



**UNIVERSIDAD NACIONAL
"PEDRO RUIZ GALLO"**

**FACULTAD DE CIENCIAS FÍSICAS
Y MATEMÁTICAS**



ESCUELA PROFESIONAL DE MATEMÁTICA

**"MÉTODOS DE PUNTOS INTERIORES APLICADO
A PPL CON OPTIMOS ALTERNATIVOS"**



UNIVERSIDAD NACIONAL "PEDRO RUIZ GALLO"	
OFICINA CENTRAL DE BIBLIOTECA PROCESOS TECNICOS	
N° DE INGRESO:	
CÓD. DE CLASIFICACIÓN:	

TESIS

**PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN MATEMÁTICAS**

Presentado por:

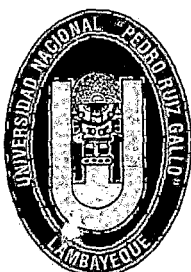
BACH. MAT. ESTRADA HUANCAS, IRIS ESTRELLA

Asesor:

LIC. MAT. CORNETERO CAPITÁN, JUAN ANTONIO

LAMBAYEQUE – PERÚ

2015



UNIVERSIDAD NACIONAL
“PEDRO RUIZ GALLO”



FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS

ESCUELA PROFESIONAL DE MATEMATICAS

**“MÉTODOS DE PUNTOS INTERIORES APLICADO A PPL CON
OPTIMOS ALTERNATIVOS”**

TESIS

**PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN MATEMÁTICAS**

PRESENTADO POR:

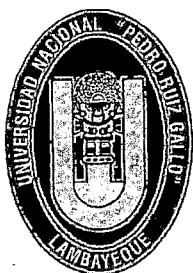
BACH. MAT. ESTRADA HUANCAS, IRIS ESTRELLA

ASESOR:

LIC.MAT. CORNETERO CAPITÁN, JUAN ANTONIO

**LAMBAYEQUE – PERÚ
2015**





UNIVERSIDAD NACIONAL
"PEDRO RUIZ GALLO"



FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS

ESCUELA PROFESIONAL DE MATEMATICAS

LOS FIRMANTES, POR LA PRESENTE CERTIFICAN QUE HAN LEÍDO Y RECOMIENDAN A LA FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS LA ACEPTACIÓN DE LA TESIS TITULADA "MÉTODOS DE PUNTOS INTERIORES APLICADO A PPL CON ÓPTIMOS ALTERNATIVOS", PRESENTADO POR LA BACH.MAT. ESTRADA HUANCAS IRIS ESTRELLA EN EL CUMPLIMIENTO PARCIAL DE LOS REQUISITOS NECESARIOS PARA LA OBTENCIÓN DEL TÍTULO PROFESIONAL DE LICENCIADO EN MATEMÁTICAS.



Dr. CÁRPENA VELÁSQUEZ ENRIQUE
Presidente Jurado de Tesis

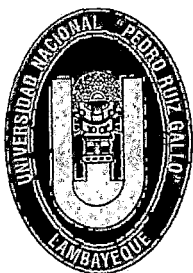


Mg. SÁNCHEZ GARCÍA DOLORES
Secretario Jurado de Tesis

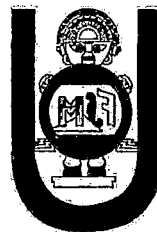


Lic. CORONADO JUÁREZ WILLIAM
Vocal Jurado de Tesis

LAMBAYEQUE, NOVIEMBRE DEL 2015



UNIVERSIDAD NACIONAL
“PEDRO RUIZ GALLO”



FACULTAD DE CIENCIAS FISICAS Y MATEMATICAS

ESCUELA PROFESIONAL DE MATEMATICAS

**“MÉTODOS DE PUNTOS INTERIORES APLICADO A PPL CON
OPTIMOS ALTERNATIVOS”**

TESIS

**PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN MATEMÁTICAS**

PRESENTADO POR:

Bach. Mat. IRIS ESTRELLA ESTRADA HUANCAS
AUTOR

Lic. JUAN ANTONIO CORNETERO CAPITÁN
ASESOR

LAMBAYEQUE – PERÚ
2015

AGRADECIMIENTOS

A Dios nuestro creador,
por su infinito amor y
guiar nuestros pasos

A mis padres Domingo y Elena por su amor,
apoyo constante, consejos y ejemplo.

Al Lic. Juan A. Cornetero Capitán
por su valioso asesoramiento en esta tesis.



DEDICATORIA

A mi hijo Juan José.

ÍNDICE GENERAL

Resumen	1
Introducción	2
1. Capítulo I. Preliminares	4
1.1. Problema De Programación Lineal (PPL)	4
1.2. Solución De Un Problema De Programación Lineal	7
1.3. Dualidad	9
1.4. El Algoritmo De Karmarkar	12
2. Capítulo II. Método De Puntos Interiores	17
2.1. Introducción a los Algoritmos de Punto Interiores	17
2.2. Exploración Algebraica	22
2.3. El Algoritmo Primal	29
2.4. El Algoritmo Dual	40
2.5. Convergencia De Los Algoritmos De Puntos Interiores	50
3. Capítulo III. Método de puntos interiores con óptimos alternativos	58
3.1. Método de Barrera Logarítmica Primal-Dual	58
3.2. Aplicación del Método a un PPL	63
3.3. Modificación al algoritmo de puntos interiores barrera logarítmica primal-dual para problemas con óptimos alternativos	68
Conclusiones	73
Bibliografía	74

RESUMEN

El método simplex nos da una solución óptima a Problemas de Programación Lineal (PPL), esta solución se encuentra en un extremo de la región factible, sin embargo puede resultar una solución no muy adecuada cuando se trata de problemas con múltiples soluciones.

De este conjunto de soluciones se puede determinar la solución más adecuada de acuerdo a determinadas condiciones del problema la que va a depender del vector de la dirección de búsqueda.

En este trabajo de tesis se presenta un análisis detallado del método de Puntos Interiores Primal y Dual mostrando sus pasos algorítmicos, que permite encontrar la solución de un Problema de Programación lineal, no necesariamente en los puntos extremos como lo hace el método Simplex. Así como se analiza su convergencia.

Así mismo se detalla el método de puntos interiores barrera logarítmica primal – dual, además la posibilidad de modificar la dirección de búsqueda en el algoritmo para ser aplicada a problemas de optimización con óptimos alternativos lo que permitirá obtener soluciones óptimas que satisfagan objetivos adicionales.

INTRODUCCIÓN

La programación lineal se desarrolló gracias a un grupo de investigadores de la Fuerza Aérea de los Estados Unidos, dirigidos por George **Dantzig**. Y su muy conocido *método Simplex* fue por más de tres décadas el único método utilizado para resolver los problemas de programación lineal (PPL) el cual se representa en su forma estándar de la siguiente forma:

$$\begin{array}{ll} \min & c^T x \\ \text{s a:} & Ax = b \\ & x \geq 0. \end{array}$$

Sin embargo el método Simplex no resulta muy eficiente para problemas con variables y parámetros numerosos, Esto por su característica combinatoria que se manifiesta en el elevado número de iteraciones.

Ante esta situación muchos investigadores se dedicaron a la búsqueda de algoritmos que no utilizaran la propiedad de combinatoria.

En 1979 Leonid **Khachiyan** desarrolló un algoritmo denominado Método elipsoidal, para determinar la solución de un sistema de ecuaciones lineales, que ya no tenía la propiedad combinatoria. Luego se han buscado otros métodos que mejoren el tiempo de ejecución polinomial del algoritmo elipsoidal.

En 1984 Narendra **Karmarkar** presenta un algoritmo de tiempo de ejecución considerablemente menor que el algoritmo elipsoidal conocido como: *Algoritmo de Karmarkar ó Método de transformación proyectiva*. A pesar de esto el algoritmo de Karmarkar requiere un tiempo de ejecución mayor que el algoritmo simplex cuando se

trata de problemas con pocas variables. Debido a esto se han propuesto muchas modificaciones.

El algoritmo de Karmarkar y las modificaciones realizadas, han dando lugar al *“Algoritmos de puntos interiores de programación lineal”*.

En la presente tesis se ha formulado el siguiente problema: ¿Es posible determinar un algoritmo del tipo de puntos interiores que permita determinar la solución de un PPL con óptimos alternativos?

En esta tesis se muestra que los algoritmos primal y dual de puntos interiores mejoran el algoritmo de Karmarkar ya que no necesitan la condición de tener un problema no degenerado.

En el Capítulo 1 se hace un resumen de los conceptos y métodos que serán utilizados para el desarrollo de esta tesis.

En el capítulo 2, se presenta las ideas básicas generales y la derivación de sus respectivos pasos algorítmicos del método de puntos interiores. Así como se analiza su convergencia.

En el capítulo III se presenta Método de puntos interiores barrera logarítmica primal-dual se determina el algoritmo y se analiza una modificación para resolver PPL con óptimos alternativos siguiendo alguna condición adicional.

Finalmente se concluye el presente trabajo con la presentación de las conclusiones y la bibliografía.

CAPÍTULO I: PRELIMINARES

1.1 Problema De Programación Lineal (PPL)

Un PPL se expresa mediante una función objetivo lineal y un conjunto de restricciones lineales en forma de desigualdades

La forma más general de un Problema de Programación Lineal (PPL) consiste en minimizar o maximizar.

$$Z = f(x) = \sum_{j=1}^n c_j x_j$$

Sujeto a:

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad i = 1, 2, \dots, p-1$$

$$\sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = 1, 2, \dots, p-1$$

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = q, \dots, m$$

Junto con las condiciones de no negatividad de x.

Donde p, q y m son enteros positivos tales que $1 \leq p \leq q \leq m$

Realizando los cambios adecuados podemos escribirlo en su **forma matricial** como sigue

$$\text{Minimizar} \quad z = c^T x$$

$$\text{Sujeto a:} \quad Ax \geq b$$

$$x_i \geq 0 \quad ; i = 1, \dots, n$$

Donde, A es una matriz de $(m \times n)$, $x \in R^n$, $c \in R^n$, $b \in R^m$.

Expresado en su **forma estándar** tiene la forma:

$$\begin{array}{ll}
 \text{Minimizar} & z = c^T x \\
 \text{Sujeto a} & Ax = b \\
 & x_i \geq 0 \quad ; i = 1, \dots, n
 \end{array} \dots\dots\dots (1.1)$$

Ejemplo: La forma estándar del PPL:

$$\begin{array}{ll}
 \text{Maximizar} & Z = 2x_1 - 3x_2 + 5x_3 \\
 \text{Sujeto a:} & x_1 + x_2 \leq 2 \\
 & 3x_1 + x_2 - x_3 \geq 3 \\
 & x_1, x_2 \geq 0
 \end{array}$$



$$\begin{array}{ll}
 \text{Minimizar} & Z = -2x_1 + 3x_2 - 5(x_6 - x_7) \\
 \text{s.a:} & x_1 + x_2 + x_4 = 2 \\
 & 3x_1 + x_2 - (x_6 - x_7) - x_5 = 3 \\
 & x_1, x_2, x_4, x_5, x_6, x_7 \geq 0
 \end{array}$$

1.1.1. Solución Factible

Es un punto $x = (x_1, x_2, \dots, x_n)$ que satisface todas las restricciones del PPL.

1.1.2. Región De Factibilidad O Polítopo.

Conjunto de todas las soluciones factibles S , es decir $S = \{x / Ax \leq b, x_i \geq 0\}$.

El **PPL**, mostrado en (1.1) puede formar parte de las siguientes tres categorías:

- a) Un **problema no factible**, es decir un problema para el cual no existen soluciones factibles, es decir, $S = \emptyset$.

Ejemplo: Considérese el siguiente PPL:

$$\begin{array}{ll}
 \text{Maximizar} & Z = 3x_1 + x_2 \\
 \text{sujeto a} & -x_1 + 2x_2 \leq 2 \\
 & 2x_1 + x_2 \leq 6 \\
 & x_1 \leq 3 \\
 & 2x_1 - 1.5x_2 \leq 4 \\
 & -x_1 - x_2 \leq -1 \\
 & x_1 + x_2 \leq 0 \\
 & x_1, x_2 \geq 0
 \end{array}$$

Este problema no es factible porque la nueva restricción $x_1 + x_2 \leq 0$ no es compatible con las restricciones previas.

b) Un **problema no acotado**, es aquel para el cual existen soluciones factibles pero la función objetivo no está acotada inferiormente.

Ejemplo: Maximizar $Z = 3x_1 + x_2$

Sujeto a:

$$\begin{array}{ll} -1.5x_1 + x_2 \leq 2 & \text{.....C1} \\ -x_1 - x_2 \leq -1 & \text{.....C2} \\ -x_1 \leq 0 & \text{.....C3} \\ -x_2 \leq 0 & \text{.....C4} \end{array}$$

El cual tiene una solución no acotada, porque como se muestra en la figura 1.1, la región factible no está acotada en la dirección de crecimiento de la función objetivo z .

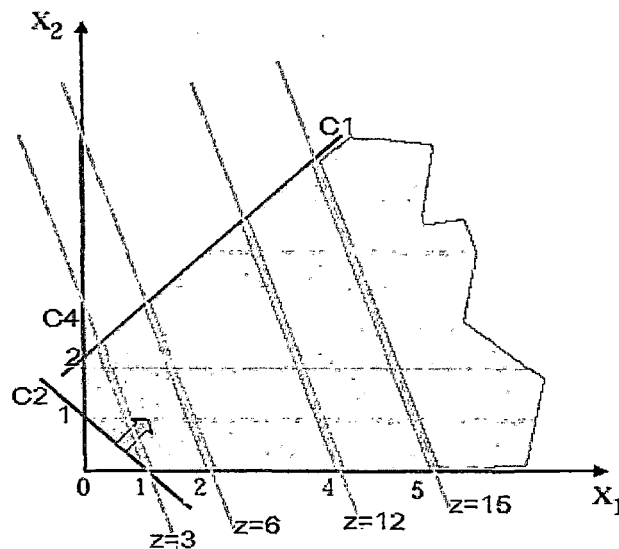


Fig. 1.1. PPL no acotado.

c) Problemas para el cual $S \neq \emptyset$ y además $c^T x$ es acotada inferiormente.

Ejemplo: PPL con solución única

Maximizar $Z = 3x_1 + x_2$

sujeto a:

$$\begin{array}{ll} -x_1 + x_2 \leq 2 & \text{.....C1} \\ x_1 + x_2 \leq 6 & \text{.....C2} \\ x_1 \leq 3 & \text{.....C3} \\ 2x_1 - x_2 \leq 4 & \text{.....C4} \\ -x_1 - x_2 \leq -1 & \text{.....C5} \\ x_1, x_2 \geq 0 & \end{array}$$

Cuya solución óptima es $x = (3,3)$, el cual se obtuvo por el método gráfico y se muestra en la figura 1.2.

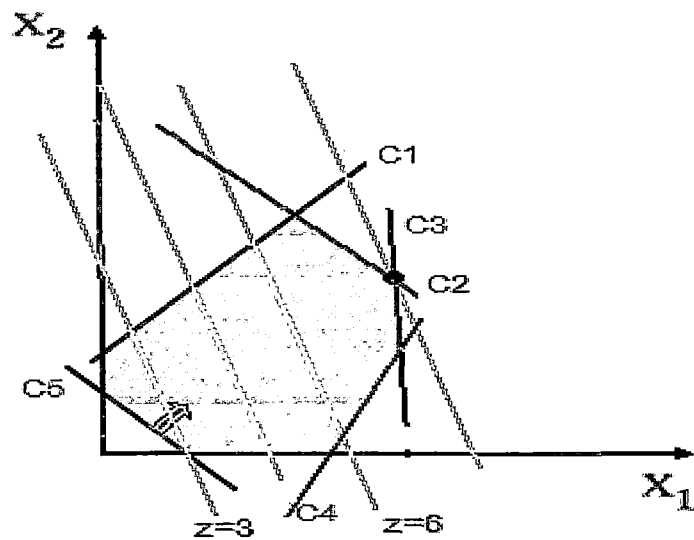


Fig. 1.2. PPL con solución única.

Ejemplo: PPL con soluciones múltiples

Maximizar $Z = x_1 + x_2$
 sujeto a:

- $-x_1 + x_2 \leq 2$ C1
- $x_1 + x_2 \leq 6$ C2
- $x_1 \leq 3$ C3
- $2x_1 - x_2 \leq 4$ C4
- $-x_1 - x_2 \leq -1$ C5
- $x_1, x_2 \geq 0$

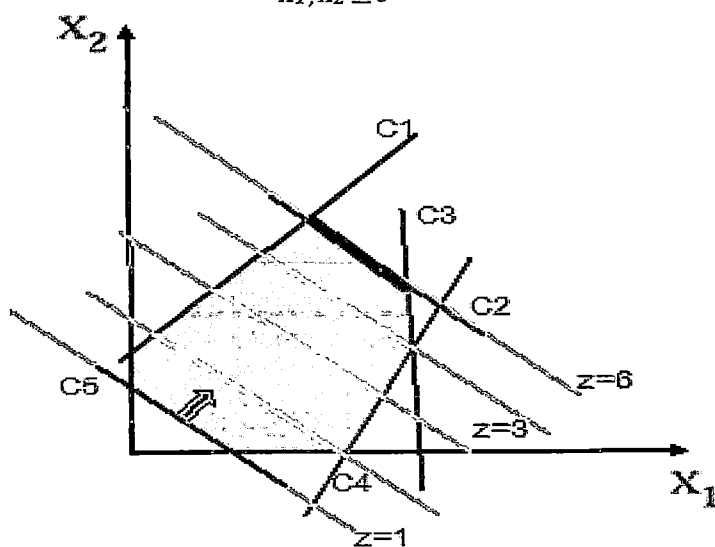


Fig. 1.3. PPL con soluciones múltiples.

1.2. Solución De Un Problema De Programación Lineal:

1.2.1. Solución óptima:

Una solución óptima del PPL (1.1) es una solución factible x^* , que cumple con la condición: $c^T x^* \leq c^T x = z, \forall x \in S$.

Si el problema es de maximización debe cumplir $z = c^T x \leq c^T x^*, \forall x \in S$.

En el problema (1.1), Sea A una matriz $m \times n$, de rango m , luego existe un número máximo m de columnas L.I. supongamos que son las columnas a_1, a_2, \dots, a_m las cuales forman una base que se denota por B, luego B es regular entonces:

$$A = [B : N],$$

B: matriz $m \times m$ formando las columnas L.I.

N: matriz $m \times (n - m)$ formando las columnas restantes.

De este modo se obtiene una partición en el vector " x " y en el vector de costo " c ", de esta manera:

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} ; \quad c = \begin{bmatrix} C_B \\ C_N \end{bmatrix}.$$

Donde x_B variables básicas, x_N variables no básicas.

Luego $Ax = b$, se escribe como: $Ax = [B : N] \begin{bmatrix} x_B \\ x_N \end{bmatrix}$

$$Ax = Bx_B + Nx_N = b$$

Como se tiene que B es regular: $x_B + B^{-1}Nx_N = B^{-1}b$

$$\text{si } x_N = 0 \Rightarrow x_B = B^{-1}b$$

Por tanto: $x = [x_B : 0]^T$ Es la solución asociada a la base B.



1.2.2. Solución Básica:

Si el vector de decisión, $x = (x_1, \dots, x_n)^T$, tiene $n-m$ componentes iguales a cero, y las m componentes están dadas por $B^{-1}b$, se llama solución básica del problema lineal.

$$x = (x_1, \dots, x_m, 0, 0, \dots, 0)^T$$

Si $x_B = B^{-1}b$ y x_B tiene todas sus componentes no negativas, tenemos una ***solución básica factible***.

Una solución básica factible, donde al menos una componente de x_B es igual a cero, se llama ***Solución Básica Factible Degenerada***. Los problemas que tienen este tipo de soluciones se llaman ***problemas degenerados***.

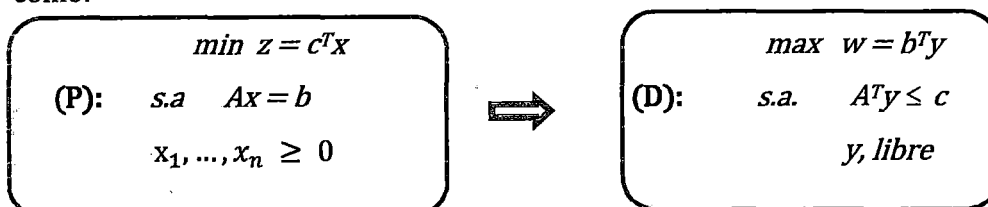
Una solución básica factible, donde m componentes del vector solución x , son positivas se llama Solución Básica Factible no Degenerada. Los problemas que tienen este tipo de soluciones se llaman ***problemas no degenerados***.

1.3. Dualidad:

Dado un PPL, llamado ***problema primal***, existe otro PPL llamado ***problema dual***, íntimamente relacionado con él. Se dice que ambos problemas son mutuamente duales. Bajo ciertas condiciones, los problemas primal y dual dan lugar al mismo valor óptimo de la función objetivo, y por tanto se puede resolver indirectamente el problema primal resolviendo el problema dual.

Muchas veces es conveniente trabajar con el Problema Dual e incluso se tienen algoritmos Primal – Dual.

Cuando el problema primal está en su forma estándar (P) el problema dual (D) se tiene como:



1.3.1. Teorema de Dualidad:

Si $x \in R^n$ es solución factible para el problema primal (P), $y \in R^n$ es solución factible para el problema dual (D), entonces $b^T y \leq c^T x$.

Demostración: De las restricciones de ambos (P) y (D) tenemos:

$$b^T y = (Ax)^T y = x^T A^T y$$

Puesto que $x_1, x_2, \dots, x_n \geq 0$ y $(A^T y \leq c)$ por la factibilidad del dual tenemos:

$$b^T y = x^T A^T y \leq x^T c = c^T x.$$

1.3.2. Teorema Fuerte de Dualidad:

Si el problema primal (P) tiene una solución óptima x^* , entonces el problema dual (D) tiene una solución óptima y^* y sus respectivos valores de las funciones objetivo son las mismas. Esto es el óptimo $c^T x^* = b^T y^*$.

Demostración: Consideremos

$$Ax^* = [B : N] \begin{bmatrix} x_B^* \\ x_N^* \end{bmatrix} = b \Rightarrow x^* = \begin{bmatrix} x_B^* \\ x_N^* \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}.$$

Usando el vector de costo $c = \begin{bmatrix} C_B \\ C_N \end{bmatrix}$ tenemos: $c^T x^* = \begin{bmatrix} C_B^T & C_N^T \end{bmatrix} \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} = C_B^T B^{-1}b.$

Como el punto es óptimo, por la condición de optimalidad, se tiene:

$$r = c_N - c_B^T B^{-1} N \geq 0. \quad \text{Donde } r \text{ es el vector de costo reducido.}$$

$$\text{El vector dual "y" es dado por } y = c_B^T B^{-1}$$

Veremos que el problema primal y dual tienen el mismo valor y que dicho vector es solución factible para el problema dual, así, $b^T y = c_B^T B^{-1} b = c^T x^*.$

Ahora para mostrar que el vector solución dual, y , es factible para el problema primal, mostraremos que $A^T y \leq c$, en efecto:

$$(A^T y)^T = y^T A = c_B^T B^{-1} [B : N] = [c_B^T : c_B^T B^{-1} N] \leq [c_B^T : c_N^T] = c^T$$

Esquema para transformar un problema primal a dual y viceversa.

PRIMAL <i>max</i>	→	DUAL <i>min</i>
restricción $i \leq b_i$	↔	variable $i \geq 0$
restricción $i = b_i$	↔	variable i , libre
restricción $i \geq b_i$	↔	variable $i \leq 0$
variable $j \geq 0$	↔	restricción $j \geq c_j$
variable j , libre	↔	restricción $j = c_j$
variable $j \leq 0$	↔	restricción $j \leq c_j$
<i>max</i> DUAL	←	<i>min</i> PRIMAL

Ejemplo:

1. Construir el Dual de:

$$\text{Max } x_1 + x_2 + 2x_3$$

$$\text{s.a. } 2x_1 - 2x_2 + x_3 \leq 12$$

$$x_1 + x_2 + x_3 = 8$$

$$x_1 \geq 3$$

$$x_1, x_2 \geq 0, x_3 \text{ libre}$$

Se tiene

$$A = \begin{bmatrix} 2 & -2 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}; \quad b = \begin{bmatrix} 12 \\ 8 \\ 3 \end{bmatrix} \quad \text{y} \quad c = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Luego el problema dual esta dado por:

$$\begin{array}{ll} \text{(D)} & \min \quad 12y_1 + 8y_2 + 3y_3 \\ & \text{s.a.} \quad 2y_1 + y_2 + y_3 \geq 1 \\ & \quad \quad -2y_1 + y_2 \geq 1 \\ & \quad \quad y_1 + y_2 = 2 \\ & \quad \quad y_1 \geq 0, y_2 \text{ libre}, y_3 \leq 0 \end{array}$$

1.4. El Algoritmo De Karmarkar.

Consideremos el problema lineal como sigue:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a.} \quad & x \in \Pi \subseteq \mathbb{R}^n. \\ & x_i \geq 0, \quad 1 \leq i \leq n \end{aligned} \tag{1.2}$$

Donde Π es la intersección de semiespacios de la forma: $Ax \geq b$.

El algoritmo de Karmarkar consiste en desplazarse por el interior de la región factible y no a lo largo de la frontera como es el caso del algoritmo simplex.

El algoritmo de Karmarkar considera en cada iteración el punto conocido como el centro de la región factible y determina el siguiente punto de costo mínimo siguiendo la dirección del vector gradiente negativo con el mejor desplazamiento posible manteniendo este nuevo punto en el interior de la región factible.

En cada paso se obtiene que el punto conocido se encuentra en el centro de la región introduciéndose una transformación proyectiva para transformar la región factible inicial en otra región, cuyo centro es la imagen del último punto conocido, luego en la nueva región factible determinamos el nuevo punto de menor costo y así sucesivamente se repite el proceso.

1.4.1. Forma estándar de karmarkar:

Karmarkar transforma todo PPL a la forma estándar de Karmarkar:

$$\text{Sea } \Delta = \left\{ x \in \mathbb{R}^n / x_1, x_2, \dots, x_n \geq 0, \sum_{j=1}^n x_j = 1 \right\}$$

El Simplex estándar (n-1) dimensional con centro $a_0 = \frac{e}{n}$, donde $e = (1, 1, \dots, 1)$

Sea $\Omega = N(A) = \{x \in \mathbb{R}^n / Ax = 0\}$ con A una matriz de orden $m \times n$ de números enteros,

$c \in \mathbb{R}^n$ y sea $x \in \Pi = \Omega \cap \Delta$ $x_1, x_2, \dots, x_n \geq 0$

Entonces el problema lineal estándar de Karmarkar es:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a} \quad & x \in \Pi = \Omega \cap \Delta. \\ & x_1, x_2, \dots, x_n \geq 0 \end{aligned}$$

o también:

$$\begin{aligned} \min \quad & z = c^T x \\ \text{s.a} \quad & Ax = 0 \\ & e \cdot x = 1 \\ & x_i \geq 0, \quad 1 \leq i \leq n \end{aligned} \quad \dots\dots\dots (1.3)$$

El punto $a_0 = (1/n, 1/n, \dots, 1/n)$, viene a ser el centro de la región factible de este problema.

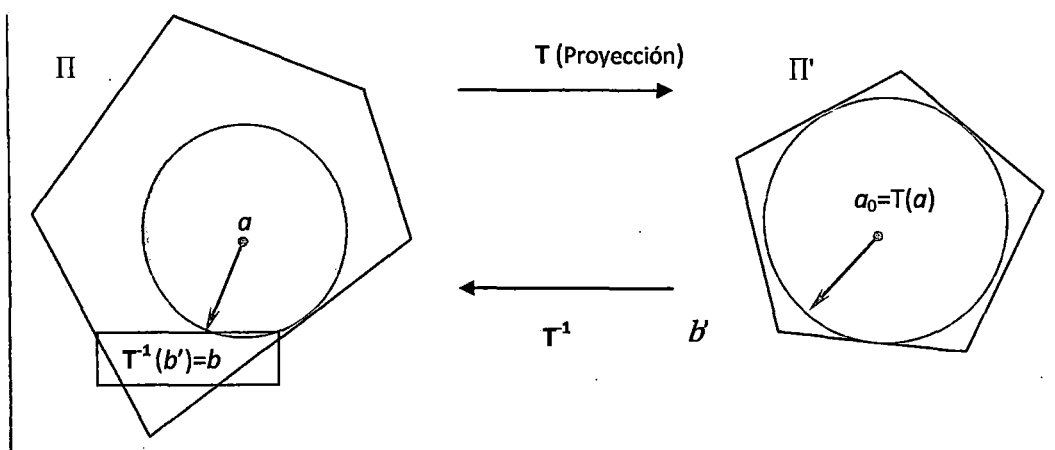


fig. 1.4 Punto inicial del Algoritmo de Karmarkar

El tamaño de un problema lineal dado en (1.1) está dado por

$$L = m n + \lceil \log_2 |P| \rceil + n \lceil \log_2 n \rceil.$$

Donde P es producto de todos los coeficientes no nulos de la función objetivo y de las restricciones y $\lceil \cdot \rceil$ es el número entero inmediato superior.

Al conjunto de operaciones para obtener un nuevo punto b a partir del punto a conocido en cada iteración se le denota por Φ , es decir $b = \Phi(a)$.

1.4.2. Etapas del algoritmo de Karmarkar:

1. Etapa inicial: Sea $x_0 = a_0 = \frac{e}{n}$ y sea $K = [2nL] / \delta$ (el número de iteraciones).
2. Etapa Iterativa: Para $k = 1$ a K , hallar $x(k) = \phi(x(k-1))$ la solución óptima que se obtiene es el punto v .
3. Criterios de parada:
 - Si $k < K$ y $cx(k) = 0$, $v = x(k)$
 - Si $k < K$ y $f(x(k)) > f(x(k-1)) - \delta$, el costo óptimo es positivo, v es un punto de costo no mayor que $cx(k)$.

donde f es la llamada función potencial definida como $f(x) = \ln \left[\frac{(c^T x)}{x_1 \dots x_n} \right]$ y permite controlar la convergencia al punto mínimo.

Calculo de $b = \Phi(a)$.

Donde $a \in \Pi$, de componentes no negativas, el vector salida $b = \phi(a)$ está en Π y $b > 0$. Si el costo mínimo es cero, entonces se tiene $c^T b = 0$ ó $f(b) \leq f(a) - \delta$.

1. Sea:

$$D = \begin{pmatrix} a_1 & 0 & 0 \dots 0 \\ 0 & a_2 & 0 \dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & 0 \dots a_n \end{pmatrix}, \forall x, \quad T_a(x) = \frac{D^{-1}x}{e^T D^{-1}x} \quad \text{y} \quad T_a^{-1}(x') = \frac{Dx'}{e^T Dx'}$$

2. Sea $c' = Dc$

3. Construir B aumentando una fila de unos a la matriz AD:

$$B = \begin{bmatrix} AD \\ e^T \end{bmatrix}$$

4. Sea c_p la proyección de c' sobre el espacio nulo de B,

$$c_p = c' - B^T(BB^T)^{-1}Bc'$$

5. Si $c_p = 0$, sea $\bar{c}_p = 0$, de otro modo sea $\bar{c}_p = \frac{c_p}{\|c_p\|}$ un vector unitario en la dirección de c_p .

6. Tome una longitud de paso $\frac{\alpha}{n}$ en dirección de $-c_p$ para obtener b' :

$$b' = a_0 - \frac{\alpha}{n} c_p$$

7. Finalmente se aplica T_a^{-1} a b' , para resolver b .

Retornar a:

$$b = T_a^{-1}(b') = \frac{Db'}{e^T Db'}.$$

Ejemplo: La forma estándar de Karmarkar.

Ejemplo:

$$\begin{array}{ll} \min. & x + y \\ \text{s.a} & x + y \geq 1 \\ & x - y \leq 1 \\ & x, y \geq 0 \end{array}$$

$$\begin{array}{ll} \min. & x + y \\ \text{s.a} & x + y - z = 1 \\ & x - y + u = 1 \\ & x, y, z, u \geq 0 \end{array}$$

donde A es de orden 2×4 , entonces $m=2$, $n=4$, por lo tanto su tamaño es $L=20$

Puesto que $L=20$, se tiene $x + y + z + u \leq n2^L = 4.2^{20}$, e introduciendo una variable de holgura se obtiene el nuevo problema.

$$\begin{array}{ll} \min & x + y \\ \text{s.a} & x + y - z = 1 \\ & x - y + u = 1 \\ & x + y + z + u + v = 2^{22} \\ & x, y, z, u, v \geq 0 \end{array}$$

Luego sustituyendo $2^{22} x' = x$, $2^{22} y' = y$, $2^{22} z' = z$, $2^{22} u' = u$, $2^{22} v' = v$, se tiene:

$$\begin{array}{ll} \min & 2^{22} x' + 2^{22} y' \\ \text{s.a} & 2^{22} x' + 2^{22} y' - 2^{22} z' = 1 \\ & 2^{22} x' - 2^{22} y' + 2^{22} u' = 1 \\ & x' + y' + z' + u' + v' = 1 \\ & x', y', z', u', v' \geq 0 \end{array}$$

Luego multiplicando la tercera restricción por " b " y restando del sistema anterior se obtiene:

$$\begin{array}{ll} Gx = 0 \\ x + y + z + u + v = 1 \\ x, y, z, u, v \geq 0 \end{array}$$

donde $G = \begin{pmatrix} 2^{22}-1 & 2^{22}-1 & -2^{22}-1 & -1 & -1 \\ 2^{22}-1 & -2^{22}-1 & -1 & 2^{22}-1 & -1 \end{pmatrix}$, y en lugar de x', y', z', u', v' se ha

escrito x, y, z, u, v .

y el problema se transforma en:

$$\begin{aligned} \min \quad & 2^{22}x + 2^{22}y \\ \text{s.a} \quad & (2^{22}-1)x + (2^{22}-1)y + (-2^{22}-1)z - u - v = 0 \\ & (2^{22}-1)x + (-2^{22}-1)y - z + (2^{22}-1)u - v = 0 \\ & x + y + z + u + v = 1 \\ & x, y, z, u, v \geq 0 \end{aligned}$$

Como $\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$ no es solución factible del sistema de restricciones mostrado anteriormente, para que lo sea se agrega una variable λ en las ecuaciones y se transforma en:

$$\begin{aligned} \min \quad & 2^{22}x + 2^{22}y + \lambda \\ \text{s.a} \quad & (2^{22}-1)x + (2^{22}-1)y + (-2^{22}-1)z - u - v + (-2^{22}+5)\lambda = 0 \\ & (2^{22}-1)x + (-2^{22}-1)y - z + (2^{22}-1)u - v + (-2^{22}+5)\lambda = 0 \\ & x + y + z + u + v + \lambda = 1 \\ & x, y, z, u, v, \lambda \geq 0 \end{aligned}$$

Que constituye el problema en la forma estándar de Karmarkar.

CAPÍTULO II. MÉTODO DE PUNTOS INTERIORES.

Los algoritmos de puntos interiores buscan aproximarse a la solución óptima a partir de un punto conocido que se encuentra en el interior de la región factible (polígono).

Consideremos un PPL en su forma estándar:

$$\begin{array}{ll} \min & z = c^T x \\ \text{s.a} & Ax = b \\ & x_i \geq 0, 1 \leq i \leq n \end{array} \dots\dots\dots (2.1)$$

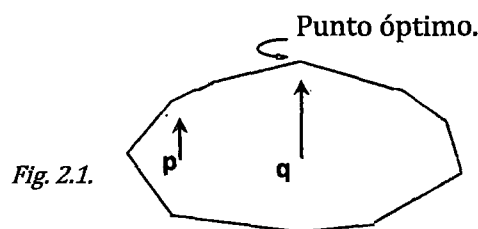
Donde $x \in R^n$, $b \in R^m$ y A es una matriz $m \times n$ de rango m .

Dado el sistema en (2.1) y un punto inicial factible e interior, lo que buscamos es obtener el punto mínimo, por tanto el desplazamiento debe ser en dirección descendente hasta llegar al siguiente punto. Este proceso debe repetirse hasta alcanzar un punto interior suficientemente cercano a la solución óptima.

2.1. INTRODUCCIÓN A LOS ALGORITMOS DE PUNTO INTERIORES:

Considerando que el problema es minimizar la función objetivo, y dado un punto inicial factible, entonces para conseguir una mejor aproximación al punto óptimo debemos seguir la dirección del vector gradiente negativo (la dirección del descenso más rápido) porque así se obtendrá el mejor decremento en el valor de la función objetivo.

Debemos tener en cuenta que el punto de partida afecta el progreso que se logra en cada paso, es decir, si uno empieza cerca de las caras del polígono el progreso esperado es pequeño, sin embargo si se empieza cerca del centro el progreso es considerable como se observa en la figura 2.1:





El polígono de la figura 2.1 es una ilustración de un PPL. y muestra el progreso hacia la solución óptima dentro de la región factible desplazándose en dirección opuesta al gradiente. Cualquier punto interior del polígono satisface las restricciones del problema y proporcionan una solución factible. Por tanto los límites de la región restringen todas las soluciones al interior de ella.

Además se observa que partiendo del punto q se llegue, mediante un desplazamiento más largo, rápidamente a la solución óptima; mientras que si partimos de p necesitamos mayor número de pasos cortos para llegar a la solución óptima.

A continuación consideremos un vector de paso (gradiente negativo de la función objetivo) con punto inicial en este punto. Luego efectuamos un desplazamiento a una nueva posición en la dirección del vector de paso tal que se reduzca el valor de la función objetivo. Esto es, tratamos de encontrar un vector de dirección para que el desplazamiento en busca de un nuevo punto sea en una dirección descendente y que sea lo más cercano posible al punto óptimo pero sin salirse de la región factible. Una forma de conseguir esto es inscribiendo una esfera en el octante $x \geq 0$ con centro en el punto p y con un radio lo más grande posible, pero manteniendo la factibilidad (que satisfagan las restricciones y el problema en general, esto se consigue con la esfera tangente por lo menos a un plano coordenado).

El punto actual está en el centro de esta esfera y se pueden hacer desplazamientos considerables hacia otro punto factible mientras se reduce el valor de la función objetivo. Puesto que el gradiente para éste problema es simplemente el vector c , el desplazamiento hacia la solución óptima (que se hace a lo largo de direcciones descendentes cuyos puntos están en dirección opuesta al gradiente) se efectúa en la dirección del vector $-c$, por lo tanto el nuevo punto es el punto donde un conjunto de nivel definido por $c^T x = \text{constante}$, es tangente al círculo inscrito.

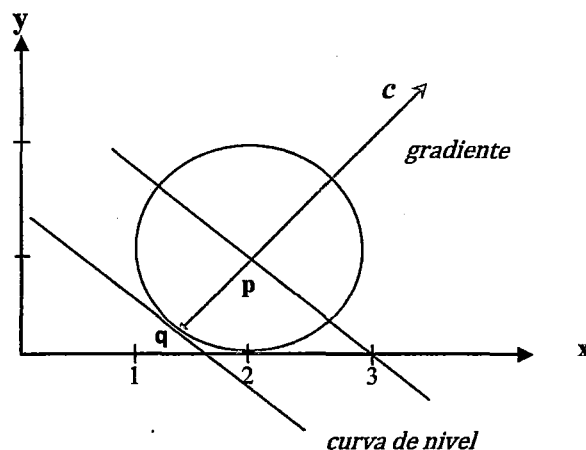


Figura 2.2: El círculo inscrito y el vector de paso.

Este paso se muestra en la figura 2.2, donde el actual punto se denota por p y el nuevo punto obtenido mediante éste procedimiento es denotado por q y examinando las componentes del nuevo punto q , vemos que el progreso es considerable a lo largo del eje vertical mientras que a lo largo del eje horizontal el progreso no es significativo. Es evidente que el mayor progreso se hace si el punto inicial estuviera a igual distancia de ambos ejes (igual distancia hacia las aristas en el caso general). Esto puede llevarse a cabo haciendo un cambio de escala (escalando) en el sistema de coordenadas de la figura 2.2 de tal manera que el círculo inscrito sea tangente a ambas paredes de la región factible.

En este espacio con una nueva *escala*, el movimiento en la dirección original descendente no es válida puesto que emplea el espacio original y no el nuevo espacio (escalado). Este cambio sin embargo es mínimo y puede ser manejado fácilmente.

En el ejemplo sólo el eje horizontal tuvo que escalarse por un factor 2 para trasladar el círculo a la posición deseada. Empleando el mismo factor para la componente horizontal del vector descendente llegamos a la situación que representa la figura 2.3.

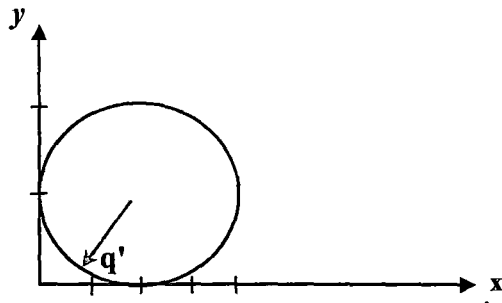


Figura 2.3: El círculo inscrito y su vector descendente en el espacio a escala.

Como se mostrará luego, escalar los ejes también equivale a escalar las restricciones de un problema de programación lineal. En realidad aplicando un escalamiento a los PPL nos resulta un nuevo problema, *"El problema de programación lineal escalado"*.

Aquí demostramos la importancia de que el punto que se elija esté a igual distancia de las paredes de la región factible, para poder obtener un buen desplazamiento hacia la solución óptima, todo esto mediante los conceptos de escalamiento y la transformación de los PPL.

No obstante, desde que los problemas originales se formularon en un diferente espacio tenemos que desarrollarlos dando un significado a la solución en ése espacio original, para hacer esto **reescalamos** la solución para llevarla de vuelta al espacio de origen, tal operación semejante a una operación de escalamiento es realizada a través de una *transformación afín*.

En nuestro caso se tomó un círculo en el espacio a escala y escalando sus ejes resulta una elipse en el espacio original (sin escala). Estas operaciones de reescalamiento se muestran en la figura 2.4, en dimensiones superiores, este reescalamiento transforma una esfera en un elipsoide. La ilustración representa un PPL en el espacio original con vectores dirección descendente del espacio escalado.

Volviendo al ejemplo, de acuerdo a la figura 2.2 en el sistema original, dado el punto inicial en la posición p y siguiendo la dirección del vector gradiente negativo, lo mejor

que se puede hacer es desplazarse al punto q , mientras que haciendo un cambio de escala, el mejor desplazamiento es al punto c , como se aprecia en la figura 2.4 después del reescalamiento. Luego si el siguiente punto en el espacio original no está a igual distancia de las paredes de la región factible, nuevamente realizar la operación anterior (escalamiento - reescalamiento) y así sucesivamente se repite el paso hasta llegar a la solución óptima.

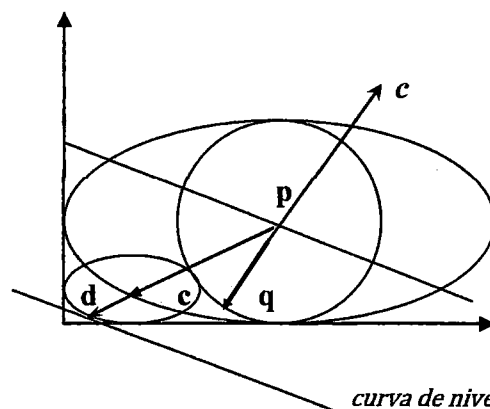


Figura 2.4: El elipsoide inscrito en el espacio original.

En la figura 2.4 la elipse más grande es la que se forma como consecuencia del reescalamiento que se hace. El punto d muestra la iteración siguiente que se realiza inscribiendo otro elipsoide centrado en c .

La discusión en éste capítulo, motiva e ilustra la importancia de centrar el punto y la elección de una dirección descendente. El centrado y una buena dirección descendente nos proporciona el gran potencial para lograr una buena reducción de costo en un solo paso. En el ejemplo se tomó un punto arbitrario (interior y factible) y se escaló hasta que estuvo centrado. Para llegar a la solución óptima se repite este proceso.

En resumen los algoritmos de puntos interiores consiste en llevar a cabo una secuencia de pasos donde un punto inicial, primero es centrado mediante un escalamiento, luego sigue un paso de mayor longitud posible en dirección descendente, obteniendo así un nuevo punto que finalmente mediante una operación de reescalamiento se trae de vuelta

el nuevo punto al espacio original. Esta clase de algoritmos se denomina también algoritmos de escala-afín.

2.2. ANÁLISIS ALGEBRAICO DEL MÉTODO:

Asumimos que x_0 de *componentes positivas* es un punto inicial (interior y factible) para el problema (2.1), entonces: $Ax_0 = b$

Enseguida queremos encontrar un vector de dirección, dx , el que tomamos en una dirección descendente hacia un nuevo punto x_N , manteniendo la factibilidad, como ya vimos se puede tomar en la dirección negativa del gradiente.

Veamos: Si el nuevo punto se obtiene mediante: $x_N = x_0 + dx$, Entonces por la factibilidad se deduce que: $Ax_N = A(x_0 + dx) = Ax_0 + Adx = b$.

Por la factibilidad de x_0 ($Ax_0 = b$), se obtiene $Adx = 0$

Si el vector de dirección se toma en dirección descendente se cumple la siguiente condición:

$$c^T x_N = c^T (x_0 + dx) = c^T x_0 + c^T dx \leq c^T x_0 \Rightarrow c^T dx \leq 0$$

Así, tenemos dos requisitos que el vector de paso (de dirección), dx , debe satisfacer.

$$Adx = 0 \quad \text{y} \quad c^T dx \leq 0 \quad (2.2)$$

Considerando que se debe dar un paso en la dirección opuesta al gradiente, claro está que esta dirección es válida solo si la factibilidad del punto actual se mantiene, y si esto proporciona una dirección descendente, es decir debe satisfacer (2.2) de lo contrario tiene que ser modificado.

A continuación veamos como modificar el gradiente de dirección de modo que se cumpla (2.2) para esto incluimos el operador proyección de la siguiente manera:

Un operador proyección hacia el espacio nulo de A ($N(A) = \{x / Ax = 0; x \in R^n\}$) está dado por la matriz " P ", $n \times n$, donde:

$$P = I_n - A^T(AA^T)^{-1}A \quad (2.3)$$

Este se denomina un operador proyección puesto que cuando opera sobre algún vector n -dimensional le corresponde otro vector que pertenece al espacio nulo de la matriz A , Veamos:

$$\begin{aligned} \text{Si } y = Pz, \Rightarrow Ay = APz &= A[I_n - A^T(AA^T)^{-1}A]z = A[z - A^T(AA^T)^{-1}Az] \\ &= Az - AA^T(AA^T)^{-1}Az = Az - Az = 0. \end{aligned}$$

$\therefore Ay = 0$, es decir $y \in N(A)$.

La matriz P por satisfacer $AP = 0$, tiene las siguientes propiedades:

$$P = P^T \text{ y } P = P^2 \quad (2.4)$$

En efecto

$$\begin{aligned} P^T &= [I_n - A^T(AA^T)^{-1}A]^T = I_n - [A^T(AA^T)^{-1}A]^T = I_n - A^T[(AA^T)^{-1}]^T A \\ &= I_n - A^T[(AA^T)^T]^{-1}A = I_n - A^T(AA^T)^{-1}A = P \end{aligned}$$

• Probaremos que

$$P - P^2 = 0; \forall A \neq 0.$$

$$A(P - P^2) = AP - AP^2 = AP - APP = 0 - 0P = 0, \text{ Como } A \neq 0 \Rightarrow P - P^2 = 0$$

$$\therefore P = P^2.$$

Ahora usamos este operador proyección para proyectar $-c$ sobre el espacio nulo de A , así encontramos el vector de dirección a través de:

$$dx = P(-c)$$

Así usando las dos propiedades de (2.4), tenemos:

$$c^T dx = -c^T P c = -c^T P^2 c = -c^T P P c = -c^T P^T P c = -(P c)^T P c = -\|Pc\|^2 \leq 0$$

Esto indica que es una dirección descendente como requiere (2.2).

Por tanto dado un punto inicial, proyectando el gradiente sobre el espacio nulo de A y tomando un paso en la dirección opuesta que viene a ser una dirección descendente se

obtiene de acuerdo a (2.2) un punto que satisface la factibilidad. Esta idea se ilustra en la figura 2.5.

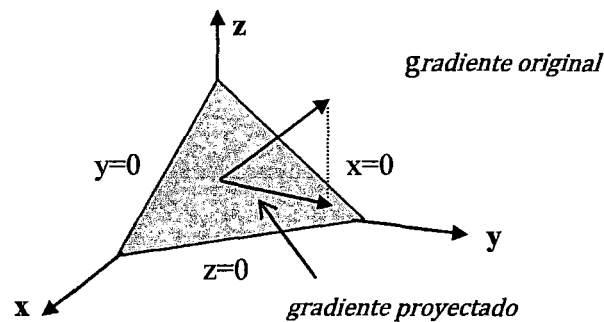


Figura 2.5: Proyección del gradiente.

Hasta ahora hemos desarrollado los requisitos necesarios para que un vector de dirección efectúe un único paso.

Ahora desarrollaremos los elementos de escalamiento.

Empecemos con un vector, $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, luego sus componentes son escaladas de algún modo para ser sustituidos por el vector escalado \mathbf{x}_1 , de tal manera que se ubiquen a igual distancia de todas las paredes del polítopo.

Retomando el problema (2.1) donde las paredes del polítopo son los planos coordenados, conseguiremos nuestro objetivo utilizando la matriz:

$$\mathbf{D} = \text{diag}(\mathbf{x}) = \begin{pmatrix} x_1 & 0 & 0 \dots 0 \\ 0 & x_2 & 0 \dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & 0 \dots x_n \end{pmatrix}_{n \times n}$$

Pues si definimos:

$$\mathbf{x}_1 = \mathbf{D}^{-1} \mathbf{x}$$

Se obtiene el nuevo punto $x_1 = (1, 1, \dots, 1)^T$ que se encuentra a la misma distancia de todas las paredes. Esto nos indica que la matriz diagonal viene a ser la matriz de cambio de escala, es decir, el escalamiento se forma a partir de las componentes del vector original (sin escalar) x . Puesto que el vector solución x está en el interior del polígono, es decir, $x > 0$, los elementos de la diagonal de la matriz D son estrictamente positivos, por tanto D es invertible.

Utilizando esta nueva matriz de cambio de escala en el PPL (2.1), obtenemos un *problema lineal escalado*, dado por:

$$\begin{aligned} \min \quad & c_1^T x_1 \\ \text{s.a.} \quad & A_1 x_1 = b \\ & x_1 \geq 0 \end{aligned} \quad (2.5)$$

Donde, $x_1 \in R^n$, $b \in R^m$, $A_1 = AD$ y $c_1 = Dc$

Con esto, cualquier vector interior factible inicial del problema (2.1) es transformado a un vector inicial factible de (2.5), ubicado a una unidad de distancia de las paredes del polígono (planos coordenados).

Ahora mediante un ejemplo ilustraremos el procedimiento expuesto anteriormente:

EJEMPLO:

$$\begin{aligned} \max \quad & z = x_1 + x_2 \\ \text{s.a.} \quad & 2x_1 + x_2 \leq 24 \\ & x_2 \leq 8 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Convirtiendo a la forma estándar y minimización se tiene:

$$\begin{aligned} \min \quad & h = -x_1 - x_2 \\ \text{s.a.} \quad & 2x_1 + x_2 + \bar{x}_3 = 24 \\ & x_2 + \bar{x}_4 = 8 \\ & x_1, x_2, \bar{x}_3, \bar{x}_4 \geq 0 \end{aligned}$$

donde

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 24 \\ 8 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

Se elige un punto inicial factible para este problema de forma arbitraria, una posible elección es empezar el problema con:

$$\mathbf{x}_0 = \begin{bmatrix} 2 \\ 4 \\ 16 \\ 4 \end{bmatrix}$$

\mathbf{x}_0

Centramos este punto, para esto se hace uso de la matriz de escalamiento \mathbf{D} , dado por:

$$\mathbf{D} = \text{diag}(\mathbf{x}_0) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

y el nuevo punto inicial escalado \mathbf{x}_1 , está a una unidad de distancia de las paredes del politopo, además \mathbf{x} y \mathbf{x}_1 se relacionan de la siguiente manera:

$$\mathbf{x} = \mathbf{D} \mathbf{x}_1$$

$$\mathbf{A}_1 = \mathbf{A}\mathbf{D} = \begin{pmatrix} 4 & 4 & 16 & 0 \\ 0 & 4 & 0 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 24 \\ 8 \end{pmatrix}, \quad \mathbf{c}_1 = \mathbf{D}\mathbf{c} = \begin{pmatrix} -2 \\ -4 \\ 0 \\ 0 \end{pmatrix}$$

Entonces resulta un problema de programación lineal escalado dado por:

$$\begin{aligned} \min \quad & h = -2x_1^1 - 4x_1^2 \\ \text{s.a} \quad & 4x_1^1 + 4x_1^2 + 16x_1^3 = 24, \\ & 4x_1^2 + 4x_1^4 = 8 \end{aligned}$$

cuyo punto inicial es: $\mathbf{x}_1^0 = (1, 1, 1, 1)^T$

Nótese que la región factible en el espacio original se muestra en la fig. 2.6 mientras que la región factible en el espacio escalado se muestra en la fig. 2.7.

Para determinar la dirección y la longitud del desplazamiento hacia el nuevo punto, a partir del punto inicial \mathbf{x}_1^0 , en el espacio escalado, proyectamos el vector gradiente sobre el espacio nulo de la matriz escalada \mathbf{A}_1 .

El operador proyección P_1 , y el gradiente proyectado, c_p , se muestran a continuación:

$$P_1 = I_n - A_1^T (A_1 A_1^T)^{-1} A_1 = \begin{pmatrix} 0.9429 & -0.0286 & 0.2286 & 0.0286 \\ -0.0286 & 0.4857 & -0.1143 & -0.4857 \\ -0.2286 & -0.1143 & 0.0857 & 0.1143 \\ 0.0286 & -0.4857 & 0.1143 & 0.4857 \end{pmatrix}$$

$$c_p = P_1 c_1 = \begin{pmatrix} -1.7714 \\ -1.8857 \\ 0.9143 \\ 1.8857 \end{pmatrix}, \text{ además } A_1 P_1 = 0.$$

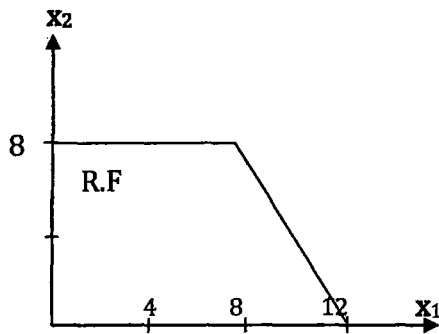


Figura 2.6:

La región factible en el espacio original.

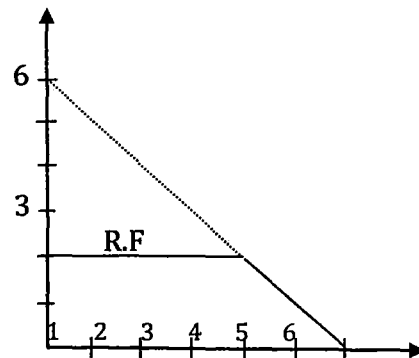


Figura 2.7:

La región factible en el espacio escalado.

Puesto que estamos resolviendo un problema de minimización, nos desplazamos en dirección opuesta al gradiente proyectado c_p , por tanto el vector de dirección dx_1 , es dado

por:

$$dx_1 = -c_p = \begin{pmatrix} 1.7714 \\ 1.8857 \\ -0.9143 \\ -1.8857 \end{pmatrix}$$

Para movernos de nuestra posición actual x_1^0 , al nuevo punto x_1^1 se tiene $x_1^1 = x_1^0 + dx_1$ pero como además queremos que:

$$x_1^1 = x_1^0 + dx_1 > 0, \text{ es decir } \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1.7714 \\ 1.8857 \\ -0.9143 \\ -1.8857 \end{pmatrix} > 0, \text{ nunca cero, porque de lo contrario}$$

el punto estaría en la frontera y no en el interior, introducimos para esto un factor de paso α , que controla la no negatividad de las restricciones.

Entonces el nuevo punto es: $x_1^1 = x_1^0 + \alpha dx_1$

donde α es una constante aún indeterminada, cuya finalidad se explica a continuación.

Nótese que podemos tomar un paso tan grande como queramos a lo largo de dx_1 , -esto significa que el operador proyección hace que el vector de paso sea lo más grande posible- y siguiendo la dirección del gradiente proyectado mantenemos la factibilidad, pero todavía debemos ocuparnos de mirar la no negatividad de las restricciones, para asegurarnos de esto se incluye *el factor de paso* α .

Observando las componentes del vector de dirección dx_1 vemos que hay componentes negativas, lo que causa al menos una componente cero. Para evitar esta posibilidad tomamos una fracción del máximo paso admisible. Esta fracción es llamada "*factor de paso*" y para este ejemplo consideramos al 95% del máximo paso admisible.

Con esta modificación tenemos $\alpha = \frac{0.95}{1.8857} = 0.5038$,

Ahora el nuevo punto sería: $x_1^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 0.5038 \begin{pmatrix} 1.7714 \\ 1.8857 \\ -0.9143 \\ -1.8857 \end{pmatrix} = \begin{pmatrix} 1.8294 \\ 1.9500 \\ 0.5394 \\ 0.0050 \end{pmatrix}$

Usando la matriz de escalamiento D , reescalamos este vector para tener el nuevo punto en el espacio original, para encontrar el nuevo valor del vector solución, x_1 , al final de la primera iteración.

$$x_1 = D x_1^1 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1.8294 \\ 1.9500 \\ 0.5394 \\ 0.0050 \end{pmatrix} = \begin{pmatrix} 3.7848 \\ 7.8000 \\ 8.6303 \\ 0.2000 \end{pmatrix}$$

Esto es importante para investigar los cambios en el valor de la función objetivo durante esta primera iteración, el valor de z en el punto inicial x_0 , es $c^T x_0 = 6$, mientras que con el nuevo vector solución x_1 , el valor de la función objetivo es: $c^T x = 11.5848$.

Puesto que la solución óptima para este problema es:

$x^* = [8 \ 8 \ 0 \ 0]^T$, con $c^T x^* = 16$, observamos que el progreso es considerable. Luego el proceso a seguir empieza con una nueva matriz escalada creada con el actual vector solución x_1 y se sigue con la misma secuencia de pasos realizada anteriormente.

2.3. EL ALGORITMO PRIMAL:

Una variante del algoritmo de Karmarkar viene a ser el algoritmo primal (escala-afín) de puntos interiores. Este algoritmo es fácil de implementar y sirve para resolver problemas mas generales de programación lineal.

2.3.1 ESCALAMIENTO Y DETERMINACIÓN DEL VECTOR DE PASO:

Como se mostró anteriormente, el escalamiento nos permite ubicar el punto actual de modo que esté a igual distancia de las paredes del polítopo. Este escalamiento se realiza usando una transformación afín. Tal transformación se define como sigue:

Sea Q una matriz no singular $n \times n$ y " v " un vector n -dimensional. La transformación definida como $T = v + Qx$, para algún $x \in \mathbb{R}^n$ es llamada una transformación afín.

Puesto que Q es no singular, T es una transformación invertible, y su inversa es también una transformación afín. En la sección anterior discutimos los efectos de tal transformación sobre el círculo inscrito. Recordar que la esfera unitaria es definida por:

$$X = \{x / x^T x \leq 1\}.$$

Entonces si T es una transformación afín aplicada sobre X el resultado $T(X)$ es un elipsoide. Dado un vector inicial, $x = [x_1, x_2, \dots, x_n]^T$, sus componentes son escaladas de algún modo para producir el vector escalado x_1 , cuyas componentes están a igual

distancia de las paredes del polítopo. Situando el vector escalado x_1 , a una unidad de distancia de las paredes nos lleva a la siguiente relación de escalamiento:

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad D = \begin{pmatrix} x_1 & 0 & 0 \dots 0 \\ 0 & x_2 & 0 \dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & 0 \dots x_n \end{pmatrix}$$

Con estas definiciones es fácilmente verificable que el vector original x y el vector escalado x_1 se relacionan como sigue:

$$x_1 = D^{-1}x$$

Desde que el vector solución x es siempre interior al polítopo, esto es $x > 0$, D es invertible.

Escalando el problema (2.1) inducimos un problema lineal escalado dado por:

$$\begin{aligned} \min \quad & c_1^T x_1 \\ \text{s.a.} \quad & A_1 x_1 = b \\ & x_1 \geq 0 \end{aligned} \quad (2.5)$$

donde: $x_1 \in R^n, b \in R^m$.

$$A_1 = AD, \quad c_1 = Dc$$

El vector inicial, x_1 , para el problema escalado se encuentra a una unidad de distancia de las paredes del polítopo, para algún vector inicial interior y factible x del problema original sin escalar.

El operador proyección para el problema lineal escalado es dado por una matriz simétrica P_1 $n \times n$, donde:

$$P_1 = I_n - A_1^T (A_1 A_1^T)^{-1} A_1 \quad (2.6)$$

El vector de dirección dx_1 para el problema escalado está dado por el negativo del vector que se obtiene proyectando el gradiente c_1 del problema escalado sobre el espacio nulo de A_1 , y luego para expresar en términos de los datos iniciales se usa (2.19), es decir:

$$\begin{aligned} dx_1 &= -P_1 c_1 = -[I_n - A_1^T (A_1 A_1^T)^{-1} A_1] c_1 = -[I_n - A_1^T (A_1 A_1^T)^{-1} A_1] Dc \\ &= -Dc + [(AD)^T (AD(AD)^T)^{-1} AD] Dc = -Dc + DA^T [AD DA^T]^{-1} AD^2 c \end{aligned}$$



$$= -Dc + DA^T [AD^2A^T]^{-1} AD^2c = -D [c - A^T (AD^2A^T)^{-1} AD^2c]$$

Haciendo:

$$(AD^2A^T) y = AD^2c \Rightarrow y = (AD^2A^T)^{-1} AD^2c \quad (2.7)$$

se tiene:

$$dx_1 = -D [c - A^T y]$$

Volviendo al espacio original (sin escalar) el vector de dirección dx , para la nueva iteración es dado por:

$$dx = D dx_1 = -D^2 [c - A^T y]$$

y si ahora se define una nueva variable $z = c - A^T y$, la expresión para el vector de dirección será:

$$dx = -D^2 z \quad (2.8)$$

Note que las operaciones de escalamiento y reescalamiento son parte de la secuencia de pasos algorítmicos conduciendo a la derivación del vector de dirección dx , mostrado en (2.8).

El vector " y " obtenido en (2.7) proporciona un estimado del vector solución para el problema dual, y el vector z proporciona un estimado del vector de costo reducido.

2.3.2 DETERMINACIÓN DE UN NUEVO PUNTO EN EL INTERIOR:

Con el vector dirección dado por $dx = -D^2 z$ tomamos un paso en esa dirección y se obtiene la siguiente iteración del vector solución x , esto se observa de la fórmula dada por:

$$x = x_0 + dx$$

Desde que el vector de dirección satisface $c^T dx \leq 0$ (condición descendente), la nueva iteración del vector solución x , satisface las restricciones $Ax = b$ y para cumplir con las restricciones de no negatividad del vector solución x , establecemos un tamaño de paso α , entonces x resulta:

$$x = x_0 + \alpha dx, \quad \alpha > 0$$

Denotando como $x_0(i)$ y dx_i la i -ésima componente de x_0 y dx respectivamente, tenemos que encontrar un α que cumpla para cada componente con:

$$x_i = x_0(i) + \alpha dx_i \geq 0, \quad \forall \quad 1 \leq i \leq n$$

Nótese que solo las componentes negativas del vector de paso dx tienen la potencialidad de violar la no negatividad de las restricciones, entonces este máximo paso admisible se encuentra mediante la siguiente expresión:

$$\alpha = \min \left\{ -\frac{x_i}{dx_i} : \forall \quad 1 \leq i \leq n \text{ y } dx_i < 0 \right\} \quad (2.9)$$

donde x_i es la i -ésima componente del actual vector solución x y dx_i es la i -ésima componente del vector de paso dx

Para evitar que alguna componente de la nueva iteración llegue a alcanzar el límite ó frontera del politopo dando lugar a que el vector solución no esté en el interior de la región factible, tomamos solo una fracción del tamaño de paso, multiplicando a α por un factor ρ , así la nueva iteración queda como sigue:

$$x = x_0 + \rho \alpha dx, \quad \alpha > 0 \text{ y } 0 < \rho < 1. \quad (2.10)$$

En la práctica el valor de ρ está en el rango 0.95 a 0.995. Tenemos que mostrar una secuencia de pasos necesarios para reducir el valor de la función objetivo manteniendo la factibilidad.

2.3.3 PASO INICIAL :

El algoritmo primal necesita un vector inicial x_0 positivo y que satisfaga las restricciones del problema primal en su forma estándar es decir que cumpla:

$$A x_0 = b, \quad x_0 > 0$$

Pero cuando se da un vector inicial positivo x_0 que no satisfaga las restricciones, se procede a evaluar el *error residual* ρ definido por:

$$\rho = b - A x_0 \quad (2.11)$$

Luego añadimos este error residual como una última columna de la matriz A y se obtiene la matriz modificada:

$$A_1 = [A \mid \rho]$$

Con esta modificación el número de columnas de A se incrementa en uno, y el vector solución inicial x_0 tiene una componente adicional 1. Si se denota por x_1 al nuevo vector inicial, se verifica fácilmente que:

$$A_1 x_1 = [A \mid \rho] \begin{pmatrix} x_0 \\ 1 \end{pmatrix} = A x_0 + \rho = b \quad (2.12)$$

Con lo cual tenemos un camino para encontrar un vector inicial factible para el sistema aumentado de restricciones $A_1 x_1 = b$, para algún vector inicial positivo x_0 .

Luego tenemos que modificar el problema de optimización inicial para reflejar estos cambios. Al añadir una columna a la matriz A , debemos agregar otra componente al vector de costos c :

Ya que la columna extra añadida a la matriz A representa el error residual, queremos poner una penalización sobre aquella componente del vector objetivo para conducirla a cero. Con esta penalización denotada por un escalar M resulta un problema de optimización modificado (conocido también como el método de penalización ó de gran M) resultando:

$$\begin{aligned} \min \quad & [c^T \mid M] x_1 \\ \text{s.a.} \quad & [A \mid \rho] x_1 = b, \\ & x_1 \geq 0 \end{aligned} \quad (2.13)$$

donde $x_1 \in R^{n+1}$, $b \in R^m$, $\rho = b - Ax_0$ y el nuevo vector solución inicial y factible, x_1^0 , para este problema es dado por:

$$x_1^0 = \begin{pmatrix} x_0 \\ 1 \end{pmatrix}, \text{ para algún } x_0 > 0, \text{ donde}$$

La constante de penalización M es considerada un número positivo grande, el problema modificado tiene una columna adicional en la matriz A y la solución así como sus vectores objetivos tienen una componente extra. Si el problema primal es factible, entonces los pasos seguidos por el algoritmo primal nos lleva a que la última

componente del vector solución modificado x_1^0 es cero. Sin embargo, si el proceso de solución no conduce a un vector con la última componente cero, implica que el problema primal no es factible.

EJEMPLO: Dado

$$\begin{aligned} \max \quad & z = x_1 + x_2 \\ \text{s.a.} \quad & 2x_1 + x_2 \leq 24 \\ & x_2 \leq 8 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Obtenemos

$$\begin{aligned} \max \quad & z = x_1 + x_2 \\ \text{s.a.} \quad & 2x_1 + x_2 + \bar{x}_3 = 24 \\ & x_2 + \bar{x}_4 = 8 \\ & x_1, x_2, \bar{x}_3, \bar{x}_4 \geq 0 \end{aligned}$$

Si $x_0 = (1 \ 1 \ 2 \ 2)^T$ arbitrario, entonces:

$$b - Ax_0 = \begin{pmatrix} 24 \\ 8 \end{pmatrix} - \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 19 \\ 5 \end{pmatrix} = \rho$$

$$\text{Si } A_1 = \begin{pmatrix} 2 & 1 & 1 & 0 & 19 \\ 0 & 1 & 0 & 1 & 5 \end{pmatrix}, \quad x_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \text{ entonces } A_1 x_1 = \begin{pmatrix} 24 \\ 8 \end{pmatrix} = b$$

Luego resulta un nuevo problema, el problema modificado:

$$\begin{aligned} \max \quad & z = x_1 + x_2 + M \bar{x}_5 \\ \text{s.a.} \quad & 2x_1 + x_2 + x_3 + 19\bar{x}_5 = 24 \\ & x_2 + x_4 + 5\bar{x}_5 = 8 \\ & x_1, x_2, x_3, x_4, \bar{x}_5 \geq 0 \end{aligned}$$

Donde $x_1 = (1 \ 1 \ 2 \ 2 \ 1)^T$ es una solución inicial factible para este problema modificado.

2.3.4 CRITERIO DE PARADA:

El problema original de optimización (primal) es dado por :

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

La formulación del Dual para este problema es dado por:

$$(D) \quad \begin{array}{ll} \max & y_0 = b^T y \\ \text{s.a.} & A^T y \leq c \\ & y, \text{ libre} \end{array} \quad (2.14)$$

Usando las restricciones del problema dual, tenemos:

$$b^T y = (Ax)^T y = x^T A^T y \leq x^T c = c^T x. \quad \Rightarrow \quad c^T x \geq b^T y$$

Vemos que cuando el problema primal y el dual son factibles, es decir, satisfacen las restricciones, la función objetivo del primal tiende a disminuir mientras que la función objetivo del dual va en aumento, la diferencia entre estos dos valores se determina por el *intervalo de dualidad* (I_{gap}).

La igualdad $c^T x = b^T y$, se reserva solo para la optimalidad. Cuando x^* , y^* son las soluciones óptimas del problema primal y dual respectivamente, entonces se cumple $c^T x^* = b^T y^*$. Este resultado importante proporciona el medio para “parar” el algoritmo, controlando el *intervalo de dualidad* y parando cuando se llegue a un valor aceptable.

Puesto que los vectores solución permanecen en el interior, el intervalo no llega a ser idénticamente cero, pero se obtiene muy cerca, luego:

$$I_{\text{gap}} = \frac{\|c^T x - b^T y\|}{1 + \|c^T x\|} \leq \varepsilon \quad (2.15)$$

donde los valores de ε , están en el rango 10^{-6} a 10^{-8} . Por lo tanto el algoritmo termina cuando $I_{\text{gap}} < \varepsilon$.

Resumen de algoritmo primal

Paso 0: Dado un vector inicial (interior y factible) en \mathbb{R}^n , $x_0 > 0$, inicialice la iteración contando, $k=0$ e inicialice el actual vector solución, $x(k) = x_0$

Paso 1: Defina la matriz de escalamiento $D(k)$ a través de:

$D(k) = \text{diag}[x_1(k), x_2(k), \dots, x_n(k)]$, donde $x_i(k)$, es la i -ésima componente de la solución actual $x(k)$.

Paso 2: Encuentre el estimado del dual, $y(k)$, donde $y(k) \in \mathbb{R}^m$ resolviendo:

$$[AD^2(k)A^T]y(k) = AD^2(k)c$$

Paso 3: Evalúe el estimado del vector de costo reducido, $z(k) \in \mathbb{R}^n$ de $z(k) = c - A^T y(k)$ y luego

$$\text{Calcule } dx(k) = -D^2(k) z(k); \quad d(x) \in \mathbb{R}^n$$

Paso 4: Hallar el nuevo vector solución a través de la fórmula:

$$x(k+1) = x(k) + \rho \alpha \quad dx(k), \quad 0 < \rho < 1 \quad y \quad \alpha = \min \left\{ -\frac{x_i(k)}{dx_i(k)} : \forall 1 \leq i \leq n \text{ y } dx_i(k) < 0 \right\}$$

Paso 5: Si el criterio de parada se satisface, fin del algoritmo, de otro modo incrementar k e ir al paso 1.

EJEMPLO:

(siguiendo los pasos del algoritmo).

$$\begin{array}{ll} \min & x_1 + x_2 + x_3 \\ \text{s.a.} & x_1 + 2x_2 + 3x_3 \geq 10 \\ & 2x_1 + x_2 - x_3 \geq 20 \\ & x_3 \geq 4 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

Convirtiendo a la forma estándar se tiene:

$$\begin{array}{ll} \min & x_1 + x_2 + x_3 \\ \text{s.a.} & x_1 + 2x_2 + 3x_3 - x_4 = 10 \\ & 2x_1 + x_2 - x_3 - x_5 = 20 \\ & x_3 - x_6 = 4 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{array}$$

$$A = \begin{bmatrix} 1 & 2 & 3 & -1 & 0 & 0 \\ 2 & 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \quad b = \begin{pmatrix} 10 \\ 20 \\ 4 \end{pmatrix} \quad c = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Sea

$$x_0 = \begin{pmatrix} 30 \\ 5 \\ 15 \\ 75 \\ 30 \\ 11 \end{pmatrix} \text{ el vector solución inicial, (uno de muchas posibilidades hallado dando valores).}$$

Luego seguiremos en detalle la secuencia de la primera iteración.

1ª Iteración. Dado el vector inicial x_0 , la matriz de escalamiento D en esta iteración es:

$$D = \begin{pmatrix} 30 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 15 & 0 & 0 & 0 \\ 0 & 0 & 0 & 75 & 0 & 0 \\ 0 & 0 & 0 & 0 & 30 & 0 \\ 0 & 0 & 0 & 0 & 0 & 11 \end{pmatrix}$$

Luego hallamos la matriz AD^2A^T y el vector AD^2c

$$AD^2A^T = \begin{pmatrix} 8650 & 1175 & 675 \\ 1175 & 4750 & -225 \\ 675 & -225 & 346 \end{pmatrix} \quad AD^2c = \begin{pmatrix} 1625 \\ 1600 \\ 225 \end{pmatrix}$$

De $(AD^2A^T)y = AD^2c$ resulta:

$$y = \begin{pmatrix} 0.0848 \\ 0.3596 \\ 0.7122 \end{pmatrix}$$

El vector de costo reducido z y el vector de paso dx se encuentran de:

$$z = c - A^T y = \begin{pmatrix} 0.2160 \\ 0.4808 \\ 0.3830 \\ 0.0848 \\ 0.3496 \\ 0.7122 \end{pmatrix} \quad dx = -D^2 z = \begin{pmatrix} -194.3997 \\ -12.0202 \\ -86.1722 \\ -476.9716 \\ -314.6424 \\ -86.1772 \end{pmatrix}$$

y usando la medida:

$$\max \left\{ -\frac{194.3997}{30}, -\frac{12.0202}{5}, -\frac{86.1722}{15}, -\frac{476.9716}{75}, -\frac{314.6424}{30}, -\frac{86.1772}{11} \right\} = 10.4881.$$

y usando un factor de paso de 95%, la nueva iteración es:

$$x = x_0 + \left(\frac{0.95}{\alpha} \right) dx = \begin{pmatrix} 30 \\ 5 \\ 15 \\ 75 \\ 30 \\ 11 \end{pmatrix} + \left(\frac{0.95}{10.4881} \right) \begin{pmatrix} -194.3997 \\ -12.0202 \\ -86.1722 \\ -476.9716 \\ -314.6424 \\ -86.1772 \end{pmatrix} = \begin{pmatrix} 12.3915 \\ 3.9112 \\ 7.1942 \\ 31.7964 \\ 1.5000 \\ 3.1942 \end{pmatrix}$$

Con los estimados primal, x , y el dual y , los respectivos valores de la función objetivo son dados por:

$$J_p = c^T x = 23.4968.$$

$$J_d = b^T y = 10.8868.$$

Utilizando el programa matlab se obtuvieron las primeras 10 iteraciones para los vectores dual y primal respectivamente los que se muestran en las tablas (2.1) y (2.2) respectivamente.

Tabla 2.1: Soluciones del vector dual y su valor objetivo.

Iteración	y_1	y_2	y_3	valor objetivo
1	0.0848	0.3496	0.7122	10.6888
2	0.0493	0.4462	0.1012	14.2217
3	0.0111	0.4970	0.4617	15.8975
4	0.0007	0.4994	0.4958	15.9781
5	0.0001	0.5000	0.4996	15.9984
6	0.0000	0.5000	1.5000	16.0000
7	0.0000	0.5000	1.5000	16.0000
8	0.0000	0.5000	1.5000	16.0000
9	0.0000	0.5000	1.5000	16.0000
10	0.0000	0.5000	1.5000	16.0000

Tabla 2.2: Soluciones del vector primal y su valor objetivo.

Iteración	x_1	x_2	x_3	x_4	x_5	x_6	v. objetivo
1	12.3915	3.9112	7.1942	31.7964	1.5000	3.1942	23.4968
2	11.6317	2.1131	4.1597	18.3369	1.2167	0.1597	17.9045
3	12.2740	0.1057	4.1249	14.8599	0.5288	0.1249	16.5045
4	11.9909	0.0856	4.0410	14.2850	0.0264	0.0410	16.1175
5	11.9971	0.0290	4.0020	14.0611	0.0210	0.0020	16.0281
6	12.0034	0.0014	4.0016	14.0112	0.0065	0.0016	16.0064
7	11.9998	0.0011	4.0005	14.0035	0.0003	0.0005	16.0014
8	12.0000	0.0003	4.0000	14.0006	0.0003	0.0000	16.0003
9	12.0000	0.0000	4.0000	14.0001	0.0000	0.0000	16.0000
10	12.0000	0.0000	4.0000	14.0000	0.0000	0.0000	16.0000

Como se observa, el valor objetivo dual disminuye con dirección al valor objetivo primal, la igualdad solo se dará en el caso de existir optimalidad.

2.4 EL ALGORITMO DUAL:

Este algoritmo puede considerarse como el algoritmo primal aplicado al problema dual, esto lo obtendremos a partir de las reglas similares a la obtención de dicho algoritmo.

Consideremos el problema primal en su forma estándar :

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

y su respectivo problema dual dado por:

$$(D) \quad \begin{array}{ll} \max & y_0 = b^T y \\ \text{s.a.} & A^T y \leq c \\ & y, \text{ libre} \end{array} \quad (2.16)$$

Agregando a la restricción del problema dual un vector de holgura z , n - dimensional no negativo se obtiene:

$$\begin{array}{ll} \max & y_0 = b^T y \\ \text{s.a.} & A^T y + z = c \\ & y, \text{ libre } z \geq 0 \end{array} \quad (2.17)$$

Asuma que (y_0, z_0) es un punto inicial factible en el interior para el problema (2.17). Esto es $z_0 > 0$ y (y_0, z_0) satisface el sistema de ecuaciones.

Para determinar el nuevo punto, a partir del punto inicial (y_0, z_0) tenemos que encontrar el vector dirección (dy, dz) . Manteniendo la factibilidad de (2.17), desplazamos la solución del problema dual según el par de vectores dirección dy, dz hacia un nuevo par que mejore el valor de la función objetivo. Con los vectores de dirección y algún tamaño de paso α , el nuevo punto y y z se expresan como:

$$\begin{array}{ll} y = y_0 + \alpha dy & y, \text{ libre} \\ z = z_0 + \alpha dz & z \geq 0 \end{array}$$

Estos nuevos vectores satisfacen (2.17) y mejoran el valor de la función objetivo, es decir:

$$b^T y \geq b^T y_0 \quad (2.18)$$

Por la factibilidad del problema dual se tiene que:

$$A^T y + z = A^T (y_0 + \alpha dy) + (z_0 + \alpha dz) = c$$

Pero como $A^T y_0 + z_0 = c$ en virtud de la factibilidad de los vectores iniciales obtenemos :

$$A^T dy + dz = 0 \quad \Rightarrow \quad dz = -A^T dy \quad (2.19)$$

2.4.1 ESCALAMIENTO:

Necesitamos desplazarnos en una dirección que mejore el valor de la función objetivo, para este fin consideraremos una matriz de escalamiento **D**:

$$D = \begin{pmatrix} 1/z_1 & 0 & 0 \dots 0 \\ 0 & 1/z_2 & 0 \dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & 0 \dots 1/z_n \end{pmatrix}$$

donde la diagonal está formada por el recíproco de las componentes del vector **z**, según (2.17). Además dichas componentes son estrictamente positivas. Con esta matriz se consideran las siguientes transformaciones de escalamiento:

$$z_1 = Dz \quad z = D^{-1} z_1 \quad dz = D^{-1} dz_1 \quad (2.20)$$

Usando estas transformaciones y por la condición de factibilidad planteada en (2.19) observamos que:

$$A^T dy + dz = A^T dy + D^{-1} dz_1 = 0$$

de lo cual se deduce:

$$DA^T dy + dz_1 = 0,$$

Y multiplicando ambos miembros por **AD** llegamos a:

$$AD^2 A^T dy + AD dz_1 = 0, \quad \Rightarrow \quad dy = -(AD^2 A^T)^{-1} AD dz_1 \quad (2.21)$$

Recordar que el valor de la función objetivo en el nuevo punto **y** es dado por $b^T(y_0 + \alpha dy)$

Por tanto para efectuar el desplazamiento en una dirección que mejore el valor de la función objetivo, deberíamos escoger dz_1 de modo que se cumpla la siguiente condición:

$$b^T dy \geq 0 \quad (2.22)$$

Definamos ahora una matriz R , $(m \times n)$:

$$R = (AD^2A^T)^{-1}AD \quad (2.23)$$

La condición de factibilidad (2.43) se escribe, $dy = -R dz_1$ y (2.44) se convierte en :

$$-b^T R dz_1 \geq 0 \quad (2.24)$$

Para satisfacer esto, seleccionamos el vector dz_1 como sigue:

$$dz_1 = -R^T b \quad (2.25)$$

así (2.46) llega a ser:

$$b^T dy = b^T (-R dz_1) = b^T R R^T b = \|R^T b\|^2 \geq 0. \quad (2.26)$$

Observamos que eligiendo dz_1 como en (2.25) resulta un mejoramiento en el valor de la función objetivo y reemplazando (2.25) en (2.21) se obtiene una relación para el vector de paso

$$dy = [(AD^2A^T)^{-1}AD] [DA^T(AD^2A^T)^{-1}] b = (AD^2A^T)^{-1} b. \quad (2.27)$$

Es interesante hacer notar que todos estos comentarios sostienen las variantes del algoritmo de Karmarkar.

Hasta aquí, todo está listo para generar el nuevo punto (hacer una iteración), mediante las expresiones vistas antes:

$$\begin{aligned} y &= y_0 + \alpha dy & y, \text{ libre} \\ z &= z_0 + \alpha dz & z \geq 0 \end{aligned}$$

donde α es el *máximo paso admisible*, necesario para exigir la no negatividad impuesta por el vector z y al igual que en el algoritmo primal se determina por:

$$\alpha = \min \left\{ -\frac{z_i}{dz_i} : \forall dz_i < 0, 1 \leq i \leq n \right\} \quad (2.28)$$

Este tamaño de paso se garantiza la factibilidad y la no negatividad de las restricciones pero podría ocurrir que el nuevo vector z alcance la frontera del politopo, es por eso la necesidad de incluir un *factor* ρ , con esto, la nueva iteración del vector z llega a ser:

$$z = z_0 + \rho \alpha dz, \quad 0 < \rho < 1 \quad (2.29)$$

generalmente ρ está entre 0.95 y 0.9995.

Por otro lado, usando la matriz de escalamiento D , ¿que ocurre cuando hacemos $w = -D^2 dz$, donde w es un vector arbitrario?.

Veamos:

$$Aw = A(-D^2 dz) = AD^2 A^T dy, \quad (dz \text{ definido en (2.19)})$$

$$Aw = [AD^2 A^T] [(AD^2 A^T)^{-1}] b = b.$$

Esto significa que w definido así, solo satisface las restricciones del problema primal, pero aún no es solución porque no necesariamente cumple con las restricciones de no negatividad impuesta por el vector de decisión x ; es decir el punto es factible pero podría ser que alguna componente sea negativa, cuando este sea el caso el estimado primal $x = w$ es no factible a pesar de que satisface las restricciones de igualdad.

Resumen del algoritmo Dual

Paso 0: Dado un par factible de vectores iniciales y_0 y z_0 tal que:

$A^T y_0 + z_0 = c, z_0 \geq 0$ inicie la iteración contando, $k = 0$ y poner:

$$y(k) = y_0, \quad z(k) = z_0$$

Paso 1: Defina una matriz diagonal $D(k)$:

$$D(k) = \text{diag} \left[\frac{1}{z_1(k)}, \frac{1}{z_2(k)}, \dots, \frac{1}{z_n(k)} \right],$$

donde $z_i(k)$, es la i -ésima componente del vector de de costo reducido $z(k)$

Paso 2: Encuentre el vector de paso $dz(k) \in \mathbb{R}^m$, resolviendo:

$$[AD^2(k)A^T] dy(k) = b$$

Paso 3: Evalúe el estimado del vector de costo reducido, $z(k) \in \mathbb{R}^n$ mediante:

$$dz(k) = -A^T dy(k).$$

Paso 4: Hallar la siguiente iteración para el vector de costo reducido y el vector dual mediante la fórmula actualizada:

$$z(k+1) = z(k) + \rho \alpha dz(k), \quad y(k+1) = y(k) + \rho \alpha dy(k)$$

donde:

$$\alpha = \min \left\{ -\frac{z_i(k)}{dz_i(k)} : \forall dz_i(k) < 0, 1 \leq i \leq n \right\} \quad \text{y} \quad 0 < \rho < 1$$

los estimados actuales del primal se encuentran de:

$$x(k+1) = -D^2 dz(k)$$

Paso 5: Si el criterio de parada $\frac{\|c^T x - b^T y\|}{1 + \|c^T x\|} \leq \varepsilon$, se satisface, fin del algoritmo, de otro modo incrementar k e ir al paso 1.

Ahora ilustraremos este algoritmo con el mismo ejemplo anterior.

EJEMPLO:

$$\begin{aligned} \text{(P)} \quad & \min x_1 + x_2 + x_3 \\ & \text{s.a. } x_1 + 2x_2 + 3x_3 \geq 10 \\ & \quad 2x_1 + x_2 - x_3 \geq 20 \\ & \quad x_3 \geq 4 \\ & \quad x_1, x_2, x_3 \geq 0 \end{aligned}$$

Transformando a la forma estándar se tiene:

$$\begin{aligned} \text{(P)} \quad & \min x_1 + x_2 + x_3 \\ & \text{s.a. } x_1 + 2x_2 + 3x_3 - \bar{x}_4 = 10 \\ & \quad 2x_1 + x_2 - x_3 - \bar{x}_5 = 20 \\ & \quad x_3 - \bar{x}_6 = 4 \\ & \quad x_1, x_2, x_3, \bar{x}_4, \bar{x}_5, \bar{x}_6 \geq 0. \end{aligned}$$

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 & 0 & 0 \\ 2 & 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 10 \\ 20 \\ 4 \end{pmatrix} \quad c = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

El dual para este problema es dado por:

$$\begin{aligned} & \max 10y_1 + 20y_2 + 4y_3 \\ \text{(D)} \quad & \text{s.a.} \quad y_1 + 2y_2 \leq 1 \\ & \quad 2y_1 + y_2 \leq 1 \\ & \quad 3y_1 - y_2 + y_3 \leq 1 \\ & \quad -y_1 \leq 0 \\ & \quad -y_2 \leq 0 \\ & \quad -y_3 \leq 0 \\ & \quad y_1, y_2, y_3: \text{libres} \end{aligned}$$

y agregando el vector z n - dimensional se tiene:

$$\begin{aligned} & \max 10y_1 + 20y_2 + 4y_3 \\ \text{(D)} \quad & \text{s.a.} \quad y_1 + 2y_2 + z_1 = 1 \\ & \quad 2y_1 + y_2 + z_2 = 1 \\ & \quad 3y_1 - y_2 + y_3 + z_3 = 1 \\ & \quad -y_1 + z_4 = 0 \\ & \quad -y_2 + z_5 = 0 \\ & \quad -y_3 + z_6 = 0 \\ & \quad y_1, y_2, y_3: \text{libres}, z_1, \dots, z_6 \geq 0 \end{aligned}$$

Una solución factible inicial para este problema es dado por:

$$y_0 = [1/4, 1/4, 1/4]^T, \quad z_0 = [1/4, 1/4, 1/4, 1/4, 1/4, 1/4]^T$$

1ª Iteración: Dado el vector inicial z_0 , la matriz de escalamiento D en esta iteración es:

$$D = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}_{(6 \times 6)} \quad \text{luego,} \quad AD^2A^T = \begin{pmatrix} 240 & 16 & 48 \\ 16 & 112 & -16 \\ 48 & -16 & 32 \end{pmatrix}_{(3 \times 3)}$$

hallando el vector de paso dy de $(AD^2A^T) dy = b$, resulta

$$dy = (AD^2A^T)^{-1}b = \begin{pmatrix} -0.0292 \\ 0.2228 \\ 0.2802 \end{pmatrix}$$

y el vector de paso para el vector de costo reducido, dz , se halla de

$$dz = -A^T dy = \begin{pmatrix} -0.4163 \\ -0.1643 \\ 0.0302 \\ -0.0292 \\ 0.2228 \\ 0.2802 \end{pmatrix}$$

Solo para las componentes negativas de dz encontramos que:

$$\alpha = \max \left\{ -\frac{-0.4163}{0.25}, -\frac{-0.1643}{0.25}, -\frac{-0.0292}{0.25} \right\} = \max \{ 1.6652 \quad 0.6572 \quad 0.1168 \} = 1.6652$$

el tamaño de paso $\alpha_\delta = \rho/\alpha$, usando el 95% del máximo paso admisible, y z es obtenido de:

$$z = z_0 + \alpha_\delta dz = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{pmatrix} + \left(\frac{0.95}{1.6652} \right) \begin{pmatrix} -0.4163 \\ -0.1643 \\ 0.0302 \\ -0.0292 \\ 0.2228 \\ 0.2802 \end{pmatrix} = \begin{pmatrix} 0.0125 \\ 0.1563 \\ 0.2673 \\ 0.2333 \\ 0.3771 \\ 0.4099 \end{pmatrix}$$

el mismo α_δ se usa para obtener el nuevo valor del vector dual, y .

$$y = y_0 + \alpha_\delta dy = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.25 \end{pmatrix} + \left(\frac{0.95}{1.6652} \right) \begin{pmatrix} 0.0292 \\ 0.2228 \\ 0.2802 \end{pmatrix} = \begin{pmatrix} 0.2333 \\ 0.3771 \\ 0.4099 \end{pmatrix}$$

el vector solución primal, x , es dado por:

$$x = -D^2 dz = -D^2 \begin{pmatrix} -0.4163 \\ -0.1643 \\ 0.0302 \\ -0.0292 \\ 0.2228 \\ 0.2802 \end{pmatrix} = \begin{pmatrix} 6.6613 \\ 2.6290 \\ -0.4839 \\ 0.4677 \\ -3.5645 \\ -4.4839 \end{pmatrix}$$

Observamos que mientras se satisface $Ax = b$, algunas de sus componentes son negativas por tanto la factibilidad primal aún no se alcanza.

Tabla 2.3: Iteraciones del vector dual, y , con su valor objetivo.

Iteración	y_1	y_2	y_3	valor objetivo
1	0.2333	0.3771	0.4099	11.5145
2	0.1216	0.3448	1.0568	14.1388
3	0.0061	0.4930	1.4619	15.7675
4	0.0010	0.4993	1.4901	15.9564
5	0.0005	0.4996	1.4979	15.9881
6	0.0000	0.4999	1.4996	15.9969
7	0.0000	0.5000	1.4998	15.9992
8	0.0000	0.5000	1.5000	15.9998
9	0.0000	0.5000	1.5000	15.9999
10	0.0000	0.5000	1.5000	16.0000

Tabla 2.4: Iteraciones del vector primal, x , y su valor objetivo.

Iteración	x_1	x_2	x_3	x_4	x_5	x_6	v. objetivo
1	6.66135	2.6290	-0.4839	0.4677	-3.5645	-4.4839	8.8065
2	12.6843	-3.6677	1.9900	1.10899	-0.2192	-2.0800	10.9366
3	12.5937	-1.9870	3.5673	9.32169	-0.3669	-0.4329	14.1740
4	11.9988	-0.0016	3.9987	13.9916	-0.0026	-0.0013	15.9959
5	12.0000	-0.0001	3.9999	13.9995	0.0000	-0.0001	15.9998
6	12.0000	0.0000	4.0000	14.0000	0.0000	0.0000	16.0000
7	12.0000	0.0000	4.0000	14.0000	0.0000	0.0000	16.0000
8	12.0000	0.0000	4.0000	14.0000	0.0000	0.0000	16.0000
9	12.0000	0.0000	4.0000	14.0000	0.0000	0.0000	16.0000
10	12.0000	0.0000	4.0000	14.0000	0.0000	0.0000	16.0000

Al concluir la primera iteración evaluamos los valores objetivos primal y dual.

$$J_p = c^T x = 8.8065$$

$$J_d = b^T y = 11.5149$$

Con las nuevas soluciones para los vectores z y y , se puede iniciar nuevamente otra iteración definiendo una nueva matriz de escalamiento. Las primeras 10 iteraciones para el vector dual " y " se muestra en la tabla 2.3, y los estimados primal x , se muestran en la tabla 2.4.

EJEMPLO 2: Problema de Klee-Minty:

En 1972 Klee y Minty construyeron un problema de 50 variables y 50 restricciones cuya solución mediante el algoritmo simplex requería 300,000 años . La forma general de este problema es la siguiente:

$$\begin{aligned}
 & \max \quad x_1 + x_2 + x_3 + \dots + x_{n-1} + x_n \\
 & \text{s.a:} \\
 & \quad x_1 \leq 1 \\
 & \quad 2x_1 + x_2 \leq 3 \\
 & \quad 2x_1 + 2x_2 + x_3 \leq 9 \quad (*) \\
 & \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\
 & \quad 2x_1 + 2x_2 + 2x_3 + \dots + 2x_{n-1} + x_n \leq 3^{n-1} \\
 & \quad x_1, \dots, x_n \geq 0
 \end{aligned}$$

Para obtener la solución óptima mediante el algoritmo simplex se requieren 2^{n-1} iteraciones.

Resolviendo (*) en un programa de computación (matlab) usando los algoritmos primal y dual, para distintos valores de " n " se obtiene la siguiente tabla:

Nº variables	Número de iteraciones para encontrar la solución óptima				
N	Primal	Dual	Simplex	Karmarkar*	Barnes*
2	228	15	3	189960	12
5	232	18	31	∞	42
10	275	20	1023	∞	202
15	285	21	32767	∞	421
20	289	22	1048575	∞	459
50	320	24	112589990684262	∞	477
			3		

En esta tabla se observa la gran ventaja de los algoritmos de puntos interiores primal y dual en comparación con el algoritmo simplex, Karmarkar y Barnes, pues cuando el número de variables del problema de Klee- Minty se aumenta, el número de iteraciones en los algoritmos de puntos interiores aumenta paulatinamente mientras que en los algoritmos simplex y Karmarkar aumentan de manera exponencial (alarmante).

2.5. CONVERGENCIA DE LOS ALGORITMOS DE PUNTOS INTERIORES:

Como hemos visto en la deducción de los algoritmos no es necesario que los problemas primal y dual tengan soluciones no degeneradas, estos algoritmos trabajan igualmente bien en problemas que no satisfacen tales postulados, sin embargo para la convergencia asumimos algunas restricciones y probaremos entre otros la convergencia a la solución óptima.

2.5.1. CONDICIONES DEL PROBLEMA:

Recordemos que nuestro objetivo es resolver el siguiente problema:

$$\begin{aligned}
 &\min c^T x \\
 &\text{s.a. } Ax = B \\
 &\quad x \geq 0.
 \end{aligned}
 \tag{2.30}$$

Sea $\Omega = \{x/ Ax = b, x \geq 0\}$.

Consideremos que este problema tiene las siguientes condiciones:

- 1) *El problema es acotado y es factible.*
- 2) *El problema es no degenerado (problema primal no degenerado).*
- 3) *El dual del problema es no degenerado (problema dual no degenerado).*

Para iniciar la solución de este problema mediante el método de puntos interiores necesitamos de un punto inicial x_0 . Luego para hacer que este punto sea centrado aplicamos una transformación afín D_{x_0} . Al aplicar esta transformación al problema (2.30) se obtiene un nuevo problema de la forma siguiente:

$$\begin{aligned} \min \quad & c_1^T x \\ \text{s.a.} \quad & A_1 x = B \\ & x \geq 0, \end{aligned} \quad (2.31)$$

donde:

$$A_1 = A D_{x_0}, \quad c_1 = D_{x_0} c$$

Sabemos que el siguiente punto (mejor solución) está dado por:

$$x_1 = x_0 + \rho \alpha dx$$

donde $dx = -c_p = -P_{x_0} c_1$, con $P_{x_0} = [I_n - D_{x_0} A^T (A D_{x_0}^2 A^T)^{-1} A D_{x_0}]$

$$\alpha = \frac{1}{\max_i \{c_p^i\}} = \frac{1}{\max_i (e_i \cdot c_p)} = \frac{1}{\gamma},$$

con $e_i = (1, 1, \dots, 1, \dots, 1)$ (e_i es el i -ésimo vector unitario) y $0 < \rho < 1$.

Así sucesivamente se construye la sucesión $\{x_k\}$ de soluciones aproximadas.

Por simplicidad este algoritmo lo resumimos de la siguiente forma:

$$x_{k+1} = T(x_k) \quad (2.32)$$

donde $T(x) = x - \frac{\rho}{\gamma} D_x P_x D_x c$.

PROPOSICIÓN 1:

Sea $c_p(x) = P_x D_x c$. Las siguientes propiedades se cumplen:

- (a) *Si la función objetivo no es constante sobre Ω , entonces la sucesión $c \cdot x^k$ es estrictamente decreciente.*

(b) Si para algún $x_0 \in \Omega^0$, tenemos, $c_p(x_0) \leq 0$ y $c_p(x_0) \neq 0$, entonces el problema es no acotado ($c_p(x_0)$ representa c_p en x_0).

(c) Si para algún $x_0 \in \Omega^0$ se cumple $c_p(x_0) = 0$, entonces todo punto factible es óptimo.

Prueba:

(a) Es inmediato de la misma construcción del algoritmo, que determina el siguiente punto

de menor costo.

(b) Sabemos que $x_{k+1} = x_k + \rho \alpha dx_k$, $\alpha > 0$ (arbitrario), $0 < \rho < 1$, es decir cual fuere el valor de α , x_{k+1} siempre es factible.

Por otro lado $c \cdot x_{k+1} \leq c \cdot x_k$, pero,

$$\begin{aligned} c \cdot x_{k+1} &= c \cdot x_k + c \cdot \rho \alpha dx_k \\ &= c \cdot x_k + c \cdot \rho \alpha (-c_p) = c \cdot x_k - \rho \alpha (c \cdot c_p) \end{aligned}$$

Puesto que $\rho \alpha > 0$, $c \cdot c_p > 0$.

Por lo tanto, si $\alpha \rightarrow \infty$, $c \cdot x_{k+1} \rightarrow -\infty$.

(c) En ausencia de degeneración tenemos $\Omega^0 = \{x > 0 / Ax = b\}$.

Si $x \in \Omega^0$ entonces $x_N = x + dx \in \Omega^0$

Además: $c(x+dx) < cx$.

Pero $x_N = x - \rho \alpha c_p$ es el punto con costo menor que x .

Pero si $c_p = 0$, entonces $x_N = x$ es el punto óptimo. Puesto que c_p es constante en todo punto $x \in \Omega^0$, entonces todo punto $x \in \Omega^0$ es óptimo.

2.5.2 CONVERGENCIA:

Las condiciones (2) y (3) enunciadas anteriormente son respectivamente equivalentes a:

(2') La matriz AD_x tiene rango "m", $\forall x \in \Omega$.

(3') El vector $(c \cdot A^T y)$, tiene a lo mas "m" ceros, $\forall y \in R^m$.

Se sabe que si A es una matriz de rango "m", la proyección sobre el espacio $N(AD_x)$ es:

$$P_x = I - D_x A^T (A D_x^2 A^T)^{-1} A D_x$$

entonces:

$$c_p = P_x D_x c$$

$$= [I_n - D_x A^T (A D_x^2 A^T)^{-1} A D_x] D_x c = D_x [c - A^T (A D_x^2 A^T)^{-1} A D_x^2 c],$$

luego haciendo: $y = (A D_x^2 A^T)^{-1} A D_x^2 c$ y $r = c - A^T y$ se obtiene $c_p = D_x [c - A^T y]$ ó,
 $c_p = D_x r$

Reescribiendo $T(x)$ en términos de (5.3) " y ", " r " y c_p se tiene:

$$\begin{aligned} T(x) &= x - \frac{\rho}{\gamma} D_x P_x D_x c \\ &= x - \frac{\rho}{\gamma} D_x c_p \\ &= x - \frac{\rho}{\gamma} D_x^2 r \end{aligned}$$

Para enfatizar la dependencia de " y ", " r " sobre x escribiremos $r(x)$, $y(x)$.

Ahora empezaremos por mostrar que los puntos límite de la sucesión x^k , $k = 0, 1, 2, \dots$ pertenecen al conjunto:

$$S = \{x \in \Omega / D_x r(x) = 0\} \quad (2.33)$$

(2') y (3') implican que S es el conjunto de vértices sobre Ω .

(3') implican que $r(x)$ tiene a lo mas " m " ceros, pero si x no es vértice de (2') se sigue que x tiene a lo mas $n-m-1$ ceros y $D_x r(x) \neq 0$.

PROPOSICIÓN 2.

$$\lim_{k \rightarrow \infty} D_{x^k} r(x^k) = 0$$

Prueba:

Por la condición 1 y por construcción c_{x^k} es monótona y acotada, por lo que c_{x^k} converge; entonces la sucesión de diferencias tiende a cero, es decir

$$0 = \lim_{k \rightarrow \infty} (c_{x^k} - c_{x^{k+1}})$$

de donde se tiene

$$0 = \lim_{k \rightarrow \infty} \frac{\rho \|c_p(x^k)\|_2^2}{\gamma(x^k)} \quad (2.34)$$

Sea, $y = c_p(x^k)$, entonces:

$\max y_i \leq \|y\|_2$, pero sabemos que $\max y_i = \gamma(x^k)$, luego:

$$0 \leq \|c_p(x^k)\|_2 \leq \frac{\|c_p(x^k)\|_2^2}{\gamma(x^k)},$$

tomando $\lim k \rightarrow \infty$

$$0 \leq \lim_{k \rightarrow \infty} \|c_p(x^k)\|_2 \leq \lim_{k \rightarrow \infty} \frac{\|c_p(x^k)\|_2^2}{\gamma(x^k)},$$

de (2.34), se tiene:

$$\lim_{k \rightarrow \infty} \frac{\|c_p(x^k)\|_2^2}{\gamma(x^k)} = 0,$$

$\Rightarrow \lim_{k \rightarrow \infty} \|c_p(x^k)\|_2 = 0$, de donde, por la continuidad de la norma se obtiene

$\Rightarrow \lim_{k \rightarrow \infty} c_p(x^k) = 0$ ó equivalentemente

$$\lim_{k \rightarrow \infty} D_{x^k} r(x^k) = 0.$$

Ahora necesitamos introducir el siguiente conjunto:

$$F = \{x \in \Omega / r(x) \geq 0\} \quad (2.35)$$

Es decir, F es el conjunto formado por aquellos puntos que mantienen factibilidad dual.

PROPOSICIÓN 3. Si $x \in S \cap F$, $\Rightarrow x$ es un punto óptimo.

Prueba:

Puesto que $x \in S$, entonces $D_x c = D_x A^T y$ donde " y " es el vector dual, entonces:

$$\begin{aligned} c \cdot x &= c \cdot D_x 1 \\ &= D_x c \cdot 1 \\ &= D_x A^T y \cdot 1 \\ &= y \cdot (D_x A^T)^T 1 \\ &= y \cdot A D_x 1 \\ &= y \cdot A x \\ &= y \cdot b \end{aligned}$$

Sea z arbitrario en Ω tal que $z \in F \Rightarrow r(z) \geq 0$.

Como

$$\begin{aligned}
 r(x) &= c - A^T y(x) \\
 r(x) + A^T y(x) &= c \Rightarrow A^T y(x) \leq c \\
 &\Rightarrow c \cdot z \geq A^T y(x) \cdot z \\
 &= y(x) \cdot A z \\
 &= y(x) \cdot b, \quad \text{pues } z \text{ factible} \\
 \therefore c \cdot z &\geq y(x) \cdot b = c \cdot x \\
 &\Rightarrow c \cdot x \leq c z, \quad \text{para cualquier } z \in \Omega
 \end{aligned}$$

Luego x es óptimo.

PROPOSICIÓN 4. Si la sucesión $\{x^k\}$ es convergente $\Rightarrow x^\infty = \lim_{k \rightarrow \infty} x^k$ es óptimo,

Prueba:

Supóngase lo contrario que x^∞ no es un punto óptimo.

$$\lim_{k \rightarrow \infty} c \cdot x^k > \min_{\Omega} c \cdot x \quad (2.36)$$

por otro lado de la continuidad de $r(x)$ y por la proposición 2, se tiene:

$$D_{x^\infty} r(x^\infty) = \lim_{k \rightarrow \infty} D_{x^k} r(x^k) = 0$$

de modo que $x^\infty \in S$.

como x^∞ no es óptimo y $x^\infty \in S$, por la proposición 3, $x^\infty \notin F$, luego debe existir un " i ", tal que $r_i(x^\infty) < 0$, además como $x^\infty \in S \Rightarrow x_i^\infty = 0$, para algún i .

Por la continuidad de $r(x)$, \exists un entero $K / \forall k \geq K, r_i(x^k) < 0$.

De $x^{k+1} = T(x^k) = x^k - \frac{\rho}{\gamma(x^k)} D_{x^k}^2 r(x^k)$ vemos que :

$$x_i^{k+1} = x_i^k - \rho [\gamma(x^k)]^{-1} (x_i^k)^2 r_i(x^k)$$

Como $(x_i^k)^2 r_i(x^k) < 0 \Rightarrow x_i^{k+1} > x_i^k > 0, \forall k \geq K$ lo cual contradice a la hipótesis

$$x_i^k \rightarrow x_i^\infty = 0.$$

\therefore si $\{x^k\}$ converge $\Rightarrow x^\infty$ es un punto óptimo.

PROPOSICIÓN 5. *La sucesión $\{x^k\}$, generada por (2.32), es convergente.*

Prueba:

Las condiciones de no degeneración y acotamiento implican que $\Omega \cap \{x/ c.x \leq M\}$ es compacto para cualquier valor de M.

Luego la sucesión $\{x^k\}$ tiene un punto límite \hat{x} .

De la proposición 2, se deduce que $\hat{x} \in S$; es decir \hat{x} es un vértice de Ω .

Por otro lado sea N el conjunto de índices no básicos de \hat{x} (es decir índices que corresponden a las componentes de \hat{x} que tienen valor cero).

Para $\delta > 0$, sea:

$$B_\delta = \{x \in \Omega / x_N < \delta\}$$

Se sigue de la condición (3') que podemos fijar un ε que satisface:

$$0 < \varepsilon < \min_{i \in N} |r_i(\hat{x})|$$

Como S es conjunto finito y $r(x)$ es continua, \exists un $\delta > 0$ tal que :

$$B_{2\delta} \cap S = \{\hat{x}\}$$

y

$$\min_{i \in N} |r_i(x)| > \varepsilon, \quad \forall x \in B_\delta \quad (2.37)$$

Sustituyendo $c_p = D_x r$ en (2.55) de la proposición 2 se tiene:

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\gamma(x^{-1}) \left\| D_{x^k} r(x^k) \right\|_2^2}{x^k} &= 0 \\ \Rightarrow \lim_{k \rightarrow \infty} \gamma(x^k)^{-1} [(x_i^k) r_i(x^k)]^2 &= 0. \end{aligned} \quad (2.38)$$

Considere el conjunto:

$$C_{\varepsilon, \delta} = \{x \in B_\delta / \rho \gamma(x)^{-1} x_i^2 r_i(x)^2 < \varepsilon \delta \quad \forall i \in N\}$$

Por (2.38) siempre existe un $\delta > 0$ tal que $\forall x \in B_\delta, x \in C_{\varepsilon, \delta}$ (\hat{x} es un elemento de B_δ).

Fijamos $x \in C_{\varepsilon, \delta}$. Puesto que $C_{\varepsilon, \delta} \subset B_\delta$, vemos de (2.58) que:

$$\rho \gamma(x)^{-1} x_i^2 |r_i(x)| < \delta, \quad \forall i \in N.$$

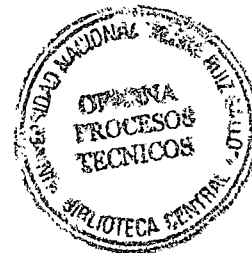
$$\text{De } T(x) = x - \frac{\rho}{\gamma} D_x^2 r$$

$$T_i(x) = x_i - \rho \gamma(x)^{-1} x_i^2 r_i(x) < 2\delta, \forall i \in \mathbb{N} \quad (2.39)$$

de donde $T(C_{\varepsilon, \delta}) \subset B_{2\delta}$.

Puesto que \hat{x} es un punto límite de la sucesión $\{x^k\}$ y además $\hat{x} \in C_{\varepsilon, \delta}$, entonces $\{x^k\}$ está en $C_{\varepsilon, \delta}$ para infinitos valores de k suficientemente grandes.

Ahora supongamos que la sucesión $\{x^k\}$ tiene dos o mas puntos límite entonces existen infinitos elementos de la sucesión que no pertenecen a $C_{\varepsilon, \delta}$, pero por (2.39) dichos elementos pertenecen a $B_{2\delta} \setminus C_{\varepsilon, \delta}$. Como este conjunto tiene clausura compacta, la subsucesión $\{x^k\} \in B_{2\delta} \setminus C_{\varepsilon, \delta}$ debe tener un punto límite \bar{x} , pero por la proposición 2, $\bar{x} \in S$, esto implica que S y la clausura de $B_{2\delta} \setminus C_{\varepsilon, \delta}$ no son disjuntos, lo cual contradice a la construcción de $B_{2\delta}$ y $C_{\varepsilon, \delta}$. Por lo tanto el punto límite de la sucesión $\{x^k\}$ es única, es decir esta sucesión es convergente.



CAPÍTULO III. Método de puntos interiores con óptimos alternativos.

Se tienen PPL en los que existe un óptimo que no es un punto extremo (*óptimos alternativos*). En este caso, el conjunto de soluciones óptimas es dado por el conjunto de combinaciones convexas de los puntos extremos del problema.

En problemas donde existen óptimos alternativos, es posible determinar una solución que procure un objetivo adicional. Por ejemplo, distribución uniforme de los recursos o agotar un recurso para preservar otro que se tiene en menor cantidad.

En este capítulo se describe y analiza una estrategia para resolver PPL con óptimos alternativos, que modifica el algoritmo de Puntos Interiores Primal Dual, permitiendo obtener soluciones óptimas que satisfagan objetivos adicionales.

3.1. Método de Barrera Logarítmica Primal-Dual

El algoritmo de Karmarkar generó una gran variedad de trabajos alrededor de su idea, la cual ha sido mejorada en muchos aspectos. Una de las más fructíferas variantes es el algoritmo de barrera logarítmica Primal-Dual.

Este método resuelve el PPL primal estándar:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a:} \quad & Ax = b \\ & x \geq 0 \end{aligned} \quad (3.1)$$

Donde x es el vector de variables primales y su problema dual estándar en función del vector de variables duales y ,

$$\begin{aligned} \max \quad & b^T y \\ \text{s.a:} \quad & A^T y \leq c \\ & y, \text{ libre} \end{aligned} \quad (3.2)$$

Incluyendo el
vector de
variables de
holgura z



$$\begin{aligned} \max \quad & b^T y \\ \text{s.a:} \quad & A^T y + z = c \\ & z \geq 0 \end{aligned} \quad (3.3)$$

Para eliminar las restricciones de no negatividad ($z \geq 0$), se emplea la función barrera logarítmica; entonces se tenemos el PPNL:

$$\begin{aligned} \max \quad & b^T y + \mu^k \sum_{i=1}^n \ln z_i \\ \text{s.a:} \quad & A^T y + z = c \end{aligned} \quad (3.4)$$

donde μ^k es el parámetro de barrera en la iteración k que durante el proceso iterativo debe ser disminuido a cero.

Se considero la función de penalización $\ln z_i$ la cual está bien definida ya que así debe cumplir con $z_i > 0$.

Los métodos de punto interior resuelven el problema (3.4) para diferentes valores del parámetro μ . Este parámetro se define de manera tal que $\mu_0 > \mu_1 > \mu_2 > \dots > \mu_\infty = 0$.

El teorema siguiente muestra el porqué de esta elección.

Teorema 3.1 (convergencia). *La sucesión de parámetros $\{\mu_t\}_{t=0}^\infty$ genera una sucesión de problemas del tipo (3.4). La sucesión de soluciones a estos problemas, $\{x_t\}_{t=0}^\infty$, converge a la solución del problema (3.3).*

La función Lagrangiana del problema (3.4) tiene la forma

$$L(x, y, z, \mu) = b^T y + \mu \sum_{i=1}^n \ln z_i - x^T (A^T y + z - c) \quad (3.5)$$

Los multiplicadores de Lagrange, x , son las variables del problema original (primal).

Empleando el Lagrangiano, las condiciones de optimalidad de primer orden del problema (3.4) son

$$\left. \begin{aligned} \nabla_x L(x, y, z, \mu) &= A^T y + z - c = 0 \\ \nabla_y L(x, y, z, \mu) &= Ax - b = 0 \\ \nabla_z L(x, y, z, \mu) &= XZe - \mu e = 0 \end{aligned} \right\} \quad (3.6)$$

Donde: $X \equiv \text{diag}(x_1, x_2, \dots, x_n)$; $Z \equiv \text{diag}(z_1, z_2, \dots, z_m)$; $e = (1, 1, \dots, 1)^T$

Veamos un ejemplo de PPL (3.1) en su forma matricial.

$$\begin{aligned}
 \text{(P)} \quad & \text{Min} \quad z = (-3 \quad -5 \quad 0 \quad 0 \quad 0) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \\
 & \text{s.a.} \quad \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} \\
 & \quad \quad \quad x \geq 0
 \end{aligned}$$

Su dual (D) según (3.3) está dado por:

$$\begin{aligned}
 \text{Max} \quad & z = (4 \quad 6 \quad 18) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\
 \text{s.a:} \quad & \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
 & \quad \quad \quad z_i \geq 0
 \end{aligned}$$

La función Lagrangiana según (3.5) es:

$$\begin{aligned}
 \mathcal{L}(x, y, z, \mu) &= (4 \quad 6 \quad 18) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \mu \sum_{i=1}^5 \ln z_i \\
 &\quad - \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}^T \left\{ \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix} - \begin{pmatrix} -3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}
 \end{aligned}$$

Las condiciones de optimalidad según (3.6) son :

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix} - \begin{pmatrix} -3 \\ 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} - \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x_1 & & & & \\ & x_2 & & & \\ & & x_3 & & \\ & & & x_4 & \\ & & & & x_5 \end{pmatrix} \begin{pmatrix} z_1 & & & & \\ & z_2 & & & \\ & & z_3 & & \\ & & & z_4 & \\ & & & & z_5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} \mu \\ \mu \\ \mu \\ \mu \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

3.1.1 El algoritmo general de solución es:

1. Calcular un punto inicial.

- Vector primal: $x^0 = \eta \hat{x}$ donde :

$$\eta = \frac{\|b\|_2 + 1}{\|A\hat{x}\|_2 + 1} \quad y \quad \hat{x}_j = \frac{1}{\|A_j\|_2 + 1}$$

- Vector dual. $z^0 = \begin{cases} si \quad c_j < 1, & z_j^0 = 1 \\ caso contrario, & z_j^0 = c_j \end{cases}$

2. Calcular el parámetro de barrera y hacer $k = 0$.

$$\mu^0 = \frac{(x^0)^T z^0}{n}$$

3. Verificar convergencia.

- Factibilidad primal:
$$\frac{\|Ax^k - b\|}{1 + \|x^k\|} \leq \varepsilon_f$$
- Factibilidad dual:
$$\frac{\|A^T y^k + z^k - c\|}{1 + \|y^k\| + \|z^k\|} \leq \varepsilon_f$$
- Condición de optimalidad:
$$\frac{\|c^T x^k - b^T y^k\|}{1 + \|b^T y^k\|} \leq \varepsilon_o$$

donde ε_f es el error de factibilidad y ε_o es el error de optimalidad.

En caso de que los criterios de convergencia se cumplan pare.

4. Calcular los errores para el punto actual.

$$r_p^k = c - A^T y^k - z^k$$

$$r_d^k = b - Ax^k$$

$$r_c^k = \mu^k e - X^T Z^k e$$

donde :

$$X = \text{diag}[x_1 \ x_2 \ \dots \ x_n]$$

$$Z = \text{diag}[z_1 \ z_2 \ \dots \ z_n]$$

$$e = [1 \ 1 \ \dots \ 1]^T$$

5. Obtener las direcciones de búsqueda.

$$\Delta x^k = \Delta x_{ctr}^k + \Delta x_{obj}^k + \Delta x_{fac}^k$$

- *Dirección central* $\Delta x_{ctr}^k = \mu^k D_k^{-1} P^k D_k^{-1} (X^k) e$

- *Dirección objetivo* $\Delta x_{obj}^k = D_k^{-1} P^k D_k^{-1} c$

- *Dirección factible* $\Delta x_{fac}^k = D_k^{-2} A^T (AD_k^{-2} A^T)^{-1} r_p^k$

$$\Delta y^k = (AD_k^{-2} A^T)^{-1} [AD_k^{-2} r_d^k - A(Z^k)^{-1} r_c^k + r_p^k]$$

$$\Delta z^k = r_d^k - A^T \Delta y^k$$

donde

$$P^k = I - D_k^{-1} A^T (AD_k^{-2} A^T)^{-1} AD_k^{-1} ; \quad D_k^{-2} = (z^k)^{-1} x^k$$

6. Calcular los tamaños de los pasos para las variables primal y dual.

$$\alpha_p^k = \min \left\{ 1, \min_{\Delta x_i^k < 0} \left\{ \frac{-x_i^k}{\Delta x_i^k} : i = 1, \dots, n \right\} \right\}$$

$$\alpha_d^k = \min \left\{ 1, \min_{\Delta z_i^k < 0} \left\{ \frac{-z_i^k}{\Delta z_i^k} : i = 1, \dots, n \right\} \right\}$$

7. Actualizar las variables Primal y Dual:

$$x^{k+1} = x^k + \gamma \alpha_p^k \Delta x^k$$

$$y^{k+1} = y^k + \gamma \alpha_d^k \Delta y^k$$

$$z^{k+1} = z^k + \gamma \alpha_d^k \Delta z^k$$

donde $0.95 \leq \gamma \leq 0.99995$

8. Reducir el parámetro de barrera: $\mu^{k+1} = \sigma^k \frac{(z^k)^T x^k}{n}$

donde n es el número de variables primales.

9. Volver al paso 3.

3.2. Aplicación del Método a un PPL.

Una empresa fabrica 2 tipos de fertilizantes, llamados fertilizantes A y B. Se emplean 3 tipos de materia prima en la preparación de estos fertilizantes como se muestra a continuación:

	Toneladas de materia prima necesarias para preparar 1 tonelada de		Cantidad mensual máxima de materia prima disponible
Materia prima	Fertilizantes tipo A	Fertilizantes tipo B	Toneladas
1	2	1	1500
2	1	1	1200
3	1	0	500
Lucro por tonelada	\$ 20	\$ 10	

El modelo matemático de este problema es:

$$\begin{aligned}
 \max \quad & z(x) = 20x_1 + 10x_2 \\
 \text{s.a.} \quad & 2x_1 + x_2 \leq 1500 \quad (1) \\
 & x_1 + x_2 \leq 1200 \quad (2) \\
 & x_1 \leq 500 \quad (3) \\
 & x_1 \geq 0 ; x_2 \geq 0
 \end{aligned}$$

$$(P): \begin{cases} \text{Min} & z = -20x_1 - 10x_2 \\ & 2x_1 + x_2 + x_3 = 1500 \\ \text{s.a.} & x_1 + x_2 + x_4 = 1200 \\ & x_1 + x_5 = 500 \\ & x_i \geq 0 \end{cases}$$

$$(D): \begin{cases} \text{Max} & z = 1500y_1 + 1200y_2 + 500y_3 \\ & 2y_1 + y_2 + y_3 + z_1 = -20 \\ & y_1 + y_2 + z_2 = -10 \\ \text{s.a.} & y_1 + z_3 = 0 \\ & y_2 + z_4 = 0 \\ & y_3 + z_5 = 0 \\ & z_i \geq 0 \end{cases} ;$$

La representación gráfica del espacio de soluciones se muestra en la fig. 3.1.

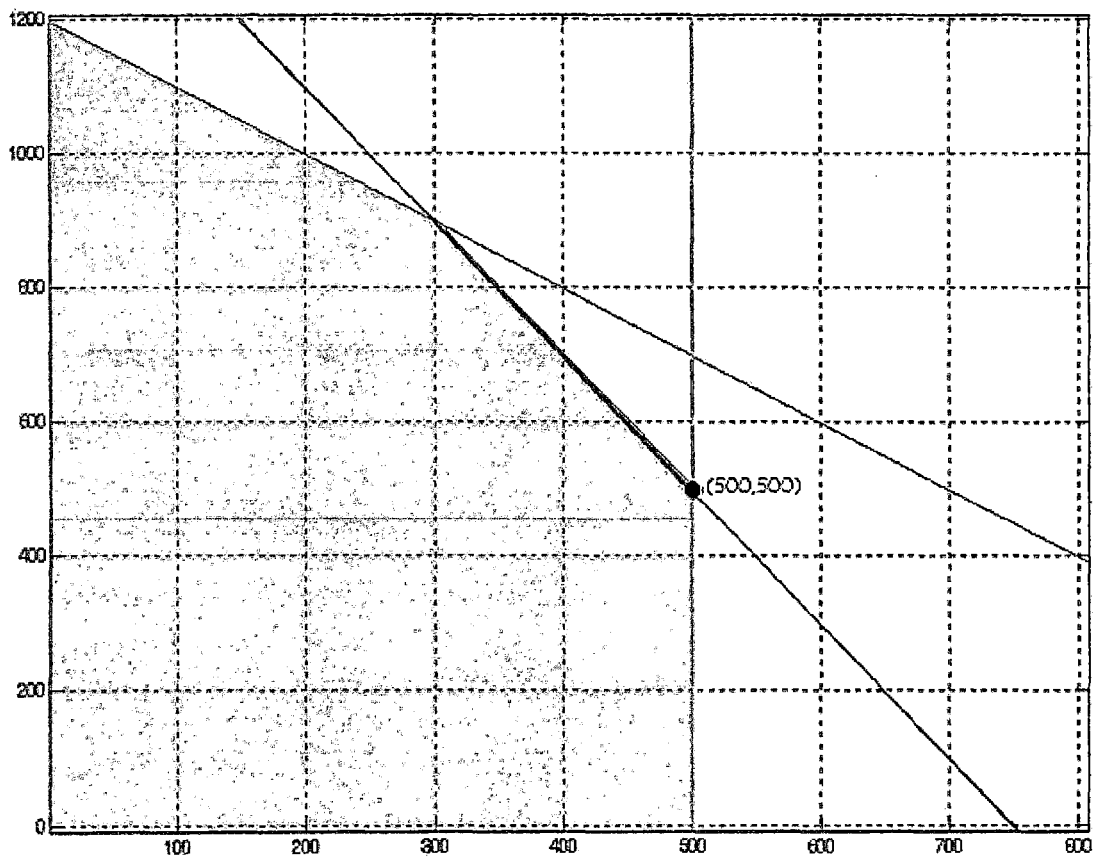


Fig. 3.1 Espacio de soluciones.

El segmento de recta resaltado es el conjunto de soluciones óptimas del problema, es decir todos los puntos sobre este segmento tienen el mismo valor para la función objetivo.

La solución x^* del problema por el método Simplex, es:

$$z(x) = 15000$$

$$x_1 = 500$$

$$x_2 = 500$$

Esta solución corresponde al vértice resaltado en la fig.3.1.

La solución del problema usando el método de Puntos Interiores Barrera Logarítmica

Primal-Dual es:

$$z(x) = 15000$$

$$x_1 = 396.2926$$

$$x_2 = 707.4144$$

La fig. 3.2 muestra el punto inicial calculado por el algoritmo puntos interiores y el camino seguido hasta encontrar el óptimo del problema.

Veamos: Hallemos el punto inicial.

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}; \quad b = \begin{pmatrix} 1500 \\ 1200 \\ 500 \end{pmatrix}; \quad c^T = (-20 \quad -10 \quad 0 \quad 0 \quad 0)$$

• Vector primal:

$$\hat{x}_j = \frac{1}{\|A_j\|_2 + 1} \Rightarrow \begin{cases} \hat{x}_1 = \frac{1}{\|A_1\|_2 + 1} = \frac{1}{\sqrt{6} + 1} \\ \hat{x}_2 = \frac{1}{\|A_2\|_2 + 1} = \frac{1}{\sqrt{2} + 1} \\ \hat{x}_3 = \frac{1}{2} \quad \hat{x}_4 = \frac{1}{2} \quad \hat{x}_5 = \frac{1}{2} \end{cases} \Rightarrow \hat{x} = \begin{pmatrix} 0,2898979 \\ 0,4142135 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix}$$

$$A\hat{x} = \begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0,2898979 \\ 0,4142135 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 1,4940093 \\ 1,2041113 \\ 0,7898979 \end{pmatrix}$$

$$\eta = \frac{\|b\|_2 + 1}{\|A\hat{x}\|_2 + 1} = \frac{1985,943324}{3,07506} = 645,822$$

$$x^0 = \eta \hat{x} = 645,822 \begin{pmatrix} 0,2898979 \\ 0,4142135 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 187,22244 \\ 267,50819 \\ 322,911 \\ 322,911 \\ 322,911 \end{pmatrix}$$

- Vector dual.

$$z^0 = \begin{cases} \text{si } c_j < 1, & z_j^0 = 1 \\ \text{caso contrario,} & z_j^0 = c_j \end{cases}$$

$$z^0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

De esta forma deben calcularse si se desea hacerlo de manera manual de lo contrario

utilizar matlab.

Nótese que la solución no corresponde a un punto extremo del conjunto de soluciones.

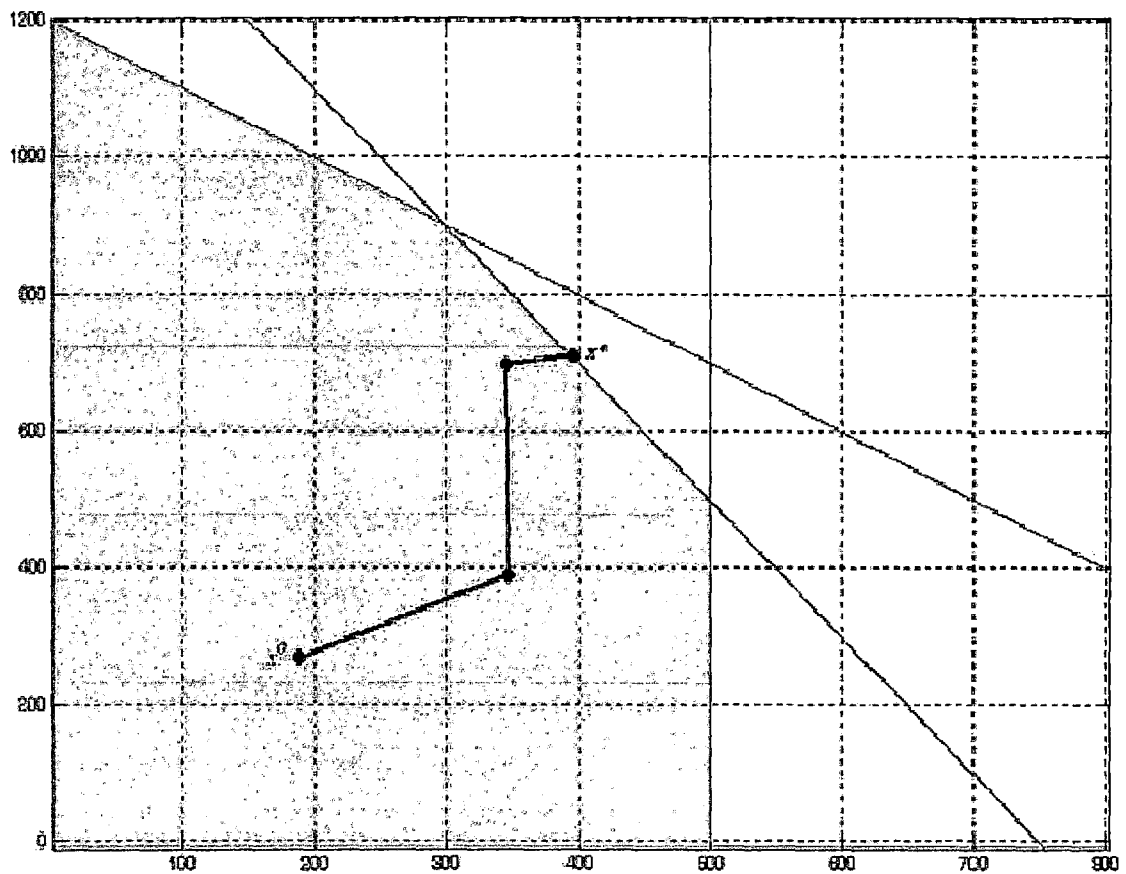


Figura 3.2. Solución mediante puntos interiores.

Como podemos observar el método de puntos interiores da la posibilidad de tener respuestas diferentes a los vértices del problema donde las restricciones se encuentran activas implicando el consumo total de los recursos.

Cuando un problema de programación lineal es resuelto por el método simplex, la solución es un vértice del conjunto factible de soluciones. Esto implica que las restricciones que forman ese vértice están activas (su relación es de igualdad). Para el caso del ejemplo mostrado, las soluciones posibles encontradas con el método simplex, podrían ser:

$$a. \ x_1 = 500 ; x_2 = 500$$

$$b. \ x_1 = 300 ; x_2 = 900$$

La solución (a) corresponde al caso donde las restricciones (1) y (3) se encuentran activas, es decir, la cantidad mensual de materia prima disponible 1 y 3 (Tabla 1) se agotan totalmente.

En la solución (b) las restricciones (1) y (2) se encuentran activas, es decir, la cantidad mensual de materia prima 1 y 2 se agotan totalmente.

De lo anterior se puede concluir que la materia prima 1 es indispensable y debe gastarse totalmente. Las materias primas 2 y 3 pueden ser manejadas de tal manera que se utilice una y se conserve la otra, o viceversa.

Se pretende por lo tanto, modificar el algoritmo de Puntos Interiores de Barrera Logarítmica Primal-Dual con el fin de que la respuesta obtenida persiga un objetivo adicional al de obtener el óptimo del problema.

Este objetivo adicional está directamente relacionado con las restricciones del problema, puesto que si existen óptimos alternativos, se pueden introducir elementos adicionales que permitan manipular de forma adecuada los recursos asociados a cada una de dichas restricciones.

3.3 Modificación al algoritmo de puntos interiores barrera logarítmica primal-dual para problemas con óptimos alternativos.

En el algoritmo de Barrera Logarítmica Primal-Dual se observa que la dirección de búsqueda del método está dado por:

$$\begin{aligned}\Delta x^k &= \Delta x_{ctr}^k + \Delta x_{obj}^k + \Delta x_{fac}^k \\ \Delta y^k &= (AD_k^{-2} A^T)^{-1} [AD_k^{-2} r_d^k - A(Z^k)^{-1} r_c^k + r_p^k] \\ \Delta z^k &= r_d^k - A^T \Delta y^k\end{aligned}$$

Donde Δy y Δz son las direcciones de búsqueda dual. Δx es la dirección de búsqueda primal y está compuesta por tres componentes: Δx_{ctr} dirección central, Δx_{obj} dirección

objetivo y Δx_{fac} dirección factible. Esta dirección de búsqueda nos permite obtener una variación directa sobre la solución del problema primal.

Las direcciones central y factible tienen como objetivo encontrar un punto factible cerca del centro del politopo primal, estas direcciones inicialmente son predominantes debido a que los parámetros μ y r_p son grandes.

Cuando se encuentra un punto factible, r_p tiende a cero y la dirección objetivo es la que domina.

Debido a que el punto inicial dado en el algoritmo se encuentra normalmente en la región factible, la dirección objetivo predomina desde el inicio del proceso de búsqueda, por lo tanto, se propone modificar dicha dirección de tal forma que persiga un segundo objetivo dado por las restricciones.

La modificación en dicha dirección está directamente relacionada con los gradientes de las restricciones que se pretenden favorecer; por tanto la dirección objetivo tendrá la siguiente forma:

$$\Delta x_{obj}^k = D_k^{-1} P^k D_k^{-1} c + \sum_i K_i \mu \nabla R_i$$

donde:

i : es el numero de restricciones que se incluirán para el ajuste de la dirección objetivo.

K_i : es un factor entre cero y uno, que indica el porcentaje de contribución del parámetro de barrera.

μ : es el parámetro de barrera, debe incluirse debido a que la dirección dada por el gradiente debe disminuir durante el proceso.

∇R_i : Es el gradiente de las restricciones que se pretenden favorecer.

3.3.1 Ejemplo de aplicación

Se aplicará la metodología expuesta al ejemplo tratado en la sección 3.1. Para realizar la comparación se parte del resultado mostrado en la fig 3.2

$$z(x) = 15000$$

$$x_1 = 396.2926$$

$$x_2 = 707.4144$$

En este, los recursos de las restricción 1 fueron agotados debido a que sobre esta recta se encuentra el espacio de soluciones, en cuanto a las restricciones 2 y 3, cuentan con recursos debido a que la solución obtenida no activa dichas restricciones.

Se pretende repartir los recursos de las restricciones 2 y 3 de tal forma que la solución siga siendo óptima.

Caso 1: Conservar una buena cantidad de la materia prima 2, y a la vez, mantener disponible en bajo porcentaje la materia prima 3. La dirección positiva del gradiente de la restricción 2 indica gastar recursos mientras que el gradiente negativo implica ahorro de dicho recurso. La proporción que se desee conservar está sujeto a otro tipo de restricciones que pueden ser adicionadas al problema.

Para este caso solo se utilizará información de la restricción 2. Se fijará el factor K igual a 0.6 que indica una contribución del 60% del parámetro de barrera sobre la dirección de ajuste. Este valor debe ser sintonizado de acuerdo al objetivo que se persigue.

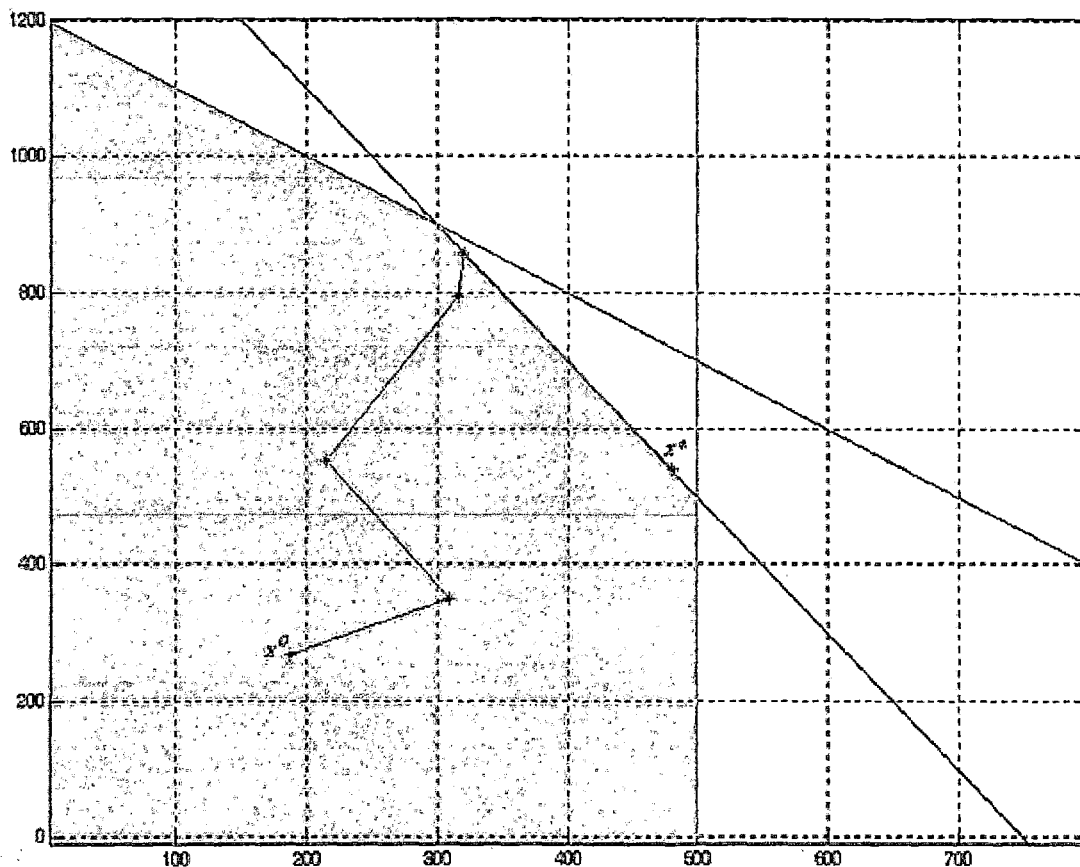


Figura 3.3. Ejemplo de aplicación caso 1.

$$z(x) = 14997.54$$

$$x_1 = 479.9680$$

$$x_2 = 539.8185$$

Caso 2: Conservar una buena cantidad de la materia prima 3, y a la vez, mantener disponible en bajo porcentaje la materia prima 2. La dirección del gradiente de la restricción 3 indica gastar recursos mientras que el gradiente negativo implica ahorro de dicho recurso.

Para este ejemplo se incluirá información de ambas restricciones (2 , 3) para ver el comportamiento de la dirección de búsqueda. Se fijará el factor K_1 en 0.6 y K_2 en 1 para los porcentajes de contribución del parámetro de barrera sobre la dirección de ajuste.

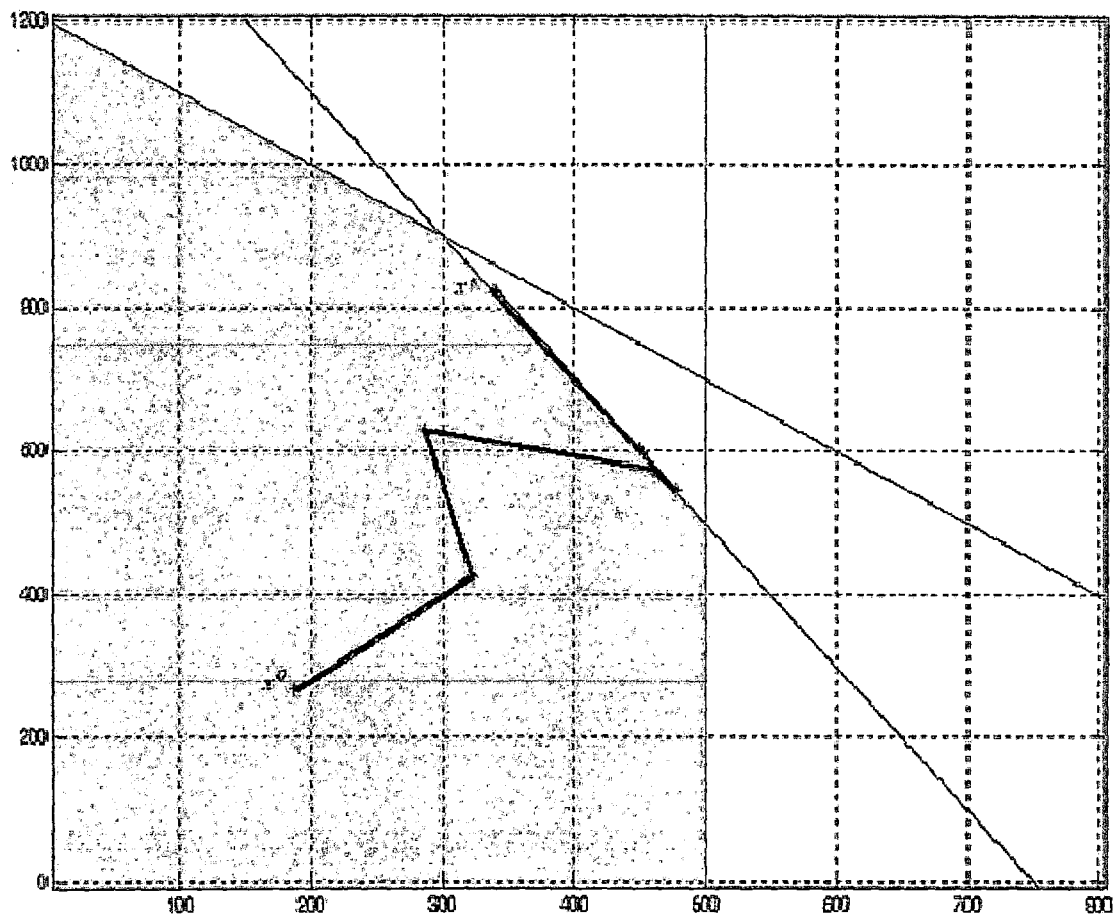


Figura 3.4. Ejemplo de aplicación caso 2.

$$z(x) = 15000$$

$$x_1 = 339.2772$$

$$x_2 = 821.4303$$

Lo anterior muestra que si se considera el gradiente de las restricciones que acotan el espacio de soluciones, de tal forma que contribuyan al cambio de la dirección objetivo, se obtendrá una convergencia más exacta del método.

CONCLUSIONES

Se Analiza y detalla los pasos del método de puntos interiores obteniéndose el algoritmo que permite encontrar la solución de un PPL

Para PPL con óptimos alternativos el método de Puntos Interiores obtiene soluciones diferentes a los puntos extremos. del polítopo.

El método de Puntos Interiores Primal-Dual busca un punto factible del problema, ubicado en el centro del polítopo y desde allí avanza hacia el óptimo actualizando las variables del problema.

Debido al proceso de búsqueda se puede modificar esta priorizando un objetivo adicional.

Una modificación posible consiste en el ajuste de la variación de la dirección objetivo del vector de búsqueda primal, considerando objetivos adicionales, pues durante el proceso del algoritmo predomina sobre las direcciones central y factible.

BIBLIOGRAFÍA

- [1] Bronson, R. (2009) *Investigación de Operaciones*. México: Mc Graw Hill.
- [2] Castillo, E., Conejo, A., Pedregal, P., García, R., Alguacil, N. (2002) *Formulación y Resolución de Modelos de Programación Matemática en Ingeniería y Ciencia*.
- [3] Quintana, V., Torres, G., Medina, P. (2000) "Interior point methods and their applications to power systems: A classification of publications and software codes", IEEE transactions Power Systems, vol 15, No. 1.
- [4] Rider, M. (2004) *Método de Puntos Interiores para Optimización en Sistemas de Potencia*. Universidad Tecnológica de Pereira.
- [5] Salamanca, A., Ruiz, J. (2002) *Algoritmo del elipsoide interior para programación lineal*. Universidad Politécnica de Madrid.
- [6] Taha, H. (2012) *Investigación de operaciones*. novena edición. Mexico: editorial Pearson
- [7] Wright, S. (1997) *Primal-Dual Interior Point Methods*. Philadelphia: Siam.