



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

Programación lineal: un algoritmo primal-dual de paso largo usando el método de la función barrera

TESIS

Para optar el Grado Académico de Magíster en Matemática
Aplicada con mención en Matemática Computacional

AUTOR

Juan Honorato LUNA VALDEZ

ASESOR

Edinson Raúl MONTORO ALEGRE

Lima, Perú

2020



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Luna, J. & Montoro, E. (2020). *Programación lineal: un algoritmo primal-dual de paso largo usando el método de la función barrera*. Tesis para optar el grado de Magíster en Matemática Aplicada con mención en Matemática Computacional. Unidad de Posgrado, Facultad de Ciencias Matemáticas, Universidad Nacional Mayor de San Marcos, Lima, Perú.

Hoja de metadatos complementarios

Código ORCID del autor	0000-0001-7885-3255
DNI o pasaporte del autor	09792418
Código ORCID del asesor	0000-0003-1327-2315
DNI o pasaporte del asesor	09627181
Grupo de investigación	Villarreal
Agencia financiadora	
Ubicación geográfica donde se desarrolló la investigación	Ciudad universitaria – UNMSM https://goo.gl/maps/eMu8XEKt6pFGzYi68
Año o rango de años en que se realizó la investigación	2019 - 2020
Disciplinas OCDE	Matemáticas aplicadas https://purl.org/pe-repo/ocde/ford#1.01.02



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS MATEMÁTICAS VICEDECANATO DE INVESTIGACIÓN Y POSGRADO UNIDAD DE POSGRADO

ACTA DE SUSTENTACIÓN DE TESIS EN MODALIDAD VIRTUAL

PARA OPTAR EL GRADO ACADÉMICO DE MAGISTER EN MATEMÁTICA APLICADA CON MENCIÓN EN MATEMÁTICA COMPUTACIONAL

Autorizado por R.D. N° 0325/FCM-D/2020

1. **FECHA DE LA SUSTENTACIÓN:** 29 de mayo del 2020

HORA INICIO: 3:05

HORA TÉRMINO: 3:40

2. MIEMBROS DEL JURADO

PRESIDENTE:	Mg. Carlos Rubén Guerrero Moncada
MIEMBRO:	Dr. José Edmundo Esparta Rodríguez
MIEMBRO:	Mg. Willy David Barahona Martínez
ASESOR:	Mg. Edinson Raúl Montoro Alegre

3. DATOS DEL TESISISTA

NOMBRES Y APELLIDOS: Bachiller Juan Honorato Luna Valdez

CÓDIGO N°: **4147020**

TÍTULO DE LA TESIS: « PROGRAMACIÓN LINEAL: UN ALGORITMO PRIMAL-DUAL DE PASO LARGO USANDO EL MÉTODO DE LA FUNCIÓN BARRERA »

4. RECOMENDACIONES

1. Hacer una referencia de cada uno de los Teoremas utilizados
2. Ilustrar geoméricamente algunos conceptos, definiciones

Datos de la plataforma virtual institucional del acto de sustentación:

<https://meet.google.com/fdd-rfao-txz?authuser=1&hs=122>

ID:

Grabación archivada en: Sustentación de Tesis: Juan Luna Valdez (2020-05-29 at 12:58 GMT-7)

NOTA OBTENIDA: en números y en letras (15) **QUINCE**



5. PÚBLICO ASISTENTE: (Nombre, apellido y DNI)

N°	Nombres y Apellidos	N° DNI
1	Juan Julio Toledo Rodríguez	06702840
2	Andrés Guardia Cayo	09406969
3	Enrique Maximiliano	09313163
4	Jesús Rule Flores	09349248
5	Oscar Benito Pacheco	06242587
6	Rosa Luz Medina Aguilar	10048247
7	Víctor Tarazona Miranda	09264893
8	Alfonso Pérez Salvatierra	06445739
9	Victoriano Yauri Luque	07371885

6. FIRMAS DE LOS MIEMBROS DEL JURADO

Dr. José Edmundo Esparta Rodríguez
MIEMBRO

Mg. Carlos Rubén Guerrero Moncada
PRESIDENTE

--	--

Mg. Willy David Barahona Martínez
MIEMBRO

Edinson Raúl Montoro Alegre
MIEMBRO ASESOR

RESUMEN

Programación lineal: un algoritmo primal-dual de paso largo usando el método de la función barrera

Juan Honorato LUNA VALDEZ

Marzo, 2020

Orientador : Mg. Edinson Montoro Alegre
Grado Obtenido : Magister en Matemática Aplicada

En el presente Trabajo de Tesis se tratará de desarrollar y describir el Método Punto Interior Primal-Dual para resolver el problema de Programación Lineal. Dicho método se caracteriza por utilizar funciones barrera, para el Problema Primal y para el Dual y así deducir el sistema no lineal Primal-Dual, cuya solución define la trayectoria central del Método de Punto Interior.

Otra característica es que se utiliza una matriz de escalamiento para deducir dos direcciones de descenso, una para el espacio Primal y otra para el espacio Dual, y que forman la descomposición ortogonal de la versión escalada de la matriz asociada a las restricciones lineales del problema Primal.

Se presenta un algoritmo denominado de "Paso Largo", que implementa el método y se demuestra que el número total de iteraciones que ejecuta es de orden polinomial.

PALABRAS CLAVES: Programación convexa, trayectoria central, método de punto interior, algoritmo Primal - dual.

ABSTRACT

Linear programming: a long-pass primal-dual algorithm using the barrier function method

Juan Honorato LUNA VALDEZ

Marzo, 2020

Advisor : Mg. Edinson Montoro Alegre
Degree qualification : Master in Applied Mathematics

In this Thesis Work, we will try to developed and described the Primal-Dual Interior Point Method to solve the Linear Programming problem. This method is characterized by using barrier functions, for the Primal and Dual Problems and thus deducing the non-linear Primal-Dual system, whose solution defines the central path of the Inner Point Method.

Another feature is that a scaling matrix is used to deduce two directions of descent, one for the Primal space and another for the Dual space, and that they form the orthogonal decomposition of the scaled version of the matrix associated with the linear constraints of the Primal problem.

An algorithm called "Long Step" is presented, which implements the method and proof that the total number of iterations that it executes is of a polynomial order.

KEY WORDS : Convex programming, central path, interior point method, Primal - dual algorithm.

Índice general

Introducción	1
1. Preliminares	2
1.1. Conjuntos Convexo	2
1.2. Poliedros: caracterización y propiedades	3
1.3. Direcciones y direcciones extremas	3
1.4. Proyección sobre conjuntos convexos	5
1.5. Funciones convexas	6
1.6. Funciones convexas diferenciables	8
1.7. Caracterización y condiciones de optimalidad	9
1.8. Condiciones de Karush-Kuhn-Tucker	11
1.9. Función lagrangeana	14
2. Dualidad y condiciones de Optimalidad	16
2.1. Dualidad	16
2.2. Direcciones de Búsqueda	20
2.3. Escalamiento de Direcciones	23
3. Algoritmo Primal-Dual de Paso Largo	27
3.1. Introducción	27
3.2. Variación de la función Barrera	30
3.2.1. Algoritmo de Paso Largo	33
3.3. Lemas importantes	33
4. Análisis de convergencia	51
5. Conclusiones	58
Bibliografía	59

Introducción

Khachiyan en 1979 planteo un algoritmo denominado Método del Elipsoide y probó que su comportamiento era de tiempo polinomial y que el algoritmo se movía a través del interior de la región factible.

A partir de 1984 con la publicación de los artículos de Khachiyan [10] y Kamarkar [9] se marcó un hito en la historia de los métodos para resolver Programa Matemático Lineal. El muy conocido método Simplex (1945) que resuelve PPL tiene la desventaja que a nivel teórico tiene un comportamiento exponencial; pero su comportamiento práctico reflejaba que estos programas se podían resolver en tiempo polinomial pero no existía prueba alguna que lo respaldara.

Posteriormente Kamarkar [9] presentó otro algoritmo denominado Algoritmo Proyectivo y que tenía un comportamiento polinomial pero mucho mejor que el anterior.

A partir de estos dos trabajos muchos investigadores del área han ido generando nuevos algoritmos considerando variantes de las ideas utilizadas por Khachiyan [10] y Kamarkar abriendo una nueva corriente de métodos denominados Método de Punto Interior, debido a que todos estos métodos se mueven por el interior del Región Factible.

En el presente trabajo de Tesis se presenta una variante de los métodos de punto interior y desarrolla el Método de Punto Interior que se caracteriza por usar dos direcciones de descenso uno para el espacio Primal y otro en el espacio Dual y que forman la descomposición ortogonal de la versión escalada de la matriz, asociada a las restricciones del programa lineal Primal.

En el primer capítulo, se presenta los conceptos básicos de programación convexa, programación lineal.

En el capítulo 2 presentamos en forma breve la teoría de dualidad y las condiciones de optimalidad. Así mismo se presenta y desarrolla el método y sus direcciones de descenso.

En el capítulo 3 se presenta el algoritmo denominado Algoritmo de Paso Largo, así mismo se presentan los lemas y teoremas que fundamentan dicho algoritmo.

Finalmente, en el capítulo 4 se hace un análisis de convergencia y se demuestra el teorema que prueba la polinomialidad del algoritmo.

Capítulo 1

Preliminares

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

1.1. Conjuntos Convexo

Definición 1.1 Sea $S \subseteq R^n$, $S \neq \phi$. Se dice que S es convexo si

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

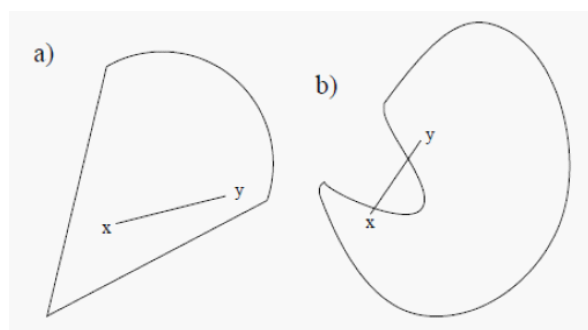


Fig 1. Conjunto convexo (a), conjunto no convexo (b)

Definición 1.2 Sean $x_1, x_2, \dots, x_k \in R^n$, $\lambda_1, \lambda_2, \dots, \lambda_k \in R^+$ tales que $\sum_{i=1}^k \lambda_i = 1$. El vector $x = \sum_{i=1}^k \lambda_i x_i$ se dice **combinación convexa** de los k vectores x_1, x_2, \dots, x_k .

1.2. Poliedros: caracterización y propiedades

Definición 1.3 Se llama conjunto poliédrico a un conjunto de la forma $S = \{x \in \mathbb{R}^n / Ax \leq b\}$ con $A \in M_{m \times n}(\mathbb{R})$ y $b \in \mathbb{R}^m$.

Ejemplo 1.1 Sea

$$S = \{(x_1, x_2); -x_1 + x_2 \leq 2, x_2 \leq 4, x_1 \geq 0, x_2 \geq 0\}$$

El conjunto S es un poliedro convexo y cerrado pero no acotado, tal como se muestra en la figura 2.

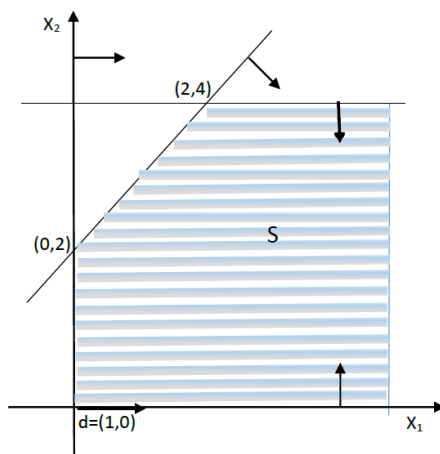


Fig 2. Conjunto poliédrico S convexo cerrado pero no acotado

Proposition 1 El conjunto $S' = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\}$ es un poliedro.

Proposition 2 Todo poliedro es un conjunto convexo.

1.3. Direcciones y direcciones extremas

Definición 1.4 Sea $S \subseteq \mathbb{R}^n$, un conjunto convexo. Un vector $d \in \mathbb{R}^n$, $d \neq 0$, se **dice dirección de S** , si $\forall x \in S$ se tiene que $x + \lambda d \in S$, $\forall \lambda \geq 0$.

Obsevación 1.1 Considerese el poliedro $S = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\}$, una direc-

ción de S debe satisfacer,

$$\begin{aligned} A(x + \lambda d) &= b \quad \forall \lambda \geq 0, \quad \forall x \in S \\ x + \lambda d &\geq 0 \quad \forall \lambda \geq 0, \quad \forall x \in S \end{aligned}$$

luego, d es la dirección de S si y solamente si satisface el sistema $Ad = 0$, $d \neq 0$.

Definición 1.5 Sea S un conjunto poliédrico y $d \in R^n$ una dirección de S . Diremos que d es **dirección extrema**, si dadas d_1 y d_2 , direcciones de S , tales que $d = \alpha d_1 + \beta d_2$ para algún $\alpha, \beta > 0$, entonces se tiene que $d = d_1 = d_2$.

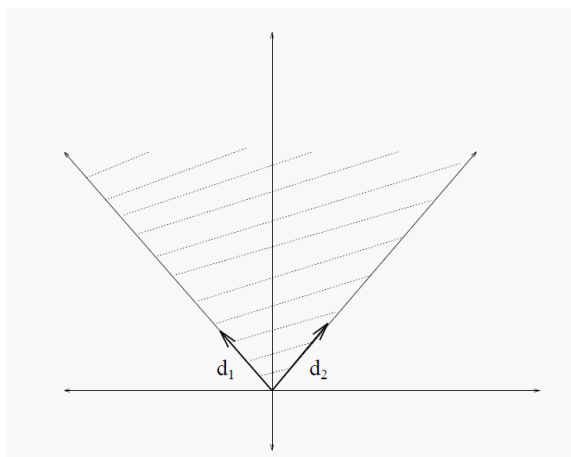


Fig. 3. El origen único extremo del poliedro, d_1 y d_2 son sus únicas direcciones extremas.

Teorema 1.1 (*Representación de conjuntos poliédricos*)

Sea $S = \{x \in R^n \mid Ax = b, x \geq 0\}$, donde $A \in M_{m \times n}(R)$ es de rango m y $b \in R^m$.

Sean x_1, x_2, \dots, x_k los puntos extremos y d_1, d_2, \dots, d_l las direcciones extremas de S .

Entonces, $x \in S$ si y sólo si, se puede ser escribir de la siguiente manera,

$$x = \sum_{i=1}^k \lambda_i x_i + \sum_{j=1}^l \mu_j d_j$$

donde $\lambda_i \geq 0 \quad \forall i = 1, \dots, k, \quad \sum_{i=1}^k \lambda_i = 1, \quad \mu_j \geq 0, \quad \forall j = 1, \dots, l.$

1.4. Separación entre un punto y un conjunto

Teorema 1.2 Sea $S \subseteq R^n$ un conjunto convexo, cerrado, no vacío, $y \in R^n$, $y \notin S$.

Entonces, existe un único $\bar{x} \in S$ que minimiza la función φ_y , donde

$$\begin{aligned} \varphi_y : S &\longrightarrow R \\ x &\longmapsto \varphi_y(x) = \|y - x\| \end{aligned}$$

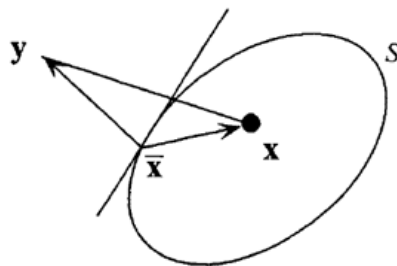


Fig 4. Distancia que minimiza

Definición 1.6 Sea $S \subseteq R^n$ un conjunto convexo cerrado no vacío.

- i) Se define la distancia de y a S , por $d(y, S) = \min \{\varphi_y(x) / x \in S\} = \|y - \bar{x}\|$, para $y \in R^n$.
- ii) Se define la proyección de y sobre S , para $y \in R^n$, mediante

$$P_S(y) = \arg \min \{\varphi_y(x) / x \in S\}$$

siendo \bar{x} el único que satisface $\varphi_y(\bar{x}) \leq \varphi_y(x)$, $\forall x \in S$.

Geoméricamente está representado por la Fig 5.

Obsevación 1.2 La notación $\arg \min$ se lee como: **El argumento que minimiza**, es decir, si $y \in S$ se tiene que $d(y, S) = 0$ y $P_S(y) = y$.

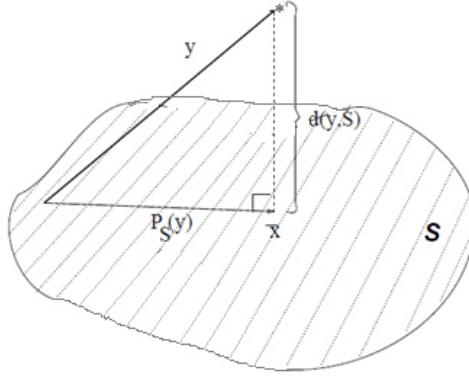


Fig 5. Distancia y proyección del punto y al conjunto S .

Teorema 1.3 Sea S un conjunto convexo cerrado no vacío e $y \notin S$. Se tiene que

$$\langle y - \bar{x}, x - \bar{x} \rangle \leq 0, \quad \forall x \in S$$

si y solamente si \bar{x} minimiza $\varphi_y(x)$.

Teorema 1.4 Sea S un conjunto convexo cerrado no vacío, entonces

$$\|P_S(x) - P_S(y)\| \leq \|x - y\|, \quad \forall x, y \in R^n$$

El teorema (1.4) equivale a decir que si S un conjunto convexo cerrado no vacío, la función de proyección P_S es de Lipschitz.

1.5. Funciones Convexas

Definición 1.7 Sea $S \subseteq R^n$ un conjunto convexo. Se dice que $f : S \rightarrow R$ es una *función convexa* si

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

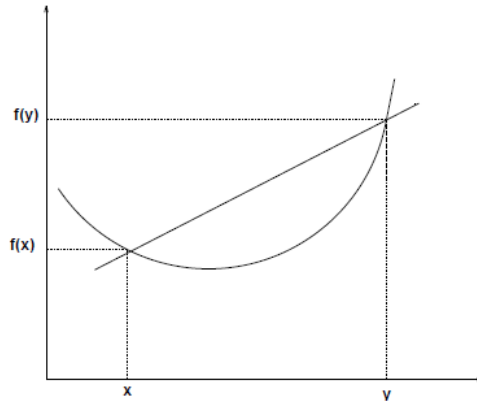


Fig 6. Función convexa

Una definición equivalente de convexidad de funciones, es la siguiente.

Definición 1.8 Sea $S \subseteq \mathbb{R}^n$ un conjunto convexo. Se dice que $f : S \rightarrow \mathbb{R}$ es una **función convexa** si $\forall k \in \mathbb{N}$, $x_1, \dots, x_k \in S$ y $\lambda_1, \dots, \lambda_k \geq 0$ tales que $\sum_{i=1}^k \lambda_i = 1$, se tiene

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

Definición 1.9 Una función f , definida sobre un convexo S , se dice **estrictamente convexa** si para todo $x, y \in S$, $x \neq y$, $0 < \lambda < 1$, se tiene

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

Definición 1.10 Se dice que la función $f : S \rightarrow \mathbb{R}$ es **cóncava** si $-f$ es convexa, o equivalentemente, si:

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

De la misma forma, la función f es **estrictamente cóncava** si $-f$ es estrictamente convexa.

1.6. Funciones convexas diferenciables

Definición 1.11 Sea $S \subseteq R^n$, no vacío, $f : S \rightarrow R$, $\bar{x} \in S$, $d \neq 0$ tal que $\bar{x} + \lambda d \in S$, $\forall \lambda \in [0, \eta]$, algún $\eta > 0$. Se define la **derivada direccional** de f en el punto \bar{x} , en la dirección d , por,

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

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

Definición 1.12 Sea $S \subseteq R^n$, no vacío, una función $f : S \rightarrow R$, es **diferenciable** en $\bar{x} \in \text{int}(S)$ si existe $\nabla f(\bar{x}) \in R^n$ tal que

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

donde $\lim_{x \rightarrow \bar{x}} \frac{o(x - \bar{x})}{\|x - \bar{x}\|} = 0$.

Teorema 1.5 Si la función f es diferenciable en el interior de S , entonces

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

Teorema 1.6 (Caracterización de convexidad en el caso diferenciable)

Sea $S \subseteq R^n$ y $f : S \rightarrow R$ una función diferenciable en S convexo. Entonces f es convexa si y sólo si

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}), \quad \forall x, \bar{x} \in S.$$

Definición 1.13 Sea $S \subseteq R^n$, no vacío, una función $f : S \rightarrow R$ es dos veces diferenciable en \bar{x} , si existe $\nabla f(\bar{x}) \in R^n$ y $H(\bar{x}) \in M_{n \times n}(R)$ tal que

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^T H(\bar{x}) (x - \bar{x}) + o(x - \bar{x}), \quad \forall x \in S$$

donde $\lim_{x \rightarrow \bar{x}} \frac{o(x - \bar{x})}{\|x - \bar{x}\|^2} = 0$.

La matriz $H(\bar{x})$ se llama **matriz Hessiana** de f en \bar{x} y está dado por:

$$H(\bar{x}) = \begin{bmatrix} \frac{\partial^2 f(\bar{x})}{\partial x_1^2} & \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j} & \vdots \\ \vdots & & & \vdots \\ \frac{\partial^2 f(\bar{x})}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_n^2} \end{bmatrix}$$

Teorema 1.7 (*Caracterización de convexidad para una función dos veces diferenciable*)

Sea $S \subseteq R^n$ un abierto, convexo, y sea una función $f : S \rightarrow R$ dos veces diferenciable en S . Entonces f es convexa si y sólo si $H(x)$ es semi-definida positiva $\forall x \in S$.

1.7. Caracterización y condiciones de optimalidad

Consideremos el programa:

$$(P) \quad \begin{cases} \text{mín } f(x) \\ x \in S \end{cases}$$

Donde $S \subseteq R^n$, en general es definido por ecuaciones no necesariamente lineales, la cual denominaremos región factible y f una función cualquiera.

Si S es un poliedro y f lineal, se tiene problemas de Programación Lineal.

Definición 1.14 Un punto $x^* \in S$ es **mínimo local** de f si existe $\epsilon > 0$ que cumple $f(x^*) \leq f(x)$, $\forall x \in S$ tal que $\|x^* - x\| < \epsilon$.

Es decir, existe una vecindad de x^* donde éste punto es mínimo.

Definición 1.15 Sea $S \subseteq R^n$, un punto $x^* \in S$ se dice es **mínimo global** de f en S si $f(x^*) \leq f(x)$, $\forall x \in S$. Es decir, x^* es solución de (P).

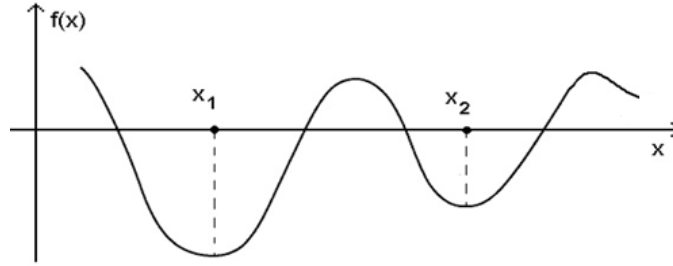


Fig 6. Mínimo global x_1 , mínimo local x_2

Theorem 3 (*Teorema de Weierstrass*)

Sea $S \subseteq \mathbb{R}^n$ un conjunto compacto y no vacío, y sea $f : S \rightarrow \mathbb{R}$ una función continua en S . Entonces el problema $\min \{f(x); x \in S\}$ alcanza su mínimo. Es decir, existe una solución minimizadora para el problema dado.

Un elemento $\bar{x} \in S$ se llama solución factible de (P) y si \bar{x} resuelve (P) se puede decir que es **solución**, **solución óptima**, **solución mínima**, o **solución global** del problema.

Teorema 1.8 Sea la función $f : S \rightarrow \mathbb{R}$, con S conjunto convexo no vacío, y sea \bar{x} una solución local del problema (P). Entonces

- (i) Si f es convexa, \bar{x} es mínimo global
- (ii) Si f es estrictamente convexa, \bar{x} es el único mínimo global.

1.7.1. Optimización sin restricciones

Aquí, obtendremos las condiciones necesarias y suficientes para resolver problemas no lineales irrestrictos (cuando $S = \mathbb{R}^n$), es decir, programas del tipo:

$$(P) \quad \begin{cases} \min f(x) \\ x \in S' \end{cases}$$

En general, supondremos que la función f es una o dos veces, continuamente diferenciable.

Teorema 1.9 (*condiciones suficientes de optimalidad*)

Sea $f \in C^2(S)$, con S abierto. Supongamos que $x^* \in S$ satisface:

- a) $\nabla f(x^*) = 0$
- b) La matriz Hessiana $H(x^*)$ es definida positiva.

Entonces, x^* es un mínimo local estricto de f en S .

1.7.2. Optimización con restricciones

Definición 1.16 Sea $f : R^n \rightarrow R$, se dice que $d \in R^n$ es **dirección de descenso** de f en \bar{x} si $\nabla f(\bar{x})^T d < 0$.

Definición 1.17 Sea $S \subseteq R^n$, $S \neq \emptyset$ y $\bar{x} \in S$. Se denomina **cono de direcciones admisibles** de S en \bar{x} al conjunto:

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

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

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

Teorema 1.10 Sea $S \subseteq R^n$ y $f : S \rightarrow R$, diferenciable, y consideremos el programa

$$(P_1) \quad \left\{ \begin{array}{l} \text{mín } f(x) \\ x \in S \end{array} \right.$$

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

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) &= 0 \\ \mu_i g_i(\bar{x}) &= 0, \quad i = 1, 2, \dots, m \\ \mu_i &\geq 0, \quad i = 1, 2, \dots, m \end{aligned}$$

Interpretación geométrica del teorema de Karush-Kuhn-Tucker

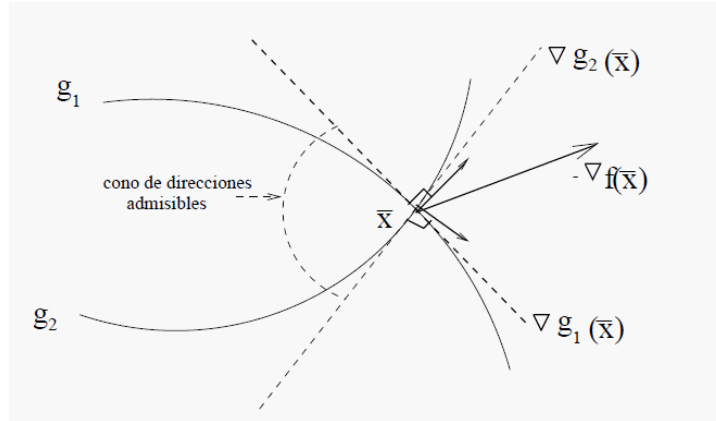


Fig 7. Condiciones del teorema de KKT para dimensión dos.

Teorema 1.11 Sean las funciones diferenciables: $f : R^n \longrightarrow R$, $g_i : R^n \longrightarrow R$ $i = 1, 2, \dots, m$ y $h_j : R^n \longrightarrow R$ $j = 1, 2, \dots, l$

Consideremos el programa

$$(P_2) \quad \begin{cases} \text{mín } f(x) \\ g_i(x) \leq 0 & i = 1, 2, \dots, m \\ h_j(x) = 0 & j = 1, 2, \dots, l \end{cases}$$

Sea \bar{x} un punto factible, supongamos que el conjunto $\{\nabla g_i(\bar{x}), \nabla h_j(\bar{x}), i \in I, j = 1, 2, \dots, l\}$ es linealmente independiente.

Entonces, si \bar{x} es solución de (P_2) , existen $\mu_i \in R$, $i = 1, 2, \dots, m$, $v_j \in R$, $j = 1, 2, \dots, l$ tales que

$$\begin{aligned} \nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) + \sum_{j=1}^l v_j \nabla h_j(\bar{x}) &= 0 \\ \mu_i g_i(\bar{x}) &= 0, \quad i = 1, 2, \dots, m \\ \mu_i &\geq 0, \quad i = 1, 2, \dots, m \end{aligned}$$

Ejemplo 1.2 *Aplicación del teorema de KKT.*

Consideremos el problema

$$(P) \quad \begin{cases} \text{mín } c^T x \\ Ax = b \\ x \geq 0 \end{cases}$$

donde la matriz $A \in M_{m \times n}(R)$ *y los vectores* $c \in R^n$, $b \in R^m$.

El programa (P) se puede escribir en su forma equivalente dado por,

$$(P') \quad \begin{cases} \text{mín } c^T x \\ -x \leq 0 \\ Ax - b = 0 \end{cases}$$

de tal manera que las funciones f , g_i , h_j *del problema general están dadas por*

$$\begin{aligned} f(x) &= c^T x \\ g_i(x) &= -x_i \\ h_j(x) &= \sum_{k=1}^n a_{jk} x_k - b_j \end{aligned}$$

Sus gradientes respectivos son:

$$\begin{aligned} \nabla f(x) &= c \\ \nabla g_i(x) &= -e_i \\ \nabla h_j(x) &= (a_{j1}, \dots, a_{jn})^T \end{aligned}$$

donde e_i *es el* i -ésimo *vector de la base canónica de* R^n .

Las ecuaciones del teorema de KKT se pueden escribir como:

$$c + \sum_{i=1}^n u_i (-e_i) + \sum_{j=1}^m v_j (a_{j1}, \dots, a_{jn})^T = 0$$

$$u_i (-x_i) = 0 \quad i = 1, \dots, n$$

$$u_i \geq 0 \quad i = 1, \dots, n$$

Haciendo el cambio de variable $y = -v$ se obtiene

$$\begin{cases} A^T y + u = c \\ u^T x = 0 \\ u \geq 0 \end{cases}$$

La caracterización de la solución del problema de Programación Lineal esta dada de la siguiente manera: x es solución del problema (P) si y sólo si existen vectores $u \in R^n$, $y \in R^m$, tales que

$$\begin{cases} Ax = b \\ A^T y + u = c \\ u^T x = 0 \\ x, u \geq 0 \end{cases}$$

Resolver el sistema anterior es complicado, puesto que algunas variables deben ser positivos y además deben satisfacer una ecuación no lineal.

1.9. Función lagrangeana

En los problemas de Programación Matemática de la forma

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

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

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

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 $L(x, \lambda)$.

Capítulo 2

Dualidad y condiciones de Optimalidad

2.1. Dualidad

El dual de un programa matemática lineal es otro programa matemática lineal, el cual se obtiene a partir del primero, y cuya solución proporciona información valiosa del problema primal.

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

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

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

El programa dual (P), será:

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

Expresado en su forma estándar,

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

donde $s \in R^n$.

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

Definimos:

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

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

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

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

Para cualquier par de soluciones primal-dual factible tenemos:

$$\begin{aligned} x^T s &= x^T (c - A^T y) = x^T c - x^T A^T y \\ &= c^T x - (y^T A x)^T \\ &= c^T x - (y^T b)^T \\ &= c^T x - b^T y \end{aligned}$$

Entonces,

$$x^T s = c^T x - b^T y \tag{2.1}$$

Generando así, los problemas barreras para el primal y dual

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

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

(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) :

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

Aplicando las condiciones de KKT:

- i) Existencia de $y \geq 0$
- ii) $y(Ax - b) = 0$
- iii) $\nabla_x L_P(x, s; \mu) = 0$

De (i) y (ii) por ser x viable en (P_μ) , se tiene:

$$Ax - b = 0, \quad x > 0$$

De (iii),

$$c - s + \mu X^{-1} e + A^T y = 0$$

Siendo X matriz diagonal.

La función lagrangeana de (D_μ) :

$$L_D(x, s; \mu) = b^T y + x^T s - \mu \sum_{j=1}^n \log(x_j s_j) - x^T (A^T y + s - c) \quad (a)$$

$$\stackrel{(2,1)}{=} 2c^T x - x^T A^T y - x^T s - \mu \sum_{j=1}^n \log(x_j s_j) \quad (b)$$

Aplicando las condiciones de KKT:

- i) Existencia de $x \geq 0$
- ii) $x (A^T y + s - c) = 0$
- iii) $\nabla_{(y,s)} L_D(x, s; \mu) = 0$

De (i) y (ii) por ser y viable en (D_μ) , se tiene:

$$A^T y + s - c = 0, \quad s > 0$$

De (iii),

$$\nabla_y L_D(x, s; \mu) \stackrel{de (a)}{=} b^T - x^T A^T = 0 \quad \rightarrow \quad Ax = b, \quad x > 0$$

$$\nabla_s L_D(x, s; \mu) \stackrel{de (b)}{=} x - \mu S^{-1} e = 0 \quad \rightarrow \quad x = \mu S^{-1} e$$

$$\rightarrow Sx = \mu e$$

$$\rightarrow Xs = \mu e$$

Obtenemos las condiciones de optimalidad de KKT:

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

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

Donde:

$$X = \text{diag}(x_1, x_2, \dots, x_n)$$

$$S = \text{diag}(s_1, s_2, \dots, s_n)$$

$$e = (1, 1, \dots, 1)^T$$

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 $S^0 \times T^0$.

2.2. 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:

$$J(F(x, y, s)) dw = -F(x, y, s) \tag{2.2}$$

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.

Notemos que:

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

De (2.2),

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

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

$$\begin{aligned} Ax - b &= 0 \\ A^T y + s - c &= 0 \end{aligned}$$

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:

$$(*) \begin{cases} Ad_x = 0 & (1') \\ A^T d_y + d_s = 0 & (2') \\ Sd_x + Xd_s = \mu e - Xs & (3') \end{cases}$$

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

De $(2')$,

$$d_s = -A^T d_y \quad (2.3)$$

En $(3')$,

$$Sd_x - XA^T d_y = \mu e - Xs$$

$$Sd_x = \mu e - Xs + XA^T d_y$$

Multiplicamos por S^{-1} ambos miembros,

$$d_x = S^{-1}(\mu e - Xs) + S^{-1}XA^T d_y \quad (2.4)$$

Multiplicamos por A ambos miembros,

$$\begin{aligned} Ad_x &= AS^{-1}(\mu e - Xs) + AS^{-1}XA^T d_y \\ 0 &\stackrel{de (1')}{=} AS^{-1}(\mu e - Xs) + AS^{-1}XA^T d_y \end{aligned}$$

Entonces,

$$\begin{aligned} AS^{-1}XA^T d_y &= AS^{-1}(Xs - \mu e) \\ d_y &= (AS^{-1}XA^T)^{-1}(AS^{-1}Xs - AS^{-1}\mu e) \\ &= -(AS^{-1}XA^T)^{-1}(\mu AS^{-1}e - AS^{-1}Xs) \\ &= -(AS^{-1}XA^T)^{-1}(\mu AS^{-1}e - AS^{-1}Sx) \\ &= -(AS^{-1}XA^T)^{-1}(\mu AS^{-1}e - Ax) \\ &\stackrel{de (*)}{=} -(AS^{-1}XA^T)^{-1}(\mu AS^{-1}e - b) \end{aligned}$$

Luego,

$$d_y = -(AXS^{-1}A^T)^{-1}(\mu AS^{-1}e - b) \quad (2.5)$$

En (2.3),

$$d_s = A^T(AXS^{-1}A^T)^{-1}(\mu AS^{-1}e - b) \quad (2.6)$$

(2.5) en (2.4),

$$\begin{aligned} d_x &= S^{-1}(\mu e - Xs) - S^{-1}XA^T(AXS^{-1}A^T)^{-1}(\mu AS^{-1}e - b) \\ &= \mu S^{-1}e - S^{-1}Xs - \underbrace{S^{-1}XA^T(AXS^{-1}A^T)^{-1}(\mu AS^{-1}e - b)}_{d_s} \\ &= \mu S^{-1}e - S^{-1}Sx - S^{-1}Xd_s \end{aligned}$$

Luego,

$$d_x = \mu S^{-1}e - x - S^{-1}Xd_s \quad (2.7)$$

De esta manera, de (2.5), (2.6) y (2.7) obtenemos las direcciones de búsqueda d_x , d_y , d_s .

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

2.3. Escalamiento de las Direcciones

Definamos las matrices diagonales a partir de las matrices diagonales X y S :

$$D := (XS^{-1})^{\frac{1}{2}} \quad , \quad V := (XS)^{\frac{1}{2}} \quad (2.8)$$

Escalando las direcciones de búsqueda de la siguiente manera:

$$\bar{d}_x := D^{-1}d_x \quad , \quad \bar{d}_s := Dd_s \quad (2.9)$$

Entonces,

$$\begin{aligned} Xd_s + Sd_x &= XD^{-1}\bar{d}_s + SD\bar{d}_x \\ &\stackrel{de (2.8)}{=} X(XS^{-1})^{-\frac{1}{2}}\bar{d}_s + S(XS^{-1})^{+\frac{1}{2}}\bar{d}_x \\ &= X^{\frac{1}{2}}S^{\frac{1}{2}}\bar{d}_s + X^{\frac{1}{2}}S^{\frac{1}{2}}\bar{d}_x \\ &\stackrel{de (2.8)}{=} (XS)^{\frac{1}{2}}(\bar{d}_s + \bar{d}_x) \\ &= V(\bar{d}_s + \bar{d}_x) \end{aligned}$$

Luego,

$$Xd_s + Sd_x = V(\bar{d}_s + \bar{d}_x)$$

Ahora, definamos:

$$\bar{d}_y := d_y \quad \bar{A} := (AD) \quad u := \frac{1}{\sqrt{\mu}}v \quad (2.10)$$

En el espacio escalado, vamos a determinar las direcciones de Newton:

- (De (1')) $Ad_x = 0 \xrightarrow{de (2,9)} AD\bar{d}_x = 0$
 $\xrightarrow{de (2,10)} \bar{A} \bar{d}_x = 0$
- (De (2')) $A^T d_y + d_s = 0 \xrightarrow{de (2,10)} D(A^T d_y + d_s) = 0$
 $\longrightarrow DA^T d_y + Dd_s = 0$
 $\xrightarrow{de (2,8),(2,9)} \bar{A}^T d_y + \bar{d}_s = 0$
- (De (3')) $Xd_s + Sd_x = \mu e - Xs$

entonces

$$\begin{aligned} Xd_s + Sd_x &\stackrel{de (2,9)}{=} XD^{-1}\bar{d}_s + SD\bar{d}_x \\ &\stackrel{de (2,8)}{=} X(XS^{-1})^{-\frac{1}{2}}\bar{d}_s + S(XS^{-1})^{-\frac{1}{2}}\bar{d}_x \\ &= S^{\frac{1}{2}}X^{\frac{1}{2}}\bar{d}_s + S^{\frac{1}{2}}X^{\frac{1}{2}}\bar{d}_x \\ &= (SX)^{\frac{1}{2}}(\bar{d}_s + \bar{d}_x) \\ \underbrace{Xd_s + Sd_x}_{\mu e - Xs} &= (SX)^{\frac{1}{2}}(\bar{d}_s + \bar{d}_x) \\ \mu e - Xs &= (SX)^{\frac{1}{2}}(\bar{d}_s + \bar{d}_x) \\ \bar{d}_s + \bar{d}_x &= (SX)^{-\frac{1}{2}}(\mu e - Xs) \end{aligned}$$

Pero

$$u := \frac{1}{\sqrt{\mu}}(XS)^{\frac{1}{2}}e \longrightarrow (XS)^{\frac{1}{2}}e = \sqrt{\mu}u \longrightarrow (XS)^{-\frac{1}{2}}e = \frac{1}{\sqrt{\mu}}u^{-1}$$

entonces,

$$\begin{aligned}
\overline{d}_s + \overline{d}_x &= \mu (SX)^{-\frac{1}{2}} e - (SX)^{-\frac{1}{2}} Xs \\
&= \mu (SX)^{-\frac{1}{2}} e - (SX)^{-\frac{1}{2}} Sx \\
&= \mu (SX)^{-\frac{1}{2}} e - (SX)^{-\frac{1}{2}} (SX) e \\
&= \frac{\mu}{\sqrt{\mu}} u^{-1} - (SX)^{\frac{1}{2}} e \\
&= \sqrt{\mu} u^{-1} - \sqrt{\mu} u \\
&= \sqrt{\mu} (u^{-1} - u)
\end{aligned}$$

De esta manera, obtenemos en el Espacio Escalado

$$\left\{ \begin{array}{l} \overline{A} \overline{d}_x = 0 \\ \overline{A}^T \overline{d}_s + \overline{d}_s = 0 \\ \overline{d}_s + \overline{d}_x = \sqrt{\mu} (u^{-1} - u) \end{array} \right.$$

Las dos primeras ecuaciones muestran que \overline{d}_x y \overline{d}_s pertenecen a los subespacios ortogonales $N(\overline{A})$ y $R(\overline{A}^T)$, respectivamente. De la última ecuación, llegamos a la conclusión que \overline{d}_x y \overline{d}_s forman la descomposición ortogonal de \overline{A} sobre los subespacios. Por lo tanto:

$$\begin{aligned}
\overline{d}_x &= \sqrt{\mu} P_{N(\overline{A})} (u^{-1} - u) \\
\overline{d}_s &= \sqrt{\mu} P_{R(\overline{A}^T)} (u^{-1} - u)
\end{aligned}$$

Aquí, $P_{N(\overline{A})} z$ y $P_{R(\overline{A}^T)} z$ denotan las proyecciones ortogonales de z sobre el espacio nulo de \overline{A} y el espacio rango de \overline{A}^T respectivamente. Por lo tanto, también tenemos:

$$\|\overline{d}_x\|^2 + \|\overline{d}_s\|^2 = \mu \|u^{-1} - u\|^2 \tag{2.11}$$

Los desplazamientos d_x , d_s , (y d_y , también) son cero, si y sólo si, $\|u^{-1} - u\|^2 = 0$, es decir x , y , s coincide con el μ -centros respectivos. Debido a esto se utiliza $\|u^{-1} - u\|$ como la cantidad que mide la cercanía al par de μ -centros.

Así, definimos

$$\delta(x, s; \mu) := \|u^{-1} - u\| \quad (2.12)$$

En el método primal-dual se utiliza una medida de proximidad diferente pero estan estrechamente relacionados, es decir,

$$\begin{aligned} \|U(u^{-1} - u)\| &= \|Uu^{-1} - Uu\| \\ &= \|UU^{-1}e - Uu\| \\ &= \|e - Uu\| \end{aligned}$$

donde,

$$U := \frac{1}{\sqrt{\mu}} (XS)^{\frac{1}{2}}$$

Pero,

$$\begin{aligned} \|e - Uu\| &= \left\| e - \frac{1}{\sqrt{\mu}} (XS)^{\frac{1}{2}} \frac{1}{\sqrt{\mu}} (XS)^{\frac{1}{2}} e \right\| \\ &= \left\| e - \frac{1}{\mu} (XS) e \right\| \\ &= \left\| e - \frac{1}{\mu} X (Se) \right\| \\ &= \left\| e - \frac{1}{\mu} Xs \right\| \end{aligned}$$

Entonces,

$$\|U(u^{-1} - u)\| = \|e - Uu\| = \left\| e - \frac{1}{\mu} Xs \right\| \quad (2.13)$$

Se prefiere la medida de proximidad tal como se define por (2.12).

Capítulo 3

Algoritmo Primal-Dual de Paso

Largo

3.1. Introducción

Dado que cada iteración del algoritmo implica un problema de mínimos cuadrado (que es una operación tediosa) el procedimiento general es ineficiente, sin embargo es obligatorio reducir drásticamente el número total de iteraciones. Esto es la complejidad computacional.

La idea detras del algoritmo de paso largo es simple, pero más profundo para las actualizaciones del parámetro de barrera, lo que hace que sean más eficientes. Por supuesto, después de la actualización, la iteración actual se convierte en el nuevo par central. No hay garantía de que el paso de Newton asociado haga que la iteración esté cerca del par central.

Por otra parte, como un paso Newton completo es casi seguro que conduce a un punto invariable, y es necesario recurrir a algún tipo de función para "Detener, atajar" los pasos. La función de barrera logarítmica primal-dual

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

es un candidato.

Esta afirmación se sustenta en una propiedad fundamental 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 3.1

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

donde

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

Demostración.

Calculemos

$$\begin{aligned} \nabla_x f &= \frac{s}{\mu} - X^{-1}e = \frac{s}{\mu} - x^{-1} = \frac{Se}{\mu} - X^{-1}e \\ &= \frac{1}{\mu} S^{1/2} X^{-1/2} X^{1/2} S^{1/2} e - X^{-1/2} S^{1/2} S^{-1/2} X^{-1/2} e \\ &= \frac{1}{\mu} (XS^{-1})^{-1/2} (XS)^{1/2} e - (XS^{-1})^{-1/2} (XS)^{-1/2} e \\ &= \frac{1}{\mu} D^{-1} v - D^{-1} v^{-1} \\ &= D^{-1} \left(\frac{v}{\mu} - v^{-1} \right) \\ &= D^{-1} \left(\frac{\sqrt{\mu}}{\mu} u - \frac{1}{\sqrt{\mu}} u^{-1} \right) \\ &= D^{-1} \left(\frac{u}{\sqrt{\mu}} - \frac{u^{-1}}{\sqrt{\mu}} \right) \\ &= \frac{1}{\sqrt{\mu}} D^{-1} (u - u^{-1}) \end{aligned}$$

$$\rightarrow \nabla_x f = \frac{1}{\sqrt{\mu}} D^{-1} (u - u^{-1})$$

Obsevación 3.1

$$D = (XS^{-1})^{1/2}, \quad v = (XS)^{1/2} e, \quad u = \frac{1}{\sqrt{\mu}} v$$

$$\begin{aligned}
\nabla_s f &= \frac{x}{\mu} - S^{-1}e = \frac{x}{\mu} - s^{-1} = \frac{Xe}{\mu} - S^{-1}e \\
&= \frac{1}{\mu} X^{1/2} S^{-1/2} \underbrace{S^{1/2} X^{1/2}} e - \underbrace{S^{-1/2} X^{1/2}} \underbrace{X^{-1/2} S^{-1/2}} e \\
&= \frac{1}{\mu} X^{1/2} S^{-1/2} X^{1/2} S^{1/2} e - X^{1/2} S^{-1/2} S^{-1/2} X^{-1/2} e \\
&= \frac{1}{\mu} (XS^{-1})^{1/2} (XS)^{1/2} e - (XS^{-1})^{1/2} (XS)^{-1/2} e \\
&= \frac{1}{\mu} Dv - Dv^{-1} \\
&= D \left(\frac{v}{\mu} - v^{-1} \right) \\
&= D \left(\frac{\sqrt{\mu}}{\mu} u - \frac{1}{\sqrt{\mu}} u^{-1} \right) \\
&= D \left(\frac{u}{\sqrt{\mu}} - \frac{u^{-1}}{\sqrt{\mu}} \right) \\
&= \frac{1}{\sqrt{\mu}} D (u - u^{-1})
\end{aligned}$$

$$\rightarrow \nabla_s f = \frac{1}{\sqrt{\mu}} D (u - u^{-1})$$

Luego,

$$\begin{aligned}
(\nabla_x f)^T dx + (\nabla_s f)^T ds &= \frac{1}{\sqrt{\mu}} D^{-1} (u - u^{-1})^T dx + \frac{1}{\sqrt{\mu}} D (u - u^{-1})^T ds \\
&= \frac{1}{\sqrt{\mu}} (u - u^{-1})^T D^{-1} dx + \frac{1}{\sqrt{\mu}} (u - u^{-1})^T D ds \\
&= \frac{1}{\sqrt{\mu}} (u - u^{-1})^T (D^{-1} dx + D ds) \\
&= \frac{1}{\sqrt{\mu}} (u - u^{-1})^T (\overline{dx} + \overline{ds}) \\
&= \frac{1}{\sqrt{\mu}} (u - u^{-1})^T (\sqrt{\mu} (u^{-1} - u)) \\
&= (u - u^{-1})^T (u^{-1} - u) \\
&= - (u - u^{-1})^T (u - u^{-1}) \\
&= - \|u - u^{-1}\|^2 = -\delta^2
\end{aligned}$$

entonces

$$(\nabla_x f)^T dx + (\nabla_s f)^T ds = -\delta^2$$

■

3.2. Variación de la Función Barrera

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

Lema 3.2 *Sea h un vector en R^n tal que $\|h\| < 1$. Entonces*

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

Demostración.

Afirmación

$$\sum_{i=1}^n |h_i|^k \leq \|h\|^k \quad \forall k \in N$$

En efecto, probaremos por inducción sobre k .

Para $k = 1$

$$\sum_{i=1}^n |h_i| \leq \|h\| = \sqrt{\sum_{i=1}^n h_i^2} \quad (\text{cumple})$$

Para $k = m$

$$\sum_{i=1}^n |h_i|^m \leq \|h\|^m \quad (\text{Hipótesis Inductiva})$$

Para $k = m + 1$

Debemos probar

$$\sum_{i=1}^n |h_i|^{m+1} \leq \|h\|^{m+1}$$

En efecto,

$$\begin{aligned}
\sum_{i=1}^n |h_i|^{m+1} &= \sum_{i=1}^n |h_i|^m |h_i| \leq \sum_{i=1}^n |h_i|^m \|h\|_\infty = \|h\|_\infty \sum_{i=1}^n |h_i|^m \\
&\leq \|h\|_\infty \|h\|^m \quad (\text{por H.I}) \\
&\leq \|h\| \|h\|^m \quad (\text{pues } \|h\|_\infty \leq \|h\|) \\
&\leq \|h\|^{m+1}
\end{aligned}$$

$$\rightarrow \sum_{i=1}^n |h_i|^{m+1} \leq \|h\|^{m+1}$$

Por tanto,

$$\sum_{i=1}^n |h_i|^k \leq \|h\|^k \quad \forall k \in \mathbb{N}$$

Sabemos

$$\log(1 - z) = - \sum_{k=1}^{\infty} \frac{z^k}{k!}, \quad |z| < 1$$

como $|h_i| \leq \|h\| < 1 \quad i = 1, 2, \dots, n$

$$\rightarrow |h_i| < 1$$

Luego

$$\begin{aligned}
\log(1 + h_i) &= h_i - \frac{h_i^2}{2} + \frac{h_i^3}{3!} - \dots \\
- \sum_{i=1}^n \log(1 + h_i) &= - \sum_{i=1}^n \left(h_i - \frac{h_i^2}{2} + \frac{h_i^3}{3!} - \dots \right) \\
&= - \left(\sum_{i=1}^n h_i - \sum_{i=1}^n \frac{h_i^2}{2} + \sum_{i=1}^n \frac{h_i^3}{3!} - \dots \right) \\
&= - \sum_{i=1}^n h_i + \frac{1}{2} \sum_{i=1}^n h_i^2 - \sum_{i=1}^n \left(\frac{h_i^3}{3!} - \frac{h_i^4}{4!} + \dots \right)
\end{aligned}$$

$$\begin{aligned}
&= -e^T h + \frac{1}{2} \|h\|^2 - \sum_{i=1}^n \left(\frac{h_i^3}{3!} - \frac{h_i^4}{4!} + \dots \right) \\
&= -e^T h + \frac{1}{2} \|h\|^2 - \sum_{i=1}^n \frac{h_i^3}{3!} + \sum_{i=1}^n \frac{h_i^4}{4!} - \dots \\
&\leq -e^T h + \frac{1}{2} \|h\|^2 - \sum_{i=1}^n \frac{h_i^3}{3!} + \sum_{i=1}^n \frac{|h_i|^4}{4!} - \dots \\
&\leq -e^T h + \frac{1}{2} \|h\|^2 + \sum_{i=1}^n \frac{|h_i|^3}{3!} + \sum_{i=1}^n \frac{|h_i|^4}{4!} + \dots \\
&\stackrel{AFIR}{\leq} -e^T h + \frac{1}{2} \|h\|^2 + \frac{1}{3!} \|h\|^3 + \frac{1}{4!} \|h\|^4 + \dots \\
&\leq -e^T h - \|h\| + \|h\| + \frac{1}{2} \|h\|^2 + \frac{1}{3!} \|h\|^3 + \frac{1}{4!} \|h\|^4 + \dots \\
&\leq -e^T h - \log(1 - \|h\|) - \|h\|
\end{aligned}$$

- $h_i \leq |h_i| \quad \rightarrow \quad \sum h_i \leq \sum |h_i|$
- $-|h_i| \leq h_i \quad \rightarrow \quad -|h_i|^3 \leq h_i^3 \quad \rightarrow \quad -h_i^3 \leq |h_i|^3$
 $\rightarrow \quad -\sum h_i^3 \leq \sum |h_i|^3$

■

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 barrera por un término constante que depende de la proximidad de la iteración actual de la trayectoria central, pero eso no depende del número n de las variables no negativas.

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

Presentamos a continuación el algoritmo primal-dual de paso largo.

3.2.1. Algoritmo de Paso Largo

PASO 0. ϵ es un parámetro de precisión;

$1 - \theta$ es el factor de reducción de μ , $0 < \theta < 1$.

ξ es un parámetro de proximidad (predeterminado $\xi = \frac{1}{2}$)

(x^0, s^0) es un par dado de soluciones factibles interiores.

$\mu^0 = \frac{(x^0)^T s^0}{n}$ es el valor inicial para el parámetro de la barrera.

x^0, s^0, μ^0 satisface $\delta(x^0, s^0; \mu^0) \leq \xi$.

PASO 1. Sea $x := x^0$, $s := s^0$, $\mu := \mu^0$.

PASO 2. Si $x^T s < \epsilon$, entonces parar.

PASO 3. Sea $\mu := (1 - \theta) \mu$

PASO 4. Si $\delta(x, s; \mu) < \xi$, entonces ir al PASO 2.

PASO 5. Buscar $\alpha > 0$ tal que

$$f(x + \alpha d_x, s + \alpha d_s; \mu) \leq f(x, s; \mu) - \frac{1}{17}$$

PASO 6. Sea $x := x + \alpha d_x$, $s := s + \alpha d_s$

PASO 7. Ir al PASO 4.

3.3. Lemas importantes

Tenga en cuenta la situación en 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, a continuación, 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

$$\begin{aligned}
\Delta f(\alpha) &:= f(x + \alpha d_x, s + \alpha d_s; \mu) - f(x, s; \mu) \\
&= \frac{1}{\mu} (x + \alpha d_x)^T (s + \alpha d_s) - \sum_{j=1}^n \log \left[(x + \alpha d_x)_j (s + \alpha d_s)_j \right] \\
&\quad - \frac{x^T s}{\mu} + \sum_{j=1}^n \log (x_j s_j) \\
&= \frac{1}{\mu} (x^T + \alpha d_x^T) (s + \alpha d_s) - \sum_{j=1}^n \log \left(x_j \left(1 + \alpha \frac{dx_j}{x_j} \right) s_j \left(1 + \alpha \frac{ds_j}{s_j} \right) \right) \\
&\quad - \frac{x^T s}{\mu} + \sum_{j=1}^n \log (x_j s_j) \\
&= \frac{1}{\mu} \left(x^T s + \alpha x^T d_s + \alpha \underbrace{d_x^T s}_0 + \alpha^2 \underbrace{d_x^T d_s}_0 \right) - \sum_{j=1}^n \log \left(x_j s_j \left(1 + \alpha \frac{dx_j}{x_j} \right) \left(1 + \alpha \frac{ds_j}{s_j} \right) \right) \\
&\quad - \frac{x^T s}{\mu} + \sum_{j=1}^n \log (x_j s_j) \\
&= \frac{1}{\mu} (x^T s + \alpha x^T d_s + \alpha s^T d_x) - \sum_{j=1}^n \log (x_j s_j) - \sum_{j=1}^n \log \left(1 + \alpha \frac{dx_j}{x_j} \right) \left(1 + \alpha \frac{ds_j}{s_j} \right) \\
&\quad - \frac{x^T s}{\mu} + \sum_{j=1}^n \log (x_j s_j) \\
&= \frac{\alpha}{\mu} (x^T d_s + s^T d_x) - \sum_{j=1}^n \log \left[\left(1 + \alpha \frac{dx_j}{x_j} \right) \left(1 + \alpha \frac{ds_j}{s_j} \right) \right]
\end{aligned}$$

Lema 3.3 Sea $\delta := \delta(x, s; \mu) > 0$, Entonces

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

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 $\Gamma := \frac{2\delta}{\delta + \sqrt{4 + \delta^2}}$, Entonces

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

Demostración.

Por definición, la iteración en x -espacio,

Veamos el siguiente resultado:

$$x + \alpha d_x = Xe + \alpha XX^{-1}d_x = X(e + \alpha X^{-1}d_x)$$

como $x > 0$, $s > 0$ se mantiene la positividad de la nueva iteración en el x -espacio, resultando $x^{-1} > 0$, $s^{-1} > 0$.

Además $\|\bar{d}_x\|^2 + \|\bar{d}_s\|^2 = \mu \|u^{-1} - u\|^2$

entonces

$$\begin{aligned} \|\bar{d}_x\|^2 + \|\bar{d}_s\|^2 &= \mu \|u^{-1} - u\|^2 \\ &= \mu\delta^2 = \mu\delta^2(x, s; \mu), \quad \mu > 0, \quad \delta > 0 \end{aligned}$$

$$\rightarrow \|\bar{d}_x\|^2 + \|\bar{d}_s\|^2 > 0 \quad \rightarrow \bar{d}_x \neq 0 \quad \text{y} \quad \bar{d}_s \neq 0$$

Siendo $D = (XS^{-1})^{1/2}$, $D^{-1} = (XS^{-1})^{-1/2} \rightarrow D, D^{-1} > 0$

Concluyendo

$$d_x = D\bar{d}_x > 0 \quad \text{y} \quad d_s = D^{-1}\bar{d}_s > 0 \quad \rightarrow \quad d_x \neq 0 \quad \text{y} \quad d_s \neq 0$$

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

Veamos que $\alpha w < 1$

En efecto,

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

pero $\delta^2 < \delta^2 + w$, $w > 0$

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

Ahora

$$\begin{aligned} \alpha w < 1 &\rightarrow \alpha \sqrt{\|X^{-1}d_x\|^2 + \|S^{-1}d_s\|^2} < 1 \\ &\rightarrow \sqrt{\alpha^2 (\|X^{-1}d_x\|^2 + \|S^{-1}d_s\|^2)} < 1 \\ &\rightarrow \sqrt{(\alpha \|X^{-1}d_x\|)^2 + (\alpha \|S^{-1}d_s\|)^2} < 1 \\ &\rightarrow (\alpha \|X^{-1}d_x\|)^2 + (\alpha \|S^{-1}d_s\|)^2 < 1 \end{aligned}$$

Luego,

$$(\alpha \|X^{-1}d_x\|)^2 < (\alpha \|X^{-1}d_x\|)^2 + (\alpha \|S^{-1}d_s\|)^2 < 1$$

y

$$(\alpha \|S^{-1}d_s\|)^2 < (\alpha \|X^{-1}d_x\|)^2 + (\alpha \|S^{-1}d_s\|)^2 < 1$$

entonces

$$(\alpha \|X^{-1}d_x\|)^2 < 1 \quad \text{y} \quad (\alpha \|S^{-1}d_s\|)^2 < 1$$

$$\alpha \|X^{-1}d_x\| < 1 \quad \text{y} \quad \alpha \|S^{-1}d_s\| < 1$$

En:

$$x + \alpha d_x = X (e + \alpha X^{-1}d_x)$$

$$s + \alpha d_s = S (e + \alpha S^{-1}d_s)$$

Faltaría ver que, $e + \alpha X^{-1}d_x > 0 \quad \wedge \quad e + \alpha S^{-1}d_s > 0$ pues $X > 0$, $S > 0$

Pero

$$e + \alpha X^{-1}d_x = \left[1 + \alpha \frac{dx_1}{x_1}, \dots, 1 + \alpha \frac{dx_n}{x_n} \right]^T$$

$$e + \alpha S^{-1}d_s = \left[1 + \alpha \frac{ds_1}{s_1}, \dots, 1 + \alpha \frac{ds_n}{s_n} \right]^T$$

y

$$\alpha \|X^{-1}d_x\| < 1 \quad \wedge \quad \alpha \|S^{-1}d_s\| < 1$$

$$\alpha \|X^{-1}d_x\| < 1 \quad \wedge \quad \alpha \|S^{-1}d_s\| < 1$$

$$\begin{array}{ccc} \downarrow & & \downarrow \\ \alpha \left| \frac{dx_i}{x_i} \right| < 1 & \wedge & \alpha \left| \frac{ds_i}{s_i} \right| < 1 \end{array}$$

$$\begin{array}{ccc} \downarrow & & \downarrow \\ -1 < \alpha \frac{dx_i}{x_i} \rightarrow 0 < \alpha \frac{dx_i}{x_i} + 1 & \wedge & -1 < \alpha \frac{ds_i}{s_i} \rightarrow 0 < \alpha \frac{ds_i}{s_i} + 1 \end{array}$$

Entonces

$$x + \alpha d_x > 0 \quad , \quad s + \alpha d_s > 0$$

Ahora, por demostrar que

$$\Delta f(\alpha) \leq -\Gamma + \log(1 + \Gamma) \quad , \quad \text{donde} \quad \Gamma = \frac{2\delta}{\delta + \sqrt{4 + \delta^2}}$$

Deducimos anteriormente que

$$\begin{aligned} \Delta f(\alpha) &= \frac{\alpha}{\mu} (x^T d_s + s^T d_x) - \sum_{j=1}^n \log \left(1 + \alpha \frac{(dx)_j}{x_j} \right) \left(1 + \alpha \frac{(ds)_j}{s_j} \right) \\ &= \frac{\alpha}{\mu} (x^T d_s + s^T d_x) - \sum_{j=1}^n \log \left(1 + \alpha \frac{(dx)_j}{x_j} \right) - \sum_{j=1}^n \log \left(1 + \alpha \frac{(ds)_j}{s_j} \right) \end{aligned}$$

Por Lema 3.2

$$\begin{aligned} \Delta f(\alpha) &= \frac{\alpha}{\mu} (S^T d_x + x^T d_x) - (e^T \alpha X^{-1}d_x + \log(1 - \alpha \|X^{-1}d_x\|)) \\ &\quad - (e^T \alpha S^{-1}d_s + \log(1 - \alpha \|S^{-1}d_s\|)) \end{aligned}$$

$$\begin{aligned}
&= \frac{\alpha}{\mu} \left((Se)^T d_x + (Xe)^T d_s \right) - (e^T \alpha X^{-1} d_x + \log(1 - \alpha \|X^{-1} d_x\|)) \\
&- (e^T \alpha S^{-1} d_s + \log(1 - \alpha \|S^{-1} d_s\|))
\end{aligned}$$

entonces

$$\begin{aligned}
\Delta f(\alpha) &= (\alpha e^T) \left(\frac{(e^T S d_x + e^T X d_s)}{\mu} - (X^{-1} d_x + S^{-1} d_s) \right) \\
&- \log(1 - \alpha \|X^{-1} d_x\|) - \log(1 - \alpha \|S^{-1} d_s\|)
\end{aligned}$$

donde

$$\begin{aligned}
\alpha w + \log(1 - \alpha w) &\leq \log(1 - \alpha \|X^{-1} d_x\|) + \log(1 - \alpha \|S^{-1} d_s\|) \\
-\log(1 - \alpha \|X^{-1} d_x\|) - \log(1 - \alpha \|S^{-1} d_s\|) &\leq -\alpha w - \log(1 - \alpha w)
\end{aligned}$$

Luego

$$\begin{aligned}
\Delta f(\alpha) &\leq (\alpha e^T) \left(\frac{(e^T S d_x + e^T X d_s)}{\mu} - (X^{-1} d_x + S^{-1} d_s) \right) \\
&- \alpha w - \log(1 - \alpha w)
\end{aligned}$$

El coeficiente α en el primer término, el miembro del lado derecho en la última desigualdad puede ser identificada como la derivada de $f(x + \alpha d_x, s + \alpha d_s; \mu)$ en $\alpha = 0$. Esta derivada, sin embargo, es simplemente la derivada direccional de $f(x, s; \mu)$ en la dirección (d_x, d_s) que por el Lema 3.1 es igual a δ^2 .

Así, se obtiene lo siguiente: recordando que $\alpha w < 1$

$$\alpha < \frac{1}{w} \rightarrow \delta^2 \alpha < \delta^2 \frac{1}{w} \rightarrow -\delta^2 \alpha > -\delta^2 \frac{1}{w},$$

si

$$\Gamma_1 = -\delta^2 \frac{1}{w} \rightarrow \Gamma_1 < -\delta^2 \alpha$$

entonces

$$-\alpha\delta^2 - \alpha w - \log(1 - \alpha w) \leq -\Gamma_1 + \log(1 + \Gamma_1)$$

$$\Delta f(\alpha) \leq -\alpha\delta^2 - \alpha w - \log(1 - \alpha w) \leq -\Gamma_1 + \log(1 + \Gamma_1)$$

Definimos la función

$$F : (-1, +\infty) \rightarrow \mathbb{R} \quad / \quad F(\Gamma_1) = -\Gamma_1 + \log(1 + \Gamma_1)$$

$$F'(\Gamma_1) = -1 + \frac{1}{1 + \Gamma_1} = -\frac{\Gamma_1}{1 + \Gamma_1}, \quad 1 + \Gamma_1 > 0$$

- $\Gamma_1 < 0$ creciente en $(-1, 0)$
- $\Gamma_1 > 0$ decreciente en $(0, +\infty)$

Por otro lado

Afirmación:

$$\|X^{-1}d_x\|^2 + \|S^{-1}d_s\|^2 \leq \|u^{-1}\|_\infty^2 \delta^2$$

En efecto,

Recordemos:

$$\bar{d}_s := Dd_s \rightarrow d_s = D^{-1}\bar{d}_s \tag{I}$$

$$D = (XS^{-1})^{1/2} \tag{II}$$

$$\bar{d}_x := D^{-1}d_x \rightarrow d_x = D\bar{d}_x \tag{III}$$

$$U := \left(\frac{XS}{\mu} \right)^{1/2} \rightarrow (XS)^{-1/2} = \frac{1}{\sqrt{\mu}}U^{-1} \tag{IV}$$

Veamos

$$X^{-1}d_x \underset{\substack{= \\ \downarrow \\ \text{De (III)}}}{=} X^{-1}D\bar{d}_x \underset{\substack{= \\ \downarrow \\ \text{De (II)}}}{=} X^{-1}(XS^{-1})^{1/2}\bar{d}_x = X^{-1}X^{1/2}S^{-1/2}\bar{d}_x$$

$$\begin{aligned}
&= X^{-1/2} S^{-1/2} \bar{d}_x \\
&= (XS)^{-1/2} \bar{d}_x \\
&= (SX)^{-1/2} \bar{d}_x \\
&\stackrel{\substack{\uparrow \\ \text{De (IV)}}}{=} \frac{1}{\sqrt{\mu}} U^{-1} \bar{d}_x
\end{aligned}$$

Entonces

$$X^{-1} d_x = \frac{1}{\sqrt{\mu}} U^{-1} \bar{d}_x$$

Ahora

$$\begin{aligned}
S^{-1} d_s &\stackrel{\substack{= \\ \downarrow \\ \text{De (I)}}}{=} S^{-1} D \bar{d}_s \stackrel{\substack{= \\ \downarrow \\ \text{De (II)}}}{=} S^{-1} (XS^{-1})^{-1/2} \bar{d}_s = S^{-1} S^{1/2} X^{-1/2} \bar{d}_s \\
&= S^{-1/2} X^{-1/2} \bar{d}_s \\
&= (XS)^{-1/2} \bar{d}_s \\
&\stackrel{\substack{= \\ \uparrow \\ \text{De (IV)}}}{=} \frac{1}{\sqrt{\mu}} U^{-1} \bar{d}_s
\end{aligned}$$

Entonces

$$S^{-1} d_s = \frac{1}{\sqrt{\mu}} U^{-1} \bar{d}_s$$

Luego,

$$\begin{aligned}
\|X^{-1} d_x\|^2 + \|S^{-1} d_s\|^2 &= \frac{1}{\mu} \|U^{-1} \bar{d}_x\|^2 + \frac{1}{\mu} \|U^{-1} \bar{d}_s\|^2 \\
&= \frac{1}{\mu} \left(\|U^{-1} \bar{d}_x\|^2 + \|U^{-1} \bar{d}_s\|^2 \right) \\
&\leq \frac{1}{\mu} \left(\|U^{-1}\|^2 \|\bar{d}_x\|^2 + \|U^{-1}\|^2 \|\bar{d}_s\|^2 \right) \\
&\leq \frac{1}{\mu} \|U^{-1}\|^2 \left(\|\bar{d}_x\|^2 + \|\bar{d}_s\|^2 \right) \\
&\leq \frac{1}{\mu} \|u^{-1}\|_\infty^2 \left(\|\bar{d}_x\|^2 + \|\bar{d}_s\|^2 \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{\|u^{-1}\|_{\infty}^2}{\mu} \left(\mu \|u^{-1} - u\|^2 \right) \\
&\leq \|u^{-1}\|_{\infty}^2 \|u^{-1} - u\|^2 \\
&\leq \|u^{-1}\|_{\infty}^2 \delta^2
\end{aligned}$$

Observación:

Se sabe que

$$\begin{aligned}
\|Ax\| &\leq \|A\| \|x\| \\
\|\bar{d}_x\|^2 + \|\bar{d}_s\|^2 &= \mu \|u^{-1} - u\|^2 \\
\delta &:= \|u^{-1} - u\|
\end{aligned}$$

Ahora, como $\alpha w < 1$ entonces $w \leq \|u^{-1}\|_{\infty} \delta$

Además $\Gamma_1 = -\delta^2 \frac{1}{w}$

$$\begin{aligned}
\rightarrow \Gamma_1 &\geq \frac{\delta^2}{\|u^{-1}\|_{\infty} \delta} = \frac{\delta}{\|u^{-1}\|_{\infty}} \geq \frac{2\delta}{\delta + \sqrt{4 + \delta^2}} = \Gamma \\
\rightarrow -\Gamma_1 + \log(1 + \Gamma_1) &\leq -\Gamma + \log(1 + \Gamma)
\end{aligned}$$

Por tanto,

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

■

- **Corolario 3.1** Si $\delta \geq \frac{1}{2}$, entonces del Lema 3.3, se tiene

$$\Delta f \leq -0,0608 \leq -\frac{1}{17}$$

Demostración.

Como $\delta \geq \frac{1}{2} \rightarrow \delta \geq 0$

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

donde $\Gamma = \frac{2\delta}{\delta + \sqrt{4 + \delta^2}} = \frac{\delta(\sqrt{4 + \delta^2} - \delta)}{2}$

$$\delta \geq \frac{1}{2} \rightarrow \sqrt{4 + \delta^2} \geq \sqrt{\frac{17}{4}} = \frac{\sqrt{17}}{2}$$

$$\rightarrow \sqrt{4 + \delta^2} > \frac{\sqrt{17}}{2} \quad \text{para valores cercanos } \delta = \frac{1}{2}$$

$$\rightarrow \sqrt{4 + \delta^2} - \delta \geq \frac{\sqrt{17}}{2} - \frac{1}{2}$$

$$\rightarrow \frac{\delta}{2} (\sqrt{4 + \delta^2} - \delta) \geq \frac{1}{2} \left(\frac{\sqrt{17}}{2} - \frac{1}{2} \right) = \frac{\sqrt{17} - 1}{8}$$

$$\rightarrow \Gamma = \frac{2\delta}{\delta + \sqrt{4 + \delta^2}} \geq \frac{\sqrt{17} - 1}{8} = 0,390388\dots$$

Defino $h : \langle -1, \infty \rangle \rightarrow \mathbb{R} / h(\Gamma) = -\Gamma + \log(1 + \Gamma)$

$$\rightarrow h'(\Gamma) = \frac{-\Gamma}{1 + \Gamma}, \quad 1 + \Gamma > 0$$

- Si $\frac{-\Gamma}{1 + \Gamma} > 0 \rightarrow -\Gamma > 0 \rightarrow \Gamma < 0$ creciente en $(-1, 0)$
- Si $\frac{-\Gamma}{1 + \Gamma} < 0 \rightarrow -\Gamma < 0 \rightarrow \Gamma > 0$ decreciente en $(0, +\infty)$

Como $\frac{\sqrt{17} - 1}{8} = 0,390388\dots > 0$

$$\Gamma \geq \frac{\sqrt{17} - 1}{8} \rightarrow h(\Gamma) \leq h\left(\frac{\sqrt{17} - 1}{8}\right)$$

$$\rightarrow -\Gamma + \log(1 + \Gamma) \leq h\left(\frac{\sqrt{17} - 1}{8}\right) \approx -0,006\dots \leq -\frac{1}{17}$$

$$\therefore -\Gamma + \log(1 + \Gamma) \leq -0,006\dots \leq -\frac{1}{17}$$

■

Lema 3.4 Si $\delta : \delta(x, s; \mu) < \frac{1}{\sqrt{2}}$, entonces

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

$$\text{donde } \sigma := \frac{\delta}{2} (\delta + \sqrt{4 + \delta^2})$$

Demostración.

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

Por otro lado notemos que

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

Denotemos $h = \frac{X s}{\mu} - e$

Afirmación 1: $\|h\| < 1$

En efecto,

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

Afirmación 2:

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

En efecto,

Sabemos que $|u_i^{-1} - u_i| \leq \|u^{-1} - u\|$ para cada $i = 1, 2, \dots, n$

Como $\delta(x, s; \mu) := \|u^{-1} - u\|$

entonces $|u_i^{-1} - u_i| \leq \delta$

$$\rightarrow -\delta \leq u_i^{-1} - u_i \leq \delta \tag{3.1}$$

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

En (3.1), multiplicamos por u_i ambos miembros

$$-u_i \delta \leq 1 - u_i^2 \leq u_i \delta$$

$$u_i^2 - \delta u_i - 1 \leq 0 \leq u_i^2 + \delta u_i - 1$$

$$\left(u_i - \frac{\delta}{2}\right)^2 - \frac{\delta^2}{4} - 1 \leq 0 \leq \left(u_i + \frac{\delta}{2}\right)^2 - \frac{\delta^2}{4} - 1$$

$$\left(u_i - \frac{\delta}{2}\right)^2 \leq \frac{\delta^2}{4} + 1 \leq \left(u_i + \frac{\delta}{2}\right)^2$$

$$\sqrt{\left(u_i - \frac{\delta}{2}\right)^2} \leq \sqrt{\frac{\delta^2}{4} + 1} \leq \sqrt{\left(u_i + \frac{\delta}{2}\right)^2}$$

$$\left|u_i - \frac{\delta}{2}\right| \leq \frac{\sqrt{\delta^2 + 4}}{2} \leq \left|u_i + \frac{\delta}{2}\right| \tag{3.2}$$

pero

$$u_i - \frac{\delta}{2} \leq \left| u_i - \frac{\delta}{2} \right|$$

y

$$\left| u_i + \frac{\delta}{2} \right| = u_i + \frac{\delta}{2}, \quad \text{pues } u_i + \frac{\delta}{2} > 0$$

En (3.2),

$$u_i - \frac{\delta}{2} \leq \frac{\sqrt{\delta^2 + 4}}{2} \leq u_i + \frac{\delta}{2} \quad (3.3)$$

De (3.3),

$$\begin{aligned} u_i - \frac{\delta}{2} \leq \frac{\sqrt{\delta^2 + 4}}{2} &\rightarrow u_i \leq \frac{\sqrt{\delta^2 + 4}}{2} + \frac{\delta}{2} \\ &\rightarrow u_i \leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \end{aligned} \quad (3.4)$$

De (3.3),

$$\begin{aligned} \frac{\sqrt{\delta^2 + 4}}{2} \leq u_i + \frac{\delta}{2} &\rightarrow \frac{\sqrt{\delta^2 + 4}}{2} - \frac{\delta}{2} \leq u_i \\ &\rightarrow \frac{-\delta + \sqrt{\delta^2 + 4}}{2} \leq u_i \end{aligned} \quad (3.5)$$

De (3.4) y (3.5)

$$\frac{-\delta + \sqrt{\delta^2 + 4}}{2} \leq u_i \leq \frac{\delta + \sqrt{\delta^2 + 4}}{2} \quad (3.6)$$

Luego,

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

Entonces

$$\|u\|_\infty \delta \leq \frac{\delta}{2} (\delta + \sqrt{\delta^2 + 4}) \quad (3.7)$$

con esto queda probado la afirmación 2.

Por otra parte, en (3.6)

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

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

Luego,

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

$$\|u^{-1}\|_\infty \leq \frac{\delta + \sqrt{\delta^2 + 4}}{2}$$

Por otro lado, como

$$\delta < \frac{1}{\sqrt{2}} \quad \rightarrow \quad 0 \leq \delta^2 < \frac{1}{2}$$

$$\rightarrow \delta^2 + 4 < \frac{1}{2} + 4 = \frac{9}{2} \quad \rightarrow \quad \sqrt{4 + \delta^2} < \frac{3}{\sqrt{2}}$$

De (II) y afirmación 2,

$$\|h\| \leq \frac{\delta}{2} (\delta + \sqrt{4 + \delta^2})$$

$$< \frac{1}{2\sqrt{2}} (\delta + \sqrt{4 + \delta^2})$$

$$< \frac{1}{2\sqrt{2}} \left(\frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} \right) = \frac{1}{2\sqrt{2}} \left(\frac{4}{\sqrt{2}} \right) = 1$$

entonces,

$$\|h\| < 1$$

Con esto queda probado la afirmación 1.

Ahora, de (I),

$$\begin{aligned}
 f(x, s; \mu) - f(x(\mu), s(\mu); \mu) &= e^T h - \sum_{j=1}^n \log(1 + h_j) \\
 \text{por lema (3.2)} &\leq e^T h - e^T h - \|h\| - \log(1 - \|h\|) \\
 &\leq -\|h\| - \log(1 - \|h\|)
 \end{aligned} \tag{3.8}$$

Pero sabemos que $h = \frac{Xs}{\mu} - e$
entonces

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

entonces

$$\begin{aligned}
 h_j &= \frac{x_j s_j}{\mu} - 1 \quad \forall j = 1, 2, \dots, n \\
 \rightarrow h_j + 1 &= \frac{x_j s_j}{\mu}
 \end{aligned}$$

Por otro lado, si

Definimos $F : \langle 0, 1 \rangle \rightarrow \mathbb{R} / F(x) = -x - \log(1 - x)$

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

■ Si $\frac{x}{1-x} > 0 \rightarrow x > 0$

Entonces $F(x)$ es creciente.

■ Si $\frac{x}{1-x} < 0 \rightarrow x < 0$ (pero no esta definida en el dominio)

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

Así mismo de afirmación 1 y (II) 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,

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

En (3.8)

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

■

Corolario 3.2 Si $\delta(x, s; \mu) \leq \frac{1}{2}$. Entonces,

$$f(x, s; \mu) - f(x(\mu), s(\mu); \mu) \leq 0,4674 < \frac{1}{2}$$

Demostración.

Por lema 3.4

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

donde $\sigma = \frac{\delta}{2} (\delta + \sqrt{4 + \delta^2})$

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 $F'(x) < 0 \rightarrow \frac{x}{1-x} < 0 \rightarrow x < 0$

Entonces $F(x)$ es decreciente.

■ Si $F'(x) > 0 \rightarrow \frac{x}{1-x} > 0 \rightarrow x > 0$

Entonces $F(x)$ es creciente.

Como $\delta \leq \frac{1}{2}$

$$\rightarrow \sigma \leq \frac{1}{4} \left(\frac{1}{2} + \sqrt{4 + \frac{1}{4}} \right) = \frac{1 + \sqrt{17}}{8} \approx 0,640388\dots$$

$$\rightarrow F(\sigma) \leq F\left(\frac{1 + \sqrt{17}}{8}\right)$$

$$\rightarrow -\sigma - \log(1 - \sigma) \leq -\left(\frac{1 + \sqrt{17}}{8}\right) - \log\left(1 - \left(\frac{1 + \sqrt{17}}{8}\right)\right) < \frac{1}{2}$$

$$\rightarrow -\sigma - \log(1 - \sigma) < \frac{1}{2}$$

$$\therefore f(x, s; \mu) - f(x(\mu), s(\mu); \mu) < \frac{1}{2}$$

■

Lema 3.5 Sea $\mu^+ := (1 - \theta)\mu$. Entonces

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

Demostración.

$$\begin{aligned} f(x(\mu), s(\mu); \mu) - f(x(\mu^+), s(\mu^+); \mu^+) &= \frac{(x(\mu))^T s(\mu)}{\mu} - \sum_{j=1}^n \log\left(\left(x(\mu)\right)_j \left(s(\mu)\right)_j\right) \\ &\quad - \frac{(x(\mu^+))^T s(\mu^+)}{\mu^+} + \sum_{j=1}^n \log\left(\left(x(\mu^+)\right)_j \left(s(\mu^+)\right)_j\right) \\ &= \frac{\mu n}{\mu} - \sum_{j=1}^n \log(\mu) - \frac{\mu^+ n}{\mu^+} + \sum_{j=1}^n \log(\mu^+) \end{aligned}$$

$$\begin{aligned}
&= n - n \log \mu - n + \log \mu^+ \\
&= n \log \mu^+ - n \log \mu \\
&= n (\log \mu^+ - \log \mu) \\
&= n \log \left(\frac{\mu^+}{\mu} \right) \\
&= n \log \left(\frac{(1 - \theta) \mu}{\mu} \right) \\
&= n \log (1 - \theta)
\end{aligned}$$

■

Lema 3.6 *Supongamos que (x, s) es un par de soluciones factibles primal-dual tal que $\delta := \delta(x, s; \mu) < \frac{1}{\sqrt{2}}$. Entonces*

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

Demostración.

Sea

$$\begin{aligned}
\left| \frac{x^T s}{\mu} - n \right| &= \left| \frac{e^T X s}{\mu} - n \right| = \left| e^T \left(\frac{X s}{\mu} - e \right) \right| \\
&\leq \|e\| \left\| \left(\frac{X s}{\mu} - e \right) \right\| && \text{(Cauchy Schwarz)} \\
&< \|e\| 1 && \text{(por lema 3.4)} \\
&< \sqrt{n}
\end{aligned}$$

pero

$$\begin{aligned}
\frac{x^T s}{\mu} - n &\leq \left| \frac{x^T s}{\mu} - n \right| < \sqrt{n} \\
\rightarrow \frac{x^T s}{\mu} - n < \sqrt{n} &\rightarrow \frac{x^T s}{\mu} < n + \sqrt{n}
\end{aligned}$$

■

Capítulo 4

Análisis de Convergencia

Lema 4.1 *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

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

Demostración.

Sea k el número de iteraciones exteriores. Entonces al actualizar k -veces el parámetro de barrera, del lema (3.5) se obtiene lo siguiente

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

$$\rightarrow \mu_k = (1 - \theta)^k \mu_0 \tag{4.1}$$

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

Luego, por lema 3.6

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

Entonces

$$\begin{aligned} (x^k)^T s^k &\leq 2n\mu^k \\ &\leq 2n(1-\theta)^k \mu_0 \quad \text{por (4.1)} \end{aligned}$$

Aplicando logaritmo en ambos miembros

$$\begin{aligned} \log\left((x^k)^T s^k\right) &\leq \log\left(2n(1-\theta)^k \mu_0\right) \\ &= \log(2n\mu_0) + \log(1-\theta)^k \\ -\log(1-\theta)^k &\leq \log(2n\mu_0) - \log\left((x^k)^T s^k\right) \\ -k \log(1-\theta) &\leq \log\left(\frac{2n\mu_0}{(x^k)^T s^k}\right) \end{aligned} \tag{4.2}$$

Utilizando $\theta \leq -\log(1-\theta)$, $0 < \theta < 1$

En (4.2)

$$k\theta \leq \log\left(\frac{2n\mu_0}{(x^k)^T s^k}\right) \tag{4.3}$$

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

Aplicando logaritmo a ambos miembros

$$\begin{aligned}
& \log \left((x^k)^T s^k \right) > \log \varepsilon \\
& -\log \left((x^k)^T s^k \right) < -\log \varepsilon \\
& \log (2n\mu_0) - \log \left((x^k)^T s^k \right) < \log (2n\mu_0) - \log \varepsilon \\
& \log \left(\frac{2n\mu_0}{(x^k)^T s^k} \right) < \log \left(\frac{2n\mu_0}{\varepsilon} \right)
\end{aligned} \tag{4.4}$$

De (4.3) y (4.4)

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

por tanto,

$$(x^k)^T s^k \leq \varepsilon$$

■

Lema 4.2 *El número de iteraciones interiores en una iteración exterior está limitada por*

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

Demostración.

Sea (x^0, s^0) el punto inicial en el algoritmo del paso largo en las iteraciones internas

$$\delta(x^0, s^0, \mu^0) \leq \frac{1}{2}, \quad \mu > 0 \quad (\text{PASO 0})$$

Entonces

$$\mu^+ = (1 - \theta) \mu \quad (\text{PASO 3})$$

(valor actualizado del parámetro de barrera)

Ahora la siguiente iteración interna será tomado con respecto μ^+ .

Supongamos que k iteraciones internas se han realizado sin actualizar μ^+ .

Entonces

$$\delta(x^j, s^j, \mu^+) > \frac{1}{2}, \quad \forall j \leq k$$

Por el corolario 3.1

$$\Delta f(\alpha) \leq -\frac{1}{17}$$

$$f(x^k, s^k, \mu^+) - f(x^{k-1}, s^{k-1}, \mu^+) \leq -\frac{1}{17}$$

⋮

$$f(x^1, s^1, \mu^+) - f(x^0, s^0, \mu^+) \leq -\frac{1}{17}$$

sumando se tiene

$$f(x^k, s^k, \mu^+) - f(x^0, s^0, \mu^+) \leq -\frac{k}{17}$$

$$f(x^0, s^0, \mu^+) - f(x^k, s^k, \mu^+) \geq \frac{k}{17} \tag{4.5}$$

Además

$$\begin{aligned} f(x^k, s^k, \mu^+) &\geq f(x(\mu^+), s(\mu^+); \mu^+) \\ -f(x^k, s^k, \mu^+) &\leq -f(x(\mu^+), s(\mu^+); \mu^+) \end{aligned}$$

de (4.5)

$$\frac{k}{17} \leq f(x^0, s^0, \mu^+) - f(x^k, s^k, \mu^+) \leq f(x^0, s^0, \mu^+) - f(x(\mu^+), s(\mu^+), \mu^+)$$

Luego

$$k \leq 17(f(x^0, s^0, \mu^+) - f(x(\mu^+), s(\mu^+), \mu^+)) \tag{4.6}$$

Pero

$$\begin{aligned}
f(x^0, s^0, \mu^+) - f(x(\mu^+), s(\mu^+), \mu^+) &= \underbrace{f(x^0, s^0, \mu^+) - f(x^0, s^0, \mu)}_A \\
&+ \underbrace{f(x^0, s^0, \mu^+) - f(x(\mu), s(\mu), \mu)}_B \\
&+ \underbrace{f(x(\mu), s(\mu), \mu) - f(x(\mu^+), s(\mu^+), \mu^+)}_C
\end{aligned}$$

Debemos acotar $f(x^0, s^0, \mu^+) - f(x(\mu^+), s(\mu^+), \mu^+)$,
entonces acotaremos A, B y C.

Veamos A:

$$\begin{aligned}
f(x^0, s^0, \mu^+) - f(x^0, s^0, \mu) &= \frac{(x^0)^T s^0}{\mu^+} - \sum_{j=1}^n \log\left(\frac{(x^0)_j (s^0)_j}{\mu^+}\right) \\
&- \frac{(x^0)^T s^0}{\mu} - \sum_{j=1}^n \log\left(\frac{(x^0)_j (s^0)_j}{\mu}\right) \\
&= \frac{(x^0)^T s^0}{\mu^+} - \frac{(x^0)^T s^0}{\mu} \\
&= (x^0)^T s^0 \left(\frac{1}{\mu^+} - \frac{1}{\mu}\right) \\
&= (x^0)^T s^0 \left(\frac{1}{(1-\theta)\mu} - \frac{1}{\mu}\right) \\
&= \frac{(x^0)^T s^0}{\mu} \left(\frac{1}{1-\theta} - 1\right) \\
&= \frac{(x^0)^T s^0}{\mu} \left(\frac{\theta}{1-\theta}\right) \\
&\leq \frac{(n + \sqrt{n})\theta}{1-\theta} \quad (\text{por lema 3.6})
\end{aligned}$$

Observaci3n: $\mu^+ = (1 - \theta)\mu$, $\delta(x^0, s^0, \mu) \leq \frac{1}{2} < \frac{1}{\sqrt{2}}$

Veamos B:

Como $\delta(x^0, s^0, \mu) \leq \frac{1}{2}$, entonces por corolario 3.2

$$f(x^0, s^0, \mu^+) - f(x(\mu), s(\mu), \mu) \leq \frac{1}{2}$$

Veamos C:

Como $\mu^+ = (1 - \theta)$, entonces por corolario 3.5

$$f(x(\mu), s(\mu), \mu) - f(x(\mu^+), s(\mu^+), \mu^+) = n \log(1 - \theta) \quad (4.7)$$

Utilizando $\theta \leq -\log(1 - \theta)$, $0 < \theta < 1$

En (4.7)

$$f(x(\mu), s(\mu), \mu) - f(x(\mu^+), s(\mu^+), \mu^+) \leq -\theta n$$

Reemplazando las cotas superiores de A, B y C en (4.6), se tiene

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

entonces

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

■

Teorema 4.1 *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 $x^T s < \varepsilon$.

Demostración.

Por el lema 4.1, el algoritmo de paso largo tiene a lo más

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

Por otra parte, del lema 4.2, se tiene

$$17 \left(\frac{\theta\sqrt{n} + \theta^2 n}{1 - \theta} + \frac{1}{2} \right) \text{ iteraciones interiores en una iteraci3n exterior}$$

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 $x^T s < \varepsilon$. ■

Capítulo 5

Conclusiones

Al final del trabajo podemos arribar a las siguientes conclusiones:

- El algoritmo de Paso Largo genera iteraciones (x^k, y^k, s^k) 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 \bar{d}_x y \bar{d}_s que resultan ser la descomposición ortogonal del espacio de \bar{A} .
- Se utiliza una medida de proximidad del tipo $\delta(x, s, \mu) := \|u^{-1} - u\|$ 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.

Bibliografía

- [1] Bonnán J. F. and Gonzaga C.C. *Convergence of interior-point algorithms for the monotone linear con elementary problem*. Mathematics of Operations Research, 21 (1996), pp. 1-25.
- [2] Cottle R. W.; Pang J. S. and Stone R. E. *The Linear Complementary Problem*. Academic Press, San Diego, 1992.
- [3] Fiacco, A. V. and McCormick, G. P., R. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Wiley, New York. 1968.
- [4] Frisch, R. *The Logarithmic Potencial Method of Convex Programming*. Technical Report, University Institute of Economics. Oslo, Norway, 1995.
- [5] Gill C. C., Murray W., Saunders M., Tomlin J. and Wright M. *On Projected Newton Barrier Methods for Linear Programming and an Equivalence to Karmarkar's Projective Method*. Mathematical Programming, 36 (1996), pp. 183-209.
- [6] Gonzaga C. C. *Polynomial affine algorithms for linear programming*, *Mathematical Programming*, 49 (1990), pp. 7-21.
- [7] Gonzaga C. C. and Todd M. J. *An $o(\sqrt{n}L)$ - iteration large-step primaldual affine algorithm for linear programming*. SIAM Journal on Optimization, 2 (1992), pp. 349-359.
- [8] Guler O., Roos C., Terlaky T., and Vial J. P. *A survey of the implications of the behavior of the central path for the duality theory of linear programming*. Management Science, 41 (1995), pp. 1922-1934.
- [9] Karmarkar N. *A new polynomial-time algorithm for linear programming*, *Combinatorica*, 4 (1984), pp. 373-395.

- [10] Khachiyan L. G. *A polynomial algorithm in linear programming*, Soviet Mathematics Doklady, 20 (1979), pp. 191-194.
- [11] Kojima M. *Basic lemmas in polynomial-time infeasible-interior-point methods for linear programs*. Annals of Operations Research, 62 (1996), pp. 1-28.
- [12] 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, (1989), 29-47.
- [13] Mcshane K. A., Monma C. L., and Shanno D. F. *An implementation of a primal-dual interior-point method for linear programming*. ORSA Journal on Computing, 1 (1989), pp. 70-83.
- [14] Monteiro R. D. and Adler I. *Interior path-following primal-dual algorithms. Part I: Linear programming*. Mathematical Programming, 44 (1989), pp. 27-41.
- [15] Nesterov Y. E. and Nemirovskii A. S. *Interior Point Polynomial Methods in Convex Programming*. SIAM Publications, Philadelphia, 1994.
- [16] 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.
- [17] Todd M. J. *A centered projective algorithm for linear programming*. Mathematics of Operations Research , 15 (1990), pp. 508-529.
- [18] Xu X. *An $o(\sqrt{n}L)$ - iteration large-step infeasible path-following algorithm for linear programming*. Technical Report, Department of Management Sciences, University of Iowa, Iowa City, 1994.
- [19] Wright, S. J., *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.
- [20] Wu Y. Ch., Debs A. S., Marsten R. E., *A direct predictor-corrector primal dual interior point algorithm for optimal power flows*. IEEE Transactions on Power systems, 9 (1994), pp 876-883.