0} X(u) = X(0)$$

Es una solución óptima del problema programación lineal.

4 Dualidad:

Consideremos el problema de programación lineal en forma simétrica (En sentido de la forma de la trayectoria dual)

$$(P)\begin{cases} máx\ c^t x \\ s.a. Ax \leq b \\ x \geq 0 \end{cases}$$

$$\begin{aligned} donde&: A \in R^{mn} \\ b &\in R^m \\ c, x &\in R^n \end{aligned}$$

Realizamos la siguiente sustitución de $\{Ax \leq b\}$ por $\{Ax + u = b; u \geq 0\}$. El cual define la región factible primal de

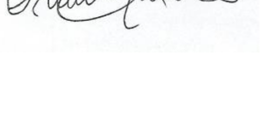
$$(P^*)\begin{cases} máx\ c^t x \\ s.a. Ax + u = b \\ x, u \geq 0 \end{cases}$$

Por la teoría de los métodos de función barrera, podemos transformar un problema con restricciones a una familia de problemas sin restricciones, definiendo:

$$f(x) = c^t x + \mu\left(\sum_i Lnx_j + \sum_i Lnu_i\right)$$

Una función barrera logarítmica de parámetro $\mu > 0$. Entonces (P) se puede aproximar por:

$$(P_u)\begin{cases} máx \quad c^t x + \mu\left(\sum_i Lnx_j + \sum_i Lnu_i\right) \\ s.a: Ax + u = b \\ x, u > 0 \end{cases}$$



$$(P_u)\begin{cases} máx \quad c^t x + \mu\left(\sum_{j}^{n} Lnx_j + \sum_{i}^{m} Lnu_i\right) \\ s.a: Ax + u = b \\ x, u > 0 \end{cases}$$

Y considerando

$$D_x = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$$

$$D_x^{-1} = \begin{pmatrix} \frac{1}{x_1} & & \\ & \ddots & \\ & & \frac{1}{x_n} \end{pmatrix}$$

(P_u) es un problema de optimización cóncava no lineal (donde μ es un número real positivo fijo).

Notar que el gradiente de la función $\emptyset(x) = \sum_i^n Lnx_j$ es igual a $D_x^{-1}e$ y también a $D_x^{-2}x$.

En efecto:

$$\nabla\emptyset(x) = \underbrace{\left(\frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_3}, \dots, \frac{1}{x_n}\right)}_{[D_x^{-1}e]^T}$$

- $D_x^{-2}.x = (D_x^{-1})^2.x$

$$D_x^{-1} = \begin{pmatrix} \frac{1}{x_1^2} & & \\ & \ddots & \\ & & \frac{1}{x_n^2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} \frac{1}{x_1} & & \\ & \ddots & \\ & & \frac{1}{x_n} \end{pmatrix} = D_x^{-1}e$$

$$\boxed{\therefore\ D_x^{-2}.x = D_x^{-1}e}$$

Consideremos un par de vectores $x \in \mathbb{R}_+^n$, y $u \in \mathbb{R}_+^m$, de tal forma que $Ax + u = b$, constituye la solución óptima para (P_u) si y solo si existe un vector $y \in \mathbb{R}^m$ tal que

$$uD_x^{-2}x - A^T y = -c$$

$$uD_u^{-1}e - y = 0$$

En efecto: tomemos (P_u)

$$\text{Max } c^t x + \mu\left(\textstyle\sum_j Lnx_j + \sum_i Lnu_i\right)$$

$$\text{s.a: } Ax + u = b$$

$$x, u > 0$$

Que es equivalente a:

$$\min \quad -c^t x - \mu\left(\sum_j Lnx_j + \sum_i Lnu_i\right)$$

$$\text{s.a: } Ax + u = b$$

$$x, u > 0$$

Tomando el Lagrangeno asociado al problema equivalente de (P_u) :

$$L(x,y,u) = -c^t x - \mu\left(\sum_j Ln\, x_j + \sum_i Ln\, u_i\right) - y^t(b - Ax - u)$$

Aplicando $K - K - T$

i) $\nabla x L(x,y,u) = 0$

$$\nabla x\left(-c^t x - \mu\left(\sum_j Lnx_J + \sum_i Lnu_i\right) - y^t b + y^t Ax + y^t u\right) = 0$$

$$-c^t - \mu\left(\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n}\right) + y^t A = 0$$

$$-c^t - \mu[D_x^{-2}.x]^T + y^t A = 0$$

$$-c - uD_x^{-2}x + A^t y = 0$$

$$\boxed{=> \mu D_x^{-2}x - A^t y = -c} \ldots\ldots(\beta)$$

$\nabla u L(x,y,u) = 0$

$$\nabla u\left(-c^t x - \mu\left(\sum_j Ln\, x_j + \sum_i Ln\, u_i\right) - y^t(b - Ax - u)\right) = 0$$

$$\nabla u\left(-c^t x - \mu\left(\sum_j Ln\, x_j + \sum_i Ln\, u_i\right) - y^t b + y^t Ax + y^t u\right) = 0$$

$$-\mu\left(\frac{1}{u_1},\frac{1}{u_2},\dots,\frac{1}{u_n}\right)+y^t=0$$

$$-\mu(D_u^{-2}.u)^T+y^T=0$$

$$\boxed{\mu(D_u^{-2}u)-y=0} \quad \dots(\propto)$$

Se deduce que un vector debe satisfacer $u=uD_y^{-1}e$: y por tanto x es óptimo en (P_u) si y sólo si existe $y\in\mathbb{R}_+^m$ tanto

$$(0)\qquad \begin{aligned} &\mu D_x^{-2}x-A^ty=-c\\ &Ax+\mu D_u^{-2}-y=b \end{aligned}$$

De $(\propto)$ podemos ver que:

$$\mu\begin{pmatrix}\frac{1}{u_1^2} & & \\ & \ddots & \\ & & \frac{1}{u_n^2}\end{pmatrix}\begin{pmatrix}u_1\\u_2\\ \cdot\\ \cdot\\ u_n\end{pmatrix}-\begin{pmatrix}y_1\\y_2\\ \cdot\\ \cdot\\ y_m\end{pmatrix}=0$$

$\Longrightarrow \mu.\frac{1}{u_i}=y_i \qquad \forall\ i=\overline{\imath,m}$

$\Longrightarrow \mu=u_i.y_i$

De manera general:

$$\mu D_y^{-1}=u$$

$$\mu D_y^{-2}y=u$$

El sistema (O) tiene algunas propiedades de simetría, consideremos el dual de (P), donde:

$$(D)\qquad \min b^ty$$

$$\text{s.a: } A^ty\geq 0$$

$$y\geq 0$$

Hacemos la siguiente sustitución $\{A^t y \geq 0, y \geq 0\}$ por $\{A^t y - v = c; v \geq 0\}$ y agregándole una función barrera logarítmica de parámetro $\mu > 0$

$$(D_u)\qquad \begin{aligned} &\min b^t y - \mu\left(\sum_i Ln\, y_i + \sum_j Lnv_i\right) \\ &\text{s.a: } A^t y - v = c \\ &y, v > 0 \end{aligned}$$

Veamos la verificación de las igualdades en (O).

Tomemos el Lagrangeano asociado a (D_u)

$$L(x, y, v) = b^t y - \mu\left(\sum_i Lny_i + \sum_j Lnv_j\right) + x^t(c - A^t y + v)$$

Aplicando $K - K - T$

$i)$ $\nabla yL(x, y, v) = 0$

$$\nabla y\left(b^t y - \mu\left(\sum_i Lny_i + \sum_j Lnv_j\right) + x^T c - x^T A^T y + x^T v\right) = 0$$

$$b^T - \mu\left(D_y^{-2} y\right)^T - x^T A^T = 0$$

$$b - \mu D_y^{-2} y - Ax = 0$$

$$\boxed{\Longrightarrow \ \mu D_y^{-2} y + Ax = b}$$

Es fácil comprobar que “y” es óptimo en (D_u) si y solo si existe un $x \in \mathbb{R}_+^n$ tal que el sistema (0) sea válido.

Tengamos en cuenta que las funciones objetivas lineales de (P_u) y (D_u) tiene como máximo una solución óptima.

La relación entre los problemas (P_u) y (D_u) se resumen en el siguiente teorema.

Proposición 3:

1. Si el problema (P_u) es no acotado entonces el problema (D_u) es no factible, si el problema (D_u) es no acotado entonces el problema (P_u) es no factible.
2. El problema (P_u) tiene solución optima si y solo si el problema (D_u) tiene solución óptima
3. Si $x\,, y$ son soluciones óptimas para (P_u) y (D_u) respectivamente, entonces la diferencia entre los valores de las funciones objetivos de (P_u) y (D_u) es igual a $(m+n)u(1+Lnu)$, mientras que la diferencia entre $c^t x$ y $b^t y$ es igual a $(m+n)u$

Demostración:

1. Del problema primal tenemos (P_u)

   $Ax + u = b$

   $(Ax)^T + u^T = b^T$

   multiplicando por la derecha por y:

   $(Ax)^T y + u^T y = b^T y$

   $y^T(Ax) + u^T y = b^T y \;\ldots\; (\propto)$

- Del problema dual tenemos (D_u)

  $A^T y - v = c$

  $(A^T y)^T - v^T = c^T$

  multiplicando por la derecha por x:

  $(A^T y)^T x - v^T x = c^T x$

  $y^T Ax - v^T x = c^T x \ldots\; (\beta)$

Restando $(\propto)$ y (β), teniendo en cuenta que para $\mu \leq 1$

$$b^T y - c^T x = u^T y + v^T x \geq \mu(u^T y + v^T x)$$

$$= \mu\left(\sum_{i=1}^{m} u_i y_i + \sum_{j=1}^{n} v_j x_j\right)$$

$$b^t y - c^t x > \mu\left(\sum_{i=1}^{m} Lnu_i y_i + \sum_{j=1}^{n} Lnv_j x_j\right)$$

$$b^t y - c^t x > \mu\left(\sum_{i=1}^{m} Lnu_i + \sum_{i=1}^{m} y_i + \sum_{j=1}^{n} Lnv_j + \sum_{j=1}^{n} Lnx_j\right)$$

$$b^t y - \mu\left(\sum_{i=1}^{m} Lny_i + \sum_{j=1}^{n} Lnv_j\right) > c^t x + \mu\left(\sum_{j} Lnx_j + \sum_{i} Lnu_i\right)$$

Esto implica que si (P_u) es ilimitado entonces(D_u) es no factible, y si (D_u) es ilimitado entonces (P_u) es no factible.

Sabemos por lo anterior del sistema (0) que (P_u) tiene una solución óptima si y solo sí (D_u) tiene una única solución. Las soluciones optimas son únicas. Si $x\ e\ y$ son soluciones optimas para (P_u) y (D_u) respectivamente, implica que:

$$u_i = \mu.\frac{1}{y_i} \qquad \wedge \qquad v_i = \mu.\frac{1}{x_j}$$

$$u_i.y_i = \mu \qquad \wedge \qquad v_i.x_j = \mu$$

$$\Longrightarrow\ v_j.x_j = u_i.y_i = \mu$$

Además

$$b^T y - c^T x = u^T y + v^T x$$

$$b^T y - c^T x = \sum_{i=1}^{m} u_i y_i + \sum_{j=1}^{n} v_j x_j$$

$$b^T y - c^T x = \sum_{i=1}^{m} \mu + \sum_{j=1}^{n} \mu$$

$$b^T y - c^T x = m\mu + n\mu$$

$$\boxed{b^T y - c^T x = (m+n)\mu}$$

Ahora:

$$\left[b^T y - \mu\left(\sum_i Lny_i + \sum_j Lnv_j\right)\right] - \left[c^T x + \mu\left(\sum_j Lnx_j + \sum_i Lnu_i\right)\right]$$

$$= b^T y - c^T x - \mu\left(\sum_i^m Lny_i + \sum_j^n Lnv_j + \sum_j^n Lnx_j + \sum_i^m Lnu_i\right)$$

$$= b^T y - c^T x - \mu\left(\sum_i^m (Lny_i + Lnu_i) + \sum_j^n (Lnv_j + Lnx_j)\right)$$

$$= b^T y - c^T x - \mu\left(\sum_i^m Lny_i u_i + \sum_j^n Lnv_j x_j\right)$$

$$b^T y - c^T x - \mu\left(\sum_i^m Ln\mu + \sum_j^n Ln\mu\right)$$

$$= \underbrace{b^T y - c^T x}_{} - \mu(mLn\mu + nLn\mu)$$

$$= (m+n)\mu - \mu(m+n)Ln\mu$$

$$= (m+n)(1 - Ln\mu)\mu$$

Esto prueba que, la brecha entre los valore óptimos de (P_u) y (D_u) depende solo de las dimensiones m y n, mas no de los datos. Se deduce entonces de la Proposición 3 que las soluciones optimas $x = x(u) \wedge y = y(u)$ son tales que $c^t x(u) \wedge b^t y(u)$ tiendan al valor optimo de (P) (el cual, por supuesto es igual al valor optimo de (D)) a medida que $\mu \rightarrow 0$

Para el enfoque de barrera primal - dual simétrica, en el trabajo necesitamos tanto (P) y (D) para tener dominios factibles del plano dimensional. Recordemos que todos los problemas de programación lineal puede ser reformulados de modo que tanto el primal y dual tengan dominios factibles de plano dimensional. Dado un problema en la forma (P), consideremos el siguiente problema, donde M es suficientemente grande:

Maximizar $c^T \propto -M\xi$

(P^*) s.a:

$$Ax + (b - Ae - e)\xi \leq b$$

$$(c - A^T e + e)^T x \leq M$$

$$x, \xi \geq 0$$

Es fácil verificar que si M es suficientemente grande, entonces x es una solución óptima para (P) si y solo si $(x, 0)$ es una solución optima para (P^*). El punto $e \in \mathbb{R}^{n+1}$ se encuentra en el interior del domino feasible (P^*).

Además, el punto $e \in \mathbb{R}^{m+n}$ se encuentra en el interior del dominio factible del dual (P^*)

$$\min b^T y + M\eta$$

(D^*) s.a: $A^T y + (c - A^T e + e)\eta \geq c$

$$(b - Ae - e)^T y \geq -M$$

$$y, \eta \geq 0$$

Esta estrategia conocida como el “Big M” es justamente usada en Programación Lineal. Alternativamente, para evitar problemas numéricos con valores grandes de “M", se puede usar la estrategia equivalente conocida como “Fase I” en Programación Lineal.

Para simplificar la notación podemos escribir (x, y) para el vector de la columna obtenida mediante la carectirización de dos vectores columna x e y . Esto será interesante al considerar el mapeto: $\Psi: \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$, definido por:

$$\Psi(x, y) = \left(uD_x^{-2}x - A^T y; Ax + uD_y^{-2}y\right)$$

Este mapeo se basa en el sistema (0), el cual puede ser escrito como $\Psi(x,y) = (-c,b)$. La matriz jacobiana de Ψ en (x,y) es igual a

$$H = \nabla\Psi(x,y) = \begin{pmatrix} \dfrac{\partial\,\Psi(x,y)}{\partial x} & \dfrac{\partial\,\Psi(x,y)}{\partial y} \\ \dfrac{\partial\,\Psi_2(x,y)}{\partial x} & \dfrac{\partial\,\Psi_2(x,y)}{\partial y} \end{pmatrix}$$

$$H = \nabla\Psi(x,y) = \begin{pmatrix} -\mu D_x^{-2} & -A^t \\ A & -\mu D_y^{-2} \end{pmatrix}$$

Asumimos que $x \wedge y$ son positivos, entonces la matriz H es definida negativa pues para algún $w \in \mathbb{R}^n$ y $Z \in \mathbb{R}^n$

$$(z,w)^T H(z,w) = (z^T, w^T) \begin{pmatrix} -\mu D_x^{-2} & -A^T \\ A & -\mu D_y^{-2} \end{pmatrix} \binom{z}{w}$$

$$\left(-z^T \mu D_x^{-2} + w^T A, -z^T A^T - w^T \mu D_y^{-2}\right) \binom{z}{w}$$

$$(z,w)^T H(z,w) = \left(-z^T \mu D_x^{-2} z + w^T A z - z^T A^T w - w^T \mu D_y^{-2} w\right)$$
$$-u\left(z^T \mu D_x^{-2} z + w^T \mu D_y^{-2} w\right)$$

En particular H es no singular. Ahora, también consideremos una matriz simétrica relacionada:

$$\widetilde{H} = \begin{pmatrix} -\mu D_x^{-2} & -A^t \\ A & -\mu D_y^{-2} \end{pmatrix}$$

Obiamente, $\widetilde{H}$ es la matriz Hessiana de la función

$$L_\mu(x,y) = c^T x + \mu \sum_j Lnx_j - y^T A x - \mu \sum_i Lny_i + y^T b$$

El cual esta bien definido para $x, y > 0$, y notemos que L es estrictamente cóncava en x para cualquier y , y es estrictamente convexa en y para cualquier x.
El par $\left(x_{(\mu)}, y_{(\mu)}\right)$ (es decir el punto donde la gradiente de $L_\mu(x,y)$ se anula) constituye el unico punto silla de $L(x,y)$, en el sentido que x es el máximo e "y" es el mínimo.

La sumatoria de logaritmos añadido a la función objetivo lineal $c^t x$ cumple el rol de “barrera”.

Supongase que se tiene un algoritmo para optimización irrestricta y comienza en el interior de la región factible e itera buscando en linea recta a partir del punto actual. La barrera “evita” que los puntos iterados se acerquen a la frontera es decir hace que permanezcan en el interior de la región factible.

Otra clasica estrategia usada en Programación No Lineal es el uso de funciones “Penalidad” (donde una penalidad es impuesta si un punto iterado cae fuera de la región factible)

Aquí vamos a considerar, algoritmos generales que iteran sobre puntos interiores del primal y del dual respectivamente. Vamos a denotar por μ un parámetro que determina en el interior factible del primal y del dual una solución $x(\mu) > 0 \wedge y(\mu) > 0$ respectivamente:

$$u(\mu) = b - Ax(\mu) > 0$$

$$v(\mu) = A^T y(\mu) - c > 0$$

Si $x(\mu) \wedge y(\mu)$ tienden (como $\mu \to 0$) a las soluciones optimas de los problemas primal y dual, respectivamente, entonce necesariamente los productos:

$$x_i(\mu) v_i(\mu) \to 0 \text{ cuando } \mu \to 0$$
$$y_i(\mu) u_i(\mu) \to 0 \text{ cuando } u \to 0$$

En otras palabras, existen funciones $u_i(\mu) \wedge v_j(\mu)$ que tienden a cero en μ tal que:

Del (P_u): $$Ax + u = b$$

$$b - Ax = u$$

$$\rightsquigarrow \; b_i - A_i x = u_i(\mu) = \mu \frac{1}{y_i}$$

Del (D_u): $$A^T y - v = c$$

$$A^T y - c = v$$

$$A_j^T y - c_j = v_j(\mu) = \mu \frac{1}{x_j}$$

El método de la función de barrera logaritmica con pesos uniformes se caracteriza por las ecuaciones.

$$y_i(\mu) u_i(\mu) = x_j(\mu) v_j(\mu) = \mu$$

Por lo general (no necesariamente uniforme) los pesos de las funciones $u_i \wedge v_i$ permanecen de manera lineal en u.

En busca de la forma natural, de la función barrera o penalización, consideremos el siguiente problema de forma general

$$(P_{f,u}) \qquad \begin{array}{l} max\colon c^t x + \mu \sum_j f(x_j) + \mu \sum_i f(u_i) \\ \text{s.a: } Ax + u = b \end{array}$$

donde $f(\xi)$ es estrictamente concava.

Sea $g(\xi) = f'(\xi)$ es decir la derivada de f , y para cualquier d-vector a sea:

$$Ga = Diag\big(g(a_1), \dots, g(a_d)\big)$$

El par (x, u) es optimo para $(P_{f,u})$ si y solo si existe un vector $y \in \mathbb{R}^m$ tal que:

$$\mu G_x e - A^T y = -c$$

$$\mu G_\mu e - y = 0$$

Ahora, nos gustaría tener condiciones de optimalidad “simétricas Primal-Dual” es decir similares a las obtenidas al sistema (0). Es decir estamos interesados en funciones $f(\xi)$

donde la solución optimal para el problema dual aproximado proporcione multiplicadores de lagrange que soporten la solución optimal del primal aproximado y viceversa.

Para esto, tomemos el problema equivalente de $(P_{f,u})$

$$min: -c^T x - \mu \sum_j f(x_j) - \mu \sum_i f(u_i)$$

s.a: $Ax + u = b$

procederemos del mismo modo que en $(\propto \wedge \beta)$.
Tomemos el Lagrangeano

$$L(x,y,u) = -c^T x - \mu \sum_j f(x_j) - \mu \sum_i f(u_i) - y^T(b - Ax - u)$$

Y calculamos

$i)$ $\nabla x\, L(x,y,u) = 0$

$$\nabla x \left(-c^T x - \mu \sum_j f(x_j) - \mu \sum_i f(u_i) - y^T(b - Ax - u) \right) = 0$$

$$-c^T x - \mu \sum_j f'(x_j) + y^T A = 0$$

$$-c - \mu \sum_j g(x_j) + A^T y = 0$$

$$\mu G_x e - A^T y = -c$$

$ii)$ $\nabla u L(x,y,u) = 0$

$$\nabla u \left(-c^T x - \mu \sum_j f(x_j) - \mu \sum_i f(u_i) - y^T(b - Ax - u) \right) = 0$$

$$-\mu \sum_i f'(u_i) + y^T = 0$$

Tomando componente a componente

$$\mu g(u_i) = y_i$$

$$g(u_i) = \frac{y_i}{\mu}$$

$$g^{-1} o g(u_i) = g^{-1}(\frac{v_i}{\mu})$$

$$u_i = g^{-1}(\frac{v_i}{\mu})$$

$$-\mu \sum_{i} g(u_i) + y^T = 0$$

$$-\mu G_u e + y = 0$$

$$\mu G_u e - y = 0$$

Como estamos interesados en funciones $f(\xi)$ donde la solución optima para el problema dual aproximada proporcione los multiplicadores de Lagrange que caractericen a la solución optima de la aproximación primal, y viceversa. Entonces, tales funciones darían lugar a teoremas de dualidad similar al teorema anterior.

Cuando μ es ilimitada, por la sustitución.

$$u_i = g^{-1}\left(\frac{y_i}{\mu}\right)$$

Se obtiene un conjunto de ecuaciones que nos gustaría tener la misma forma como:

$$\mu G_x e - A^T y = -c$$

En otras palabras, necesitamos la función $g(\xi)$ que cumpla:

$$\mu g(\xi) = g^{-1}\left(\frac{\xi}{\mu}\right)$$

para cada ξ y $\mu > 0$.

El último requisito es muy restrictivo, lo cual implica que $g(\xi) = g^{-1}(\xi)$ para que $\mu g(\xi) = g\left(\frac{\xi}{\mu}\right)$, de donde resulta que $g(\xi) = \frac{g_{(1)}}{\xi}$.

Además, se llega a la única condición de restricción de barrera o de penalidad que son primal-dual de la forma en que la $f(\xi) = KLn\big((\xi)\big)$, donde k es constante.
En efecto:

$$g(\xi) = g^{-1}(\xi)$$

$$g(\xi) = \frac{1}{g_{(\xi)}}$$

$$g(\xi) = \frac{1}{\xi}$$

$$g(\xi) = \frac{g(1)}{\xi}$$

$$f'(\xi) = \frac{k}{\xi}$$

$$f(\xi) = kLn(|\xi|)$$

Tales funciones son apropiadas solo como funciones de barrera, es decir, funciones de penalización interior (y no como procedimientos de punto interior).

Sabemos que para cualquier par (x^0, y^0) de soluciones factibles interiores (de $(P)\ y\ (D)$ respectivamente) existen pesos que determinan un par de puntos de barrera ponderados de $x^0 \wedge y^0$ de los conjuntos optimos. La caracterización de estos caminos es simple.

Para simplificar la notación, los indices de columnas y filas variarán en conjuntos distintos de modo que podamos utilizar w_i^* para denotar un peso asociado con una fila y w_j denota un asociado a una columna. Teniendo en cuenta los puntos interiores $x^0 \wedge y^0$ en:

$$w_j = v_j . x_j^0$$

$$w_j = [A_j^T y^0 - c_j] y_i^0$$

$$w_i^* = u_i y_i^0$$

$$w_i^* = [b_i - A_i X^0] y_i^0$$

Tomamos la función:

$$C^T x + \mu \left(\sum_j w_j Lnx_j + \sum_i w_i^* Lnv_i \right)$$

El cual tiene un máximo en el interior de la región factible primal.

Tomemos también la función:

$$b^T y - \mu\left(\sum_i w_i^* \, Lny_i + \sum_j w_j Lnv_j\right)$$

El cual también tiene un mínimo sobre el interior de la región dual factible.

Si w es el total de los pesos, la diferencia entre los valores de las funciones lineales es igual a $W_\mu(1 - Ln\mu)$

En efecto,

$$c^t x + \mu\left(\sum_j w_j Lnx_j + \sum_i w_i^* Lnu_i\right) - \left(b^t y - \mu\left(\sum_i w_i^* Lny_i + \sum_j w_j Lnv_j\right)\right)$$

$$= c^T x - b^T y - \mu\left(\sum_j w_j Lnx_j + \sum_i w_i^* Lnu_i + \sum_i w_i^* Lny_i + \sum_j w_j Lnv_j\right)$$

$$= c^T x - b^T y - \mu\left(\sum_j w_j Lnx_j v_j + \sum_i w_i^* Lnu_i y_i\right)$$

$$= c^T x - b^T y - \mu\left(\sum_j w_j Ln\mu + \sum_i w_i^* Ln\mu\right)$$

$$= w_\mu - w_\mu . Ln\mu$$

$$= w_\mu(1 - Ln\mu)$$

Los caminos se caracterizan por la propiedad de que a lo largo de cada uno de los productos de las variables $x_j v_j \wedge y_i u_i$ estos son proporcionales a μ. En otras palabras, las radios entre estos productos se mantienen constantes, es decir:

$$x_j v_j = \mu w_j$$

$$y_i u_i = \mu w_i^*$$

se mantienen a lo largo de los caminos.

En la siguiente sección vamos a estudiar las descripciones diferenciales de estas trayectorias.

Hemos argumentado que una trayectoria solución determina un par de soluciones interiores factibles (para el problema dual y primal x_0 e y_0) y es el lugar geométrico de los puntos factibles interiores con las razones de variables complementarias. Estos mismos productos a través de nuestra interpretación sugiere una generalización natural.

Consideremos el siguiente conjunto de ecuaciones:

$$(*) \qquad \begin{aligned} x_j\left(A_j^t y - C_j\right) &= \mu w_j \\ y_i(b_i - A_i x) &= \mu w_i^* \end{aligned}$$

El problema original (P) requiere que $Ax < b$ y $x \geq 0$. Sin embargo, podemos considerar 2^{n+m} problemas diferentes que corresponden a 2^{n+m} formas diferentes de elegir las restricciones sobre los signos de las variables

$$x_j \wedge u_i = b_i - A_i x$$

El problema Dual para cada una de ellas es obtenida por aprovechables cambios de signo de las variables complementarias duales. Para tales pares de problemas Primal-Dual, el producto de las variables complementarias son no negativas. En otras palabras, si todos los productos $x_j v_j$ y $y_i u_i$ son no negativas, entonces x e y son soluciones factibles del primal y dual respectivamente para al menos uno de esos pares de problemas. En cualquier caso, x e y son soluciones factibles de algún par de problemas (no necesariamente dual) que puede ser obtenido de los originales cambiando las direcciones de algunas desigualdades. El sistema (*) define trayectorias solución para todas las combinaciones factibles de problemas Primal-Dual.

UNA DESCRIPCIÓN GENERAL DE LAS TRAYECTORIAS SOLUCIÓN

Una interesante descripción de las trayectorias solución se puede hacer de la siguiente manera. En primer lugar, consideremos el problema

$$(\mathrm{P}\infty) \qquad \begin{array}{l} \max \ \sum_j Lnx_j + \sum_i Lnu_i \\ \\ \text{s.a: } Ax + u = b \\ x, u > 0 \end{array}$$

Que es, en cierto sentido, el límite de (P_μ) cuando $\mu \to \infty$ (en el sentido de problema equivalente)

Si (P_∞) tiende a la solución optima x^∞, entonces $x(\mu)$ tiende a x^∞ a medida que μ a infinito.

En segundo lugar, consideremos el problema de minimización:

$$\min c^T x - \mu\left(\sum_j Lnx_j + \sum_i Lnu_i\right)$$

Sujeto a las mismas limitaciones, y es fácil de ver que a medida que u tiende al infinito, el camino de acceso de este último también se acerca a x^∞.

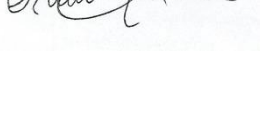
Parece apropiado aplicar en este punto un cambio de parámetro, de forma que las trayectorias de los problemas de optimización se puedan describir de una manera unica o unificada.

Consideremos la sustitución $\mu = tg\theta$,

en la siguiente función objetivo no lineal:

$$c^T x - \mu\left(\sum_j Lnx_j + \sum_i Lnu_i\right) \ ; \ 0 < \theta < \frac{\pi}{2}$$

$$c^T x - tg\theta\left(\sum_j Lnx_j + \sum_i Lnu_i\right)$$



$$cos\theta(c^t x) - sen\theta(\sum_j Lnx_j + \sum_i Lnu_i)$$

Lo obtenido es la parte de la ruta de acceso correspondiente al problema de minimización, mientras que el intervalo de $\frac{\pi}{2} < \theta < \pi$ corresponde al problema de maximización.
Al evaluar $\theta = \frac{\pi}{2}$ corresponde a la suma máxima de logaritmos.

Si se limitó las intersecciones de conjuntos de nivel de $c^t x$ con el poliedro factible y el problema lineal tiene un mínimo, entonces el camino está bien definido para $0 \leq \theta \leq \frac{\pi}{2}$.
Si el poliedro factible no está acotado, entonces el camino no está definido en $\theta = \frac{\pi}{2}$. De hecho, que diverge hasta el infinito como $\theta \to \frac{\pi}{2}$.

Las ecuaciones definen la trayectoria que tienen la forma:

$$(sen\theta)D_x^{-1} - A^t y = -(cos\theta)c$$

$$Ax + (sen\theta)D_y^{-1} = (cos\theta)b$$

En efecto:

tomemos el Lagrangeano en la función objetivo:

$$min: -c^t x - \mu\left(\sum_j f(x_j) + \sum_i f(u_i)\right)$$

$$L(x, y, u_i) = -c^T x - \mu \sum_i f(u_i) - y^T(b - Ax - u)$$

$i)$ $\nabla x L(x, y, u_i) = 0$

$$\nabla x\left(-c^T x - \mu \sum_j f(x_j) - \mu \sum_i f(u_i) - y^T(b - Ax - u)\right) = 0$$

$$-c^T - \mu \sum_j f'(x_j) + y^T A = 0$$

$$-c - \mu \sum_j f'(x_j) + A^T y = 0$$

$$-c - \mu D_x^{-1} + A^t y = 0$$

Sabemos que $\mu = tg\theta = \frac{sen\theta}{cos\theta}$

$$-c - \left(\frac{sen\theta}{cos\theta}\right) D_x^{-1} + A^t y = 0$$

$$-(cos\theta)c - (sen\theta)D_x^{-1} + (cos\theta)A^T y = 0$$

$$(sen\theta)D_x^{-1} - (cos\theta)A^T y = -(cos\theta)c$$

Proposición 4

Sean $\bar{x}, \bar{y}, \bar{u}$ y $\bar{v}$. los valores óptimos de las variables en (P_u) y (D_u) al final de los caminos (es decir, cuando $\mu \to 0$, suponiendo que el problema tiene una solución óptima). Entonces, para cada uno de las pares de las variables complementarias $(\bar{x_i}, u_i)$ y $(\bar{y_j}, v_j)$, uno de ellos es positivo mientras que el otro es igual a cero.

Demostración:

Obviamente al menos uno de los miembros de cada par es igual a cero. Es bien conocido que en los vertices degenerados (de la región factible) algunos tienen pares donde ambos miembros pueden ser igual a cero. Sin embargo, degeneración significa que ya sea el problema primal o el dual tiene una cara optima de dimension mayor que cero (es decir infinitas soluciones).

Afirmamos que las vias de solución a los puntos convergen en el interior relativo de las caras óptimas de los problemas primal y dual. El punto $\bar{x}$, es donde la suma $\sum_j Lnx_j + \sum_i Lnu_i$ (ponderado de todas las variables que no son igual a cero en la cara del óptimo) se maximiza respecto a la cara óptima. Obviamente cada variable que no es identico a cero en la cara óptima no se anula en $\bar{x}$. Esto implica nuestra proposición.

Suponiendo que los límites de $x(0)$ y $y(0)$ existen, y si consideramos que las variables se anulan en este punto, entonces también se anularan en cualquier otro punto del conjunto optimo.

Denotemos por I el conjunto de indices i tales que:

$$A_i x = b_i$$

en cada solución óptima del primal.

Además, sea J el conjunto de los índices j tal que

$$x_j = 0$$

establecido en cada solución optima dual y

$$con\ y_j = 0 \text{ para } i \in I \text{ y } A_j^t y = c_j \text{ para } j \notin J$$

Consideremos el siguiente problema ($\tilde{P}$):

$$Min\ c^t x$$

$$\begin{aligned}
\text{s.a: } A_i x &\geq b_i & (i \in I) \\
A_i x &\leq b_i & (i \notin I) \\
x_j &\leq 0 & (j \in J) \\
x_j &\geq 0 & (j \notin J)
\end{aligned}$$

Este problema puede ser aproximado por ($\widetilde{P_\mu}$):

$$Min\ c^T x - \mu\left(\sum_{j\in J} Ln(-x_j) + \sum_{j\notin J} Lnx_j + \sum_{i\in I} Ln(-u_i) + \sum_{(i\in I)} Lnu_i\right)$$

s.a: $Ax + u = b$

$x_j < 0 \;\; (j \in J);\; x_j > 0 \;\; (j \notin J);\; u_i < 0 \;\; (i \in I) \quad u_j > 0 \;\; (i \notin I)$

De esto se deduce que las condiciones de optimalidad para $(\widetilde{P_u})$ son los mismo que para (D_u) en el sentido de que las vias de solución (asumiendo que existen en ambos lados) se puede unir de forma continua a la cara óptima cómun a los problemas (P) y $(\tilde{P})$.

Por otro lado, recordemos que la función $C^T x$ aumenta monotonamente cuando $u \to 0$. De ello se desprende que siempre y cuando el camino se pueda continuar, se puede extender a traves de la disposición del hiperplano de manera que en cada célula que se desplaza (monotonamente en terminos de $C^T x$) a partir de un mínimo de células a un máximo de células ya que también es un mínimo de una celda adyacente, y así sucesivamente. La sustitución $u = tg\theta$ produce una representación continua de una trayectoria combinada que viaja a través de las células delimitadas, se puede extender en ambos lados con las células no acotadas en que la trayectoria, tiende a infinito.

Salvo el caso patologico, los caminos no visitan las células en las que la función no tiene ni un máximo, ni un mínimo. Un caso patológico es por ejemplo, un cilindro poliedrico que la función lineal, es ilimitada para ambos

4.Tangentes a la Trayectoria (o Tangente a la Curva)

Veamos ahora, las propiedades de diferenciabilidad.
Sea (x^0, y^0) un par de soluciones factibles interiores (de los problemas (P)y(D) respectivamente). Y sea u^0, v^0 los vectores de holgura correspondientes a los problemas primal y dual respectivamente y usando los productos $w_j = x_j^0 v_j^0$ y $w_i^* = x_i^0 v_i^0$ para definir un par de caminos como se explicó anteriormente, vamos a examinar ahora la tangente a la trayectoria en el punto de partida.
La trayectoria está determinada por las siguientes ecuaciones:

$$\mu w_j \frac{1}{x_j} - A_j^T y = -c_j \qquad\qquad x = x(\mu)$$

$$A_i x + u\mu.\frac{1}{y_i} = b_i \qquad\qquad y = y(\mu)$$

Aplicando diferenciación con respecto a µ.

- $\mu w_j \frac{1}{x_j} - A_j^T y = -c_j$

$$w_j \frac{1}{x_j} + \mu w_j x \left(\frac{-\dot{x}_j}{x_j^2}\right) - A_j^T \dot{y} = 0$$

$$-\mu w_j.\left(\frac{-\dot{x}_j}{x_j^2}\right) - A_j^T \dot{y} = -w_j.\frac{1}{x_j} \ \ldots\ldots\ldots. \boxed{1}$$

- $A_i x + \mu w_i^*.\frac{1}{y_i} = b_i$

$$A_i \dot{x} + w_i^*.\frac{1}{y_i} + (\mu w_i^*)\left(\frac{-\dot{y}_j}{y_j^2}\right) = 0$$

$$A_i \dot{x} - \mu w_i^*.\frac{\dot{y}_j}{y_j^2} = - u w_i^*.\frac{1}{y_i} \ldots\ldots\ldots\ldots\ldots. \boxed{2}$$

Luego obtenemos las siguientes ecuaciones de $\boxed{1}$ y $\boxed{2}$

$$\left.\begin{aligned} -\mu w_j.\left(\frac{-\dot{x}_j}{x_j^2}\right) - A_j^T \dot{y} &= -w_j.\frac{1}{x_j} \\ A_i \dot{x} - \mu w_i^*.\frac{\dot{y}_j}{y_j^2} &= -u w_i^*.\frac{1}{y_i} \end{aligned}\right\} \boxed{*}$$

Luego considerando un punto (x', y') con: $x' = x^\circ - \delta\dot{x}$

$$y' = y^\circ - \delta\dot{y}$$

Donde: $x^0 y\, y^0$ constituyen la solución de este último sistema de ecuaciones en $x = x^0$y $y = y^0$ y $\mu = 1$ y δ es cualquier número positivo

Obviamente (x', y')se encuentra en la tangente a la curva en $x°$ y $y°$.

Luego:

$$b^T y' - C^T x' = b^T(\dot{y} - \delta\dot{y}) - C^T(x° - \delta\dot{x})$$

$$= b^T y^0 - \delta b^T \dot{y} - C^T x^0 + \delta C^T \dot{x}$$

$$= (b^T y^0 - C^T x^0) + \delta C^T \dot{x} - \delta b^T \dot{y}$$

$$= (b^T y^0 - C^T x^0) - \delta(b^T \dot{y} - C^T \dot{x})$$

$$= (b^T y^0 - C^T x^0)(1 - \delta)$$

Denotamos los vectores de holgura correspondientes al par (x', y') por $u' y\, v'$y sea w_i^* y w_j' los que denotan los productos correspondientes de variables complementarias.

Luego tenemos:

- $w_j' = x_j' v_j' = (x_j^0 - \delta\dot{x}_j)(v_j^0 - \delta A_j^T \dot{y}) \qquad x_j^0, v_j^0 = w_j$

$$\boxed{v_j = A_j^T y - c_j} \qquad \boxed{A_j^T \dot{y} = -\mu w_j . \frac{\dot{x}_j}{x_j^2} + w_j . \frac{1}{x_j}}$$

$$\rightarrow v_j' = A_j^T y - c_j$$

$$= A_j^T(y° - \delta\dot{y}) - c_j$$

$$= A_j^T y^0 - \ \delta\dot{y}A_j^T - c_j = A_j^T y^{\circ} - c_j - \delta A_j^T \dot{y} = v_j^0 - \delta A_j^T \dot{y} \ \dots(\alpha)$$

- $w_j' = (x_j^0 v_j^0 - x_j^0 \delta A_j^T \dot{y} - \delta \dot{x}_j . v_j^0 + \delta^2 \dot{x}_j A_j^T \dot{y})$

$$= w_j + \frac{x_j^0 \delta \mu y \dot{x}_j}{x_j^2} - x_j^0 \delta . w_j . \frac{1}{x_j}$$

$$- \delta \dot{x}_j . \frac{w_j}{x_j^0} - \delta^2 \left(\dot{x}_j\right)^2 \mu \frac{w_j}{x_j^2} + \delta^2 \dot{x}_J . w_j \frac{1}{x_j}$$

$$= w_j \left[1 + \frac{x_j^0 \delta \, \mu \, \dot{x}_J}{x_j^2} - x_j^0 \delta . \frac{1}{x_j} - \frac{\delta}{x_j^0} . \dot{x}_J - \frac{\delta^2 \left(\dot{x}_j\right)^2 \mu}{x_j^2} + \delta^2 . \frac{\dot{x}_J}{x_j} \right]$$

$$\boxed{x_j = x_j^0} \qquad \boxed{\mu = 1}$$

$$= w_j \left[1 + x_j^0 . \delta \dot{x}_J - x_j^0 . \delta . \frac{1}{x_j^0} - \delta \frac{\dot{x}_J}{x_j^0} + \delta^2 \frac{\dot{x}_J}{x_j^0} \left(1 - \frac{\dot{x}_J}{x_j^0} \right) \right]$$

$$= w_j \left[1 + \delta \frac{\dot{x}_J}{x_j^0} - \delta - \delta \frac{\dot{x}_J}{x_j^0} + \delta^2 \frac{\dot{x}_J}{x_j^0} \left(1 - \frac{\dot{x}_J}{x_j^0} \right) \right]$$

$$= w_j \left[1 - \delta + \delta^2 \frac{\dot{x}_J}{x_j^0} \left(1 - \frac{\dot{x}_J}{x_j^0} \right) \right]$$

De igual modo:

$$(w_i^*)' = y_i' . u_i' = w_i^* \left[1 - \delta + \delta^2 \frac{\dot{y}_j}{y_j^0} \left(1 - \frac{\dot{y}_j}{y_j^0} \right) \right]$$

Para que los puntos x' y y' permanezcan viables en sus respectivos problemas, es necesario y suficiente que las siguientes cantidades sean menores o igual a 1.

$$\delta\frac{\dot{x}_j}{x_j^0},\ \delta\left(1-\frac{\dot{x}_j}{x_j^0}\right) \qquad ; \qquad \delta\frac{\dot{y}_j}{y_j^0},\ \delta\left(1-\frac{\dot{y}_j}{y_j^0}\right)$$

Es importante examinar las propiedades de las tangentes para establecer algunas conecciones entre los métodos de punto interior.

Para esto, volvamos por un momento al problema de forma estándar. Ahora trabajaremos con el problema de la forma:

$$\begin{array}{ll} & \text{Maximizar } C_x^T \\ \text{(P)} & \text{s.a.} \quad Ax = b \\ & x \geq 0 \end{array}$$

Dado un par (x^0, y^0) donde:

$$Ax^0 = b \qquad\qquad \text{y } A^T y^0 > C$$

$$x^0 > 0$$

El par del camino $x = x(u), y = y(u)$ a través de (x^0, y^0) (Es decir con $x(1) = x^0$y $y(1) = y^0$), se determina por el siguiente sistema de ecuaciones:

$$\left(A_j^T y - c_j\right)x_j = u\left(A_j^T y^\circ - c_j\right)x_j^0 \quad ; \ \mathrm{j} = 1, \ldots \mathrm{n}$$

$$Ax = b$$

$$w_j = x_j^0, v_j^0$$

$$w_j = x_j^0(A_j^T y^\circ - c_j)$$

Por diferenciación en $\mu = 1$, tenemos del siguiente sistema (donde (x^0, y^0)fue sustituido por (x, y) para simplificar:

$$\left(A_j^T y - c_j\right)\dot{x}_j + x_j A_j^T \dot{y} = \left(A_j^T y - c_j\right)x_j$$

$$A\dot{x} = 0 \qquad\qquad j = 1, \dots, n$$

Denotamos como de costumbre

$$v_j = A_j^T y - c_j$$

Así: $\quad v_j \dot{x}_j + x_j A_j^T \dot{y} = v_j x_j$

$$A\dot{x} = 0$$

$$D_v \dot{x} + D_x A^T \dot{y} = D_v D_x e$$

$$A\dot{x} = 0$$

Donde:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$$D_v = \begin{bmatrix} v_1 & & 0 \\ & \ddots & \\ 0 & & v_n \end{bmatrix}, \; D_x = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix}$$

$$D_x^{-1} = \begin{pmatrix} \frac{1}{x_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{x_n} \end{pmatrix}$$

En matrices:

$$\begin{pmatrix} D_v & D_x A^T \\ A & O \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} D_v D_x e \\ 0 \end{pmatrix}$$

$$F_1 \times D_x^{-1} \Longrightarrow \begin{pmatrix} D_x^{-1} D_v & D_x^{-1} D_x A^T \\ A & O \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} D_x^{-1} D_v D_x e \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} D_v D_x^{-1} & A^T \\ A & O \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} D_v e \\ 0 \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

Ahora sea: $\xi \in \mathbb{R}^n$, $\xi = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

donde: $\xi_j = \sqrt{\frac{v_j}{x_i}} . \dot{x}_J = v_j^{1/2} . \frac{1}{x_j^{1/2}} . \dot{x}_J$

$$\longrightarrow \xi = D_v^{1/2} D_x^{-1/2} \dot{x},$$

donde $D_v^{1/2} = \begin{pmatrix} v_i^{1/2} & & 0 \\ & \ddots & . \\ 0 & & v_n^{1/2} \end{pmatrix}$, $D_x^{1/2} = \begin{pmatrix} x_i^{1/2} & & 0 \\ & \ddots & . \\ 0 & & x_n^{1/2} \end{pmatrix}$

$$\longrightarrow \dot{x} = D_v^{-1/2} D_x^{1/2} . \xi$$

$$D_x^{-1} = \begin{pmatrix} \frac{1}{x_1^{1/2}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{x_n^{1/2}} \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} D_v D_x^{-1} . D_v^{-1/2} D_x^{1/2} & A^T \\ A D_v^{-1/2} D_x^{1/2} & O \end{pmatrix} \begin{pmatrix} \xi \\ \dot{y} \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

$$F_1 . D_v^{-1/2} D_x^{1/2} \qquad \begin{pmatrix} D_v^{1/2} . D_x^{-1/2} & A^T \\ A D_v^{-1/2} D_x^{1/2} & O \end{pmatrix} \begin{pmatrix} \xi \\ \dot{y} \end{pmatrix} = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} I & D_v^{-1/2} . D_x^{1/2} A^T \\ A D_v^{-1/2} D_x^{1/2} & O \end{pmatrix} \begin{pmatrix} \xi \\ \dot{y} \end{pmatrix} = \begin{pmatrix} D_v^{1/2} & D_x^{1/2} & e \\ & 0 & \end{pmatrix}$$

De aqui, resulta que el vector $\xi = \left(\xi_{1,\dots,}\xi_n\right)^T$es la proyección ortogonal del vector $D_v^{1/2} D_x^{1/2} e$ en el espacio nulo de la matriz $A D_v^{-1/2} D_x^{1/2}$.

Así, la interpretación de la dirección en términos de x es el siguiente:

Dado un par de soluciones factibles interiores primal y dual (x, y), el problema(P) es equivalente al siguiente problema (P'), donde Z es la variable de optimización y x e y son fijos:

$$\begin{gathered}\text{Minimizar } (A^T y - C)^T Z \\ \text{s.a} \qquad Az = 0 \\ x - z \geq 0\end{gathered}$$

Notamos que, el gradiente de la función objetivo es el valor:

$$v = A^T y - C$$

Sin embargo, el algoritmo toma un “paso gradiente” sólo después de aplicar la siguiente transformación lineal.

$$T(z) = T_{x,y}\,(z) = D_v^{1/2} D_x^{-1/2} z$$

Esta transformación lleva el punto actual “x” al vector de medias geométricas de los valores de las variables complementarias:

$$x' = T(x) = D_v^{1/2} D_x^{-1/2} e$$

es decir:

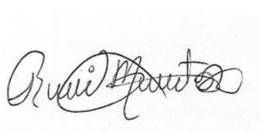
$$x_j' = \sqrt{x_j v_j} \qquad\qquad (j = 1, \dots, n)$$

Y la variable z se transforma en:

$$\xi = T(z)$$

De esta manera obtenemos un problema equivalente:

$$\begin{gathered}\text{Minimizar: } \left(D_v^{1/2} D_x^{1/2} e\right)^T \xi = 0 \\ \text{s.a} \\ A D_v^{-\frac{1}{2}} D_x^{\frac{1}{2}} \xi = 0\end{gathered}$$



$$x' - \xi \geq 0$$

Aquí el gradiente es el mismo vector $x´$ de medias geométricas y es la proyección ortogonal de $D_v^{1/2} D_x^{1/2} e$ en el espacio nulo de la matriz $A D_v^{-1/2} D_x^{1/2}$.

5. Propiedades diferenciales de las trayectorias solución

En esta sección vamos a estudiar las derivadas de orden superior asociadas con las curvas $x = x(u)$ y $y = y(u)$ de soluciones factibles interiores primal y duales discutidos en las secciones anteriores.

Por conveniencia, usaremos aquí la notación que se utiliza generalmente en el contexto del problema complementariedad lineal.

Denotamos por $z = z(u)$ el $(n+m)$ vector obtenido mediante la concatenación de $x(u)$ y $y(u)$, y también utilizar $s = s(u)$ para denotar el $(n+m)$-vector obtenido mediante la concatenación de los vectores de la holgura $v(u)$ y $u(u)$.

Dejemos que M indique la matriz:

$$M = \begin{pmatrix} 0 & -A^T \\ A & 0 \end{pmatrix}$$

Y sea "q" el que denota al $(n+m)$-vector obtenido mediante la concatenación de los vectores c y $-b$. Dejemos que $\dot{z_i}$ denote la derivada de z (como una función de u) y

$$\dot{z} = (\dot{z}_1, \ldots, \dot{z}_{n+m})$$

También "ampliamos" las operaciones aritméticas con vectores (aplicando lo componente por componente).

Por ejemplo:

$$\frac{\dot{z}}{z^2} = \left(\frac{\dot{z}_1}{z_1^2}, \ldots, \frac{\dot{z}_{n+m}}{z_{n+m}^2}\right)$$

Con la nueva anotación el sistema combinado de restricciones primal y dual es la siguiente.

$$S + Mz = -q$$
$$z, s \geq 0$$

Un par de soluciones óptimas es caracterizado por la condición de holgura complementaria

$$s_i z_i = 0$$

Una trayectoria solución que pasa por el punto (z^0, s^0) se determina por la ecuación:

$$u\left(\frac{z^o s^o}{z}\right) + Mz = \ -q$$

El cual queremos resolver para z a medida que u se próxima a 0.

Sea:

$$F(z,u) = u\ \left(\frac{z^0 s^o}{z}\right) + Mz + q$$

Entonces, queremos resolver $F(z,u)$=0

Podemos evaluar las derivadas $d^k\, z/\, du^k$ diferenciando a F

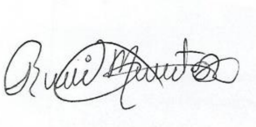
Sea $w_i = z_i^0 s_i^0$ y “w” denota al vector que consiste con las componentes $w_i's$.

Primero,

$$\frac{dF}{du} = -uw.\frac{\dot{z}}{z^2} + M\dot{z} + w.\frac{1}{z}$$

Siempre que “a” es un vector, denotamos por: $\Delta(a) = Da$ a la matriz diagonal cuyos elementos diagonales son las componentes de “a”.

Se deduce que el valor de $\dot{z}$ en u puede obtenerse resolviendo el siguiente sistema de ecuaciones lineales (donde $\dot{z}$ es la incógnita y suponiendo que z se conoce):



$$\left[-u\Delta\left(\frac{w}{z^2}\right)+M\right]\dot{z}=-\frac{w}{z}$$

Segundo:

$$\frac{d^2F}{du^2}=2uw\frac{\dot{z}^2}{z^3}-uw\frac{\ddot{z}}{z^2}+M\ddot{z}-\frac{w\dot{z}}{z^2}-\frac{w\dot{z}}{z^2}$$

Por lo tanto, el valor de $\ddot{z}$ se obtiene resolviendo el siguiente sistema de ecuaciones lineales (asumiendo u,z, $\dot{z}$ son conocidos):

$$\left[-u\Delta\left(\frac{w}{z^2}\right)+M\right]\ddot{z}=2w\frac{\dot{z}}{z^2}-2\,uw\frac{\dot{z}^2}{z^3}$$

Notemos que el coeficiente de la matriz.

$$Q=-\mu\Delta\left(\frac{w}{Z^2}\right)+M$$

es la misma que aparece en las ecuaciones que definen $\dot{z}$ y $\ddot{z}$.

Se pueden demostrar que, para cada valor de K, el valor de la k-esima derivada $z^{(k)}$, se puede obtener resolviendo un sistema lineal, donde la matriz de coeficientes es la misma matriz Q, y el vector del lado derecho es un polinomio en términos de $\frac{1}{z}$ y las derivadas:$\dot{z}$, $\ddot{z}$ ... $z^{(k-1)}$.

El hecho de que todas las derivadas de z puedan ser evaluadas como soluciones del sistemas lineales, con la misma matriz de coeficientes, es debido a la estructura particular de la función F, es decir:

$$F(z,u)=u\alpha(z)+\beta(z)$$

Donde $\alpha\ y\ \beta$ son cualquier C^{∞} de $\mathbb{R}^{n+m}$en si misma. Luego se sigue que:

$$\frac{dF}{du}=-u\frac{D_\alpha}{D_z}\dot{z}+\frac{D_\beta}{D_z}\dot{z}+\alpha(z)=0$$

Así el $\dot{z}$ se obtiene a partir de los siguiente sistema

$$\left[\frac{uD_\alpha}{D_z} + \frac{D_\beta}{D_z}\right]\dot{z} = - \alpha(z)$$

La segunda derivada tiene la forma:

$$\frac{d^2F}{du^2} = \left[\frac{uD_\alpha}{D_z} + \frac{D_\beta}{D_z}\right]\ddot{z} + u\alpha_1(z,\dot{z}) + \beta_1(z,\dot{z})$$

Y se puede demostrar por inducción sobre k que la derivada de orden k tiene la forma:

$$\frac{d^kF}{du^k} = \left[\frac{uD_\alpha}{D_z} + \frac{D_\beta}{D_z}\right]z^{(k)} + u\alpha_{k-1}\left(z,\dot{z},\dots z^{k-1}\right) + \beta_{k-1}\left(z,\dot{z},\dots z^{k-1}\right)$$

6.Comportamiento cerca de los vértices

Es conveniente considerar en esta sección el problema de programación en forma estándar, que es

$$Maximizar\ c^t x$$

Sujeto a:

$$Ax = b$$

$$x \geq 0$$

Donde $A \in \mathbb{R}^{mn} (m \geq n),\ x, c \in \mathbb{R}^n$ y $b \in \mathbb{R}^m$

Sea B una matriz cuadrada de orden m, que consiste en las primeras m columnas de A. es decir:

$$B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}$$

Asumiendo que B es no singular y $B^{-1}b > 0$
Entonces, B cumple con ser una base factible no degenerada.

Por otro lado, sea N una matriz de orden $m \times (n - m)$ que consiste en las últimas $(n - m)$ columnas de A y la denotamos por:

$$N = \begin{pmatrix} a_{1m+1} & \dots & a_{1n} \\ \vdots & & \\ a_{m(m+1)} & \dots & a_{mn} \end{pmatrix}$$

Nosotros denotaremos la restricción de algún vector v que tenga dimensión igual a "m", como: v_B y la restricción de dimensión igual a "$n - m$" como: v_N

Por lo tanto, los vectores c_B, c_N, x_B, x_N son definidos con respecto a las componentes de los vectores c y x.

Veamos

$$c_B = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}, \ c_N = \begin{pmatrix} c_{m+1} \\ \vdots \\ c_n \end{pmatrix}, \ x_B = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \ x_N = \begin{pmatrix} x_{m+1} \\ \vdots \\ x_n \end{pmatrix}$$

También denotaremos por $D = (D_{(x)})$ a la matriz diagonal de orden "n" cuyos elementos de la diagonal son las componentes del vector "x"

Es decir:

$$D_{(x)} = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$$

Y también D_B y D_N serán las matrices diagonales de orden m y $n - m$ respectivamente correspondiente a los vectores x_B y x_N.

Asumiendo que el primal y dual tienen regiones factibles de dimensión completa, entonces la trayectoria está definida siempre y cuando la solución interior factible para el problema primal y dual estén bien definido, es decir, se tenga par inicial x^0, y^0 tal que,

Se cumpla $x^0 \in \mathbb{R}^n$ tal que $Ax^0 = b$ y $x^0 > 0$ y sea $y^0 \in \mathbb{R}^m$ tal que $A^t y^0 \geq c$.

Entonces, la trayectoria comienza en (x^0, y^0) y se genera por las ecuaciones:

$$(*) \begin{cases} x_j v_j = \mu x_j^0 v_j^0 \quad , \quad \forall_j = 1,2, \dots, n \\ \qquad Ax = b \end{cases}$$

Es obvio que, para cualquier punto sobre la trayectoria, si comenzamos acorde con esta definición, entonces nada varía desde que el producto de variables complementarias permanecen en la misma proporción.
El sistema $(*)$ es equivalente a:

$$\begin{cases} x_j(A_j^t y - c_j) = u x_j^0 (A_j^t y^0 - c_j) \quad , \quad \forall_j = 1,2,\dots,n \\ Ax = b \end{cases}$$

Sea $w_j = x_j^0\left(A_j^t y^0 - c_j\right)$ $\rightarrow$ del sistema $(*)$ se tiene: $x_j v_j = \mu w_j$

Como $Ax = b$

$\rightarrow Bx_B + Nx_N = b$

Luego $x_B = B^{-1}(b - Nx_N)$

De la ecuación: $x_j v_j = \mu w_j$

Y despejando v_j

Se obtiene

$$v_j = \mu \frac{w_j}{x_j} \quad , \qquad \forall_j = 1,2,\dots,n$$

Ahora

$$v_j = \left(A_j^T y - c_j\right) = \mu \frac{w_j}{x_j} \quad , \quad \forall_j = 1,2,\dots,n$$

Considerando la restricción $j = 1,2,\dots,m$, se tiene:

$$\left(A_j^T y - c_j\right) = \left(B_j^T y - c_j\right) = \mu\left(\frac{w_j}{x_j}\right)$$

Luego:

$$(B^T y - c_B) = \mu D_{w_B} D_{x_B}^{-1}.e$$

Donde:

$$D_{w_B} = \begin{pmatrix} w_1 & & & & \\ & w_2 & & 0 & \\ & & w_3 & & \\ & & & \ddots & \\ 0 & & & & w_m \end{pmatrix}, \; D_{x_B}^{-1} = \begin{pmatrix} \frac{1}{x_1} & & & \\ & \frac{1}{x_2} & & \\ & & \ddots & \\ & & & \frac{1}{x_m} \end{pmatrix}$$

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{m\times 1}$$

Luego: $B^T y - c_B = \mu D_{x_B}^{-1}.e$

Despejando "y" se tiene:

$$y = B^{-T}[\mu D_{w_B} D_{x_B}^{-1}.e + c_B]\ldots (**)$$

Anteriormente se planteó

$$x_j v_j = \mu w_j$$

donde $x_j = \mu \frac{w_j}{v_j}$

$$x_j = \mu \frac{w_j}{A_j^t y - c_j} \ldots (\propto)$$

Luego $(**)$ en $(\propto)$ se obtiene:

$$x_j = \mu \frac{w_j}{A_j^T B^{-T}\left[\mu D_{w_B} D_{x_B}^{-1} e + c_B\right] - c_j}$$

Donde:

$$x_j = \mu \frac{w_j}{A_j^T B^{-T} \mu D_{w_B} D_{x_B}^{-1} e + A_j^T B^{-T} c_B - c_j} \ldots (\beta)$$

Luego, suponiendo que $B^{-1}b > 0$, entonces, ésta es la solución óptima primal y es única y $B^{-T}c_B$ es la única solución óptima dual y el camino formado por las $x_j's$ y $y_i's$ convergen a estos puntos respectivamente.

Como μ tiende a cero, entonces las variables no básicas, que son los x_j con $j = m + 1, \ldots, n$ están dadas por x_j y asintóticamente se comportan como:

$$x_j \sim \mu \frac{w_j}{A_j^T B^{-T} c_B - c_j} \qquad \forall_j = m+1, \ldots, n \ \ldots (\gamma)$$

Si observamos el denominador, podemos reconocer que éste es llamado algunas veces "el costo reducido" con respecto a la base B, es decir,

$$\widetilde{c_j} = c_j - A_j^T B^{-T} c_B$$

Y reemplazando en (γ) se obtiene:

$$x_j \sim -\mu \frac{w_j}{\tilde{c}_j} \qquad \forall_j = m+1, \dots, n$$

Notemos que si y^0 está cerca de la solución óptima dual, es decir $y^0 \; cerca\; de\; B^{-t} c_B$ entonces tenemos:

$$x_j \sim \mu \frac{w_j}{A_j^T B^{-T} c_B - c_j} \qquad j = m+1, \dots, n$$

entonces

$$\Rightarrow \; x_j \sim \mu \frac{w_j}{A_j^T y^0 - c_j} \qquad j = m+1, \dots, n$$

Pero sabemos que:

$$w_j = x_j^0 (A_j^T y^0 - c_j) \Rightarrow \frac{w_j}{A_j^T y^0 - c_j} = x_j^0$$

Así:

$$x_j \; tiende\; a\; \mu x_j^0 \qquad \forall_j = m+1, \dots, n$$

Es decir $x_j \sim \mu x_j^0 \quad \forall_j = m+1, \dots, n$

En otras palabras, si comenzamos lo suficientemente cerca de una solución optimal, entonces la trayectoria nos lleva aproximadamente en línea recta a la solución óptima. Es decir, esto es la diferencia con respecto a los algoritmos de reescalado lineal donde todas las rutas tienden a una sola dirección de aproximación a la solución óptima [16].

V. RESULTADOS

Hemos arribado a los siguientes resultados:

1. La proposición 1, implica que el sistema (O) posee solución única para cualquier $\mu > 0$. Es decir, se determina una única trayectoria continua $x = x(\mu)$ con μ en los reales positivos.
2. Cuando A posee rango completo también se determina la trayectoria continua $y = y(\mu)$.
3. Bajo la condición de que existan los límites de $x = x(\mu)$ $y = y(\mu)$ cuando μ tiende a cero y lo denotamos por $\bar{x}$ y $\bar{y}$. Entonces, se cumple que: $A\bar{x} = b,\ x \geq 0,$ $y\ A^t\bar{y} \geq c$. Además, si se tiene que $\bar{x}_j > 0.\ A_j^t\bar{y} = c_j$, entonces $\bar{x}$ y $\bar{y}$ son soluciones óptimas del primal (P) y del dual (D) respectivamente.
4. El x' de medias geométricas, es la proyección ortogonal de $D_v^{1/2}D_x^{1/2}e$ en el espacio nulo de la matriz $AD_v^{-1/2}D_x^{1/2}$.
5. La proposición 3, caracteriza la relación entre las variables complementarias Primal y Dual.

VI. DISCUSIÓN DE RESULTADOS

6.1 Contrastación de la Hipótesis

En el proyecto de investigación se formuló la hipótesis general de que, dado un Problema de Programación Lineal Estándar y una apropiada función barrera, entonces se puede deducir el sistema de ecuaciones no lineales que definen a las Trayectorias Solución del problema. Asimismo, se pudo generar en base a la hipótesis específica de que el Problema Barrera asociado (Pu) asociado a (P) al resolverse usando las condiciones de optimalidad, permite generar las condiciones y propiedades de continuidad de las Trayectorias Solución.

Por otro lado, ha sido posible arribar a los siguientes resultados no previstos

1. Si comenzamos lo suficientemente cerca de una solución óptima, entonces la trayectoria nos lleva aproximadamente en línea recta a la solución óptima. Es decir, esta característica es la diferencia con respecto a los algoritmos de reescalado lineal donde todas las rutas tienden a una sola dirección de aproximación a la solución óptima [16].

2. Con un cambio apropiado de parámetro $\mu = tan\theta$, se ha podido describir de una manera unica o unificada las trayectorias de los problemas de optimización.

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, 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.

VII. CONCLUSIONES

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

1. Se ha encontrado formulas explicitas para la Derivada K-ésima de la Trayectoria Solución como solución de un Sistema Lineal
2. Se ha probado la existencia de Trayectorias Solución continuas para Problemas de Programación Lineal.
3. Se ha probado la gran utilidad de la función barrera logarítmica en cuanto a la aplicación al Primal y Dual generando los problemas $(P_\mu)\ y\ (D_\mu)$ que están íntimamente relacionados como lo muestra la proposición 3

VIII. RECOMENDACIONES

Se recomienda

1. Continuar con estas ideas para aplicarlas al Problema de Complementariedad Lineal debido al punto de contacto que se encuentra con la función $F(z;\ \mu) = \mu\left(\frac{z^0 s^0}{z}\right) + Mz + q$
2. Asimismo, aplicarlas al Problema de Programación Cuadrática por estar íntimamente relacionada al problema de complementariedad.

IX. REFERENCIAS BIBLIOGRÁFICAS

[1] E. R. Barnes (1986) A Variation on Karmarkar`s Algorithm for Solving Linear Programming Problems. Mathematical Programming 36:174-182.

[2] D. A. Bayer and J. C. Lagarias, The nonlinear geometry of linear programming I: Affine and projective rescaling trajectories, AT&T Preprint.

[3] R. W. Cotle and G. B. Dantzig (1968), Complementary pivot theory of mathematical programming, in *Mathematics of Decision Sciences,* G. B. Dantzig and a. F. Veinott, Jr. (eds), American Mathematical Society, Providence, R. I., pp.115-136.

[4] G. B. Dantzig (1963) *Linear Programming and Extensions,* Princeton University Press, Princeton, N.J.,

[5] B.C. Eaves and H. Scarf (1976), The solution of systems of piecewise linear equations, *Math. Operations Res.* **1**, 1-27.

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

[7] K. R. Frish (1955), The logarithmic potential method of convex programming, unpublished manuscript, University Institute of Economics, Oslo, Norway.

[8] C. B. Garcia and W. I. Zangwill (1981), *Pathways to Solutions, Fixed Points and Equilibria,* Prentice-Hall, Englewood Cliffs, N. J.

[9] P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin, and M. H. Wright (1985), On projected Newton barrier methods for linear programming and equivalence to Karmarkar´s projective method, Technical Report SOL 85-11, Systems Optimization Laboratory, Department of Operations Research, Stanford, Calif.

[10] H. M. Greenberg, and J. E. Kalan (1972),Methods of feasible paths in nonlinear programming, Technical Report CP 72004, Computer Science/Operations Research Center, Southern Methodist University.

[11] P. Huard (1967), Resolution of mathematical programming with nonlinear constraints by the method of centers, in *Nonlinear Programming,* J. Abadie (ed.), North-Holland, Amsterdan, pp. 207-219.

[12] N. Karmarkar (1984), Anew polynomial-time algorithm for linear programming, *Combinatorica* **2** , 373-395.

[13] C. E. Lemke (1965),Bimatrix equilibrium points and mathematical programming, *Management Sci.* **11**, 681-689.

[14] C. E. Lemke (1968), On complementary pivot theory, in *Mathematics of Decision Sciences,* G. B. Dantzig and A. F. Veinott, Jr. (eds), American Mathematical Society, Providence, R. I., pp.95-114.

[15] O. Mangasarian (1984), Normal solutionsof linear programs, *Math. Programming Study* **22**, 206-216.

[16] N. Megido and M. Shub (1988), Boundary behavior of interior point algorithms for linear programming, *Math. Operations Res.* **13**,

[17] K. G. Murty (1985), A new interior variant of the gradient projection method for linear programming, Technical Paper 85-18, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor.

[18] J. L. Nazareth (1986), Homotopies in linear programming, *Algorithmica* **1,** 529-536.

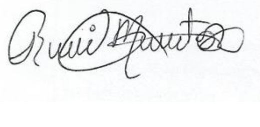
[19] H. E. Scarf (1967), The approximation of fixed points of continuous mappings, *SIAM, J. Appl. Math.* **15**, 1328-1343.

[20] S. Smale (1986), Talk at Mathematical Sciences Research Institute (MSRI), Berkeley, California.

[21] R. J. Vanderbei, M. J. Meketon, and B. A. Freedman (1986), A modification of Karmarkar´s linear programming algorithm, *Algorithmica* **1**, 395-409.

[22] J. H. Luna Valdez (2020): Programación lineal: un algoritmo primal-dual de paso largo usando el método de la función barrera, Tesis de maestría, UNMSM: https://hdl.handle.net/20.500.12672/16094

[23] P. E. Quijano Urbano (2019): Algoritmo primal - dual para el problema de programación lineal basado en el método de barrera logarítmica, Tesis de Maestría, UNMSM: https://ndle.net/20.500.12672/11456



ANEXOS

Matriz de consistencia

PROBLEMA	OBJETIVOS	HIPOTESIS	VARIABLES	METODOLOGÍA
Problema General ¿De qué manera se puede describir a las trayectorias solución de un Problema de Programación Lineal? **Problema Específico** ¿Qué propiedades tienen las trayectorias solución de un Problema de Programación Lineal?	**Objetivo general:** Describir a las trayectorias solución de un Problema de Programación Lineal **Objetivos específicos:** Deducir las propiedades que tienen las trayectorias solución de un Problema de Programación Lineal	**Hipótesis general:** Dado el Problema de Programación Lineal en forma Estandar (P) y una apropiada Función Barrera, entonces se puede deducir el Sistema de Ecuaciones no lineales que definen a las Trayectorias Solución **Hipótesis específicas:** Dado el Problema de Programación Lineal (P), su Problema de Barrera asociada $(P_u)\{$ maximizar $c^t x + u \sum \ln x_i$ sujeto a $Ax = b$ $x > 0\}$ y el Sistema No Lineal de ecuaciones que definen a la trayectoria solución, entonces el conjunto de soluciones de dicho Sistema No lineal brinda las propiedades de las Trayectorias Solución	**Variable Independiente:** Sistema de Ecuaciones No Lineales asociada al Problema de Programación Lineal **Variable Dependiente:** Trayectoria Solución	**Nivel de Investigación** Investigación científica aplicada **Tipo de Investigación:** Investigación aplicada, cuantitativa y transversal. **Diseño de la Investigación:** El proyecto está enmarcado en el tipo de investigación básica. El diseño de la investigación a desarrollar es teórico y consiste en obtener un esquema iterativo que nos permita caracterizar a las Trayectorias Solución de un Problema de Programación Lineal