PESQUIMAT 24(1): 91–105 (2021)
https://doi.org/10.15381/pesquimat.v24i1.20558

ISSN:1560-912X/ ISSN-E:1609-8439
Facultad de Ciencias Matemáticas – UNMSM

Programación lineal: convergencia del algoritmo primal-dual de paso largo por el método de la función barrera

Juan Luna Valdez [1] ***y Edinson Montoro Alegre*** [2]

Resumen: En el presente trabajo se tratará de desarrollar y describir el método punto interior primal-dual para resolver el problema de programación lineal. Este método se caracteriza por utilizar funciones barrera, para el problema primal y para el dual y asi deducir el sistema no lineal primal-dual, cuya solución define la trayectoria central del método de punto interior. Se demuestra que el número total de iteraciones que ejecuta es de orden polinomial.

Palabras clave: Programación convexa, trayectoria central, método de punto interior, algorítmo primal-dual.

Linear programming: convergence of the long-pass primal-dual algorithm by the barrier function method

Abstract: In this work we will try to develop and describe the primal-dual interior point method to solve the linear programming problem. This method is characterized by using barrier functions, for the primal problem and for the dual and thus deduce the primal-dual nonlinear system, whose solution defines the central path of the interior point method. It is shown that the total number of iterations that it executes is of polynomial order.

Keywords: Convex programming, central path, interior point method, primal-dual algorithm.

Recibido: 18/05/2021. Aceptado: 10/06/2021. Publicado online: 30/06/2021.

[1]UNMSM, Facultad de Ciencias Matemáticas. e-mail: jlunav@unmsm.edu.pe
[2]UNMSM, Facultad de Ciencias Matemáticas, e-mail:emontoroa@unmsm.edu.pe

1. Introducción

Dado un problema de programación lineal, llamado problema primal, existe otro problema de programación lineal, llamado problema dual, estrechamente relacionado con él. Se dice que ambos problemas son mutuamente duales. Bajo ciertas hipótesis, 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. Esto puede suponer una ventaja computacional relevante.

En este trabajo se estudia el método de punto interior y se resolverá un problema de programación lineal con restricción usando un algoritmo primal-dual. Para ello usaremos el métodos de barrera y las condiciones de Karush-Khun-Tucker.

Consideremos el siguiente problema de programación lineal, expresado en forma estándar, como problema primal:

$$(P) \qquad \begin{cases} \text{mín } \; c^T x \\ s.a : Ax = b \\ x \geq 0 \end{cases}$$

donde $c, \; x \in R^n, \; b \in R^m, \; A \in R^{m\times n} \;\; y \;\; rango(A) = m, \;\; m < n.$

El problema dual (P), expresado en su forma estándar,

$$(D) \qquad \begin{cases} \text{máx } \; b^T y \\ s.a : A^T y + s = c \\ s \geq 0 \end{cases}$$

donde $s \in R^n$.

Se plantea el problema dual por que permite resolver problemas lineales donde el número de restricciones es mayor que el número de variables. Gracias a los resultados obtenidos, la solución de uno de los problemas (primal o dual) nos proporciona de forma automática la solución del otro problema.

El alcance de la solución dual permite realizar importantes interpretaciones de los problemas de programación lineal. Además la solución del dual se puede obtener desde el problema primal lo cual permite generar algunos métodos como el método dual del simplex; otra de las ventajas de la dualidad, es la posibilidad de resolver gráficamente algunos problemas.

2. Preliminares

En esta sección presentaremos una serie de definiciones y resultados básicos del análisis convexo, problema de programación lineal y el método de barrera en programación lineal.

Definición 2.1 *Sea* $S \subseteq R^n$, $S \neq \phi$. *Se dice que* S *es convexo si*

$$\lambda x + (1-\lambda)\, y \in S \ , \quad \forall x, y \in S, \ \forall \lambda \in [0,1]$$

Definición 2.2 *Sea* $S \subseteq R^n$ *, un conjunto convexo. Un vector* $d \in R^n$, $d \neq 0$*, se* ***dice dirección de S,*** *si* $\forall x \in S$ *se tiene que* $x + \lambda d \in S, \ \forall \lambda \geq 0.$

Definición 2.3 *Sea* $S \subseteq R^n$ *un conjunto convexo. Se dice que* $f : S \longrightarrow R$ *es una* ***función convexa*** *si*

$$f(\lambda x + (1-\lambda)\, y) \leq \lambda f(x) + (1-\lambda) f(y)\,, \quad \forall x, y \in S, \qquad 0 \leq \lambda \leq 1$$

Definición 2.4 *Sea* $S \subseteq R^n$ *un conjunto convexo. Se dice que* $f : S \longrightarrow R$ *es una* ***función convexa*** *si* $\forall k \in N$, $x_1, ..., x_k \in S$ *y* $\lambda_1, ..., \lambda_k \geq 0$ *tales que* $\sum\limits_{i=1}^{k} \lambda_i = 1$*, se tiene*

$$f(\lambda_1 x_1 + ... + \lambda_k x_k) \leq \lambda_1 f(x_1) + ... + \lambda_k f(x_k)$$

Definición 2.5 *Sea* $S \subseteq R^n$*, no vacío,* $f : S \longrightarrow R$, $\overline{x} \in S$, $d \neq 0$ *tal que* $\overline{x} + \lambda d \in S$, $\forall \lambda \in [0, \eta[$, *algún* $\eta > 0$*. Se define la* ***derivada direccional*** *de* f *en el punto* $\overline{x}$*, en la dirección* d*, por,*

$$f'(\overline{x}, d) = \lim_{\lambda \to 0^+} \frac{f(\overline{x} + \lambda d) - f(\overline{x})}{\lambda} \in \overline{R}$$

donde $\overline{R} = R \cup \{-\infty, +\infty\}$ *, (cuando el límite existe).*

Definición 2.6 *Sea* $S \subseteq R^n$*, no vacío, una función* $f : S \longrightarrow R$*, es* ***diferenciable*** *en* $\overline{x} \in int(S)$ *si existe* $\nabla f(\overline{x}) \in R^n$ *tal que*

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^T (x - \overline{x}) + o(x - \overline{x})\,, \qquad \forall x \in S$$

donde $\displaystyle \lim_{x \to \overline{x}} \frac{o(x - \overline{x})}{\|x - \overline{x}\|} = 0.$

Teorema 2.7 *Si la función* f *es diferenciable en el interior de* S*, entonces*

$$f'(\overline{x}, d) = \nabla f(\overline{x})^T d$$

Definición 2.8 *Sea* $f : R^n \longrightarrow R$*, se dice que* $d \in R^n$ *es* ***dirección de descenso*** *de* f *en* $\overline{x}$ *si* $\nabla f(\overline{x})^T d < 0.$

Definición 2.9 *Sea* $S \subseteq R^n$, $S \neq \phi$ *y* $\overline{x} \in S$*. Se denomina* ***cono de direcciones admisibles*** *de* S *en* $\overline{x}$ *al conjunto:*

$$A(\overline{x}) = \{d \in R^n \ / \ d \neq 0, \ \ \overline{x} + \lambda d \in S, \quad \forall \lambda \in [0, \eta[\quad \text{para algún} \ \ \eta > 0\}$$

y el conjunto $D(\overline{x}) = \left\{ d \in R^n \ / \ \ \nabla f(\overline{x})^T d < 0 \right\}$*, es el conjunto de direcciones de descenso.*

2.1. Función lagrangeana

En los problemas de Programación Matemática de la forma

$$(P) \quad \begin{cases} \text{mín } f(x) \\ s.a: g_i(x) \leq 0, \quad i = 1, ..., m \\ x \in S \end{cases}$$

por lo general se usa la función lagrangena que consiste en construir el siguiente programa

$$(P_\lambda) \quad \begin{cases} \text{mín } f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) \\ s.a: \\ x \in S \end{cases}$$

Donde $\lambda \in R^m$.

Los valores λ_j son llamados multiplicadores de Lagrange o parámetros lagrangeanos.
Por cada λ se tiene un programa de optimización (P_λ) en la variable x.

La función

$$f_\lambda(x) = f(x) + \lambda^T g(x) = f(x) + \sum \lambda_i g_i(x)$$

es llamada función lagrangeana de (P_λ) y se escribe como $L(x, \lambda)$.

2.2. Condiciones de Karush-Kuhn-Tucker (KKT)

Teorema 2.10 *Sea $S \subseteq R^n$ y $f : S \longrightarrow R$, diferenciable, y consideremos el programa*

$$(P_1) \qquad \left\{ \min_{x \in S} f(x) \right.$$

Sea $\overline{x}$ solución de (P_1) y supongamos que $\{\nabla g_i(\overline{x}), \quad i \in I\}$ es un conjunto linealmente independiente. Entonces existen $\mu_1, \mu_2, ..., \mu_m \in R$ tales que

$$\nabla f(\overline{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i(\overline{x}) = 0$$

$$\mu_i g_i(\overline{x}) = 0, \quad i = 1, 2, ..., m$$

$$\mu_i \geq 0, \quad i = 1, 2, ..., m$$

2.3. Dualidad y condiciones de optimalidad

Consideremos un problema de programación lineal, expresado en su forma estándar, como programa matemática lineal primal:

$$(P) \quad \begin{cases} \text{mín } c^T x \\ s.a: Ax = b \\ x \geq 0 \end{cases} \tag{1}$$

donde $c,\ x \in R^n,\ b \in R^m,\ A \in R^{m \times n}$ *y* $rango(A) = m, \quad m < n$.

El programa dual (P), será:

$$(PD) \quad \begin{cases} \text{máx } b^T y \\ s.a: A^T y \leq c \end{cases}$$

Expresado en su forma estándar,

$$(D) \qquad \begin{cases} \text{máx } b^T y \\ s.a : A^T y + s = c \\ s \geq 0 \end{cases} \tag{2}$$

donde $s \in R^n$.

Utilizamos el Método de Barrera Logarítmica para penalizar las restricciones de desigualdad.

Definimos:

$$S^0 = \{x \in R^n \ / \ Ax = b, \quad x > 0\} \neq \phi$$

$$T^0 = \left\{(y,s) \in R^m \times R^n \ / \ A^T y + s = c, \quad s > 0\right\} \neq \phi$$

Asociamos a los problemas (P) y (D) la siguiente función barrera logarítmica primal-dual:

$$f_{PD}(x,s;\mu) := \frac{x^T s}{\mu} - \sum_{j=1}^{n} \log(x_j s_j) \tag{3}$$

Para cualquier par de soluciones primal-dual factible tenemos:

$$x^T s = c^T x - b^T y$$

Generando así, los problemas barreras para el primal y dual

$$(P_\mu) \qquad \begin{cases} \min\limits_{s.a.} c^T x - x^T s + \mu \sum_{j=1}^{n} \log(x_j s_j), \quad \mu > 0 \\ x \in S^0 \end{cases}$$

$$(D_\mu) \qquad \begin{cases} \max\limits_{s.a.} b^T y + x^T s - \mu \sum_{j=1}^{n} \log(x_j s_j), \quad \mu > 0 \\ (y,s) \in T^0 \end{cases}$$

(P_μ) es un programa convexo y (D_μ) es un programa cóncavo, como $S^0 \times T^0$ son no vacíos, entonces (P_μ) y (D_μ) son superconsistentes, por tanto los problemas poseen solución única. Aplicamos las condiciones de KKT para caracterizar una solución de (P_μ) y (D_μ). Pero antes,

calculemos la función lagrangeana de cada uno.

La función lagrangeana de (P_μ):

$$L_P(x,s;\mu) = c^T x - x^T s + \mu \sum_{j=1}^{n} \log(x_j s_j) + y^T (Ax - b) \tag{4}$$

Aplicando las condiciones de KKT obtenemos las condiciones de optimalidad:

$$(PD) \qquad \begin{cases} A^T y + s - c = 0, \quad s > 0 \\ Ax - b = 0, \quad x > 0 \\ Xs - \mu e = 0 \end{cases}$$

El sistema (PD) es el sistema KKT para (P) y (D).

Donde:

$$X = diag\,(x_1, x_2, ..., x_n)$$
$$S = diag\,(s_1, s_2, ..., s_n)$$
$$e = (1, 1, ..., 1)^T$$

Sea $(\ x(\mu), y(\mu), s\,(\mu))$ la solución del sistema (PD) para cada $\mu > 0$. Con éstas variables se define la Curva Paramétrica (o Trayectoria Central).

$$C = \{(\ x(\mu), y(\mu), s\,(\mu))\ \ /\ \mu > 0\}$$

que se encuentra en $S^0 \times T^0$.

2.4. Direcciones de Búsqueda

Para resolver (PD) utilizamos el método de Newton.
Dado $w = (x, y, s)$ punto inicial, donde $x > 0,\ s > 0$, resolvemos el sistema:

$$J\,(F\,(x, y, s))\,dw = -F\,(x, y, s) \tag{5}$$

donde $dw = (d_x, d_y, d_s)^T$ direcciones de Newton que son calculados a partir de un punto inicial $w = (x, y, s)$, tal que (x, y, s) es un punto factible interior Primal-Dual.

Se tiene que:

$$F\,(x, y, s) = \begin{bmatrix} Ax - b \\ A^T y + s - c \\ Xs - \mu e \end{bmatrix} \quad \Longrightarrow \quad J\,(F\,(x, y, s)) = \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix}$$

Entonces,

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = - \begin{bmatrix} Ax - b \\ A^T y + s - c \\ Xs - \mu e \end{bmatrix}$$

Como (x, s) es un par factible interior primal-dual, i.e., $x > 0,\ s > 0$, y se cumple

$$Ax - b = 0$$
$$A^T y + s - c = 0$$

Luego tenemos el siguiente sistema:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ Xs - \mu e \end{bmatrix}$$

Los vectores $d_x,\ d_y,\ d_s$ satisfacen:

$$(*) \qquad \begin{cases} Ad_x = 0 \\ A^T d_y + d_s = 0 \\ Sd_x + Xd_s = \mu e - Xs \end{cases}$$

Resolvemos el sistema (*) para obtener las direcciones de búsqueda en el espacio primal y dual que son denotadas por $d_x,\ d_y,\ d_s$.

De esta manera, obtenemos las direcciones de búsqueda $d_x,\ d_y,\ d_s$.

$$\begin{aligned} d_x &= \mu S^{-1}e - x - S^{-1}Xd_s \\ d_y &= -\left(AXS^{-1}A^T\right)^{-1}\left(\mu AS^{-1}e - b\right) \\ d_s &= A^T\left(AXS^{-1}A^T\right)^{-1}\left(\mu AS^{-1}e - b\right) \end{aligned}$$

Ahora definimos la medida de proximidad de x, y, s con el μ-centros como

$$\delta\left(x, s; \mu\right) := \left\|u^{-1} - u\right\| \tag{6}$$

2.5. Algoritmo de Paso Largo

El procedimiento del algoritmo de paso largo es que para las actualizaciones del parámetro de barrera, hace que sean más eficientes. Esto es, después de la actualización, la iteración actual se convierte en el nuevo par central.

La función de barrera logarítmica primal-dual

$$f_{PD}\left(x, s; \mu\right) := \frac{x^T s}{\mu} - \sum_{j=1}^{n} \log\left(x_j s_j\right)$$

es un candidato.

Esta afirmación se sustenta en una propiedad importante de f_{PD} en relación con la dirección de búsqueda que se define en el algoritmo primal-dual de paso corto. Esta dirección de búsqueda resulta ser una dirección de descenso para f_{PD}.

Lema 2.11

$$\left(\nabla_x f\right)^T d_x + \left(\nabla_s f\right)^T d_s = -\left\|u - u^{-1}\right\|^2 = -\delta^2$$

donde

$$f := f_{PD} = \frac{x^T s}{\mu} - \sum_{j=1}^{n} \log\left(x_j s_j\right)$$

2.6. Variación de la Función Barrera

Para probar la convergencia polinomial del algoritmo, es fundamental que se pueda garantizar una reducción importante de la función barrera en cada iteración.

Lema 2.12 *Sea h un vector en R^n tal que $\|h\| < 1$. Entonces*

$$\sum_{i=1}^{n} \log\left(1 + h_i\right) \geq e^T h + \|h\| + \log\left(1 - \|h\|\right)$$

Es posible mostrar que existe una longitud de paso a lo largo de la dirección de búsqueda, que garantiza una mínima disminución de la función de barerra por un término constante que depende de la proximidad de la iteración actual de la trayectoria central.

Este argumento se utiliza para limitar el número de iteraciones que se requiere para llegar a un punto cerca de un μ-centro dado.

3. Lemas importantes de convergencia

Teniendo en cuenta la que hacemos un movimiento desde x hasta $x + \alpha d_x$ en el $x-$espacio y un movimiento desde s hasta $s + \alpha d_s$ en el $s-$espacio, donde $0 < \alpha < 1$.

En términos generales, nos dice que un tamaño de paso α se toma desde (x, s) a lo largo de la dirección (d_x, d_s). La variación inducida de la función de barrera para un μ fijo, es dado por

$$\begin{aligned}
\Delta f(\alpha) &:= f(x + \alpha d_x,\ s + \alpha d_s;\ \mu) - f(x, s; \mu) \\
&= \frac{1}{\mu}(x + \alpha d_x)^T (s + \alpha d_s) - \sum_{j=1}^{n} \log\left[(x + \alpha d_x)_j (s + \alpha d_s)_j\right] \\
&- \frac{x^T s}{\mu} + \sum_{j=1}^{n} \log(x_j s_j) \\
&= \frac{1}{\mu}\left(x^T + \alpha d_x^T\right)(s + \alpha d_s) - \sum_{j=1}^{n} \log\left(x_j\left(1 + \alpha\frac{dx_j}{x_j}\right) s_j\left(1 + \alpha\frac{ds_j}{s_j}\right)\right) \\
&- \frac{x^T s}{\mu} + \sum_{j=1}^{n} \log(x_j s_j) \\
&= \frac{\alpha}{\mu}\left(x^T d_s + s^T d_x\right) - \sum_{j=1}^{n} \log\left[\left(1 + \alpha\frac{dx_j}{x_j}\right)\left(1 + \alpha\frac{ds_j}{s_j}\right)\right] \qquad (7)
\end{aligned}$$

Lema 3.1 *Sea* $\delta := \delta(x, s; \mu) > 0$, *Entonces*

$$w = \sqrt{\left\|X^{-1}d_x\right\|^2 + \left\|S^{-1}d_s\right\|^2} > 0$$

Tomando

$$\alpha = \frac{1}{w} - \frac{1}{\delta^2 + w}$$

se tiene $x + \alpha d_x > 0$ *y* $s + \alpha d_s > 0$.

Por otra parte, si $\Gamma := \dfrac{2\delta}{\delta + \sqrt{4 + \delta^2}}$, *Entonces*

$$\Delta f(\alpha) \leq -\Gamma + \log(1 + \Gamma)$$

Lema 3.2 *Si* $\delta : \delta(x, s; \mu) < \dfrac{1}{\sqrt{2}}$, *entonces*

$$f(x, s; \mu) - f(x(\mu), s(\mu); \mu) \leq -\sigma - \log(1 - \sigma)$$

donde $\sigma := \dfrac{\delta}{2}\left(\delta + \sqrt{4 + \delta^2}\right)$

Demostracion 1

$$\begin{aligned} f(x,s;\mu) - f(x(\mu), s(\mu); \mu) &= \frac{x^T s}{\mu} - \sum_{j=1}^{n} \log(x_j s_j) - \frac{(x(\mu))^t s(\mu)}{\mu} + \sum_{j=1}^{n} \log\Big((x(\mu))_j (s(\mu))_j\Big) \\ &= \frac{x^T s}{\mu} - \sum_{j=1}^{n} \log(x_j s_j) - \frac{\mu n}{\mu} + \sum_{j=1}^{n} \log(\mu) \\ &= \frac{x^T s}{\mu} - n - \left(\sum_{j=1}^{n} \log(x_j s_j) - \sum_{j=1}^{n} \log(\mu)\right) \\ &= e^T\left(\frac{Xs}{\mu} - e\right) - \sum_{j=1}^{n} \log\left(\frac{x_j s_j}{\mu}\right) \end{aligned} \tag{8}$$

Por otro lado tenemos que

$$\begin{aligned} x^T s &= e^T S x = e^T X s \\ e^t e &= \eta \end{aligned}$$

Denotemos $h = \dfrac{Xs}{\mu} - e$

Afirmación 1: $\|h\| < 1$ *En efecto,*

$$\begin{aligned} \|h\| &= \left\|\frac{Xs}{\mu} - e\right\| = \left\|U\left(u - u^{-1}\right)\right\| \leq \|U\| \left\|u - u^{-1}\right\| \\ &\leq \|u\|_\infty \left\|u - u^{-1}\right\| \\ &\leq \|u\|_\infty \ \delta \end{aligned} \tag{9}$$

Afirmación 2:

$$\|u\|_\infty \ \delta \leq \frac{\delta}{2}\left(\delta + \sqrt{4 + \delta^2}\right)$$

En efecto, Sabemos que $\left|u_i^{-1} - u_i\right| \leq \left\|u^{-1} - u\right\|$ *para cada* $i = 1, 2, ..., n$

Como $\delta(x,s;\mu) := \left\|u^{-1} - u\right\|$ *entonces* $\left|u_i^{-1} - u_i\right| \leq \delta$

$$\rightarrow \quad -\delta \leq u_i^{-1} - u_i \leq \delta \tag{10}$$

Por definición, $u = \dfrac{(Xs)^{1/2}}{\sqrt{\mu}}$, *como* $x > 0,\ s > 0 \ \rightarrow u_i > 0$

De (10), se tiene que

$$\begin{gathered} -u_i\delta \leq 1 - u_i^2 \leq u_i\delta \\ u_i^2 - \delta u_i - 1 \leq 0 \leq u_i^2 + \delta u_i - 1 \\ \sqrt{\left(u_i - \frac{\delta}{2}\right)^2} \leq \sqrt{\frac{\delta^2}{4} + 1} \leq \sqrt{\left(u_i + \frac{\delta}{2}\right)^2} \\ \left|u_i - \frac{\delta}{2}\right| \leq \frac{\sqrt{\delta^2 + 4}}{2} \leq \left|u_i + \frac{\delta}{2}\right| \end{gathered} \tag{11}$$

De (11) se tiene que,

$$u_i \leq \frac{\delta + \sqrt{\delta^2 + 4}}{2}$$

$$\frac{-\delta + \sqrt{\delta^2 + 4}}{2} \leq u_i$$

Luego,

$$\frac{-\delta + \sqrt{\delta^2 + 4}}{2} \leq u_i \leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \tag{12}$$

Entonces,

$$\begin{aligned} \underbrace{\max_{1 \leq i \leq n} |u_i|} &\leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \\ \|u\|_\infty &\leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \end{aligned}$$

Luego, se tiene que

$$\|u\|_\infty \, \delta \leq \frac{\delta}{2} \left(\delta + \sqrt{\delta^2 + 4} \right) \tag{13}$$

con esto queda probado la afirmación 2.

Por otra parte, de (12)

$$\frac{2}{\delta + \sqrt{\delta^2 + 4}} \leq u_i^{-1} \leq \frac{2}{-\delta + \sqrt{\delta^2 + 4}}$$

$$\frac{-\delta + \sqrt{\delta^2 + 4}}{2} \leq u_i^{-1} \leq \frac{\delta + \sqrt{\delta^2 + 4}}{2}$$

Luego,

$$\begin{aligned} \underbrace{\max_{1 \leq i \leq n} \left| u_i^{-1} \right|} &\leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \\ \left\| u^{-1} \right\|_\infty &\leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \end{aligned}$$

Por otro lado, como

$$\begin{aligned} \delta < \frac{1}{\sqrt{2}} \quad &\rightarrow \quad 0 \leq \delta^2 < \frac{1}{2} \\ &\rightarrow \quad \delta^2 + 4 < \frac{1}{2} + 4 = \frac{9}{2} \quad \rightarrow \quad \sqrt{4 + \delta^2} < \frac{3}{\sqrt{2}} \end{aligned}$$

De (9) y afirmación 2,

$$\begin{aligned} \|h\| &\leq \frac{\delta}{2} \left(\delta + \sqrt{4 + \delta^2} \right) \\ &< \frac{1}{2\sqrt{2}} \left(\delta + \sqrt{4 + \delta^2} \right) \\ &< \frac{1}{2\sqrt{2}} \left(\frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} \right) = \frac{1}{2\sqrt{2}} \left(\frac{4}{\sqrt{2}} \right) = 1 \end{aligned}$$

entonces,

$$\|h\| < 1 \tag{14}$$

Con esto queda probado la afirmación 1.

Ahora, de (8) y lema (2.12) se tiene que,

$$\begin{aligned} f(x, s; \mu) - f(x(\mu), s(\mu); \mu) &= e^T h - \sum_{j=1}^{n} \log(1 + h_j) \\ &\leq e^T h - e^T h - \|h\| - \log(1 - \|h\|) \\ &\leq -\|h\| - \log(1 - \|h\|) \end{aligned} \tag{15}$$

Pero sabemos que $h = \dfrac{Xs}{\mu} - e$
esto es,

$$\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} = \begin{bmatrix} x_1 s_1/\mu \\ x_2 s_2/\mu \\ \vdots \\ x_n s_n/\mu \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 s_1/\mu - 1 \\ x_2 s_2/\mu - 1 \\ \vdots \\ x_n s_n/\mu - 1 \end{bmatrix}$$

entonces

$$h_j = \frac{x_j s_j}{\mu} - 1 \qquad \forall\ j = 1, 2, ..., n$$
$$\rightarrow\ h_j + 1 = \frac{x_j s_j}{\mu}$$

Por otro lado, definimos

$F : \langle 0, 1\rangle \longrightarrow R \ \ / \ \ F(x) = -x - \log(1 - x)$

$$F'(x) = -1 + \frac{1}{1-x} = \frac{x}{1-x}$$

- *Si* $\dfrac{x}{1-x} > 0 \quad \rightarrow \quad x > 0$
  Entonces $F(x)$ *es creciente.*
- *Si* $\dfrac{x}{1-x} < 0 \quad \rightarrow \quad x < 0$ *(pero no esta definida en el dominio)*

$\therefore\ F(x)$ *es creciente en* $\langle 0, 1\rangle$

Así mismo de afirmación 1 y (9) se tiene

$$\|h\| \leq \|u\|_\infty \delta \leq \frac{\delta}{2}\left(\frac{\delta + \sqrt{\delta^2 + 4}}{2}\right) = \sigma$$

entonces

$$\|h\| \leq \sigma$$

Luego,

$$\begin{aligned} F(\|h\|) &\leq F(\sigma) \\ -\|h\| - \log(1-\|h\|) &\leq -\sigma - \log(1-\sigma) \end{aligned}$$

De (15), se tiene

$$f(x,s;\mu) - f(x(\mu), s(\mu);\mu) \leq -\sigma - \log(1-\sigma) \tag{16}$$

Lema 3.3 *“ Sea $\mu^{+} := (1-\theta)\mu$. Entonces*

$$f(x(\mu), s(\mu);\mu) - f\left(x\left(\mu^{+}\right), s\left(\mu^{+}\right);\mu^{+}\right) = n\log(1-\theta)$$

”

Lema 3.4 *Supongamos que (x,s) es un par de soluciones factibles primal-dual tal que $\delta := \delta(x,s;\mu) < \dfrac{1}{\sqrt{2}}$. Entonces*

$$\frac{x^T s}{\mu} < n + \sqrt{n}$$

3.1. Análisis de Convergencia

Lema 3.5 *Si k es el menor entero tal que*

$$k \geq \frac{1}{\theta}\log\left(\frac{2n\mu^0}{\varepsilon}\right)$$

después de k iteraciones exteriores como máximo, el algoritmo primal-dual del paso largo, se detiene y se obtiene el par (x,s) factible interior primal-dual que satisface

$$x^T s \leq \varepsilon$$

Demostracion 2 *Sea k el número de iteraciones exteriores. Entonces al actualizar k-veces el parámetro de barrera, del lema (3.3) se obtiene lo siguiente*

$$\begin{aligned} \mu_k &= (1-\theta)\mu_{k-1} \\ &= (1-\theta)(1-\theta)\mu_{k-2} \\ &\vdots \\ &= \underbrace{(1-\theta)(1-\theta)\ldots(1-\theta)}_{k-veces}\mu_0 \end{aligned}$$

$$\rightarrow \mu_k = (1-\theta)^k \mu_0 \tag{17}$$

Además en k-iteraciones exteriores, se tiene $\left(x^k, s^k\right)$ el par factible interior primal-dual.

Luego, por lema (3.4)

$$\begin{aligned} \frac{\left(x^k\right)^T\left(s^k\right)}{\mu^k} &< n + \sqrt{n} \\ &\leq n + n = 2n \end{aligned}$$

Entonces

$$\begin{aligned}\left(x^k\right)^T s^k &\leq 2n\mu^k \\ &\leq 2n\left(1-\theta\right)^k \mu_0\end{aligned}$$

Aplicando logaritmo en ambos miembros

$$\begin{aligned}\log\left(\left(x^k\right)^T s^k\right) &\leq \log\left(2n\left(1-\theta\right)^k \mu_0\right) \\ &= \log\left(2n\mu_0\right) + \log\left(1-\theta\right)^k \\ -\log\left(1-\theta\right)^k &\leq \log\left(2n\mu_0\right) - \log\left(\left(x^k\right)^T s^k\right) \\ -k\log\left(1-\theta\right) &\leq \log\left(\frac{2n\mu_0}{\left(x^k\right)^T s^k}\right)\end{aligned}$$

Utilizando $\theta \leq -\log\left(1-\theta\right)$, $\quad 0 < \theta < 1$ *Entonces*

$$k\theta \leq \log\left(\frac{2n\mu_0}{\left(x^k\right)^T s^k}\right) \tag{18}$$

(Hip Aux) supongamos que $\left(x^k\right)^T s^k > \varepsilon$

Aplicando logaritmo a ambos miembros

$$\begin{aligned}\log\left(\left(x^k\right)^T s^k\right) &> \log\varepsilon \\ -\log\left(\left(x^k\right)^T s^k\right) &< -\log\varepsilon \\ \log\left(2n\mu_0\right) - \log\left(\left(x^k\right)^T s^k\right) &< \log\left(2n\mu_0\right) - \log\varepsilon \\ \log\left(\frac{2n\mu_0}{\left(x^k\right)^T s^k}\right) &< \log\left(\frac{2n\mu_0}{\varepsilon}\right) \qquad (19)\end{aligned}$$

De (18) y (19)

$$\begin{aligned}k\theta &< \log\left(\frac{2n\mu_0}{\varepsilon}\right) \\ k &< \frac{1}{\theta}\log\left(\frac{2n\mu_0}{\varepsilon}\right) \qquad (\rightarrow\leftarrow)\end{aligned}$$

por tanto,

$$\left(x^k\right)^T s^k \leq \varepsilon \tag{20}$$

Lema 3.6 *El número de iteraciones interiores en una iteración exterior está limitada por*

$$17\left(\frac{n\theta^2+\theta\sqrt{n}}{1-\theta}+\frac{1}{2}\right)$$

Teorema 3.7 *Si ε tiene el valor predeterminado, entonces el número total de iteraciones interiores en el algoritmo de largo paso está limitada por*

$$17\left(\frac{\theta n+\sqrt{n}}{1-\theta}+\frac{1}{2\theta}\right)\log\left(\frac{2n\mu^0}{\varepsilon}\right)$$

La salida del algoritmo es un par factible primal-dual (x,s) tal que $x^Ts<\varepsilon$.

Demostracion 3 *Por el lema (3.5), el algoritmo de paso largo tiene a lo más*

$$\frac{1}{\theta}\log\left(\frac{2n\mu^0}{\varepsilon}\right) \quad \textit{iteraciones exteriores}$$

Por otra parte, del lema (3.6), se tiene

$$17\left(\frac{\theta\sqrt{n}+\theta^2 n}{1-\theta}+\frac{1}{2}\right) \quad \textit{iteraciones interiores en una iteración exterior}$$

Entonces, el total de iteraciones es

$$17\left(\frac{\theta\sqrt{n}+\theta^2 n}{1-\theta}+\frac{1}{2}\right)\left(\frac{1}{\theta}\log\left(\frac{2n\mu^0}{\varepsilon}\right)\right)$$

donde (x,s) es par factible primal-dual tal que $x^Ts<\varepsilon$.

4. Conclusión

En este trabajo llegamos a las siguientes conclusiones:

- El algoritmo de Paso Largo genera iteraciones $\left(x^k,y^k,s^k\right)$ estrictamente factible, y que en un primer momento no se garantiza que esté lo suficientemente cerca del par central.
- En cada iteración, el algoritmo usa dos direcciones $\overline{d_x}$ y $\overline{d_s}$ que resultan ser la descomposición ortogonal del espacio de $\overline{A}$.
- Se utiliza una medida de proximidad del tipo $\delta(x,s,\mu) := \left\|u^{-1}-u\right\|$ en lugar de una norma clásica $\left\|e-\frac{1}{\mu}Xs\right\|$.
- Al Algoritmo de Paso Largo que implemente el método tiene un comportamiento Polinomial.

Referencias bibliográficas

[1] Frisch, R. *The Logarithmic Potencial Method of Convex Programming.* Technical Report, University Institute of Economics. Oslo, Norway, 1995.

[2] Gonzaga C. C. *Polinomial affine algorithms for linear programming, Mathematical Programming*, 49 (1990), pp. 7-21.

[3] Karmarkar N. *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373-395.

[4] Khachiyan L. G. *A polynomial algorithm in linear programming*, Soviet Mathematics Doklady, 20 (1979), pp. 191-194.

[5] Kojima M., Yoshise A. *A primal-dual interior point algorithm for linear programming, Progress in Mathematical Programming: Interior Point and Related Methods.* Springer Verlag, New York, (989), 29-47.

[6] Monteiro R. D. and Adler I. *Interior path-following primal-dual algorithms. Part I: Linear programming.* Mathematical Programming, 44 (1989), pp. 27-41.

[7] Tapia R. A., Zhang Y., and Ye. Y. *On the convergence of the iteration sequence in primal-dual interior-point methods.* Mathematical Programming, 68 (1995), pp. 141-154.

[8] Wright, S. J., *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.