



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**INTRODUCCIÓN A LA PROGRAMACIÓN
SEMIDEFINIDA: EFICIENCIA ALGORÍTMICA E
IMPLEMENTACIÓN A PROBLEMAS DE
PROGRAMACIÓN LINEAL**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARÍA

P R E S E N T A:

ESTEBAN CASTRO RAMOS



**DIRECTOR DE TESIS:
ACT. Germán Valle
2009**



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

A mis padres Arturo Castro y Ana María Ramos, por su paciencia y apoyo a lo largo de todos estos años y los venideros.

A mi hermana Mariana Castro, por todo el arte que ha aportado a mi vida.

A mi asesor Germán Valle, por la paciencia de Santo Job que tiene y la dedicación que tuvo para la elaboración del presente trabajo.

An equation for me has no meaning unless it expresses a thought of god.
Srinivasa Ramanujan

Índice general

Agradecimientos	I
Introducción	IV
1. El Método Simplex	1
1.1. Introducción	1
1.2. Descripción del método Simplex	2
1.3. El Dual	14
1.4. La eficiencia del método Simplex	21
2. El algoritmo proyectivo de Karmarkar	30
2.1. Conceptos preliminares	30
2.2. Descripción del algoritmo	36
2.3. Análisis de convergencia	47
3. Introducción a la Programación Semidefinida	52
3.1. Definiciones previas	52
3.2. Objetivo de la PSD	54
3.3. Aplicaciones y alcance	58
3.4. Conceptos preliminares	60
4. El método primal de la barrera logarítmica	69
4.1. Conceptos preliminares	69
4.2. La búsqueda primal	73
4.3. Convergencia y eficiencia	79
Conclusiones	92
Bibliografía	92

Introducción

La presente tesis trata sobre optimización con restricciones, es decir, se plantea una función para la cual se desea calcular su máximo (mínimo) y las variables que aparecen deben de cumplir con ciertas restricciones, las cuales son, en general, desigualdades. La función puede, en principio, ser lineal o no al igual que las restricciones. Sin embargo, sólo se hablará de aquellas funciones lineales con restricciones del mismo tipo.

Existen diversos métodos para resolver este tipo de problemas. El más conocido y usado es el Método Simplex. Este método se detalla en el Capítulo 1, así como la definición de Dual y su desempeño como algoritmo. Se demostrará que este método puede ser terriblemente ineficiente (en el sentido del tiempo o del número de iteraciones).

En el Capítulo 2, se da una de las primeras alternativas al Método Simplex: El algoritmo de Karmarkar. En este Capítulo se describe a detalle en qué consiste este nuevo enfoque para resolver los problemas antes mencionados así como su eficiencia algorítmica. Se demostrará que el algoritmo es mucho más eficiente que el Simplex (nuevamente, en términos de iteraciones). En este punto del trabajo, quedará claro que tanto el Simplex como el algoritmo de Karmarkar son útiles para problemas lineales con restricciones del mismo tipo.

En el Capítulo 3 se darán una serie de conceptos y resultados preliminares para analizar un nuevo tipo de enfoque capaz de resolver problemas lineales como no lineales: la Programación Semidefinida. Se expondrá el por qué del nombre y las principales aplicaciones que tiene.

En el Capítulo 4 se hablará concretamente de un algoritmo de punto interior que resuelve problemas de Programación Semidefinida. Se expondrá a detalle la convergencia y la eficiencia de dicho algoritmo, así como sus principales aplicaciones.

Capítulo 1

El Método Simplex

1.1. Introducción

Desde que Leibniz y Newton inventaron el cálculo diferencial e integral, los modeladores descubrieron que podían encontrar óptimos a problemas prácticos: la velocidad máxima de un proyectil con movimiento parabólico, los máximos y mínimos de frecuencia de un objeto resonante, entre otros. Todos estos problemas pueden ser modelados en términos de una variable y esa variable no tiene restricción alguna, salvo pertenecer al dominio de la función. Sin embargo, con el paso del tiempo surgieron otro tipo de problemas “con restricciones” que no podían ser resueltos usando directamente la técnicas del cálculo. Por ejemplo, considere una fábrica que puede manufacturar n productos. Cada uno de ellos numerados $1, 2, \dots, n$. Suponga que la fábrica requiere de m materias primas diferentes y que en este momento, la fábrica dispone de b_i unidades de la i -ésima materia prima y que para producir una unidad del j -ésimo producto, se requiere a_{ij} unidades de la i -ésima materia prima. Asuma que se producen x_j unidades del j -ésimo producto y que cada unidad producida tiene una utilidad neta de c_j unidades monetarias. La utilidad total está dada por

$$Z = \sum_{i=1}^n c_i x_i$$

Si se deseara obtener la máxima ganancia con ese modelo, resultaría que la ganancia es infinita (bastaría con aumentar el valor de las x_i 's tanto como se quiera). Sin embargo, el planteamiento del modelo no termina con la definición de Z . Existen ciertas restricciones al modelo. Por ejemplo, las unidades producidas deben ser no negativas, es decir, $x_j \geq 0 \forall j$ (no tiene

sentido producir, por ejemplo, -10 unidades de algún producto). Por otro lado, no se puede producir más de lo que permite el total de materia prima, es decir,

$$\sum_{j=1}^n a_{ij}x_j \leq b_i \quad \forall i = 1, 2, \dots, m$$

Así pues, el modelo completo queda dado por:

$$\max Z = \sum_{i=1}^n c_i x_i$$

sujeto a:

$$\sum_{j=1}^n a_{ij}x_j \leq b_i \quad \forall i = 1, 2, \dots, m$$

$$x \geq 0$$

Al Problema anterior se le conoce como **problema de Programación Lineal**. En la actualidad existen varios métodos para resolver este tipo de problemas. El más usado, y conocido, es el llamado **Método Simplex**. En las siguientes secciones se detalla en qué consiste este método.

1.2. Descripción del método Simplex

El método Simplex es un proceso iterativo el cual comienza con una solución que satisfaga la no-negatividad y luego se busca una nueva solución, que sea mejor en el sentido de que aumente el valor de la función objetivo. Este proceso continua hasta que se llegue a una solución óptima, si es que existe. De no existir, el mismo método es capaz de detectarlo.

Antes de comenzar con un problema de programación lineal general, conviene recordar algunas definiciones.

Definición 1.1. Sea V un espacio vectorial con un producto interior definido en él. Un *semi espacio* S de V , es un subconjunto de V de la forma:

$$S = \{v \in V : p \cdot v \leq k\}$$

donde $p \in V$ y k es un escalar.

Un ejemplo de semi espacio es el conjunto $\{x \in \mathbb{R}^2 : p \cdot x \leq 1\}$ con $p = (-1, 1)$. Este conjunto describe a todos los puntos en \mathbb{R}^2 que están por debajo de la recta $y = x + 1$. Dada la naturaleza de los problemas de programación lineal, se estará trabajando con la intersección de semi espacios. Por esa razón, es conveniente dar la siguiente

Definición 1.2. Sean $\{\mathcal{E}_i \in I\}$ con I un conjunto finito de índices, un conjunto de semi espacios en \mathbb{R}^n . Al conjunto $\bigcap_{i=1}^n \mathcal{E}_i$ se le llama *poliedro*. Es decir, un poliedro es la intersección finita de semi espacios.

Más adelante se habla de una característica importante del conjunto de soluciones factibles de un problema de programación lineal. Para poder hablar de ella, se requieren de las siguientes dos definiciones.

Definición 1.3. Sea $A \subseteq \mathbb{R}^n$. Si $\forall x, y \in A$ se tiene que $\{x + t(x - y) : t \in [0, 1]\} \subset A$, se dice que A es *convexo*. A la combinación lineal $\{x + t(x - y) : t \in [0, 1]\}$ se le llama *combinación lineal convexa*. Además, si $t \in (0, 1)$, se dice que es una combinación lineal estrictamente convexa o combinación lineal convexa estricta.

El concepto de convexidad se puede generalizar a otro tipo de conjuntos que no sean necesariamente subconjuntos del espacio \mathbb{R}^n , sin embargo, en el presente trabajo los problemas a enfrentar se presentan justo en \mathbb{R}^n .

Ahora bien, considere el siguiente problema general de programación lineal:

$$\max \sum_{j=1}^n c_j x_j$$

sujeto a:

$$\begin{aligned} \sum_{j=1}^n a_{ij} x_j &\leq b_i & i = 1, 2, \dots, m \\ x_j &\geq 0 & j = 1, 2, \dots, n \end{aligned}$$

El conjunto de soluciones factibles, es decir, puntos en el espacio \mathbb{R}^n que cumplen con las n desigualdades, es un **poliedro convexo**. Si un poliedro convexo es acotado, se llama **politopo convexo**. Para cada i , la ecuación

$$\sum_{j=1}^n a_{ij} x_j = b_i$$

define un hiperplano y cada uno de éstos divide al espacio en dos semi espacios. Así, se tienen en total m hiperplanos y $n + m$ semi espacios. Conviene dar la siguiente definición para la intersección de hiperplanos:

Definición 1.4. Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo. Se dice que un punto $x \in C$ es un *vértice*, si es la intersección de n -hiperplanos linealmente independientes.

Por ejemplo, en \mathbb{R}^2 , $\{x \in \mathbb{R}^2 : p \cdot x \leq 1\} \cap \{x \in \mathbb{R}^2 : w \cdot x \leq 5\}$ con $p = (-1, 1)$, $w = (1, 1)$ definen un vértice.

Es importante hacer notar, que si se tienen n variables, se requieren de por lo menos n ecuaciones para tener una solución única. Por otro lado, tener mas de n ecuaciones resultaría redundante, pues al menos una podrá ser escrita como combinación lineal de las demás. En resumen, cada vértice queda determinado por n ecuaciones. El primer paso para solucionar el problema, es definir las variables de holgura y nombrar a la función objetivo:

$$\begin{aligned} \max Z &= \sum_{j=1}^n c_j x_j & (1.1) \\ w_i &= b_i - \sum_{j=1}^n a_{ij} x_j & i = 1, 2, \dots, m \end{aligned}$$

Donde los términos w_i son las variables de holgura, las x_i son las variables de decisión, las a_{ij} y las c_i 's son constantes. Cabe aclarar, que las variables de holgura son variables que se introducen en cada una de las desigualdades para convertirlas en ecuaciones. Conforme se avanza para encontrar la solución óptima con el método Simplex, las variables de holgura se van mezclando con las variables originales. Por esa razón, a veces es conveniente manejar una notación en la que las variables de holgura y las de decisión sean mas o menos indistinguibles entre sí. Para ello, sea $x_{n+i} = w_i$ para $i = 1, 2, \dots, m$. Con esta notación podemos reescribir 1.1 como

$$\begin{aligned} \max Z &= \sum_{j=1}^n c_j x_j \\ x_{n+i} &= b_i - \sum_{j=1}^n a_{ij} x_j & i = 1, 2, \dots, m. \end{aligned}$$

Ésta es la tabla simplex inicial. Conforme el método avanza, se va cambiando de una tabla a otra en la búsqueda de una solución óptima. Por

la construcción del método, el Simplex busca el óptimo en los vértices del poliedro. A lo largo del presente trabajo, se estará trabajando con matrices. Por esa razón, resulta conveniente dar una notación para el conjunto de matrices de tamaño $m \times n$ con entradas en los Reales. A este conjunto se le denotará por $\mathbf{M}_{m \times n}(\mathbb{R})$. A la dimensión del subespacio formado por los renglones de A se le conoce como *rango* de A y será denotado por $\text{rango}(A)$. Antes de continuar con la construcción, resulta conveniente la siguiente:

Definición 1.5. Sean $A \in \mathbf{M}_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$, $x \in \mathbb{R}^{n+m}$, con cada entrada de x no negativa. Considérese ahora el sistema:

$$[A, I]x = b$$

Suponga que el $\text{rango}(A, b) = \text{rango}(A) = m$. Después de un posible rearrreglo de las columnas de A , sea $A = [B, N]$, donde $B \in \mathbf{M}_{m \times m}(\mathbb{R})$ invertible y $N \in \mathbf{M}_{m \times (n-m)}(\mathbb{R})$. El punto $x = [x_{\mathcal{B}}, x_{\mathcal{N}}]^t$ donde

$$\begin{aligned} x_{\mathcal{B}} &= B^{-1}b \\ x_{\mathcal{N}} &= 0 \end{aligned}$$

es llamado una *solución básica* del sistema. Si cada entrada del vector $x_{\mathcal{B}}$ es no negativa, entonces a x se le llama *solución factible básica* del sistema. A la matriz B se le conoce como *matriz básica o base* y a N como *matriz no básica*. A las entradas de $x_{\mathcal{B}}$ se les llama *variables básicas* y a las entradas de $x_{\mathcal{N}}$ se les llama *variables no básicas*. Si al menos una entrada es cero, x será una *solución factible básica degenerada*.

Observe que se pide que el rango de A sea exactamente el número de renglones de A . Esta restricción garantiza que los renglones de A forman un conjunto linealmente independiente. Como los renglones de A forman un conjunto linealmente independiente, y \mathbb{R}^m tiene dimensión m , resulta que los renglones de A forman una base de \mathbb{R}^m . De ahí que la matriz B se le llame matriz básica. Existe una relación entre las soluciones básicas y los puntos del poliedro. Esa relación queda dada en el *teorema de representación*, pero antes de enunciarlo y demostrarlo, se requieren de los siguientes conceptos.

Definición 1.6. Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo y $c \in C$. Se dice que c es punto extremo de C si **no** puede ser expresado como combinación lineal estrictamente convexa de dos puntos diferentes de C , o más formalmente, c es punto extremo si

$$\forall c_1, c_2 \in C, c_1 \neq c_2, \nexists \lambda \in (0, 1) \quad \text{tal que} \quad c = \lambda c_1 + (1 - \lambda)c_2.$$

Con esta definición y tomando en cuenta la región factible, se tiene la siguiente

Afirmación 1.1. Sea $A \in \mathbf{M}_{m \times n}(\mathbb{R})$ con $\text{rango}(A) = m$. Entonces un punto extremo factible del conjunto $R = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ es un vértice (la intersección de n hiperplanos linealmente independientes). Las desigualdades denotan en realidad desigualdad entrada a entrada.

Demostración. Primero, observe que hay m hiperplanos y cada uno está dado por

$$\mathcal{H}_i = \{x \in \mathbb{R}^n : \vec{a}_i \cdot x = b_i\}$$

donde \vec{a}_i denota el i -ésimo vector renglón de A . Los hiperplanos son linealmente independientes pues el $\text{rango}(A) = m$. Así, $\bigcap_{i=1}^m \mathcal{H}_i = \{x \in \mathbb{R}^n : Ax = b\}$. Tome $r \in \bigcap_{i=1}^m \mathcal{H}_i$ y $r_1, r_2 \in R$ y $\lambda \in (0, 1)$.

Suponga que

$$\begin{aligned} r &= \lambda r_1 + (1 - \lambda)r_2 \\ \implies b &= Ax = A(\lambda r_1 + (1 - \lambda)r_2) \\ &= \lambda Ar_1 + (1 - \lambda)Ar_2 \\ &= \lambda b + (1 - \lambda)b \\ &< b \end{aligned}$$

Lo cual es claramente una contradicción. □

La afirmación anterior supone un conjunto de soluciones factibles acotado, pero no siempre se cuenta con una región acotada; además en principio puede resultar complicado probar que un cierto conjunto es no acotado. Para poder caracterizar a los poliedros acotados se requieren de las siguientes dos definiciones.

Definición 1.7. Sea $C \in \mathbb{R}^n$ un conjunto convexo. Una dirección de C es un vector $d \in \mathbb{R}^n, d \neq \vec{0}$, tal que $\forall c \in C, c + \mu d \in C, \forall \mu \geq 0$

Observe que para el caso del conjunto de soluciones factibles R , una dirección es un vector tal que $Ad \leq 0$. En efecto, d es una dirección si para

cada $x \in R$, ocurre que:

$$\begin{aligned} x + \mu d \in R \quad \forall \mu \geq 0 &\iff \\ A(x + \mu d) \leq b &\iff \\ Ax + \mu Ad \leq b &\iff \\ b + \mu Ad \leq b &\iff \\ Ad \leq 0 & \end{aligned}$$

Definición 1.8. Una dirección d de un conjunto convexo C es extrema si no puede ser expresada como combinación lineal positiva de dos direcciones distintas de C .

Ya se puede enunciar y demostrar el teorema de representación. Con este teorema no sólo se establece la relación que existe entre puntos extremos del conjunto de soluciones factibles y las soluciones factibles, sino también se detecta si dicho conjunto es acotado o no.

Teorema 1.1. Sea $R = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ un conjunto poliédrico no vacío. Las siguientes afirmaciones son válidas:

1. El conjunto de puntos extremos es no vacío y tiene un número finito de puntos $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$.
2. El conjunto de direcciones extremas es vacío si y sólo si R es acotado.
3. Si el conjunto R es no acotado entonces el conjunto de direcciones extremas es no vacío y tiene un número finito de vectores $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_\ell$.
4. $\bar{\mathbf{x}} \in R$ si y solo si $\bar{\mathbf{x}} \in R$ puede ser representado como una combinación lineal convexa de $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ más una combinación lineal positiva de $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_\ell$, esto es,

$$\begin{aligned} \bar{\mathbf{x}} &= \sum_{j=1}^k \lambda_j \mathbf{x}_j + \sum_{i=1}^{\ell} \mu_i \mathbf{d}_i \\ \sum_{j=1}^k \lambda_j &= 1 \\ \lambda_j &\geq 0 \quad j = 1, 2, \dots, k; \\ \mu_i &\geq 0 \quad j = 1, 2, \dots, \ell. \end{aligned} \tag{1.2}$$

Demostración. Se denotará con S_p y S_d a los conjuntos de puntos y direcciones extremas de R . Primero se probará que $1 \leq k < \infty$, donde $|S_p| = k$. Tome $\bar{\mathbf{x}} \in R$. Si $\bar{\mathbf{x}} \in S_p$, entonces $k \geq 1$. En otro caso, suponga que r es el número máximo de hiperplanos linealmente independientes que se intersectan en $\bar{\mathbf{x}}$ y suponga que $\bar{\mathbf{x}} = \lambda \mathbf{y}_1 + (1 - \lambda) \mathbf{y}_2$, con $\lambda \in (0, 1)$, $\mathbf{y}_1, \mathbf{y}_2 \in R$ y $\mathbf{y}_1 \neq \mathbf{y}_2$. Note que $0 \leq r < n$. Sea $\mathbf{d} = \mathbf{y}_2 - \mathbf{y}_1 \neq \mathbf{0}$ de modo que $\mathbf{y}_1 = \bar{\mathbf{x}} - (1 - \lambda) \mathbf{d}$ y $\mathbf{y}_2 = \bar{\mathbf{x}} + \lambda \mathbf{d}$. Si se desplaza desde $\bar{\mathbf{x}}$ en dirección \mathbf{d} o $-\mathbf{d}$, ambas direcciones permiten un movimiento dentro del conjunto, sin embargo al menos una no permite un movimiento infinito ya que $R \subset \{\mathbf{x} : \mathbf{x} \geq \mathbf{0}\}$. De este modo, se puede suponer, sin pérdida de generalidad, que $\bar{\gamma} = \max \{\gamma : \bar{\mathbf{x}} - \gamma \mathbf{d} \in R\} < \infty$ y tome $\bar{\mathbf{y}}_1 = \bar{\mathbf{x}} - \bar{\gamma} \mathbf{d}$. Note que el número máximo de hiperplanos que se intersectan en $\bar{\mathbf{y}}_1$ debe ser $\bar{r} \geq r + 1$ pues en $\bar{\mathbf{y}}_1$ se intersectan los mismos hiperplanos que en $\bar{\mathbf{x}}$ y al menos uno más que detenga el movimiento a lo largo de $-\mathbf{d}$. Si $\bar{r} = n$ entonces $\bar{\mathbf{y}}_1 \in S_p$ y por lo tanto $k \geq 1$. En otro caso se sustituye $\bar{\mathbf{x}}$ por $\bar{\mathbf{y}}_1$ y se repite el proceso hasta que $\bar{r} = n$; esto debe ocurrir en un número máximo de $n - r$ pasos, así se demuestra que $k \geq 1$. Ahora, como el número de formas en que se pueden seleccionar n hiperplanos de los $m + n$ totales, es finito, se tiene también que $k < \infty$.

De manera análoga, se puede hacer la demostración para el caso del conjunto de direcciones extremas.

Ahora, suponga que $\bar{\mathbf{x}}$ se puede escribir como en la expresión 1.2, en este caso no es difícil ver que $\bar{\mathbf{x}} \in R$. El problema es mostrar que si $\bar{\mathbf{x}} \in R$ entonces $\bar{\mathbf{x}}$ se puede escribir como en la expresión 1.2. Primero, sea

$$\bar{R} = R \cap \{\mathbf{x} : \mathbf{1x} \leq M\}$$

donde M es suficientemente grande de modo que $\mathbf{1x}_j \leq M$ para $j = 1, \dots, k$ y $\mathbf{1x} \leq M$. Con esto se logra acotar la región de soluciones factibles, $\bar{\mathbf{x}}$ sigue siendo factible y los puntos extremos de R son puntos extremos de \bar{R} . Sea $\bar{S}_p = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_{k+u}\}$ el conjunto de puntos extremos de \bar{R} , observe que $0 \leq u < \infty$. Se mostrará que $\bar{\mathbf{x}}$ es una combinación lineal convexa de los puntos en \bar{S}_p . Si $\bar{\mathbf{x}} \in \bar{S}_p$ no hay nada que demostrar. En otro caso, suponga que el sistema $\mathbf{Gx} = \mathbf{g}$ está determinado por los hiperplanos de \bar{R} que se intersectan en $\bar{\mathbf{x}}$, observe que $\text{rango}(\mathbf{G}) \leq n - 1$. Se puede encontrar una solución $\mathbf{d} \neq \mathbf{0}$ al sistema $\mathbf{Gd} = \mathbf{0}$, sea $\bar{\gamma}_1 = \max \{\gamma : \bar{\mathbf{x}} + \gamma \mathbf{d} \in \bar{R}\}$. Como \bar{R} es acotado $0 < \bar{\gamma}_1 < \infty$. Tome $\bar{\mathbf{y}}_1 = \bar{\mathbf{x}} + \bar{\gamma}_1 \mathbf{d}$. Lo importante es que $\bar{\mathbf{y}}_1 \in \bar{R}$ y además en $\bar{\mathbf{y}}_1$ se intersecta al menos un hiperplano más que en $\bar{\mathbf{x}}$. Si en $\bar{\mathbf{y}}_1$ se intersectan n hiperplanos $\bar{\mathbf{y}}_1$ es punto extremo de \bar{R} si no es así, se puede repetir el procedimiento aplicándolo a

$\bar{\mathbf{y}}_1$. En un número máximo de $n - \text{rango}(\mathbf{G})$ se tiene un punto extremo. Calcule ahora

$$\bar{\gamma}_2 = \max \{ \gamma : \bar{\mathbf{x}} + \gamma (\bar{\mathbf{x}} - \bar{\mathbf{y}}_1) \in \bar{R} \}$$

y haga

$$\bar{\mathbf{y}}_2 = \bar{\mathbf{x}} + \bar{\gamma}_2 (\bar{\mathbf{x}} - \bar{\mathbf{y}}_1)$$

Observe que $\bar{\gamma}_2 < \infty$ y además en $\bar{\mathbf{y}}_2$ se intersecta al menos un hiperplano más que en $\bar{\mathbf{x}}$. Lo que es más importante, $\bar{\mathbf{x}}$ es combinación lineal convexa de $\bar{\mathbf{y}}_1$ y $\bar{\mathbf{y}}_2$, $\bar{\mathbf{x}} = \delta \bar{\mathbf{y}}_1 + (1 - \delta) \bar{\mathbf{y}}_2$ con $\delta = \bar{\gamma}_2 / (1 + \bar{\gamma}_2)$. Si $\bar{\mathbf{y}}_2 \in \bar{S}_p$, entonces se ha expresado a $\bar{\mathbf{x}}$ como combinación lineal convexa de los puntos extremos de \bar{R} . En otro caso se puede repetir el procedimiento con $\bar{\mathbf{y}}_2$ y expresarlo como combinación lineal convexa estricta de un $\bar{\mathbf{y}}_3$ y $\bar{\mathbf{y}}_4$. Siguiendo este procedimiento se puede expresar a $\bar{\mathbf{x}}$ como combinación lineal convexa de los puntos extremos de \bar{R} .

$$\bar{\mathbf{x}} = \sum_{j=1}^{k+u} \delta_j \mathbf{x}_j \quad \text{con} \quad \sum_{j=1}^{k+u} \delta_j = 1, \quad \delta_j \geq 0 \text{ para } j = 1, \dots, k+u. \quad (1.3)$$

Si $\delta_j = 0$ para $j = k+1, \dots, k+u$, $\bar{\mathbf{x}}$ es combinación lineal convexa de los puntos extremos de \bar{R} . En otro caso consideren \mathbf{x}_ν tal que $\nu > k$ con $\delta_\nu > 0$. Note que \mathbf{x}_ν es un punto extremo de \bar{R} generado por el hiperplano $\mathbf{1}\mathbf{x} = M$. Los otros $n-1$ hiperplanos incidentes en \mathbf{x}_ν son hiperplanos "originales", es decir., existe un punto extremo $\mathbf{x}_{i(\nu)}$, $1 \leq i(\nu) \leq k$ adyacente a $\mathbf{x}_{i(\nu)}$. En tal caso, $\mathbf{x}_\nu - \mathbf{x}_{i(\nu)}$ es un dirección de R . De hecho, $\bar{\mathbf{d}}_i(\nu) = (\mathbf{x}_\nu - \mathbf{x}_{i(\nu)}) / \theta_\nu$ con $\theta_\nu = \mathbf{1}(\mathbf{x}_\nu - \mathbf{x}_{i(\nu)})$ es dirección extrema de R . De aquí se tiene que $\mathbf{x}_\nu = \mathbf{x}_{i(\nu)} + \theta_\nu \bar{\mathbf{d}}_i(\nu)$ sustituyendo esto en la expresión 1.3 se sigue que

$$\bar{\mathbf{x}} = \sum_{i=1}^k \delta_i \mathbf{x}_i + \sum_{\nu=k+1}^{k+u} \delta_\nu \mathbf{x}_{i(\nu)} + \sum_{\nu=k+1}^{k+u} \delta_\nu \theta_\nu \bar{\mathbf{d}}_i(\nu)$$

□

Ahora bien, cada tabla tiene m variables básicas y n no básicas. Sea \mathcal{B} el conjunto de índices correspondientes a las variables básicas y \mathcal{N} el conjunto de índices correspondientes a las variables no básicas. Inicialmente, salvo por un posible reacomodo en las columnas que se puede suponer sin perder la generalidad, se tiene $\mathcal{N} = \{1, 2, \dots, n\}$ y $\mathcal{B} = \{n+1, n+2, \dots, n+m\}$, pero esto cambia al final de la primera iteración.

Con cada iteración, exactamente una variable va de básica a no básica, y viceversa. La variable que va de no básica a básica se le llama *variable entrante*. Se elige de modo que la nueva solución siga siendo factible y se aproveche una dirección sobre la cual se mejora la función objetivo. Por esa razón, se toma alguna cuyo coeficiente \bar{c}_j sea positivo: *escoja una k del conjunto $\{j \in \mathcal{N} : \bar{c}_j > 0\}$* . Observe que se ha colocado una barra a los coeficientes de la función objetivo, pues con cada iteración éstos cambian. Si este conjunto es vacío, entonces la solución actual es la óptima. Si el conjunto contiene más de un elemento, se debe de elegir uno. Hay varios criterios de selección, pero por lo general, se escoge el índice k que tiene el mayor coeficiente y si hay más de una, se escoge la que tenga el menor subíndice. La razón para escoger la de menor subíndice quedará clara cuando se hable de tablas degeneradas y la regla de Bland. La variable que va de básica a no básica, se le llama *variable saliente*. Es escogida para preservar la no negatividad de las variables básicas actuales. Una vez que se decidió que la variable x_k debe ser la entrante, su valor deberá de ser incrementado de cero a un valor positivo. Este incremento cambia el valor de las variables básicas:

$$x_i = b_i - a_{ij}x_k, \quad i \in \mathcal{B}$$

Se debe asegurar que cada una de estas variables permanezca no negativa, esto es, se requiere que:

$$b_i - a_{ij}x_k \geq 0, \quad i \in \mathcal{B} \tag{1.4}$$

De estas expresiones, las únicas que pueden llegar a ser negativas mientras x_k aumenta son aquellas cuya a_{ik} es positiva; las demás permanecen igual o aumentan. Por esa razón, se puede restringir el análisis aquellos índices i para los cuales a_{ik} es positivo. Para ese índice i , el valor de x_k en la que la expresión se convierte en cero es:

$$x_k = \frac{b_i}{a_{ik}}.$$

Como se desea que ninguna se haga negativa, se debe aumentar x_k sólo el mínimo de estos valores:

$$x_k = \min_{\{i \in \mathcal{B} : a_{ik} > 0\}} \frac{b_i}{a_{ik}}$$

De aquí se desprende la regla para seleccionar la variable saliente: *escoja l de $\{i \in \mathcal{B} : a_{ik} > 0 \text{ y } \frac{b_i}{a_{ik}} \text{ es mínimo}\}$* . La regla que se acaba de enunciar describe exactamente el proceso que se sigue en la práctica, sin embargo, existe una

forma completamente equivalente de escribir esta regla. Para construir la expresión, se considera una excepción cuando $b_i = 0$ para poder reescribir la desigualdad 1.4 como:

$$\frac{1}{x_k} \geq \frac{a_{ik}}{b_i}, \quad i \in \mathcal{B}$$

Como se desea tener el máximo incremento en x_k , se tiene que

$$x_k = \left(\max_{i \in \mathcal{B}} \frac{a_{ik}}{b_i} \right)^{-1}.$$

De aquí que la regla para seleccionar la variable saliente, en éstos términos, es: *escoja l de* $\{i \in \mathcal{B} : \frac{a_{ik}}{b_i} \text{ es máximo}\}$. Una vez que la variable básica saliente y la variable no básica entrante han sido seleccionadas, la forma de pasar de una tabla a otra se da mediante operaciones elementales por renglones para lograr el cambio. La variable que se seleccionó con alguna de las reglas anteriores se le llama *pivote*. Todo lo anterior, funciona si $b_i > 0$ para toda $i = 1, 2, \dots, m$, pues en caso de que $b_i = 0$ para alguna i , al iniciar el método se tendría una solución factible degenerada lo cual podría llevar a que el método cicle; en caso de que $b_i < 0$ para alguna i , se tendría $x_k < 0$, lo cual caería en contradicción con la restricción de no negatividad y, por tanto, se tendría una solución inicial no factible. Si se tiene este último caso, se introduce un problema auxiliar con el cual, al resolverse, se obtiene una solución factible. El problema auxiliar es:

$$\begin{aligned} \max - \sum_{j=1}^n x_{aj} + \dots + x_{an} \\ \text{sujeto a : } \sum_{j=1}^n a_{ij}x_j + x_{aj} \leq b_i \quad i = 1, 2, \dots, m \\ x_j, x_{aj} \geq 0 \quad j = 0, 1, \dots, n. \end{aligned}$$

A las nuevas variables x_{ai} se les llama *variables artificiales*. Las variables x_{ai} estrictamente positivas, son aquellas cuya correspondiente b_i sea negativa, es decir, $x_{ai} > 0$ siempre que $b_i < 0$; las demás serán cero. Si el óptimo de la nueva función objetivo es $x_{ai} = 0 \quad \forall i$, se habrá encontrado una solución factible para el problema original. Si el punto óptimo es $x_{ai} \neq 0$ para alguna i , entonces el problema original no tiene ninguna solución factible. Al proceso de resolver este problema auxiliar y encontrar una solución factible para el problema original se le conoce como *Fase I*; y al proceso de resolver el

problema original, una vez que ya se encontró dicha solución, se le conoce como *Fase II*. Existe una alternativa para encontrar una solución factible inicial, conocido como el *método del la gran M*, que consiste en agregar a las variables artificiales con un coeficiente M común a todas ellas, donde M es un número positivo suficientemente grande. Así, el problema modificado queda como sigue

$$\begin{aligned} \max \quad & -\sum_{j=1}^n c_j x_j + M \sum_{j=1}^n x_{a_j} \\ \text{sujeto a:} \quad & \sum_{j=1}^n a_{ij} x_j + x_{a_j} \leq b_i \quad i = 1, 2, \dots, m \\ & x_j, x_{a_j} \geq 0 \quad j = 0, 1, \dots, n. \end{aligned}$$

De nuevo, sólo las variables artificiales cuya $b_i < 0$ serán distintas de cero y las demás serán cero. Puede pensarse a los términos $M \sum_{j=1}^n x_{a_j}$ como una penalización muy grande si $x_{a_i} \neq 0$ para alguna i ; de ahí que el método simplex se encargue de sacar de la base a las variables artificiales conforme va avanzando en la búsqueda del óptimo para el problema original. La pregunta natural ahora es: ¿qué tan grande debe ser M ? Ésta será respondida con detalle en el siguiente capítulo. Por ahora basta con decir que tal M siempre existe.

Ahora, si todos los cocientes $\frac{a_{ik}}{b_i}$ son no positivos, ninguna de las variables básicas se hará cero cuando la variable entrante incremente. Así, la variable entrante puede ser incrementada indefinidamente y producir un valor en la función objetivo arbitrariamente grande. En tal situación, se dice que el problema es *no acotado*.

Mientras se desarrolla el método Simplex puede ocurrir que se llegue a una tabla *degenerada*. Una tabla es degenerada si $b_i = 0$ para alguna $i \in \mathcal{B}$. Esta situación puede provocar problemas cuando se producen *pivotes degenerados*. Un pivote es degenerado, si alguno de los cocientes en el cálculo de la variable saliente es $+\infty$, es decir., si el numerador es positivo y el denominador es cero. Geométricamente quiere decir que el vértice elegido es la intersección de más de m caras del poliedro. Algebraicamente significa que existe un empate en los cocientes de las variables salientes.

Cuando se tiene esta situación, es usual que uno o más de los pivotes subsecuentes será degenerado pero que eventualmente un pivote no degenerado llevará a una tabla no degenerada, sin embargo, esta tabla puede que ya haya aparecido en algún paso anterior, en cuyo caso, el Simplex entra en un ciclo infinito. A este comportamiento se le llama *ciclo*. En la práctica, es común obtener tablas degeneradas pero es muy raro encontrarse con problemas que ciclen y se puede demostrar que para que un problema cicle es necesario que tenga por lo menos tres restricciones y seis variables, aunque

no necesariamente tenga solución óptima

Una de las técnicas usadas para resolver el problema del ciclo, es la *perturbación*. Obsérvese que se obtiene una tabla degenerada cuando $b_i = 0$ para alguna $i \in \mathcal{B}$. Perturbar quiere decir, en este contexto, introducir m variables (una por cada restricción) cuyos valores sean positivos pero lo suficientemente pequeños como para que el problema original no se altere. Estas variables se escriben como:

$$0 \ll \epsilon_m \ll \epsilon_{m-1} \ll \dots \ll \epsilon_1 \ll \text{cualquier otro dato.}$$

Una vez introducidas las variables ϵ , se procede con el método Simplex y cuando se llegue a la solución factible óptima, simplemente se eliminan las variables ϵ . Es importante hacer notar, que esta variante no afecta la selección de la variable entrante, pero si lo hace para determinar la variable saliente.

La idea es que cada ϵ_i actúe en una escala completamente diferente de todas las demás y de los datos del problema, es decir, que ninguna combinación lineal de las variables ϵ_i que surja al paso del Simplex alcance al orden de los datos del problema y así no ocurran cancelaciones. Es claro que este tipo de números no pueden existir en los reales, así que se trata a las variables ϵ_i como símbolos abstractos con las propiedades ya descritas. Geométricamente lo que ocurre es que cada vértice del poliedro se “parte” en dos muy cercanos entre sí. Para ilustrar el concepto, se verá a continuación un ejemplo. Considere la siguiente tabla degenerada:

$$\begin{array}{rcl} Z & = & 4 + 2x_1 - x_2 \\ \hline w_1 & = & 0.5 \qquad \qquad - x_2 \\ w_2 & = & \qquad - 2x_1 + 4x_2 \\ w_3 & = & \qquad \qquad x_1 - 3x_2 \end{array}$$

El primer paso es introducir los parámetros simbólicos

$$0 < \epsilon_3 \ll \epsilon_2 \ll \epsilon_1$$

para obtener el problema perturbado:

$$\begin{array}{rcl} Z & = & 4 \qquad \qquad \qquad + 2x_1 - x_2 \\ \hline w_1 & = & 0.5 + \epsilon_1 \qquad \qquad - x_2 \\ w_2 & = & \qquad \qquad \epsilon_2 \qquad - 2x_1 + 4x_2 \\ w_3 & = & \qquad \qquad \epsilon_3 + x_1 - 3x_2 \end{array}$$

Esta tabla no es degenerada. La variable entrante es x_1 y la variable saliente es sin duda w_2 . La siguiente tabla es

$$\begin{array}{r} Z = 4 + \epsilon_2 - w_2 + 3x_2 \\ \hline w_1 = 0.5 + \epsilon_1 - x_2 \\ x_1 = 0.5\epsilon_2 - 0.5w_2 + 2x_2 \\ w_3 = 0.5\epsilon_2 + \epsilon_3 - 0.5w_2 - x_2 \end{array}$$

Para el siguiente pivote, la variable entrante es x_2 y la variable saliente es w_3 . La nueva tabla es

$$\begin{array}{r} Z = 4 + 2.5\epsilon_2 + 3\epsilon_3 - 2.5w_2 - 3w_3 \\ \hline w_1 = 0.5 + \epsilon_1 - 0.5\epsilon_2 - \epsilon_3 + 0.5w_2 + w_3 \\ x_1 = 1.5\epsilon_2 + 2\epsilon_3 + 1.5w_2 - 2w_3 \\ x_2 = 0.5\epsilon_2 + \epsilon_3 - 0.5w_2 - w_3 \end{array}$$

Esta última tabla es la óptima. En este punto, simplemente se desechan las variables ϵ y se obtiene la solución óptima para el problema original:

$$\begin{array}{r} Z = 4 - 2.5w_2 - 3w_3 \\ \hline w_1 = 0.5 + 0.5w_2 + w_3 \\ x_1 = -1.5w_2 - 2w_3 \\ x_2 = -0.5w_2 - w_3 \end{array}$$

Otra técnica usada para resolver el problema de la degeneración es la *regla de Bland*, que consiste en escoger tanto la variable entrante como la saliente de sus respectivos conjuntos sean las de índice menor.

1.3. El Dual

Asociado a todo problema de programación lineal, existe otro llamado **dual**. Se demostrará que el dual del dual, es el problema original; por esta razón, al problema original es a veces llamado el **primal**. Así, los problemas vienen siempre en parejas primal/dual. Resulta que cada solución factible para uno de estos problemas da una cota superior en el valor óptimo del otro problema. A continuación se verá, con un ejemplo, la construcción del problema dual y el concepto se formalizará después.

Considere el siguiente problema:

$$\max \quad 4x_1 + x_2 + 3x_3$$

sujeto a:

$$\begin{aligned}x_1 + 4x_2 &\leq 1 \\3x_1 - x_2 + x_3 &\leq 3 \\x_1, x_2, x_3 &\geq 0\end{aligned}$$

La primera observación es que toda solución factible da una cota inferior en el valor óptimo del problema primal. En efecto, si x^* es el punto óptimo de un problema de programación lineal implica que para cualquier otro punto factible se tenga que $c^t \cdot x \leq c^t \cdot x^*$. Por ejemplo, la solución $(x_1, x_2, x_3) = (1, 0, 0)$ significa que $Z_{opt} \geq 4$ y $(x_1, x_2, x_3) = (0, 0, 3)$ significa que $Z_{opt} \geq 9$. Se averiguará qué tan acertadas son estas cotas de la siguiente manera: multiplique la primera restricción por dos y sume tres veces la segunda

$$\begin{array}{r}2 \quad (x_1 + 4x_2) \leq 2(1) \\+3 \quad (3x_1 - x_2 + x_3) \leq 3(3) \\ \hline 11x_1 + 5x_2 + 3x_3 \leq 11.\end{array}$$

Como cada variable es no negativa, se puede comparar la suma con la función objetivo y notar que

$$4x_1 + x_2 + 3x_3 \leq 11x_1 + 5x_2 + 3x_3 \leq 11$$

Hasta ahora, se sabe que $9 \leq Z_{opt} \leq 11$. Estas cotas definen un intervalo donde se encuentra el valor óptimo, sin embargo, se puede mejorar la aproximación. Para ello, se puede usar la misma técnica, pero reemplazando los números específicos que se usaron por variables y luego tratar de encontrar los valores de esas variables que den una mejor cota superior. Así, se empieza por multiplicar las dos restricciones por variables no negativas y_1, y_2 . El hecho de que sean no negativas implica que preservan las desigualdades:

$$\begin{array}{r}y_1 \quad (x_1 + 4x_2) \leq 2y_1 \\+y_2 \quad (3x_1 - x_2 + x_3) \leq 3y_2 \\ \hline (y_1 + 3y_2)x_1 + (4y_1 - y_2)x_2 + y_2x_3 \leq y_1 + 3y_2.\end{array}$$

Si se estipula que cada coeficiente de las x_i 's sea por lo menos del mismo orden que el coeficiente correspondiente de la función objetivo, se tiene

$$\begin{aligned}y_1 + 3y_2 &\geq 4 \\4y_1 - y_2 &\geq 1 \\y_2 &\geq 3\end{aligned}$$

Ahora se puede comparar esta suma con la función objetivo

$$\begin{aligned} Z &= 4x_1 + x_2 + 3x_3 \\ &\leq (y_1 + 3y_2)x_1 + (4y_1 - y_2)x_2 + y_2x_3 \\ &\leq y_1 + 3y_2. \end{aligned}$$

Ahora se tiene una cota superior, $y_1 + 3y_2$, que se debe minimizar para obtener la mejor cota superior. Por esa razón, se tiene el siguiente problema de optimización:

$$\begin{aligned} \min \quad & y_1 + 3y_2 \\ \text{sujeto a:} \quad & y_1 + 3y_2 \geq 4 \\ & 4y_1 - y_2 \geq 1 \\ & y_2 \geq 3 \\ & y_1, y_2 \geq 0 \end{aligned}$$

Este es el problema de programación lineal dual asociado con el problema original. A continuación, se define el dual de forma general.

Definición 1.9. Dado un problema de programación lineal en su forma estándar:

$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j x_j \\ \text{sujeto a:} \quad & \sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = 1, 2, \dots, m \\ & x_j \geq 0 \quad j = 1, 2, \dots, n \end{aligned} \quad (1.5)$$

el problema *dual asociado* está dado por:

$$\begin{aligned} \min \quad & \sum_{i=1}^m b_i y_i \\ \text{sujeto a:} \quad & \sum_{i=1}^m y_i a_{ij} \geq c_j \quad j = 1, 2, \dots, n \\ & y_i \geq 0 \quad i = 1, 2, \dots, m \end{aligned}$$

Al problema 1.5 se le llama *primal*. Como el objetivo de la presente no es hablar del Simplex como algoritmo, sino de la eficiencia del mismo, sólo se mencionarán los resultados más relevantes para la teoría del Dual. Para mayor referencia respecto a la teoría del dual para Programación Lineal puede consultar [7] Bazaraa, S. Mokhtar et. all o [8] Vanderbei, J. Robert. Lo primero que se hará, es demostrar que el dual del dual es de nuevo el problema original, Para ello, primero se debe escribir el problema dual en forma estándar, es decir., cambiar el problema de minimizar a maximizar y cambiar las desigualdades de mayor o igual a menor o igual. Para cambiar el problema de minimizar a maximizar, recuerde:

$$\min f = - \max -f$$

Aplicando esto a la función objetivo del dual, se tiene que

$$\min \sum_{i=1}^m b_i y_i = - \max \left(- \sum_{i=1}^m b_i y_i \right)$$

Para cambiar la dirección de las desigualdades, simplemente se multiplica por -1 . El resultado de estos dos pasos es

$$\begin{aligned} & - \max && \sum_{i=1}^m (-b_i) y_i \\ \text{sujeto a:} & \sum_{i=1}^m (-a_{ij}) y_i \leq (-c_j) && j = 1, 2, \dots, n \\ & y_i \geq 0 && i = 1, 2, \dots, m \end{aligned}$$

Ahora se toma el dual de este problema:

$$\begin{aligned} & - \min && \sum_{j=1}^n (-c_j) x_j \\ \text{sujeto a:} & \sum_{j=1}^n (-a_{ij}) x_j \geq (-b_i) && i = 1, 2, \dots, m \\ & x_j \geq 0 && j = 1, 2, \dots, n \end{aligned}$$

Esto último es claramente equivalente al problema primal.

Hay dos propiedades fundamentales respecto al dual: la *propiedad débil* y la *propiedad fuerte* del dual. Estos dos conceptos se pueden generalizar a problemas de programación convexa más generales. Aquí se enuncian como teoremas y su demostración puede consultarse en [8].

Teorema 1.2. Sean \bar{x} y \bar{y} dos soluciones factibles, la primera para el problema primal y la segunda para el dual. Entonces, se tiene que

$$\sum_j c_j x_j \leq \sum_i b_i y_i$$

Teorema 1.3. Suponga que el problema primal admite una solución óptima x^* . Entonces, el problema dual también tiene una solución óptima y^* tal que

$$\sum_j c_j x^* = \sum_i b_i y^*$$

Una observación importante respecto al dual y al primal, es la siguiente: si se escriben en forma matricial ambos programas, se observa que la matriz del dual es la matriz transpuesta negativa del problema primal. Para ilustrar

este hecho, recuerde el ejemplo que se dió al principio de esta sección:

$$\begin{array}{rcl}
 Z & = & 4x_1 + x_2 + 3x_3 \\
 \hline
 w_1 & = & 1 - x_1 - 4x_2 \\
 w_2 & = & 3 - 3x_1 + x_2 - x_3 \\
 \\
 -Z' & = & -y_1 - 3y_2 \\
 \hline
 z_1 & = & -4 + y_1 + 3y_2 \\
 z_2 & = & -1 + 4y_1 - y_2 \\
 z_3 & = & -3 + y_2
 \end{array}$$

El primer problema es el primal y el segundo el dual del mismo. Ahora, reescribiendo todo en matrices, se tiene:

$$\begin{bmatrix} 0 & 4 & 1 & 3 \\ 1 & -1 & -4 & 0 \\ 3 & -3 & 1 & -1 \end{bmatrix} \xleftrightarrow{\text{neg. trans.}} \begin{bmatrix} 0 & -1 & -3 \\ -4 & 1 & 3 \\ -1 & 4 & -1 \\ -3 & 0 & 1 \end{bmatrix}$$

Para la construcción en general de este hecho, considere la siguiente forma estándar de un un problema de programación lineal:

$$\begin{array}{l}
 \max c^t x \\
 \text{sujeto a: } Ax \leq b \\
 x \geq 0
 \end{array}$$

y su correspondiente dual

$$\begin{array}{l}
 \min b^t y \\
 \text{sujeto a: } A^t y \geq c \\
 y \geq 0
 \end{array}$$

Sean w un vector con las variables de holgura para el problema primal y z un vector con las variables de holgura para el problema dual. Reescribiendo ambos problemas, se tiene:

$$\begin{array}{l}
 \max c^t x \\
 \text{sujeto a: } Ax + w = b \\
 x, w \geq 0 \\
 \min b^t y \\
 \text{sujeto a: } A^t y - z = c \\
 y, z \geq 0
 \end{array}$$

Con la siguiente notación

$$\bar{A} = [A, I], \quad \bar{c} = \begin{bmatrix} c \\ 0 \end{bmatrix} \quad y \quad \bar{x} = \begin{bmatrix} x \\ w \end{bmatrix},$$

el problema primal puede ser escrito como

$$\begin{aligned} & \max \bar{c}^t \bar{x} \\ & \text{sujeto a: } \bar{A} \bar{x} = b \\ & \bar{x} \geq 0 \end{aligned}$$

Análogamente para el problema dual:

$$\hat{A} = [-I, A^t], \quad \hat{b} = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad y \quad \hat{y} = \begin{bmatrix} z \\ y \end{bmatrix}$$

se puede reescribir como

$$\begin{aligned} & \min \hat{b}^t \hat{y} \\ & \text{sujeto a: } \hat{A} \hat{y} = c \\ & \hat{y} \geq 0 \end{aligned}$$

Observe que $\bar{A} \in \mathbf{M}_{m \times (n+m)}(\mathbb{R})$. Las primeras n columnas son las variables no básicas iniciales y las últimas m columnas son las variables básicas iniciales. Después de algunas iteraciones, las columnas básicas y no básicas quedan revueltas, pero aún se puede escribir la igualdad

$$[A, I] = [\bar{N}, \bar{B}]$$

con la convención de que la igualdad se preserva sólo después de reorganizar las columnas de forma adecuada. De la parte dual, la matriz $\hat{A} = [-I, A^t] \in \mathbf{M}_{m \times (n+m)}(\mathbb{R})$. Las primeras n columnas son las variables iniciales básicas (para el problema dual) y las últimas m columnas representan las variables no básicas. Si los pivotes que fueron aplicados para el problema primal se aplican al dual, entonces las columnas del problema dual quedan exactamente igual a las del problema primal y se puede escribir la igualdad

$$[-I, A^t] = [\hat{B}, \hat{N}]$$

nuevamente con la convención de que la igualdad se preserva sólo después de reorganizar las columnas de forma adecuada.

La tabla dual involucra a la matriz $\bar{B}^{-1}\bar{N}$ mientras que la tabla dual involucra a la matriz $\hat{B}^{-1}\hat{N}$. No es evidente que estas dos matrices sean la transpuesta negativa mutuas. Para exhibirlo, considere lo que ocurre cuando se multiplica \bar{A} por \hat{A}^t tanto en la notación permutada como en la normal:

$$\bar{A}\hat{A}^t = [\bar{N}, \bar{B}] \begin{bmatrix} \hat{B}^t \\ \hat{N}^t \end{bmatrix} = \bar{N}\hat{B}^t + \bar{B}\hat{N}^t$$

y

$$\bar{A}\hat{A}^t = [A, I] \begin{bmatrix} -I \\ A \end{bmatrix} = -A + A = 0$$

Claramente, estas últimas dos expresiones deben coincidir, con lo cual se tiene que

$$\bar{N}\hat{B}^t + \bar{B}\hat{N}^t = 0$$

Despejando y multiplicando por la derecha por la inversa de \hat{B}^t y por la izquierda por la inversa de \bar{B} , se tiene

$$\bar{B}^{-1}\bar{N} = -\left(\hat{B}^{-1}\hat{N}\right)^t$$

que es la propiedad que se deseaba exhibir.

Otra observación importante, es que esta propiedad se conserva a través de todas las tablas conforme el método avanza. Para verlo, se analizará a continuación una sola iteración. Para mantener una notación simple, sólo se tomarán en cuenta cuatro entradas genéricas de la matriz de coeficientes: el elemento pivote a , otro elemento en la fila del elemento pivote b , otro elemento en su columna c , y un cuarto elemento para hacer un arreglo rectangular d . La iteración provoca los siguientes cambios:

- el elemento pivote es reemplazado por su recíproco.
- los elementos en la fila del pivote cambian por su inverso aditivo multiplicado por el recíproco del pivote
- los elementos en la columna del pivote son multiplicados por el recíproco del pivote
- a todos los demás elementos se les suma $-\frac{bc}{a}$

Estos cambios se pueden visualizar mejor en la siguiente tabla:

$$\begin{array}{|c|c|} \hline b & a \\ \hline d & c \\ \hline \end{array} \xrightarrow{\text{Pivote}} \begin{array}{|c|c|} \hline -\frac{b}{a} & \frac{1}{a} \\ \hline d - \frac{bc}{a} & \frac{c}{a} \\ \hline \end{array}$$

Si ahora se hace lo mismo con el dual, se tiene:

$$\begin{array}{|c|c|} \hline -b & -d \\ \hline -a & -c \\ \hline \end{array} \xrightarrow{\text{Pivote}} \begin{array}{|c|c|} \hline \frac{b}{a} & -d + \frac{bc}{a} \\ \hline -\frac{1}{a} & -\frac{c}{a} \\ \hline \end{array}$$

Obsérvese que la tabla dual resultante es la transpuesta negativa de la tabla primal. Como el procedimiento de pivotar es el mismo en cualquier iteración del algoritmo, se sigue que la propiedad se mantiene.

1.4. La eficiencia del método Simplex

Hasta este momento, se ha expuesto como funciona el método Simplex y cuales son sus variantes para poder llegar, en caso de que exista, a la solución factible óptima; sin embargo, poco se ha hablado de su eficiencia algorítmica. Antes de empezar con definiciones formales y teoremas que demuestran la eficiencia algorítmica del Simplex, se verá a continuación un ejemplo que fue publicado por V. Klee y G.J. Minty donde se muestra claramente que el Simplex, en el peor de los escenarios, puede ser indeseablemente ineficiente. Por supuesto, no es el único ejemplo de esta clase, pero es de los mas elegantes por su sencillez.

$$\begin{array}{rcl}
 \max & 2^{n-1}x_1 + 2^{n-2}x_2 + \dots + 2x_{n-1} + x_n & \\
 \text{sujeto a:} & & \\
 & x_1 & \leq 5 \\
 & 4x_1 + x_2 & \leq 25 \\
 & 8x_1 + 4x_2 + x_3 & \leq 125 \\
 & \vdots & \vdots \\
 & 2^n x_1 + 2^{n-1}x_2 + \dots + 4x_{n-1} + x_n & \leq 5^n
 \end{array} \quad (1.6)$$

$$x_i \geq 0 \quad i = 1, 2, \dots, n$$

Este problema tiene n variables, $2n$ restricciones y 2^n vértices. Este simple problema, al ser resuelto con el método Simplex empezando en el origen, pasa por cada uno de los vértices del poliedro antes de encontrar la solución factible óptima en el punto $(0, 0, \dots, 0, 5^n)$. Esto confirma que el Simplex no resuelve los problemas en tiempo polinomial. En los siguientes párrafos, se demuestra que lo hace en tiempo exponencial.

A continuación, se da una definición que resultará de utilidad para futuras explicaciones de la eficiencia del Simplex.

Definición 1.10. Dos vértices son *vecinos* si tienen $n - 1$ hiperplanos en común.

Según la descripción del método Simplex dada en la sección anterior, con cada iteración el Simplex tiene dos tareas, a saber:

1. Revisar si el vértice actual es óptimo (y de serlo detenerse)
2. Determinar a qué vértice vecino moverse.

Cuando el Simplex comienza, regularmente lo hace desde el origen, y desde ahí, es fácil revisar si es el vértice óptimo y también resulta relativamente fácil determinar a qué vértice vecino se debe mover, pero rara vez el origen es óptimo. En tal caso, el Simplex se mueve a un vecino que incrementa el valor de la función objetivo y convierte a este nuevo vértice un nuevo origen para comenzar de nuevo. En este punto de la discusión ya se puede empezar a intuir que, si el número de vértices es muy grande, el Simplex puede comportarse de forma poco eficiente. Más aún, el número de vértices no es la única razón por la cual el Simplex se pueda llegar a comportar de forma poco eficiente, también está el problema del ciclo. Sin embargo, como se demostrará más adelante, con el método de la perturbación y con la regla de Bland, el Simplex no cicla. Eso significa que el Simplex terminará necesariamente en algún momento ya sea encontrando una solución factible óptima o reportando que no existe tal solución.

El siguiente teorema dice que, sin tomar en cuenta por el momento la regla de Bland y el método de la perturbación, si el Simplex falla, debe entrar en un ciclo. Recuerde ahora que en un problema de programación lineal escrito como el programa 1.1 se tienen n variables de decisión y m restricciones.

Teorema 1.4. Si el método Simplex falla en terminar, entonces debe entrar en un ciclo.

Demostración. Una tabla queda completamente determinada al especificar cuales son las variables básicas y cuales son las no básicas. Hay un total de

$$\binom{n+m}{m}$$

posibilidades distintas. Si el Simplex falla en terminar, debe repetir alguna de estas tablas más de una vez y, por lo tanto, cicla. \square

Los siguientes dos teoremas garantizan que el Simplex siempre terminará.

Teorema 1.5. Con el método de la perturbación, el Simplex siempre terminará.

Demostración. Bastará con demostrar que nunca se produce una tabla degenerada. Como se mencionó en la sección anterior, las variables ϵ_i pueden cancelarse, por esa razón, pueden ser pensadas como variables independientes. Extrayendo a las ϵ_i de la primera tabla, se tiene el siguiente patrón:

$$\begin{array}{cccc} \epsilon_1 & & & \\ & \epsilon_2 & & \\ & & \ddots & \\ & & & \epsilon_m. \end{array}$$

Después de varios pivotes, los términos ϵ formarán un sistema de ecuaciones:

$$\begin{array}{ccccccc} r_{11}\epsilon_1 & + & r_{12}\epsilon_2 & \dots & + & r_{1m}\epsilon_m & \\ r_{21}\epsilon_1 & + & r_{22}\epsilon_2 & \dots & + & r_{2m}\epsilon_m & \\ \vdots & & \vdots & \ddots & & \vdots & \\ r_{m1}\epsilon_1 & + & r_{m2}\epsilon_2 & \dots & + & r_{mm}\epsilon_m. & \end{array}$$

Como este sistema de ecuaciones lineales fue obtenido del sistema original mediante operaciones de “pivoteo” y éstas son reversibles, entonces el rango de ambos sistemas debe ser el mismo. Como el sistema original tenía rango m , cada sistema subsecuente debe tener rango m . Entonces, para cada renglón i , existe $j = 1, 2, \dots, m$ tal que $r_{ij} \neq 0$. Esto implica que ninguno de los renglones puede ser degenerado. Por lo tanto, no hay tablas degeneradas. \square

Antes de enunciar y demostrar el segundo teorema, conviene tener en cuenta la siguiente:

Definición 1.11. Se dice que la variable x_k es *leal* si está en una base y no está en alguna otra.

Teorema 1.6. Si se sigue la regla de Bland para escoger los pivotes, el Simplex siempre terminará.

Demostración. Bastará con demostrar que esta variante no produce ciclos. La prueba se hará por contradicción. Suponga que con esta variante, el Simplex cae en un ciclo. Sin pérdida de generalidad, suponga que el Simplex cicla desde el principio. Sean T_0, T_1, \dots, T_{k-1} las tablas donde el Simplex cicla, es decir, el Simplex produce la siguiente secuencia:

$$T_0, T_1, \dots, T_{k-1}, T_0, T_1, \dots$$

Sea x_t la variable leal con el mayor subíndice y sea T la tabla de T_0, \dots, T_{k-1} de donde x_t deja la base. De nuevo, sin pérdida de generalidad, suponga que $T = T_0$. Sea x_s la correspondiente variable saliente. Suponga que T es como sigue:

$$Z = v + \sum_{j \in \mathcal{N}} c_j x_j$$

$$x_i = b_i - \sum_{j \in \mathcal{N}} a_{ij} x_j \quad i \in \mathcal{B}.$$

Como x_s es la variable saliente y x_t es la variable entrante, se tiene que $s \in \mathcal{N}$ y $t \in \mathcal{B}$. Sea T^* una tabla en T_0, T_1, \dots, T_{k-1} en la que la variable x_t entra a la base. Suponga que T^* tiene la forma:

$$Z = v^* + \sum_{j \in \mathcal{N}^*} c_j^* x_j \tag{1.7}$$

$$x_i = b_i^* - \sum_{j \in \mathcal{N}^*} a_{ij}^* x_j \quad i \in \mathcal{B}^*.$$

Como todas las tablas son degeneradas, se tiene que $v^* = v$, y por ello se puede escribir la función objetivo del programa 1.7 como

$$Z = v + \sum_{j=1}^{n+m} c_j^* x_j, \tag{1.8}$$

donde se ha extendido la notación c_j^* a todas las variables (tanto las originales como las de holgura) fijando las $c_j^* = 0$ para $j \in \mathcal{B}^*$. Ignorando por el momento la posibilidad de que algunas variables pueden ser negativas, considérense

las soluciones obtenidas al incrementar el valor de x_s mientras se mantienen en cero las demás variables en \mathcal{N} :

$$\begin{aligned} x_s &= y, \\ x_j &= 0, & j \in \mathcal{N} \setminus \{s\}, \\ x_i &= b_i - a_{is}y, & i \in \mathcal{B}. \end{aligned}$$

En este paso, la función objetivo está dada por

$$Z = v + c_s y.$$

Sin embargo, usando la ecuación 1.8 también se puede escribir como:

$$Z = v + c_s^* y + \sum_{i \in \mathcal{B}} c_i^* (b_i - a_{is} y).$$

Igualando estas dos expresiones de Z , se obtiene

$$(c_s - c_s^* + \sum_{i \in \mathcal{B}} c_i^* a_{is}) y = \sum_{i \in \mathcal{B}} c_i^* b_i.$$

Como esta expresión debe ser una igualdad para toda y , se sigue que el coeficiente de y del lado izquierdo de la igualdad debe hacerse cero (así como también el lado derecho de la igualdad):

$$c_s - c_s^* + \sum_{i \in \mathcal{B}} c_i^* a_{is} = 0$$

Ahora, el hecho de que x_s sea la variable entrante en T implica que

$$c_s > 0.$$

Recuerde que x_t es la variable leal con el mayor subíndice. Como x_s también es leal, se tiene que $s < t$. Como x_s no es la variable entrante en T^* (pues lo es x_t), se ve que

$$c_s^* \leq 0$$

De estas tres últimas ecuaciones, se tiene que

$$\sum_{i \in \mathcal{B}} c_i^* a_{is} < 0.$$

Entonces debe existir un índice $r \in \mathcal{B}$ para el cual

$$c_r^* a_{rs} < 0. \tag{1.9}$$

Consecuentemente, $c_r^* \neq 0$ y $r \in \mathcal{N}^*$. Entonces, x_r es leal y por eso $r \leq t$. De hecho, $r < t$, pues $c_t^* a_{ts} > 0$. Para mostrar que este producto es positivo, note que ambos factores lo son: c_t^* es positivo, pues x_t es la variable entrante en T^* ; y a_{st} es positivo, pues x_t es la variable saliente en T . El hecho de que $r < t$ implica que $c_r^* \leq 0$. De lo contrario, de acuerdo con el criterio de tomar el menor subíndice, x_r sería la variable entrante para T^* . Entonces 1.9 implica que

$$a_{rs} > 0$$

Como cada tabla del ciclo da la misma solución, se sigue que cada variable leal en estas tablas es cero. Claramente es cero en la tabla en la cual es no básica. En particular, $x_r = 0$. Pero en T , x_r es básica. Por eso,

$$b_r = 0$$

Las últimas dos desigualdades implican que x_r era candidato para ser la variable saliente en T , y como $r < t$, debió de ser escogida en lugar de x_t . Esta es la contradicción que se buscaba. □

Gracias a los dos teoremas anteriores, se puede estar seguro que al resolver un problema con el método Simplex el proceso es finito, en tanto que no se caiga en ciclos. La pregunta a responder ahora es, en el peor de los escenarios, cuánto tiempo se requiere para resolver un problema. Como el Simplex opera moviéndose de una solución factible a otra sin regresar a alguna previamente visitada, una cota superior para el número de iteraciones requeridas es:

$$\binom{n+m}{m}.$$

Para valores fijos de la suma $n+m$, la expresión anterior se maximiza cuando $n = m$. Esto da lugar al siguiente:

Teorema 1.7. $\forall n \in \mathbb{N}$, se tiene que:

$$\frac{1}{2n} 2^{2n} \leq \binom{2n}{n} \leq 2^{2n} \tag{1.10}$$

Demostración. Sea $n \in \mathbb{N}$. Primero se demostrará la siguiente desigualdad:

$$\binom{2n}{n} \leq 2^{2n}$$

Para probarla, recuerde el teorema del binomio:

$$(a + b)^n = \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i \quad (1.11)$$

Tomando $a = b = 1$ y el exponente como $2n$ en la Ecuación 1.11, se tiene que

$$2^{2n} = \sum_{i=0}^{2n} \binom{2n}{i} \quad (1.12)$$

Observe ahora que, $\forall n \in \mathbb{N}$, se tiene que

$$\binom{2n}{n} \leq \sum_{i=0}^{2n} \binom{2n}{i}$$

De esta última desigualdad y de la Ecuación 1.12, se sigue que

$$\binom{2n}{n} \leq \sum_{i=0}^{2n} \binom{2n}{i} = 2^{2n}$$

Lo que demuestra la desigualdad derecha de la relación 1.10. La parte izquierda de la desigualdad se demostrará por inducción. Para $n = 1$:

$$2 = \frac{1}{2(1)} 2^{2(1)} \leq \binom{2}{1} = \frac{2!}{1!1!} = 2$$

Suponga que la desigualdad es válida cuando $n = k - 1$, es decir

$$\frac{1}{2(k-1)} 2^{2(k-1)} \leq \binom{2(k-1)}{k-1} \quad (1.13)$$

A continuación, se probará para $n = k > 1$. De la Desigualdad 1.13, se tiene que:

$$\begin{aligned} \frac{1}{2(k-1)} 2^{2(k-1)} &\leq \frac{(2k-2)!}{(k-1)!(k-1)!} \iff \\ \left(\frac{1}{2(k-1)} 2^{2(k-1)} \right) \left(\frac{(2k-1)2k}{k \cdot k} \right) &\leq \left(\frac{(2k-2)!}{(k-1)!(k-1)!} \right) \left(\frac{(2k-1)2k}{k \cdot k} \right) \end{aligned}$$

pues es claro que $\frac{(2k-1)2k}{k \cdot k} > 0$, $\forall k \in \mathbb{N}$.

Simplificando ambos miembros de la desigualdad y reagrupando términos de lado izquierdo, se tiene:

$$\frac{2k-1}{2k-2} 2^{2k-2} \frac{2k}{k \cdot k} \leq \frac{2k!}{k!k!} = \binom{2k}{k}$$

Ahora, observe que:

$$\frac{2k-1}{2k-2} 2^{2k-2} \frac{2}{k} \geq \frac{2(k-1)}{2(k-1)} 2^{2k-2} \frac{2}{k} = \frac{2^{2k}}{2k} = \frac{1}{2k} 2^{2k}$$

Lo cual, completa la prueba. \square

El teorema anterior garantiza que, en el peor de los casos, se requerirán 2^{2n} iteraciones para poder encontrar la solución factible óptima. Ahora, como el Simplex es un método iterativo que va de vértice en vértice, se verá a continuación el tiempo requerido para una sola iteración. Suponga que se está en el vértice \mathbf{u} . Por definición, es el único punto donde n restricciones se cumplen en igualdad. Cada uno de sus vecinos comparte $n-1$ de estas restricciones, así que \mathbf{u} tiene a lo más nm vecinos; así, se debe escoger qué desigualdad desechar y cuál conservar.

Una forma de realizar una iteración, sería revisar cada posible vecino para confirmar si es en verdad un vértice del poliedro y determinar su valor en la función objetivo. La primera tarea es sencilla (es sólo un producto punto) pero la parte de revisar si en verdad es un vértice requiere resolver un sistema de n ecuaciones con n incógnitas, esto es, que se satisfagan exactamente las n desigualdades, y revisar si la solución es factible. Por eliminación Gaussiana esto toma $O(n^3)$ de tiempo, dando un ineficiente tiempo de $O(mn^4)$ por iteración.

El factor mn^4 puede ser mejorado a mn , haciendo del Simplex un algoritmo práctico. Resulta que el costo en tiempo de reescribir el problema en términos de coordenadas locales es de sólo $O((m+n)n)$; esto toma ventaja de que la visión local del punto cambia poco entre las iteraciones, en sólo una las desigualdades que lo definen.

Para seleccionar al mejor vecino, tome primero en cuenta que la visión local en el punto \mathbf{u} de la función objetivo es de la forma

$$\max c_{\mathbf{u}} + \tilde{c}y$$

donde $c_{\mathbf{u}}$ es el valor de la función objetivo en \mathbf{u} . Esto inmediatamente identifica una dirección a la cual moverse: se escoge cualquier $\tilde{c}_i > 0$ (si no hay alguna, entonces el vértice actual es el óptimo). En la sección anterior, se dieron algunos de los criterios más comunes para escoger las variables salientes y entrantes. Como el resto del problema ha sido reescrito en términos de las nuevas coordenadas, es fácil ver que tanto se debe aumentar en la variable correspondiente antes de que alguna restricción sea violada (y si se puede aumentar indefinidamente, se sabe que el problema es no acotado).

Se sigue que el tiempo por iteración del Simplex es de tan sólo $O(mn)$. El número de iteraciones que hay en esta expresión no sobrepasa las 2^{2n} , por el Teorema 1.7. Pero esta cota superior es exponencial en n ; mas aún, el ejemplo que se dió al principio de esta sección requiere de un número exponencial de iteraciones. En otras palabras, el *simplex es un algoritmo de tiempo exponencial*. Sin embargo, tales ejemplos exponenciales no son comunes en la práctica y por ello el simplex es tan valioso y tan usado.

Capítulo 2

El algoritmo proyectivo de Karmarkar

Este es el primer método de punto interior que se expondrá en el presente trabajo. Este método fue desarrollado en 1984 por N. Karmarkar y su principal ventaja, es que su tiempo de ejecución es polinomial.

La idea del algoritmo es la siguiente: se toma un punto en el interior del conjunto de soluciones factibles. Se construye una bola que esté contenida en el conjunto de soluciones factibles. Se optimiza la función dentro de esa bola y, en el nuevo punto encontrado, se vuelve a construir otra bola para volver a optimizar la función dentro de esa bola. Así se procede un número finito de veces. A partir del último punto encontrado, se procede a buscar un punto extremo que optimice la función.

Todas estas ideas se quedarán formalizadas a través de este capítulo. Para poder hablar formalmente del algoritmo, primero se requieren de los resultados que se exponen en la siguiente sección.

2.1. Conceptos preliminares

Antes de exponer bajo qué hipótesis trabaja el método de Karmarkar y dar la descripción formal, conviene tener en cuenta las siguientes definiciones.

Definición 2.1. Sean $x_1, x_2, \dots, x_n \in \mathbb{R}^n$ y $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ con $\sum_{i=1}^n \lambda_i = 1$. A la combinación lineal $\sum_{i=1}^n \lambda_i x_i$ se le llama *combinación afín*.

Definición 2.2. Sea $S_A \subseteq \mathbb{R}^n$. Se dice que S_A es un *subespacio afín*, si $\forall x_1, x_2 \in S_A$, toda combinación lineal afín de estos dos vectores también está en S_A .

Un buen ejemplo de subespacio afín es el siguiente. Considere el conjunto

$$K = \{x \in \mathbb{R}^2 : x_2 \leq x_1 + 1, x_2 \leq -x_1 + 5\}$$

Para probar que es un subespacio afín, sean $x, y \in K; \lambda, \beta \in \mathbb{R}$ con $\lambda + \beta = 1$. Entonces

$$\begin{aligned} \beta x + \lambda y &= (\beta x_1 + \lambda y_1, \beta x_2 + \lambda y_2) \\ \beta x_2 + \lambda y_2 &\leq \beta(x_1 + 1) + \lambda(y_1 + 1) \\ &= \beta x_1 + \beta + \lambda y_1 + \lambda \\ &= \beta x_1 + \lambda y_1 + \lambda + \beta \\ &= \beta x_1 + \lambda y_1 + 1 \end{aligned}$$

Con lo cual se cumple la primera condición del conjunto K . La segunda es análoga a la primera. Observe que, salvo por la no negatividad, las restricciones del conjunto son parecidas a las de un programa lineal.

Para poder hablar de la eficiencia (en el sentido de iteraciones) de este algoritmo, se requiere de saber qué tan complejo o qué tan grande es el problema en cuestión. Se podría pensar que el tamaño o la complejidad de un problema depende sólo del número de variables que están involucradas. Esto dejaría de lado el orden de las constantes con las que se está tratando. Por esa razón, en la definición del tamaño de un problema se deben involucrar tanto el número de variables como el orden de las constantes. Formalmente, se tiene la siguiente

Definición 2.3. Para un problema de programación lineal escrito en su forma estándar:

$$\begin{aligned} \max \quad & c^t x \\ \text{sujeto a: } & [A, I]x = b \\ & x_i \geq 0 \end{aligned} \tag{2.1}$$

con x un vector columna de n entradas y A una matriz de $m \times n$. La **longitud del problema**, denotada por L , es la cantidad de bits requeridos para almacenar todos los datos del problema 2.1; es decir:

$$\begin{aligned} L &= [1 + \log_2(1 + m)] + [1 + \log_2(1 + n)] + \sum_j [1 + \lceil \log_2(1 + |c_j|) \rceil] \\ &+ \sum_i \sum_j [1 + \lceil \log_2(1 + |a_{ij}|) \rceil] + \sum_i [1 + \lceil \log_2(1 + |b_i|) \rceil] \end{aligned}$$

Donde $\lceil \cdot \rceil$ denota el mayor entero.

A continuación, se enuncian y demuestran una serie de afirmaciones que serán de gran utilidad para la construcción del método así como para la discusión de la eficiencia del mismo.

Afirmación 2.1. Sea $x = (x_{\mathcal{B}}, x_{\mathcal{N}})$ una solución factible básica del sistema $[A, I]x = b$ con cada entrada de x no negativa, donde $A \in \mathbf{M}_{m \times n}(\mathbb{R})$ de rango m y cada entrada de x , A y b son enteras. Entonces, la k -ésima entrada de $x_{\mathcal{B}k}$ satisface que $x_{\mathcal{B}k} < \frac{2^L}{mn}$

Demostración. Por definición, $x_{\mathcal{B}} = B^{-1}b$. Por ello, se tiene que:

$$\begin{aligned}
x_{\mathcal{B}k} &= \sum_{i=1}^m B_{ki}^{-1}b_i \\
&\leq \sum_{i=1}^m |B_{ki}^{-1}b_i| \\
&\leq 1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| \iff \\
\log_2 \left(\sum_{i=1}^m |B_{ki}^{-1}b_i| \right) &\leq \log_2 \left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| \right) \\
&< \log_2(m) + \log_2(n) + \log_2 \left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| \right) \quad (2.2) \\
&< 3 + \log_2(m+1) + \log_2(n+1) + \log_2 \left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| \right) \\
&= (1 + \log_2(m+1)) + (1 + \log_2(n+1)) \\
&\quad + (1 + \log_2 \left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| \right)) \\
&\hspace{15em} (2.3) \\
&\leq \lceil (1 + \log_2(m+1)) \rceil + \lceil (1 + \log_2(n+1)) \rceil \\
&\quad + \lceil (1 + \log_2 \left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| \right)) \rceil \\
&< L
\end{aligned}$$

La última desigualdad es obvia, pues L es el número de bits que se requieren para guardar un problema completo, mientras que el lado izquierdo sólo

representa una parte del mismo. Ahora bien, usando la desigualdad 2.2 se tiene que:

$$\begin{aligned} \log_2(m) + \log_2(n) + \log_2\left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i|\right) &= \log_2\left[mn\left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i|\right)\right] \\ \log_2\left[mn\left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i|\right)\right] &< L \iff \\ mn\left(1 + \sum_{i=1}^m |B_{ki}^{-1}b_i|\right) &< 2^L \iff \\ 1 + \sum_{i=1}^m |B_{ki}^{-1}b_i| &< \frac{2^L}{mn} \end{aligned}$$

Pero $\sum_{i=1}^m B_{ki}^{-1}b_i < 1 + \sum_{i=1}^m |B_{ki}^{-1}b_i|$. Lo cual completa la prueba \square

De la afirmación anterior, se desprenden los siguientes resultados.

Corolario 2.1. Dada la desigualdad de la Afirmación 2.1, se tiene que para la norma Euclidiana

$$\|x\| < \frac{2^L}{n}$$

Demostración. Usando que $\|x\|^2 = x \cdot x$.

$$\begin{aligned} x \cdot x &< \sum_{i=1}^m \left(\frac{2^L}{mn}\right)^2 \\ &= m \left(\frac{2^L}{mn}\right)^2 \\ &< m^2 \left(\frac{2^L}{mn}\right)^2 \\ &= \frac{2^{2L}}{n^2} \iff \\ \|x\| &< \frac{2^L}{n} \end{aligned}$$

\square

El corolario anterior da paso a dar una cota para el valor óptimo de la función objetivo. Formalmente se tiene el siguiente resultado.

Afirmación 2.2. Considere el siguiente problema:

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & [A, I]x = b \\ & x_i \geq 0 \end{aligned}$$

Con L como en la definición 2.3 y suponga que el problema es acotado y factible. Entonces, $Z_{opt} \leq 2^L$

Demostración. Observe que, para cualquier problema de programación lineal

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & [A, I]x = b \\ & x_i \geq 0 \end{aligned}$$

es equivalente a

$$\begin{aligned} \min \quad & (kc)^t x \\ \text{sujeto a:} \quad & [A, I]x = b \\ & x_i \geq 0 \end{aligned}$$

con $k \in \mathbb{R}, k > 0$ constante. Sea $\bar{c} = \frac{n}{\|c\|} c$ y considérese el problema

$$\begin{aligned} \min \quad & \bar{c}^t x \\ \text{sujeto a:} \quad & [A, I]x = b \\ & x_i \geq 0 \end{aligned}$$

Por la desigualdad de Cauchy-Schwartz, se sabe que

$$\bar{c}^t x \leq \|\bar{c}\| \cdot \|x\|$$

Usando la afirmación anterior, se tiene que

$$\begin{aligned} \|\bar{c}\| \cdot \|x\| &\leq \|\bar{c}\| \frac{2^L}{n} \\ &= \frac{n}{\|c\|} \|c\| \frac{2^L}{n} \\ &= 2^L \end{aligned}$$

Observe que el razonamiento anterior es válido para cualquier solución factible, en particular para la solución óptima. Así

$$Z_{opt} \leq 2^L$$

□

Corolario 2.2. Para cualquier problema de programación lineal, se tiene que

$$-2^L \leq Z_{opt} \leq 2^L$$

Demostración. Inmediato de la afirmación anterior y de que

$$\max f = -\min -f$$

□

En el capítulo anterior, se expuso una alternativa para el método de las dos fases; dicha alternativa se conoce como el método de la gran M . Sin embargo, sólo se dijo que tal M existe, pero no se habló de su magnitud. Con la siguiente afirmación, queda bien argumentado qué tan grande debe ser la M .

Afirmación 2.3. Sea

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & [A, I]x = b \\ & x_i \geq 0 \end{aligned}$$

un problema de programación lineal acotado y factible. Sea $M = 2^{L+1}$ y considérese el siguiente problema

$$\begin{aligned} \min \quad & c^t x + M\vec{1}^t x_a \\ \text{sujeto a:} \quad & [A, I]x + x_a = b \\ & x_i, x_{ai} \geq 0 \end{aligned}$$

Entonces, en el óptimo $x_a = \vec{0}$ o el conjunto de soluciones óptimas es vacío.

Demostración. Suponga que $x_a \neq \vec{0}$. Por el corolario anterior, se tiene que

$$\begin{aligned} c^t x + M\vec{1}^t x_a & \geq -2^L + M\vec{1}^t x_a \\ & > -2^L + M(-2^L) \\ & = 2^L \end{aligned}$$

Lo cual es una contradicción. Esto completa la prueba.

□

El siguiente resultado resultará de utilidad para el análisis de convergencia del algoritmo.

Teorema 2.1. Sea $\bar{\alpha} = \frac{n\alpha}{(n-1)}$ con $\alpha \in (0, 1)$, $x \in \mathbb{R}^n$ tal que $\vec{1} \cdot x = 1$ y $x_i > 0 \quad i = 1, 2, \dots, n$ con $\vec{1} = (1, 1, \dots, 1)$. Si

$$\|nx - \vec{1}\| \leq \sqrt{\frac{n-1}{n}} \bar{\alpha} < 1$$

Entonces,

$$0 \leq -\sum_{j=1}^n \ln(nx_j) \leq \frac{\bar{\alpha}^2}{2(1-\bar{\alpha})^2}$$

Para una demostración del resultado anterior, puede consultar el Bazaraa [7].

2.2. Descripción del algoritmo

La idea básica de este algoritmo, es generar una sucesión de puntos factibles en el interior del poliedro que converja a la solución óptima. En esencia, esto se hace resolviendo una secuencia de problemas restringidos que no involucren desigualdades. El algoritmo que está por describirse, realiza una serie de cambios de variables que permiten resolver los problemas aplicando algunos conceptos del cálculo. Para que el concepto de “interior de un conjunto” no quede al aire, se formaliza a través de las siguiente

Definición 2.4. Sea $A \subseteq \mathbb{R}^n$ y $x \in \mathbb{R}^n$. Se dice que x es *punto interior* de A si $\exists \epsilon > 0$ tal que $V_\epsilon(x) \subset A$. Al conjunto de todos los puntos interiores de A se le llama el *interior de A* .

Este algoritmo resuelve problemas de la forma:

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & Ax = 0 \\ & \vec{1}^t x = 1 \\ & x_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned} \tag{2.4}$$

donde $A \in \mathbf{M}_{m \times n}(\mathbb{R})$ de rango m , $n \geq 2$, las entradas de A y c son todas enteras, $\vec{1}$ es un vector columna cuyas n -entradas son todas 1 y las siguientes dos suposiciones son válidas para el problema 2.4

A1. El punto $x_0 = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ es factible

A2. El valor óptimo es cero

Una observación importante es que no porque se tenga un problema de la forma 2.4 automáticamente se van a cumplir las hipótesis A1 y A2. Un ejemplo es el siguiente:

$$\begin{aligned} \min \quad & c \\ \text{sujeto a:} \quad & Ax = 0 \\ & \vec{1}^t x = 1 \\ & x_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned} \quad (2.5)$$

con $c \neq 0$. El problema anterior está en la forma de 2.4, pero no se cumple la hipótesis A2.

En caso de que se tenga un problema de la forma

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & Ax = b \\ & x_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned} \quad (2.6)$$

con $b \neq \vec{0}$, para convertirlo a la forma requerida y además que cumpla con las hipótesis A1 y A2, se procede a *regularizar* el problema. Para este fin, se añade una constante que cumpla que $\sum_{i=1}^n x_i \leq Q$. La Q puede manejarse como una constante entera conocida, derivada de las condiciones de factibilidad y optimalidad. Por la Afirmación 2.1 y el Corolario, en el peor de los casos se puede tomar $Q = 2^L$. Si esta restricción es activa en el óptimo, se deduce que la región es no acotada y por lo tanto no existe óptimo finito.

Añadiendo esta constante junto con una variable de holgura x_{n+1} , se reescribe el problema 2.6 como sigue:

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & Ax = b \\ & \vec{1}^t x + x_{n+1} = Q \\ & x_i \geq 0 \quad i = 1, 2, \dots, n+1 \end{aligned}$$

En este punto, se hace homogéneo el sistema $Ax = b$ al reemplazar b con $\frac{b(\vec{1}x) + x_{n+1}}{Q}$, pero esto puede afectar la densidad de A , pues generalmente b es denso. Para evitar esto, se agrega una variable *pancha* x_{n+2} junto con la

restricción $x_{n+2} = 1$, para reescribir las restricciones de manera equivalente como sigue:

$$\begin{aligned} Ax - bx_{n+2} &= 0 \\ x_{n+2} &= 1 \\ \bar{\Gamma}^t x + x_{n+1} + x_{n+2} &= Q + 1 \end{aligned}$$

más las restricciones de no negatividad. Usando la última igualdad, ya se puede escribir el Problema 2.6 de la siguiente forma:

$$\begin{aligned} \min \quad & c^t x \\ \text{sujeto a:} \quad & Ax - bx_{n+2} = 0 \\ & \bar{\Gamma}^t x + x_{n+1} - Qx_{n+2} = 0 \\ & \bar{\Gamma}^t x + x_{n+1} + x_{n+2} = (Q + 1) \\ & x_i \geq 0 \quad i = 1, 2, \dots, n + 2 \end{aligned}$$

A continuación, se usa el cambio de variables $x_j = (Q + 1)y_j$, $1 \leq j \leq n + 2$, para que en la última restricción quede igualada a 1. Así, se tiene:

$$\begin{aligned} \min \quad & c^t y \\ \text{sujeto a:} \quad & Ay - by_{n+2} = 0 \\ & \bar{\Gamma}^t y + y_{n+1} - Qy_{n+2} = 0 \\ & \bar{\Gamma}^t y + y_{n+1} + y_{n+2} = 1 \\ & y_i \geq 0 \quad i = 1, 2, \dots, n + 2 \end{aligned} \tag{2.7}$$

Las restricciones del Programa 2.7 son ahora de la forma del problema 2.4. Para que se cumpla la hipótesis A1, se introduce una variable artificial en la función objetivo y_{n+3} con coeficiente M suficientemente grande como para que sea 0 en el óptimo y con coeficientes en las restricciones tales que hagan que la solución $y = (\frac{1}{n+3}, \dots, \frac{1}{n+3})$ sea factible. Para que esto ocurra, los coeficientes de las restricciones homogéneas deben sumar cero y el coeficiente de y_{n+3} en la última igualdad debe de ser 1. Como se demostró, en el peor de los escenarios puede tomarse la M como 2^{L+1} . El problema auxiliar es como sigue:

$$\begin{aligned} \min \quad & c^t y + My_{n+3} \\ \text{sujeto a:} \quad & Ay - by_{n+2} - [A\bar{\Gamma}^t - b]y_{n+3} = 0 \\ & \bar{\Gamma}^t y + y_{n+1} - Qy_{n+2} - (n + 1 - Q)y_{n+3} = 0 \\ & \bar{\Gamma}^t y + y_{n+1} + y_{n+2} + y_{n+3} = 1 \\ & y_i \geq 0 \quad i = 1, 2, \dots, n + 3 \end{aligned} \tag{2.8}$$

El Problema 2.8 es de la forma del Problema 2.4 con $(n + 3)$ variables y además cumple con las hipótesis A1 y A2.

Como se demostró que un problema de programación lineal puede ser llevado a la forma del Programa 2.4 y para simplificar la notación, se supondrá de ahora en adelante que ya se tiene el problema escrito de la forma adecuada.

Dada cualquier solución factible x_k , se define la matriz diagonal $D_k = \text{diag}\{x_{k1}, \dots, x_{kn}\}$, donde las variables x_{ki} son las entradas del vector x_k . Considere la siguiente transformación, denotada por T

$$y = \frac{D_k^{-1}x}{\bar{1}D_k^{-1}x} \quad (2.9)$$

Observe que la región factible del Programa 2.4 está descrita por la intersección de un subespacio de dimensión $(n - m)$, definido por las igualdades homogéneas $Ax = 0$, con el conjunto $S_x = \{x \in \mathbb{R}^n : \bar{1}^t x = 1, x_i \geq 0\}$. Esta intersección da como resultado una región de dimensión $(n - m - 1)$. Bajo la transformación T , cualquier $x \in S_x$ se convierte en un punto en el conjunto $(n - 1)$ -dimensional $S_y = \{y \in \mathbb{R}^n : \bar{1}y = 1, y_i \geq 0\}$. En particular, el punto x_k es transformado en el punto $y_0 = (\frac{1}{n}, \dots, \frac{1}{n})$, que es el centro de S_y . Note que si se hubiera tomado la transformación afín $y' = D_k^{-1}x$, se tendría que $y' \notin S_y$; sin embargo, si este punto y' es proyectado en S_y sobre el rayo definido por el origen y y' , se obtiene el punto dado por la transformación T . Por esta razón, a la transformación T se le conoce como la *transformación proyectiva*.

Considere ahora la transformación proyectiva inversa, que se obtiene de resolver para x en S_x la ecuación 2.9. Esto da como resultado $x = (D_k y)(\bar{1}^t D_k^{-1} x)$. Como $\bar{1}^t x = 1$, se tiene que $(\bar{1}^t D_k y)(\bar{1}^t D_k^{-1} x) = 1$. Despejando, $\bar{1}^t D_k^{-1} x = \frac{1}{\bar{1}^t D_k y}$. Substituyendo esto en la última expresión de x , se tiene

$$x = \frac{D_k y}{\bar{1}^t D_k y} \quad (2.10)$$

En este punto es claro que la transformación 2.9 tiene una inversa, y que $T(S_x) = S_y$ y $T^{-1}(S_y) = S_x$.

Bajo la transformación T , el Problema 2.4 se convierte en

$$\begin{aligned} \min \quad & \frac{c^t D_k y}{\bar{1}^t D_k y} \\ \text{sujeto a:} \quad & AD_k y = 0 \\ & \bar{1}^t y = 1 \\ & y_i \geq 0 \end{aligned} \quad (2.11)$$

Observe que aunque las restricciones aún son lineales, la función objetivo ya no lo es, pues se ha convertido en un cociente de funciones lineales. A este tipo de problema se le conoce como *problema fraccional de programación lineal*. Sin embargo, por la hipótesis A2, el valor óptimo del problema 2.11 sigue siendo cero. Por ello, equivalentemente se puede minimizar el numerador del problema, pues el denominador es estrictamente positivo y acotado lejos el cero $\forall y \in S_y$. Tomando

$$\bar{c} = c^t D_k, \quad P = \begin{bmatrix} AD_k \\ \bar{1}^t \end{bmatrix} \quad P_0 = \begin{bmatrix} \bar{0}^t \\ \bar{1}^t \end{bmatrix}$$

se puede reescribir el problema 2.11 como

$$\begin{aligned} \min \quad & \bar{c}^t y \\ \text{sujeto a:} \quad & Py = P_0 \\ & y_i \geq 0 \end{aligned} \tag{2.12}$$

En lugar de resolver el Problema 2.12, lo cual es equivalente a resolver el problema original, se optimizará una restricción mas simple del mismo. Al resolver el problema, se encontrará un punto dentro de la región de soluciones factibles que incremente el valor en la función objetivo aunque no necesariamente de forma monótona.

Para definir esta restricción, sea $r > 0$ y considérese la vecindad $V_r(y_0)$ con centro en $y_0 = (\frac{1}{n}, \dots, \frac{1}{n})$ y cuyo radio r es tal que la intersección de esta vecindad con el conjunto $\{y \in \mathbb{R}^n : \bar{1}^t y = 1\}$ resulta en una bola $(n-1)$ -dimensional con el mismo centro y radio, inscrita en S_y . Observe que r es la distancia de y_0 a alguna de las caras de S_y . Sin pérdida de generalidad, se puede tomar el punto donde la primera componente es cero, a saber, $(0, \frac{1}{n-1}, \dots, \frac{1}{n-1}) \in S_y$; así, $r = \frac{1}{\sqrt{n(n-1)}}$.

Suponga ahora, que se desea minimizar el problema 2.12 pero *dentro* de la vecindad $V_{\alpha r}(y_0)$ con $0 < \alpha < 1$, es decir, se le agregó al problema 2.12 una restricción, a saber, las y 's deben de estar dentro de la vecindad. Note que se ha encogido la vecindad por un factor de α ; más aún, al tomar los puntos dentro de la vecindad, con cada iteración las entradas de aquéllos serán estrictamente positivas. Como $y_i \geq 0 \quad \forall i$, la intersección de $V_{\alpha r}(y_0)$ con $\{y \in \mathbb{R}^n : \bar{1}^t y = 1\}$ implica que se puede reescribir el problema como

$$\begin{aligned} \min \quad & \bar{c}^t y \\ \text{sujeto a:} \quad & Py = P_0 \\ & (y - y_0)^t (y - y_0) \leq \alpha^2 r^2 \end{aligned} \tag{2.13}$$

Note que el sistema de ecuaciones $Py = P_0$ define un espacio afín de dimensión $(n - m - 1)$ que pasa por el centro de la vecindad $V_{ar}(y_0)$. Por ello, la región factible del problema 2.13 es una esfera con centro en y_0 y de dimensión $(n - m - 1)$. En este momento, es claro que la solución del problema 2.13 se obtiene al proyectar menos el gradiente de la función objetivo $(-\bar{c}^t)$ centrado en y_0 a la superficie definida por $Py = P_0$ y moviéndose desde y_0 sobre esta dirección proyectada hacia el borde de la esfera. Tomando $c_p = \bar{c}^t$ y el óptimo del problema 2.13 como y_{act} , se tiene:

$$y_{act} = y_0 - \alpha r \frac{c_p}{\|c_p\|} \quad (2.14)$$

Si se tiene que $c_p = \vec{0}$, entonces cualquier solución factible sería óptima, pues se tomó $y_0 = (\frac{1}{n}, \dots, \frac{1}{n})$ sólo por simplicidad, es decir, es el punto factible mas fácil de construir y de trabajar, pero cualquier otro funcionaría. De esta manera, cualquier x_k sería solución óptima al problema original.

Antes de poder calcular c_p , es necesario primero demostrar la siguiente

Afirmación 2.4. Sea $A \in \mathbf{M}_{m \times n}(\mathbb{R})$ de rango m , $x_k \in \mathbb{R}^n$ con cada entrada estrictamente positiva. Entonces, la matriz PP^t es invertible, con

$$P = \begin{bmatrix} AD_k \\ \bar{1}^t \end{bmatrix}$$

$$D_k = \text{diag}\{x_{k1}, \dots, x_{kn}\}$$

Demostración. Suponga que PP^t no es invertible y sea $r_i \quad i = 1, 2, \dots, (m+1)$ el i -ésimo renglón de PP^t . Como PP^t no es invertible, entonces $\det(PP^t) = 0$. Eso significa que $\{r_i\}_{i=1}^{m+1}$ es un conjunto linealmente dependiente. Realizando operaciones elementales en PP^t , se puede hacer un renglón de ceros. Entonces, P es equivalente a una matriz que tenga algún renglón de ceros y consecuentemente el rango de $P \neq m$, lo cual es una contradicción. \square

Ahora, observe que como c_p está en la superficie definida por $Py = P_0$, el vector $(\bar{c} - c_p)$ pertenece al espacio definido por los gradientes a $Py = P_0$. De ahí que exista \bar{w} tal que $P^t \bar{w} = \bar{c} - c_p$. Multiplicando ambos lados de la ecuación por P , se tiene que $PP^t \bar{w} = P\bar{c}$ pues por definición $Pc_p = 0$. Usando la afirmación anterior, se tiene que $\bar{w} = (PP^t)^{-1} P\bar{c}$. Como $c_p = \bar{c} - P^t \bar{w}$, substituyendo el valor de \bar{w} , se tiene que:

$$c_p = [I - P^t(PP^t)^{-1}P]\bar{c} \quad (2.15)$$

Dada y_{act} en la expresión 2.14, el nuevo vector correspondiente x_{k+1} es obtenido con la transformación inversa 2.10

$$x_{k+1} = \frac{D_k y_{act}}{\bar{1}^t D_k y_{act}} \quad (2.16)$$

Esto completa una iteración. Se vuelve a comenzar después de incrementar k en 1.

Como se da una sucesión de puntos interiores, en un número finito de pasos sólo se puede garantizar que $Z_{opt} < 2^{-L}$. Para encontrar un punto extremo óptimo, se sigue el procedimiento conocido como *esquema de purificación*, pero antes de describir detalladamente dicho esquema, resulta conveniente tener la siguiente

Definición 2.5. Sea $X = \{x \in \mathbb{R}^n : Ax \leq b, \quad x_i \geq 0 \quad \forall i = 1, 2, \dots, n\}$ con $A \in \mathbf{M}_{m \times n}(\mathbb{R})$, $b \in \mathbb{R}^m$. Se dice que una restricción $\alpha x \leq \beta$ de X es *activa* en $\bar{x} \in X$ si $\alpha \bar{x} = \beta$

Empezando con x_k tal que $c^t x_k < 2^{-L}$, si n restricciones linealmente independientes son activas en x_k , entonces x_k ya es una solución factible básica (además de óptima). En otro caso, existe una dirección $d \neq \vec{0}$ en el espacio nulo de restricciones activas, es decir, d satisface al sistema de ecuaciones homogéneas definido por las restricciones activas. Se procede a moverse en la dirección d si $cd < 0$, y en la dirección $-d$ en otro caso, hasta que alguna restricción interrumpa el movimiento. Esto necesariamente debe ocurrir, pues se supone que la región factible está acotada.

Para ilustrar la forma de reescribir un problema general, considérese el siguiente problema de programación lineal escrito de forma estándar

$$\max x_1 + 3x_2$$

sujeto a:

$$2x_1 - x_2 \leq 8$$

$$x_1 + 4x_2 \leq 15$$

$$x_i \geq 0$$

Considere el problema dual dado por

$$\min 8w_1 + 15w_2$$

sujeto a:

$$\begin{aligned} 2w_1 - w_2 &\geq 1 \\ -w_1 + 4w_2 &\geq 3 \\ x_i &\geq 0 \end{aligned}$$

Si se minimiza la diferencia entre el dual y el primal sujeto tanto a las restricciones duales como primales, se asegura el tener una función objetivo cuyo óptimo sea cero ¹. Esto hace que se cumpla la hipótesis A1.

$$\min x_1 + 3x_2 - (8w_1 + 15w_2)$$

sujeto a:

$$\begin{aligned} 2x_1 - x_2 &\leq 8 \\ x_1 + 4x_2 &\leq 15 \\ 2w_1 - x_2 &\geq 1 \\ -w_1 + 4w_2 &\geq 3 \\ x_i, w_i &\geq 0 \end{aligned}$$

Se introducen variables de holgura para transformar las desigualdades en ecuaciones:

$$\min x_1 + 3x_2 - (8w_1 + 15w_2)$$

sujeto a:

$$\begin{aligned} 2x_1 - x_2 + h_1 &= 8 \\ x_1 + 4x_2 + h_2 &= 15 \\ 2w_1 - x_2 - h_3 &= 1 \\ -w_1 + 4w_2 - h_4 &= 3 \\ x_i, w_i, h_i &\geq 0 \end{aligned}$$

Se introduce la Q que es una cota para la suma de variables

$$\min x_1 + 3x_2 - (8w_1 + 15w_2)$$

¹Observe que se está suponiendo que el óptimo es finito

sujeto a:

$$\begin{aligned}
 2x_1 - x_2 + h_1 &= 8 \\
 x_1 + 4x_2 + h_2 &= 15 \\
 2w_1 - x_2 - h_3 &= 1 \\
 -w_1 + 4w_2 - h_4 &= 3 \\
 x_1 + x_2 + w_1 + w_2 + h_1 + h_2 + h_3 + h_4 + h_5 &= Q \\
 x_i, w_i, h_i &\geq 0
 \end{aligned}$$

Se agrega la variable artificial t junto con la restricción de que sea igual a 1 para homogeneizar las restricciones

$$\min x_1 + 3x_2 - (8w_1 + 15w_2)$$

sujeto a:

$$\begin{aligned}
 2x_1 - x_2 + h_1 - 8t &= 0 \\
 x_1 + 4x_2 + h_2 - 15t &= 0 \\
 2w_1 - x_2 - h_3 - t &= 0 \\
 -w_1 + 4w_2 - h_4 - 3t &= 0 \\
 x_1 + x_2 + w_1 + w_2 + h_1 + h_2 + h_3 + h_4 + h_5 + t &= Q + 1 \\
 x_i, w_i, h_i, t &\geq 0
 \end{aligned}$$

Haciendo los siguientes cambios de variables

$$\begin{aligned}
 x_1 &= (Q + 1)y_1 & x_2 &= (Q + 1)y_2 \\
 w_1 &= (Q + 1)y_3 & w_2 &= (Q + 1)y_4 \\
 h_1 &= (Q + 1)y_5 & h_2 &= (Q + 1)y_6 \\
 h_3 &= (Q + 1)y_7 & h_4 &= (Q + 1)y_8 \\
 h_5 &= (Q + 1)y_9 & t &= (Q + 1)y_{10}
 \end{aligned}$$

El problema se reescribe como

$$\min y_1 + 3y_2 - (8y_3 + 15y_4)$$

sujeto a:

$$\begin{aligned}
2y_1 - y_2 + y_5 - 8y_{10} &= 0 \\
y_1 + 4y_2 + y_6 - 15y_{10} &= 0 \\
2y_3 - y_4 - y_7 - y_{10} &= 0 \\
-y_3 + 4y_4 - y_8 - 3y_{10} &= 0 \\
y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9 - Qy_{10} &= 0 \\
y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9 + y_{10} &= 1 \\
y_i &\geq 0
\end{aligned}$$

Por último, para que se cumpla la hipótesis A1, se tiene

$$\min y_1 + 3y_2 - (8y_3 + 15y_4) + My_{11}$$

sujeto a:

$$\begin{aligned}
2y_1 - y_2 + y_5 - 8y_{10} + 6y_{11} &= 0 \\
y_1 + 4y_2 + y_6 - 15y_{10} + 9y_{11} &= 0 \\
2y_3 - y_4 - y_7 - y_{10} - y_{11} &= 0 \\
-y_3 + 4y_4 - y_8 - 3y_{10} + y_{11} &= 0 \\
y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9 - Qy_{10} - (9 - Q)y_{11} &= 0 \\
y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9 + y_{10} + y_{11} &= 1 \\
y_i &\geq 0
\end{aligned}$$

Así, el problema anterior cumple cabalmente con las hipótesis A1 y A2 requeridas para aplicar el algoritmo de Karmarkar. Observe que el tamaño del problema aumentó considerablemente. Por esa razón, para ejemplificar el algoritmo, considere el siguiente problema de programación lineal escrito de forma estándar

$$\begin{aligned}
&\min && x_2 \\
\text{sujeto a: } &x_1 + x_2 - 2x_3 &= 0 \\
&x_1 + x_2 + x_3 &= 1 \\
&x_i &\leq 0
\end{aligned}$$

En este ejemplo, $n = 3$ y $m = 1$. Claramente, el punto $x_0 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^t$ es factible y el óptimo es cero. Así, se cumplen con las hipótesis A1 y A2. Además se tiene que

$$r = \frac{1}{\sqrt{6}}, \quad \alpha = \frac{2}{9}, \quad L = 15$$

Primera iteración

Empezando con $x_0, k = 0$, se define $D_0 = \text{diag}\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$. Como $\bar{1}^t D_0 x_0 = 3$, la transformación 2.9 da $x = y$, así se tiene que $y_0 = x_0$. Según las ecuaciones 2.9 a 2.12, se tiene que:

$$\bar{c} = \left(0, \frac{1}{3}, 0\right) \quad P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \\ 1 & 1 & 1 \end{bmatrix} \quad PP^t = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 3 \end{bmatrix}$$

$$(PP^t)^{-1} = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \quad (PP^t)^{-1}P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \quad P^t(PP^t)^{-1}P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

De la ecuación 2.15, se tiene que

$$c_p = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{3} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{6} \\ \frac{1}{6} \\ 0 \end{bmatrix} \quad \|c_p\| = \sqrt{\frac{1}{6^2} + \frac{1}{6^2} + 0} = \frac{\sqrt{2}}{6}$$

Usando la ecuación 2.14, se tiene que

$$y_{act} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} - \left(\frac{2}{9}\right) \left(\frac{1}{\sqrt{6}}\right) \left(\frac{2}{\sqrt{2}}\right) \begin{bmatrix} -\frac{1}{6} \\ \frac{1}{6} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{9\sqrt{3}} + \frac{1}{3} \\ \frac{1}{3} - \frac{1}{9\sqrt{3}} \\ \frac{1}{3} \end{bmatrix}$$

Usando la transformación inversa, se tiene $x_1 = y_{act}$ y $c^t x_1 = \frac{1}{3} - \frac{1}{9\sqrt{3}}$.

Segunda iteración

Con $k = 1$, se define

$$D_1 = \text{diag} \left\{ \frac{1}{9\sqrt{3}} + \frac{1}{3}, \frac{1}{3}, \frac{1}{3} - \frac{1}{9\sqrt{3}}, \frac{1}{3} \right\} \quad \bar{c} = \left(0, \frac{1}{3} - \frac{1}{9\sqrt{3}}, 0\right)$$

$$P = \begin{bmatrix} \frac{1}{9\sqrt{3}} + \frac{1}{3} & \frac{1}{3} - \frac{1}{9\sqrt{3}} & -\frac{2}{3} \\ 1 & 1 & 1 \end{bmatrix}$$

Se calcula de nuevo c_p junto con su nueva norma

$$c_p = [I - P^t(PP^t)^{-1}P]\bar{c} = \begin{bmatrix} \frac{121}{2214\sqrt{3}} - \frac{121}{738} \\ \frac{119}{738} - \frac{1}{54\sqrt{3}} \\ \frac{1}{369} - \frac{1}{1107\sqrt{3}} \end{bmatrix} \quad \|c_p\| = \sqrt{\frac{1627}{29889} - \frac{80}{3321\sqrt{3}}}$$

Para simplificar la lectura, se usarán de aquí en adelante valores decimales. Substituyendo los valores en la ecuación 2.14 se tiene

$$y_{act} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} - 0.45006524 \begin{bmatrix} -0.1324029 \\ 0.1505548 \\ -0.0181517 \end{bmatrix} = \begin{bmatrix} 0.3930009 \\ 0.2654855 \\ 0.3415134 \end{bmatrix}$$

Finalmente, de la Ecuación 2.16, se tiene

$$D_1 y_{act} = \begin{bmatrix} 0.1562112 \\ 0.0714641 \\ 0.1138378 \end{bmatrix} \quad \vec{1}^t D_1 y_{act} = 0.3415131$$

$$x_2 = \frac{D_1 y_{act}}{\vec{1}^t D_1 y_{act}} = \begin{bmatrix} 0.457409 \\ 0.209258 \\ 0.333333 \end{bmatrix}$$

Esto completa la segunda iteración. El valor actual en la función objetivo es de $c^t x_2 = 0.209258$

Observe que el valor de las x'_k s es de la forma $(\frac{1}{3} + \delta, \frac{1}{3} - \delta, \frac{1}{3})$ donde δ incrementa gradualmente hasta alcanzar el valor de $\frac{1}{3}$ cuando $k \rightarrow \infty$. Para efectos de este ejemplo, tome $2^{-L} = .0625$. Así, las iteraciones deberían terminar cuando $c^t x_k < 0.0625$. Esto ocurra para $k = 6$, con $x_6 = (0.606509, 0.060158, \frac{1}{3})^t$ y $c^t x_6 = 0.06158$.

Finalmente, se aplica el esquema de purificación para encontrar el punto extremo óptimo. Observe que las únicas dos restricciones linealmente independientes son activas en x_6 . Se debe encontrar una solución $d \in \mathbb{R}^3$, $d \neq \vec{0}$ para el sistema homogéneo

$$\begin{aligned} d_1 + d_2 + 2d_3 &= 0 \\ d_1 + d_2 + d_3 &= 0 \end{aligned}$$

Resolviendo para d_1 y d_3 en términos de d_2 , se tiene $d_1 = -d_2$ y $d_3 = 0$. Tomando arbitrariamente $d_2 = 1$, da la dirección $(-1, 1, 0)^t$. Como $c^t d = 1 > 0$, se debe mover en la dirección $-d = (1, -1, 0)$. La restricción $x_2 \geq 0$ bloquea el movimiento, dando una longitud total de 0.060158. La nueva solución $x^* = (\frac{2}{3}, 0, \frac{1}{3})$ es punto extremo y óptimo.

2.3. Análisis de convergencia

Para poder probar la convergencia del algoritmo, se desarrollarán dos construcciones: primero, se construirá un problema *relajado* del Problema

2.12 que además complemente las restricciones del Problema 2.13; la segunda construcción compete a una función llamada la *función potencial*. Ésta última mide el progreso del algoritmo en cada iteración.

Para hacer la primera construcción, considere la vecindad n -dimensional $V_R(y_0)$ cuyo centro coincide con el de S_y y cuyo radio sea tal que al intersectar $V_R(y_0)$ con $\{y \in \mathbb{R}^n : \bar{1}^t y = 1\}$ resulte en una bola $(n-1)$ -dimensional que circunscriba S_y . Por construcción, R es la distancia desde y_0 a cualquiera de los vértices de S_y . Sin pérdida de generalidad, se puede tomar la distancia a $(1, 0, \dots, 0)$. Así,

$$R = \sqrt{\frac{(n-1)}{n}}$$

Considere ahora el problema obtenido al añadir la restricción $y \in V_R(y_0)$ al Problema 2.12, pero quitando la restricción de la no-negatividad:

$$\begin{aligned} \min \quad & \bar{c}^t y \\ \text{sujeto a:} \quad & Py = P_0 \\ & (y - y_0)^t (y - y_0) \leq R^2 \end{aligned} \quad (2.17)$$

Nuevamente, la región factible del Problema 2.17 resulta ser una bola $(n-m-1)$ -dimensional centrada en y_0 , definida por la intersección del subespacio afín $\{y \in \mathbb{R}^n : Py = P_0\}$ con $V_R(y_0)$. Una solución al Problema 2.17 es obtenida de manera similar a la expresión 2.14, y está dada por

$$\bar{y}_{act} = y_0 - \frac{Rc_p}{\|c_p\|} \quad (2.18)$$

con c_p como en la Ecuación 2.15. Ya se está en las condiciones para obtener un estimado del progreso con respecto de la función objetivo $\bar{c}^t y$. De la Ecuación 2.14, claramente se tiene que si $c_p \neq \vec{0}$, $\bar{c}^t y_{act} < \bar{c}^t y_0$ porque

$$\bar{c}^t (y_0 - y_{act}) = \frac{\alpha r \bar{c}^t c_p}{\|c_p\|} = \alpha r \|c_p\| > 0$$

pues

$$\bar{c}^t = c_p + P^t (PP^t)^{-1} P \bar{c}^t \quad \text{de la Ecuación 2.15}$$

así

$$\bar{c}^t c_p = \|c_p\|^2 + \bar{c}^t P^t (PP^t)^{-1} P c_p = \|c_p\|^2$$

pues por definición

$$Pc_p = 0$$

Denotando y^* como la solución óptima del Problema 2.12, como el Problema 2.13 (resuelto por y_{act}) es una restricción del Problema 2.12 (resuelto por \bar{y}_{act}) y el Problema 2.17 es una versión relajada del Problema 2.12, se tiene que

$$\bar{c}^t \bar{y}_{act} \leq \bar{c}^t y^* \leq \bar{c}^t y_{act} < \bar{c}^t y_0$$

Más aún, usando lo anterior junto con las ecuaciones 2.15 y 2.18, se tiene

$$0 < \bar{c}^t (y_0 - y_{act}) \leq \bar{c}^t (y_0 - y^*) \leq \bar{c}^t (y_0 - \bar{y}_{act}) = \frac{R}{\|c_p\|} \bar{c}^t c_p = \frac{R}{\alpha r} \bar{c}^t (y_0 - y_{act})$$

Las desigualdades anteriores implican que

$$\begin{aligned} \bar{c}^t (y_0 - y^*) &\leq \frac{R}{\alpha r} \bar{c}^t (y_0 - y_{act}) = \frac{R}{\alpha r} [\bar{c}^t (y_0 - y^*) - \bar{c}^t (y_{act} - y^*)] \\ &\iff 1 \leq \frac{R}{\alpha r} [1 - \frac{\bar{c}^t (y_{act} - y^*)}{\bar{c}^t (y_0 - y^*)}] \\ &\iff 1 - \frac{R}{\alpha r} \leq -\frac{R}{\alpha r} \frac{\bar{c}^t (y_{act} - y^*)}{\bar{c}^t (y_0 - y^*)} \\ &\iff \frac{\bar{c}^t (y_{act} - y^*)}{\bar{c}^t (y_0 - y^*)} \leq 1 - \frac{\alpha r}{R} = 1 - \frac{\alpha}{(n-1)} \end{aligned} \quad (2.19)$$

Por la hipótesis A2, se tiene que $\bar{c}^t y^* = 0$, así que

$$\frac{\bar{c}^t y_{act}}{\bar{c}^t y_0} \leq 1 - \frac{\alpha}{(n-1)} \quad (2.20)$$

La desigualdad 2.19 garantiza que, con cada iteración, el valor de $\bar{c}^t y$ se reduce en un $\frac{\alpha}{(n-1)}\%$. Sin embargo, $\bar{c} = D_k c$ depende de la k -ésima iteración; más aún, la Ecuación 2.19 sólo garantiza un decrecimiento en el numerador de la Ecuación 2.11. De hecho, el valor de la función objetivo del Problema 2.11, y por tanto del Problema 2.4, puede no decrecer y hasta aumentar. Por estas razones, se requiere de una función que sea estrictamente monótona y así asegure la convergencia al óptimo. Se usará una función particular llamada la *función potencial* que está dada por:

$$f(x) = \sum_{j=1}^n \ln \left[\frac{c^t x}{x_j} \right] = n \ln(c^t x) - \sum_{j=1}^n \ln(x_j) \quad (2.21)$$

De acuerdo con la Ecuación 2.9,

$$\frac{c^t x}{x_j} = \frac{c^t D_k y}{x_{kj} y_j} = \frac{c^t y}{x_{kj} y_j}$$

Así, la Función 2.21 se convierte en

$$F(y) = f \left[\frac{D_k y}{\vec{1}^t D_k y} \right] = n \ln(c^t y) - \sum_{j=1}^n \ln(y_j) - \sum_{j=1}^n \ln(x_{kj}) \quad (2.22)$$

Usando la expresión anterior, se medirá cuánto decrece la función en la k -ésima iteración en términos de las variables y :

$$F(y_{act}) - F(y_0) = n \ln \left[\frac{c^t y_{act}}{c^t y_0} \right] - \sum_{j=1}^n \ln(n y_{(act)j})$$

pues $y_{0j} = \frac{1}{n} \quad \forall \quad j = 1, 2, \dots, n$. Sin embargo, de la Ecuación 2.20, se tiene que

$$\ln \left[\frac{c^t y_{act}}{c^t y_0} \right] \leq \ln \left[1 - \frac{\alpha}{(n-1)} \right] \leq -\frac{\alpha}{(n-1)}$$

pues $\ln(1-x) \leq -x \quad \forall \quad x \in [0, 1)$. Usando esta última ecuación, se tiene que

$$F(y_{act}) - F(y_0) \leq -\frac{n\alpha}{(n-1)} - \sum_{j=1}^n \ln [n y_{(act)j}] \quad (2.23)$$

Observe ahora que y_{act} cumple con las hipótesis del Teorema 2.1 y de la Ecuación 2.14

$$\|n y_{act} - \vec{1}\| = n\alpha r = \sqrt{\frac{(n-1)}{n}} \bar{\alpha}$$

Usando el resultado del Teorema 2.1 en la Ecuación 2.23, se tiene

$$F(y_{act}) - F(y_0) \leq -\bar{\alpha} + \frac{(\bar{\alpha})^2}{2(1-\bar{\alpha})^2} \leq -\frac{1}{5} \quad \text{cuando} \quad \bar{\alpha} = \frac{n\alpha}{(n-1)} = \frac{1}{3} \quad (2.24)$$

Por tanto, cuando $\bar{\alpha} = \frac{1}{3}$, la función $F(\cdot)$ (y consecuentemente la función potencial $f(\cdot)$) decrece por $\frac{1}{5}$ en cada iteración. Esto significa que después de k iteraciones,

$$f(x_k) - f(x_0) \leq -\frac{k}{5}$$

pero

$$f(x_k) - f(x_0) = n \ln \left(\frac{c^t x_k}{c^t x_0} \right) - \sum_{j=1}^n \ln(n x_{kj})$$

pues $x_0 = (\frac{1}{n}, \dots, \frac{1}{n})$. Así

$$n \ln\left(\frac{c^t x_x}{c^t x_0}\right) \leq \sum_{j=1}^n \ln(nx_{kj}) - \frac{k}{5} \quad (2.25)$$

Usando que la media geométrica $\left[\prod_{j=1}^n (x_{kj})\right]^{\frac{1}{n}}$ es menor o igual a la media aritmética $\frac{\bar{1}^t x_k}{n}$ y que $\bar{1}^t x_k = 1$ se tiene que

$$\sum_{j=1}^n \ln(nx_{kj}) = n \ln \left[n \left(\prod_{j=1}^n x_{kj} \right)^{\frac{1}{n}} \right] \leq n \ln(\bar{1}^t x_k) = 0$$

Usando lo anterior en la desigualdad 2.25 se tiene que

$$\ln \left[\frac{c^t x_k}{c^t x_0} \right] \leq -\frac{k}{5n} \quad \forall k = 0, 1, \dots \quad (2.26)$$

Las desigualdades anteriores prueban que, aunque el valor en la función objetivo pudiera llegar a incrementar entre dos iteraciones consecutivas, un decrecimiento general queda garantizado a lo largo de las iteraciones. Más aún, de la desigualdad 2.26 se tendrá que

$$c^t x_k < 2^{-L} \quad \text{cuando } c^t x_k \leq (c^t x_0) e^{-\frac{k}{5n}} < 2^{-L}$$

De la desigualdad anterior se tiene que, cuando $k = 10nL$, se tendrá

$$c^t x_k \leq (c^t x_0) e^{-\frac{k}{5n}} = \left(\sum_{j=1}^n \frac{c_j}{n} \right) (e^{-2L}) < (2^L)(2^{-2L}) = 2^{-L}$$

Más aún, cada iteración requiere no más de $O(n^3)$ operaciones, y así el algoritmo tiene una complejidad pseudopolinomial de $O(n^4 L)$.

Capítulo 3

Introducción a la Programación Semidefinida

En este capítulo se expondrá qué es la Programación Semidefinida (PSD) y su alcance; para ello, primero se desarrollará la herramienta necesaria para dicho propósito. En el siguiente capítulo, se expondrá en detalle un algoritmo de punto interior para resolver problemas de PSD. La idea de la PSD es escribir los problemas en términos de matrices positivas semidefinidas y tratarlas como puntos del espacio de matrices simétricas positivas semidefinidas. De ahora en adelante, a menos que se especifique lo contrario, las letras mayúsculas denotarán matrices y las minúsculas vectores o escalares.

3.1. Definiciones previas

Dado que los problemas de la PSD requieren del manejo de matrices simétricas positivas semidefinidas, conviene primero dar la definición precisa de matriz simétrica y de matriz positiva semidefinida:

Definición 3.1. Sea $A \in \mathbf{M}_{n \times n}(\mathbb{R})$. A es *simétrica* si $A^t = A$. Al conjunto de matrices simétricas se le denotará como S_n

Definición 3.2. Sean $A, B \in \mathbf{M}_{n \times n}(\mathbb{R})$. Se dice que A es *positiva semidefinida*, denotado por $A \succeq 0$, si $\forall x \in \mathbb{R}^n \setminus \{\vec{0}\}, x^t A x \geq 0$. Se escribe $A \succeq B$ para denotar que $A - B \succeq 0$. Se dice además que la matriz A es *positiva definida*, denotado por $A \succ 0$, si $\forall x \in \mathbb{R}^n \setminus \{\vec{0}\}, x^t A x > 0$. Análogo si la matriz es negativa semidefinida o negativa definida. Al conjunto de matrices, de tamaño $n \times n$, simétricas positivas semidefinidas se le denotará como S_n^+ .

De la definición de matriz positiva semidefinida, se desprenden los siguientes enunciados equivalentes:

Afirmación 3.1. Sea A una matriz positiva semidefinida. Entonces

1. A es positiva semidefinida
2. Todos los valores propios de A son no negativos
3. Todos los determinantes de las submatrices superiores izquierdas de A son no negativos, es decir, si

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

las submatrices superiores izquierdas son de la forma:

$$A_i = \begin{bmatrix} a_{11} & \dots & a_{1i} \\ a_{2i} & \dots & a_{2i} \\ \vdots & \vdots & \vdots \\ a_{i1} & \dots & a_{ii} \end{bmatrix} \quad \forall i = 1, 2, \dots, n$$

y sus determinantes son no negativos

La prueba de las equivalencias anteriores se derivan directamente de la Definición 3.2. Observe además que si una matriz es simétrica positiva semidefinida, todas las submatrices superiores izquierdas también lo serán. De ahora en adelante, a menos que se indique lo contrario, $X \in S_n^+$.

Como se va trabajar con la traza de una matriz, resulta útil dar la definición de traza de una matriz junto con sus propiedades más importantes. De nuevo, se omitirá la prueba de cada una de ellas.

Definición 3.3. Sea $A \in \mathbf{M}_{n \times n}(\mathbb{R})$. La traza de la matriz A se define como

$$Tr(A) = \sum_{i=1}^n a_{ii}$$

donde a_{ij} representa la entrada del renglón i y la columna j . Además, se toma la siguiente notación

$$Tr(AB^t) = \langle A, B \rangle$$

Afirmación 3.2. Sean $A, B \in \mathbf{M}_{n \times n}(\mathbb{R}), \gamma \in \mathbb{R}$. Entonces, las siguientes relaciones son válidas:

$$\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B) \quad (3.1)$$

$$\text{Tr}(\gamma A) = \gamma \text{Tr}(A) \quad (3.2)$$

$$\text{Tr}(A^t) = \text{Tr}(A) \quad (3.3)$$

$$\text{Tr}(AB) = \text{Tr}(BA) \quad (3.4)$$

$$\text{Tr}(A) = \sum_{i=1}^n \lambda_i(A) \quad (3.5)$$

donde $\lambda_i(A)$ es el i -ésimo valor propio de A .

3.2. Objetivo de la PSD

El objetivo de la PSD es optimizar problemas de la forma

$$(P) \quad \min c^t x$$

sujeto a: $A_0 + \sum_{i=1}^n x_i A_i \succeq 0$

donde $c \in \mathbb{R}^n$ y $A_i \in \mathbf{M}_{n \times n} \mathbb{R}$. Observe que un problema de PSD es un problema de optimización convexa: sean $x, y \in \mathbb{R}^n, \lambda \in [0, 1]$, entonces

$$A_0 + \sum_{i=1}^n (\lambda x_i + (1 - \lambda y_i)) A_i = A_0 + \lambda \sum_{i=1}^n x_i A_i + (1 - \lambda) \sum_{i=1}^n y_i A_i$$

La formulación anterior es la forma general de un problema de PSD, el cual puede ser llevado a la siguiente forma:

$$(P) \quad \min \text{Tr}(CX)$$

sujeto a: $\text{Tr}(A_i X) = b_i \quad i = 1, 2, \dots, m$
 $X \succeq 0$

que tiene su correspondiente dual asociado

$$(D) \quad \max b^t y$$

sujeto a: $\sum_{i=1}^n y_i A_i + S = C$
 $S \succeq 0, y \in \mathbb{R}^n$

donde los escalares b_i y las matrices A_i son datos conocidos.

Si los problemas están como en (P) (respectivamente (D)), se dirá que están en su forma *estándar*. Los valores óptimos del problema (P) (respectivamente (D)) serán denotados por $p^*(d^*)$. Si la matriz X fuese diagonal, se tendría un problema de programación lineal. Antes de ilustrar este hecho, se requiere de la siguiente

Afirmación 3.3. Sea $x \in \mathbb{R}^n$. Entonces, $x_i \geq 0 \iff \text{diag}(x) \succeq 0$, donde $\text{diag}(x)$ denota a la matriz que resulta de colocar a las entradas del vector x en la diagonal y ceros en las demás entradas.

Demostración. Sea $x \in \mathbb{R}^n$ con cada entrada no negativa y suponga que $\text{diag}(x) \not\succeq 0$. Entonces, al menos un determinante de alguna submatriz superior izquierda es negativo, lo cual es una contradicción al punto 3 de la Afirmación 3.1. El regreso es inmediato utilizando el mismo punto de la misma afirmación. \square

Considere el siguiente programa lineal

$$\begin{aligned} & \max && c^t x \\ \text{sujeto a:} && \sum_{j=1}^n a_{ij} x_j + b_i \geq 0 \\ && x_i \geq 0, 1 \leq i \leq n \end{aligned}$$

Como un vector $x \in \mathbb{R}^n$ tiene cada una de sus componentes mayor o igual que cero si y solo si la matriz $\text{diag}(x) \succeq 0$, se puede expresar el problema anterior como

$$\begin{aligned} & \max && c^t x \\ \text{sujeto a:} && \text{diag}(Ax + b) \succeq 0 \end{aligned}$$

Así, la programación lineal es un caso particular de la PSD.

La teoría del dual es más débil para la PSD que para la programación lineal, sin embargo, aún se tiene la propiedad dual débil:

$$\text{Tr}(CX) - b^t y = \text{Tr} \left(\left(S + \sum_{i=1}^m y_i A_i \right) X \right) - \sum_{i=1}^m y_i \text{Tr}(A_i X) = \text{Tr}(SX) \geq 0 \quad (3.6)$$

de donde la desigualdad se sigue de que $X \succeq 0, S \succeq 0$. Las soluciones factibles (X, y, S) que cumplan con

$$\text{Tr}(CX) - b^t y = \text{Tr}(SX) = 0$$

son óptimas.

Más adelante se argumenta que la programación cuadrática también queda abarcada por la PSD; sin embargo, para dicha exposición se requiere saber qué es el complemento de Schur y algunas de sus propiedades. A continuación se da la definición formal del complemento de Schur.

Definición 3.4. Sea $X \in \mathbf{M}(\mathbb{R})_{n \times n}$ de la forma:

$$X = \begin{bmatrix} A & B \\ B^t & C \end{bmatrix}$$

con A invertible. El *complemento de Schur* de A en X (también llamado el complemento de Schur de X con respecto a A) es la matriz

$$S_{chur} = C - B^t A^{-1} B$$

Para poder reformular un problema de programación cuadrática en uno de PSD, se requieren los siguientes resultados.

Afirmación 3.4. Sea X como en la Definición 3.4, entonces $X \succ 0 \iff A \succ 0$ y $S_{chur} \succ 0$.

Demostración. Sea $X \succ 0$ simétrica. La simetría de A es clara; queda demostrar que $A \succ 0$. Para ello, se $x \in \mathbb{R}^n \setminus \{0\}$. Se tiene que:

$$\begin{aligned} x^t A x &= [x^t \quad 0] \begin{bmatrix} A & B \\ B^t & C \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} \\ &= [x^t \quad 0] X \begin{bmatrix} x \\ 0 \end{bmatrix} \geq 0 \end{aligned}$$

Sólo queda demostrar que $S_{chur} \succ 0$. Para ello, sea $x \in \mathbb{R}^n$. Suponga que $x = \begin{bmatrix} y \\ z \end{bmatrix}$. Así,

$$\begin{aligned} 0 < x^t X x &= (y^t, z^t) \begin{bmatrix} A & B \\ B^t & C \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} \\ &= y^t A y + y^t B z + z^t B^t y + z^t C z \\ &= y^t A y + y^t B z + z^t B^t y + z^t C z + z^t B^t (A^{-1})^t B z - z^t B^t (A^{-1})^t B z \\ &= y^t A y + y^t B z + z^t B^t y + z^t B^t (A^{-1})^t B z + z^t C z - z^t B^t (A^{-1})^t B z \\ &= (y^t + z^t B^t (A^{-1})^t) (A y + B z) + z^t (C - B^t A^{-1} B) z \\ &= (y + A^{-1} B z)^t (A y + A A^{-1} B z) + z^t (C - B^t A^{-1} B) z \\ &= (y + A^{-1} B z)^t A (y + A^{-1} B z) + z^t (C - B^t A^{-1} B) z \end{aligned} \quad (3.7)$$

Observe que se usó el hecho de que A es simétrica, es decir que $A = A^t$ y que $(A^{-1})^t = (A^t)^{-1}$. Como $x^t X x \geq 0$ para cualquier $x \neq \vec{0}$, tome $z \neq \vec{0}$ y $y = -A^{-1}Bz$. Esto hace que el primer término de la Ecuación 3.7 se anule. Reescribiendo, se tiene que:

$$0 < x^t X x = z^t (C - B^t A^{-1} B) z = z^t S_{chur} z$$

Esto completa la prueba. \square

Afirmación 3.5. Sea X como en la Definición 3.4 y $A \succ 0$. Entonces $X \succeq 0 \iff S_{chur} \succeq 0$

Demostración. La implicación hacia la derecha es inmediata de la afirmación anterior. Para la implicación hacia la izquierda, sea $x = \begin{bmatrix} y \\ z \end{bmatrix}$, $z \neq \vec{0}$ y suponga que

$$\begin{aligned} A \succ 0 \quad y \quad S_{chur} = C - B^t A^{-1} B \succeq 0 &\iff \\ 0 \leq (y + A^{-1} B z)^t A (y + A^{-1} B z) + z^t (C - B^t A^{-1} B) z & \\ = (y + A^{-1} B z)^t (A y + A A^{-1} B z) + z^t (C - B^t A^{-1} B) z & \\ = (y^t + z^t B^t (A^{-1})^t) (A y + B z) + z^t (C - B^t A^{-1} B) z & \\ = y^t A y + y^t B z + z^t B^t y + z^t B^t (A^{-1})^t B z + z^t C z - z^t B^t (A^{-1})^t B z & \\ = y^t A y + y^t B z + z^t B^t y + z^t C z + z^t B^t (A^{-1})^t B z - z^t B^t (A^{-1})^t B z & \\ = y^t A y + y^t B z + z^t B^t y + z^t C z & \\ = (y^t, z^t) \begin{bmatrix} A & B \\ B^t & C \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} & \\ = x^t X x & \end{aligned}$$

Lo que completa la prueba. \square

Usando la Definición 3.4 y las dos afirmaciones anteriores, se sigue que la restricción cuadrática

$$(Ax + b)^t (Ax + b) - (c^t x + d) \leq 0, x \in \mathbb{R}$$

puede ser reemplazada por la restricción semidefinida

$$\begin{bmatrix} I & Ax + b \\ (Ax + b)^t & c^t x + d \end{bmatrix} \succeq 0$$

De la misma manera, se puede representar el *cono de segundo orden* (o “barquillo”)

$$\left\{ (t, x) : t \geq \sqrt{\sum_{i=1}^n x_i^2} \right\}$$

con la restricción semidefinida

$$\begin{bmatrix} tI & x \\ x^t & t \end{bmatrix} \succeq 0$$

3.3. Aplicaciones y alcance

Ya se expuso que la programación lineal está es parte de la PSD, es decir, cualquier problema de programación lineal puede ser expresado en términos de la PSD. Una pregunta natural es, si también es posible expresar los problemas de PSD como problemas de programación lineal. La intuición hace pensar que no es posible y a continuación se expone un problema que puede ser expresado en términos de la Programación Semidefinida pero no en términos de programación lineal. Considérese el problema

$$\begin{aligned} \min & \quad \frac{(c^t x)^2}{d^t x} \\ \text{sujeto a:} & \quad Ax \geq b \end{aligned}$$

donde se supone que $d^t x > 0$ siempre que $Ax - b \geq 0$ y la última desigualdad significa que cada entrada del vector resultante será no negativo. Claramente este problema no puede ser expresado en términos de otro problema equivalente de programación lineal. Se procede a introducir una variable auxiliar t , con la condición de que sea una cota superior para la función objetivo:

$$\begin{aligned} \min & \quad t \\ \text{sujeto a:} & \quad Ax \geq b \\ & \quad \frac{(c^t x)^2}{d^t x} \leq t \end{aligned} \tag{3.8}$$

Nuevamente, usando las propiedades del complemento de Schur, el problema 3.8 puede ser expresado como

$$\begin{aligned} \min & \quad t \\ \text{sujeto a:} & \quad \begin{bmatrix} t & c^t x & 0 \\ c^t x & d^t x & 0 \\ 0 & 0 & \text{diag}(Ax - b) \end{bmatrix} \succeq 0 \end{aligned}$$

el cual es un problema de PSD. De manera general, dos técnicas son útiles para reescribir un problema de programación lineal o uno de programación cuadrática en uno de PSD:

- Reconocer los complementos de Schur para expresar las restricciones como una sola matriz.
- Introducir una variable auxiliar que sea una cota para la función objetivo.

Un problema típico de la PSD, es encontrar el máximo de los valores propios de $A \in S_n$ denotado por $\lambda_{\max}(A)$. La clave está en notar que $t \geq \lambda_{\max}(A) \iff tI - A \succeq 0$. Así, el problema de PSD se escribe como

$$\begin{aligned} \min \quad & t \\ \text{sujeto a :} \quad & tI - A \succeq 0 \\ & t \in \mathbb{R} \end{aligned}$$

Los problemas de este tipo se presentan principalmente en la optimización estructural, teoría de gráficas, teoría de control y optimización combinatoria. Dentro de la programación estructural, el problema más conocido es encontrar la estructura óptima de barras que conectan a otra estructura terrestre de nodos (un ejemplo famoso de tal estructura es la Torre Eiffel). El diseño queda determinado una vez que el tamaño de las barras se ha decidido. Dos variantes de este problema son:

1. Minimizar el peso de la estructura tal que su frecuencia fundamental (frecuencia a la cual la estructura resuena) permanezca por encima de un valor crítico.
2. Minimizar el peor comportamiento (“energía acumulada”) de la estructura final dado el conjunto de fuerzas que la estructura debe soportar.

El segundo problema se puede modelar como sigue:

$$\begin{aligned} \min_{t, x_1, \dots, x_k} \quad & \left\{ \max_{j=1, \dots, k} \left\{ x_j^t f_j \right\} \right\} \\ \text{sujeto a:} \quad & \left(\sum_{i=1}^m t_i \frac{E_i}{l_i^2} b_i^t b_i \right) x_j = f_j, \quad j = 1, \dots, k \\ & \sum_{i=1}^m t_i = V, \quad t \geq 0 \end{aligned}$$

donde las variables t_i son los volúmenes de las barras (variables diseño) y las variables f_i son las fuerzas que la estructura debe soportar. El desplazamiento de los nodos sujetos a la fuerza f_i está dado por el vector x_j . Los

vectores b_i dependen sólo de la disposición de los nodos y son conocidos. Las variables E_i y las l_i denotan el módulo de Young y la longitud de la barra i , respectivamente. La primera restricción da equilibrio a la barra y la segunda da el volumen total. La función objetivo mide el comportamiento de las barras en el sentido del desplazamiento de las mismas. Para este tipo de problemas, se quiere maximizar la rigidez de la estructura. La matriz

$$K(t) = \left[\sum_{i=1}^m t_i \frac{E_i}{l_i^2} b_i^t b_i \right]$$

también se le llama *matriz de rigidez*; da la relación entre los desplazamientos y la fuerza aplicada mediante $K(t)x_j = f_j$. Usando la técnica del complemento de Schur, se puede reformular el problema en uno de PSD como

$$\begin{aligned} & \min \tau \\ & \text{sujeto a:} \\ & \begin{bmatrix} \tau & f_j^t \\ f_j & \sum_{i=1}^m t_i \frac{E_i}{l_i^2} b_i^t b_i \end{bmatrix} \succeq 0, \quad j = 1, \dots, k \\ & \sum_{i=1}^m t_i = V, \quad t \geq 0 \end{aligned}$$

3.4. Conceptos preliminares

En este punto, se tiene claro que la PSD no sólo contiene a la programación lineal (LP), sino que además sus aplicaciones se dan en campos teóricos (maximización de valores propios) y prácticos (el problema de la estructura); sin embargo, hasta ahora se ha hablado del planteamiento del problema pero no se ha expuesto cómo resolverlos de manera sistemática, es decir, no se ha dado ningún algoritmo que encuentre óptimos o establezca que ninguno existe. Actualmente existen varios algoritmos para resolver problemas de PSD, todos ellos caen en la categoría de los algoritmos de punto interior. La presente expondrá uno de ellos en el siguiente capítulo, sin embargo, antes de exponerlo se requieren de los resultados y definiciones que se darán a continuación.

Recuerde que p^* y d^* son los valores óptimos del problema primal (P) y dual (D) respectivamente. Se usará la convención de que $p^* = -\infty$ si (P) es no acotado y $p^* = \infty$ si (P) es no factible. Análogamente para (D). A los conjuntos de soluciones factibles de (P) se le denotará por \mathcal{P} , análogamente

\mathcal{D} . Así mismo, $\mathcal{P}^*, \mathcal{D}^*$ denotarán los conjuntos óptimos, es decir,

$$\mathcal{P}^* = \{X \in \mathcal{P} : Tr(CX) = p^*\} \quad y \quad \mathcal{D}^* = \{(S, y) \in \mathcal{D} : b^t y = d^*\}$$

Un problema (P) se dice que se *puede resolver* si el conjunto \mathcal{P}^* es no-vacío. Las siguientes 2 suposiciones serán hechas, a menos que se especifique lo contrario:

B1. Las matrices $A_i, i = 1, 2, \dots, m$, son linealmente independientes.

B2. Los conjuntos \mathcal{P} y \mathcal{D} tienen interiores no vacíos.

La suposición B1 garantiza que y queda determinado por un dual factible S dado y además es única. Esta suposición es equivalente a la su suposición de programación lineal que la matriz de restricciones debe tener rango igual al número de renglones. A la suposición B2 se le conoce como *factibilidad estricta*. Formalmente:

Definición 3.5. Un problema primal (P) se dice que es *estrictamente factible* si

$$\exists X \in \mathcal{P} \quad \text{con} \quad X \succ 0$$

Análogamente, un problema dual (D), será estrictamente factible si

$$\exists (y, S) \in \mathcal{D} \quad \text{con} \quad S \succ 0$$

En el capítulo anterior se definió el concepto de subespacio afín. Resulta ahora conveniente dar la definición de *interior relativo*.

Definición 3.6. Sea C un espacio vectorial con alguna norma definida en él. Se define el *interior relativo* de C , denotado por $\mathbf{relint}C$, como

$$\mathbf{relint}C = \{x \in C : V_r(x) \cap \mathbf{aff}C \subseteq C\}$$

para alguna $r > 0$.

Con esta última definición, se puede ver que la Definición 3.5 es equivalente a la *condición de Slater*: Si el problema primal es convexo, es decir, de la forma:

$$\begin{aligned} \min \quad & f_0(x) \\ \text{sujeto a:} \quad & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & Ax = b \end{aligned}$$

Si $\exists x \in \text{relint}D$, con D el dominio de la función, tal que

$$f_i(x) < 0, \quad i = 1, \dots, m \quad Ax = b$$

entonces el teorema fuerte de dualidad se cumple. La equivalencia se sigue de que el interior de S_n^+ consta de matrices positivas definidas. Para poder declarar que un problema es infactible o no acotado, se requiere de la siguiente

Definición 3.7. Se dice que un problema primal (P) tiene un *rayo de mejora* si existe una matriz simétrica $\bar{H} \succeq 0$ tal que $\text{Tr}(A_i \bar{H}) = 0, \forall i$ y $\text{Tr}(C \bar{H}) < 0$. Análogamente, se dice que un problema dual (D) tiene un rayo de mejora si existe un vector $\bar{y} \in \mathbb{R}^m$ tal que $\bar{S} = -\sum_{i=1}^m \bar{y}_i A_i \succeq 0$ y $b^t \bar{y} > 0$.

Los rayos de mejora en el problema primal (P) causan que el problema dual (D) sea no factible y al revés. Formalmente se tiene el siguiente resultado

Afirmación 3.6. Si existe un rayo de mejora en el problema dual (\bar{y}) entonces (P) es no factible. Análogamente, si existe un rayo de mejora en el problema primal (\bar{X}), entonces (D) es no factible.

Demostración. Sean \bar{y} un rayo de mejora del problema dual y X una matriz factible del problema primal. Se tiene entonces que

$$0 < b^t \bar{y} = \sum_{i=1}^m \bar{y}_i \text{Tr}(A_i X) = -\text{Tr}(X \bar{S}) \leq 0$$

Lo cual es claramente una contradicción. La prueba para el caso en el que se tenga un rayo de mejora en el problema primal es análoga \square

Definición 3.8. El problema primal (P) se dice que es *fuertemente infactible* si el dual (D) tiene un rayo de mejora. Análogamente, se dice que un problema dual (D) es fuertemente infactible si el problema primal (P) tiene un rayo de mejora

Si se tratara el caso de la programación lineal, todo problema infactible es fuertemente infactible, sin embargo, en la PSD se puede dar el caso de que el problema sea *débilmente infactible*:

Definición 3.9. Un problema (P) es débilmente infactible si $\mathcal{P} = \emptyset$ y $\forall \epsilon > 0, \exists X \succeq 0$ tal que

$$|\text{Tr}(A_i X) - b_i| \leq \epsilon, \quad \forall i$$

Análogamente, un problema (D) es llamado débilmente infactible si $\mathcal{D} = \emptyset$ y $\forall \epsilon > 0, \exists y \in \mathbb{R}^m, S \succeq 0$ tal que

$$\left\| \sum_{i=1}^m y_i A_i + S - C \right\| \leq \epsilon$$

con

$$\|A\|^2 = Tr(AA^t) = \sum_i \sum_j a_{ij}^2 = \sum_i \lambda_i^2(A) \quad (\text{la norma de Frobenius})$$

Un ejemplo de un problema que es débilmente infactible es el siguiente. Considérese un problema (D) como

$$\begin{aligned} & \max && y_1 \\ \text{sujeto a:} & && y_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \preceq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

donde la solución “ ϵ -infactible” está dada por

$$S = \begin{bmatrix} \frac{1}{\epsilon} & 1 \\ 1 & \epsilon \end{bmatrix}, y_1 = \frac{1}{\epsilon}$$

Recuerde la Desigualdad 3.6. Esta desigualdad implica que el intervalo dual es siempre no negativo. La longitud del intervalo dual se dice que es cero, si

$$p^* = \inf_{X \in \mathcal{P}} Tr(CX) = \sup_{S, y \in \mathcal{D}} b^t y = d^*$$

En el capítulo 1, se dieron las definiciones y propiedades de las condiciones de optimalidad. En el caso de la PSD se dan propiedades similares. Las condiciones de optimalidad para (P) y (D) son:

$$\left. \begin{aligned} Tr(A_i X) &= b_i, & X \succeq 0, & \quad i = 1, 2, \dots, m \\ \sum_{i=1}^m y_i A_i + S &= C, & S \succeq 0 \\ XS &= 0 \end{aligned} \right\} \quad (3.9)$$

Las soluciones factibles que cumplan con la última desigualdad son llamadas *complementarias*. Como S y X son positivas semidefinidas, la condición de que $XS = 0$ es equivalente a $Tr(XS) = 0$. Esto da a lugar al siguiente resultado, llamado *ortogonalidad*:

Afirmación 3.7. Sean (X, S) y (X^0, S^0) dos pares de soluciones factibles y tome $\Delta X = X - X^0$, $\Delta S = S - S^0$. Entonces se tiene que

$$Tr(\Delta X \Delta S) = 0$$

Demostración.

$$\begin{aligned}
Tr(\Delta X \Delta S) &= Tr[(X - X^0)(S - S^0)] \\
&= Tr(XS - XS^0 - X^0S + X^0S^0) \\
&= Tr(XS) - Tr(XS^0) - Tr(X^0S) + Tr(X^0S^0) \\
&= [Tr(CX) - b^t y] - [Tr(CX) - b^t y^0] - [Tr(CX^0) - b^t y] \\
&\quad + [Tr(CX^0) - b^t y^0] \\
&= 0
\end{aligned}$$

□

A continuación se dará la definición de *solución primal complementaria maximal*, pero para ello, se requieren las siguientes definiciones.

Definición 3.10. Sea V un espacio vectorial con producto interior y sea $W \subseteq V$. Se define W^\perp como sigue:

$$W^\perp = \{v \in V : v \cdot w = 0, \forall w \in W\}$$

A W^\perp se le denomina *complemento ortogonal* de V .

Definición 3.11. Sean V un espacio vectorial con producto interior y W_1, W_2 subespacios de V . Se dice que W_1 y W_2 son *ortogonales* si $w_1 \cdot w_2 = 0, \forall w_1 \in W_1, w_2 \in W_2$.

Definición 3.12. Sean V un espacio vectorial y $W_1, W_2 \subseteq V$ no vacíos. Se define $W_1 + W_2$ como

$$W_1 + W_2 = \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$$

Tomando en cuenta las definiciones anteriores, se introducen los siguientes subespacios de \mathbb{R}^n : \mathcal{B} como el subespacio generado por las columnas de las soluciones primales óptimas X ; \mathcal{N} como el subespacio generado por las columnas de las soluciones duales óptimas S ; \mathcal{T} como el complemento ortogonal del subespacio $\mathcal{B} + \mathcal{N}$. Para cualquier par de soluciones óptimas (X, S) se tiene que $XS = 0$; por ello, los subespacios \mathcal{B}, \mathcal{N} son ortogonales. Así, los subespacios antes definidos dividen \mathbb{R}^n en 3 subespacios mutuamente ortogonales. El subespacio generado por cualquier solución factible primal (dual) será denotado por $\mathcal{R}(X), (\mathcal{R}(S))$. Ahora se está en condiciones de dar la siguiente

Definición 3.13. Sea X una solución factible óptima primal. Si $\mathcal{R}(X) = \mathcal{B}$, se dice que X es una *solución primal complementaria maximal*. Análogamente, si S es una solución factible dual y $\mathcal{R}(S) = \mathcal{N}$, se dice que S es una *solución dual complementaria maximal*. Si X y S son ambas complementarias maximales, al par (X, S) se le llama *par complementario óptimo maximal*; más aún, si $\mathcal{T} = \{0\}$ se dice que el par (X, S) es *estrictamente complementario*.

Tomando en cuenta el Teorema Espectral (consulte [2]), se introduce la siguiente notación. Como $X \succeq 0$ y $S \succeq 0$, la descomposición espectral de X y S son de la forma:

$$X = Q^t \Lambda Q, \quad S = Q^t \Sigma Q \quad (3.10)$$

donde Λ y Σ son matrices cuyas diagonales tienen los valores propios de X y S respectivamente, y Q es una matriz ortogonal. Claramente $XS = 0 \iff \Lambda\Sigma = 0$.

Suponga ahora que se tiene el siguiente problema:

$$\begin{aligned} \min \quad & Tr(CX) - \mu \log \det(X) \\ \text{sujeto a:} \quad & \\ & Tr(A_i X) = b_i \quad i = 1, \dots, m \\ & X \succeq 0 \end{aligned}$$

donde el parámetro μ decrece a cero. Las condiciones de optimalidad para este problema son:

$$\begin{aligned} Tr(A_i X) &= b_i \quad i = 1, \dots, m \\ \sum_{i=1}^m y_i A_i + S &= C \\ XS &= \mu I \\ X \succeq 0, S \succeq 0, \mu &> 0 \end{aligned} \quad (3.11)$$

Estas ecuaciones pueden ser vistas como una perturbación de las originales por un parámetro μ . La solución de este sistema para alguna μ dada será denotada por $X(\mu), S(\mu), y(\mu)$, y serán interpretadas como una representación paramétrica de una curva suave (llamada *el camino central*). Con los siguientes resultados, se demuestra que, cuando $\mu \rightarrow 0$, se converge a una solución complementaria maximal. La existencia de puntos límite, es una consecuencia la siguiente

Afirmación 3.8. Dada $\bar{\mu} > 0$, el conjunto

$$\{(X(\mu), S(\mu)), 0 < \mu < \bar{\mu}\}$$

es acotado.

Demostración. Sean $(X_0, S_0), (X(\mu), S(\mu))$ dos soluciones para alguna $\mu > 0$. Por la Afirmación 3.7, se tiene que

$$Tr((X(\mu) - X_0)(S(\mu) - S_0)) = 0$$

Las condiciones de optimalidad para el problema parametrizado implican que $Tr(X(\mu)S(\mu)) = n\mu$. Así, la ecuación anterior se simplifica a

$$Tr(X(\mu)S_0) + Tr(X_0S(\mu)) = n\mu + Tr(X_0S_0)$$

Observe que los términos del lado izquierdo de la igualdad son no-negativos, pues las matrices son positivas semidefinidas. Por ello, se tiene

$$Tr(X(\mu)S_0) \leq n\mu + Tr(X_0S_0)$$

Y para una $\bar{\mu} > 0$, se tiene que

$$Tr(X(\mu)) \leq \frac{n\mu + Tr(X_0S_0)}{\lambda_{min}(S_0)}, \quad \forall \mu < \bar{\mu} \quad (3.12)$$

Donde $\lambda_{min}(S_0)$ es el menor valor propio de S_0 . Observe que para la norma de Frobenius,

$$\|X\|^2 = \sum_i \lambda_i^2(X) \leq \left(\sum_i \lambda_i(X) \right)^2 = Tr^2(X)$$

para cualquier X positiva semidefinida. Así, usando la desigualdad anterior en la Desigualdad 3.12 se tiene que

$$\|X\| \leq \frac{n\mu + Tr(X_0S_0)}{\lambda_{min}(S_0)}, \quad \forall \mu < \bar{\mu}$$

Análogo para $\|S\|$. □

Tome ahora

$$X(\mu_t) = Q^t(\mu_t)\Lambda(\mu_t)Q(\mu_t), \quad S(\mu_t) = Q^t(\mu_t)\Sigma(\mu_t)Q(\mu_t)$$

con $\{\mu_t\}$ una sucesión de términos estrictamente positivos, que converge a 0 cuando $t \rightarrow \infty$. La afirmación anterior implica que los valores propios de $X(\mu_t)$ y $S(\mu_t)$ están acotados. Las matrices $Q(\mu_t)$ son ortonormales para toda t , y por ello restringidas a un conjunto acotado. Así, la secuencia de $(Q(\mu_t), \Lambda(\mu_t), \Sigma(\mu_t))$ tiene un punto de acumulación. Sea $(Q^*, \Lambda^*, \Sigma^*)$ ese punto de acumulación. Existe entonces una subsucesión de $\{\mu_t\}$ (denotada aún por $\{\mu_t\}$ para no complicar la notación) tal que

$$\lim_{t \rightarrow \infty} Q(\mu_t) = Q^*, \quad \lim_{t \rightarrow \infty} \Lambda(\mu_t) = \Lambda^*, \quad \lim_{t \rightarrow \infty} \Sigma(\mu_t) = \Sigma^*$$

Obsérvese que $\Lambda(\mu_t)\Sigma(\mu_t) = \mu_t I$. Definiendo

$$X^* = Q^{*t} \Lambda^* Q^* = \lim_{t \rightarrow \infty} X(\mu_t), \quad S^* = Q^{*t} \Sigma^* Q^* = \lim_{t \rightarrow \infty} S(\mu_t) \quad (3.13)$$

se tiene que $\Lambda^* \Sigma^* = 0$ y el par (X^*, S^*) es óptimo.

Ya sólo falta demostrar que el par (X^*, S^*) es complementario maximal. Esto queda demostrado en el siguiente

Teorema 3.1. El par (X^*, S^*) definido como en los límites 3.13 es complementario maximal.

Demostración. Sea (X, S) un par óptimo arbitrario. Usando la Afirmación 3.7 y $Tr(XS) = 0, Tr(X(\mu_t), S(\mu_t)) = n\mu_t$ se tiene que:

$$\begin{aligned} 0 &= Tr[(X(\mu_t) - X)(S(\mu_t) - S)] \\ &= Tr[X(\mu_t)(S - S(\mu_t)) + X(S(\mu_t) - S)] \\ &= Tr(X(\mu_t)S + XS(\mu_t) - X(\mu_t)S(\mu_t) - XS) \\ &= Tr(X(\mu_t)S) + Tr(XS(\mu_t)) - Tr(X(\mu_t)S(\mu_t)) - Tr(XS) \\ &= Tr(X(\mu_t)S) + Tr(XS(\mu_t)) - n\mu_t \iff \\ n\mu_t &= Tr(X(\mu_t)S) + Tr(XS(\mu_t)) \end{aligned}$$

Como

$$\begin{aligned} X(\mu_t)S(\mu_t) = \mu_t I &\iff \frac{1}{\mu_t} X(\mu_t) = S^{-1}(\mu_t) \\ \text{y } \frac{1}{\mu_t} S(\mu_t) &= X^{-1}(\mu_t) \end{aligned}$$

dividiendo ambos lados por μ_t , se tiene que

$$Tr(S^{-1}(\mu_t)S) + Tr(X^{-1}(\mu_t)X) = n \quad (3.14)$$

para toda t . Lo cual claramente implica que

$$\text{Tr}(X^{-1}(\mu_t)X) \leq n \quad \text{Tr}(S^{-1}(\mu_t)S) \leq n \quad (3.15)$$

pues ambos términos de 3.14 son no-negativos. Tome ahora la i -ésima columna de la matriz $Q(\mu_t)$ como $q_i(\mu_t)$ y el i -ésimo elemento de la matriz $\Lambda(\mu_t)$ por $\lambda_i(\mu_t)$. Se tiene entonces que

$$X^{-1}(\mu_t) = Q^t(\mu_t)\Lambda^{-1}(\mu_t)Q(\mu_t) = \sum_{i=1}^n \frac{1}{\lambda_i(\mu_t)} q_i^t(\mu_t)q_i(\mu_t) \quad (3.16)$$

Substituyendo el valor de $X^{-1}(\mu_t)$ en la Desigualdad 3.15, se tiene que

$$\text{Tr}(XX^{-1}(\mu_t)) = \sum_{i=1}^n \text{Tr}\left(\frac{1}{\lambda_i(\mu_t)} Xq_i^t(\mu_t)q_i(\mu_t)\right) = \sum_{i=1}^n \frac{q_i^t(\mu_t)Xq_i(\mu_t)}{\lambda_i(\mu_t)} \leq n \quad (3.17)$$

La última desigualdad implica que, para cada i , se tiene

$$q_i^t(\mu_t)Xq_i(\mu_t) \leq n\lambda_i(\mu_t)$$

Haciendo $t \rightarrow \infty$, se tiene

$$\begin{aligned} q_i^{*t}Xq_i^* &\leq n\lambda_i^*, i = 1, 2, \dots, n \implies \\ q_i^{*t}Xq_i^* &= 0 \quad \text{si } \lambda_i^* = 0 \implies \\ Xq_i^* &= \vec{0} \quad \text{si } \lambda_i^* = 0 \end{aligned}$$

Pues $q_i^{*t}Xq_i^* = \|X^{\frac{1}{2}}q_i^*\|^2$, donde $X^{\frac{1}{2}}$ es raíz cuadrada simétrica de X . La última implicación significa que el espacio vectorial formado por los renglones de X es ortogonal a cada columna q_i^* de Q^* para la cual se cumpla $\lambda_i = 0$. Por ello el espacio formado por las columnas de X es un subespacio del espacio generado por las variables q_i^* para las cuales $\lambda_i > 0$. Como X es simétrica, se tiene que $\mathcal{R}(X) \subseteq \mathcal{R}(X^*)$. Como X fue una solución primal cualquiera, se tiene entonces que $\mathcal{R}(X^*) = \mathcal{B}$. La construcción para S^* es análoga. \square

Con este último resultado, queda probado que existe una sucesión de puntos factibles que convergen a un par óptimo, pero no se ha dado un método formal para ello. En la literatura existen varios métodos para encontrar dicho par óptimo. Todos esos métodos son de punto interior. En el siguiente capítulo se analiza a detalle uno de ellos. Para ver algún método en particular se puede consultar [5].

Capítulo 4

El método primal de la barrera logarítmica

En este capítulo se expone a detalle el método primal de la barrera logarítmica. Con este método se sigue el camino central, es decir, se da una sucesión de puntos factibles a través del camino central para converger al óptimo. A pesar de que este algoritmo trabaja con puntos primales, se verá que la información que aporta el dual también es tomada en cuenta. Se probará, además, que este algoritmo es eficiente (en el sentido del número de iteraciones) que aquéllos que emulan al Simplex (cuyos desempeños peores son justo los del método Simplex para el caso de la Programación Lineal).

Como en los capítulos anteriores, antes de dar la descripción detallada del algoritmo, se requiere exponer ciertos conceptos preliminares.

4.1. Conceptos preliminares

Para poder simplificar la notación, se da la siguiente

Definición 4.1. Sea $A \in \mathbf{M}_{n \times n}(\mathbb{R})$. Se define

$$\mathbf{vec}(A) = (a_{11}, a_{21}, a_{31}, \dots, a_{n1}, a_{12}, a_{22}, \dots, a_{nn})^t$$

Para ilustrar la definición anterior, considere la matriz

$$A = \begin{bmatrix} 13 & 69 & 25 \\ 19 & 5 & 84 \\ 33 & 102 & 71 \end{bmatrix}$$

Entonces

$$\mathbf{vec}(A) = (13, 19, 33, 69, 5, 102, 25, 84, 71)^t$$

Más adelante se definirá la función logarítmica de la barrera y se requerirá dar sus derivadas. Para ello, resultan de gran utilidad los siguientes resultados.

Definición 4.2. Sean V_1, V_2 espacios normados; $f : A \subset V_1 \rightarrow V_2$; $v \in \text{int}V_1$. Se dice que f es diferenciable en v (en el sentido de Fréchet), si existe una función lineal y continua $L \in \mathcal{L}(V_1, V_2)$ tal que

$$\lim_{x \rightarrow a} \frac{f(x) - (f(a) + L(x - a))}{\|x - a\|}$$

donde $\mathcal{L}(V_1, V_2)$ denota al conjunto de todos los operadores lineales de V_1 en V_2

En el caso de funciones de una variable, la definición anterior coincide con la definición de derivada en un punto de funciones de una variable: suponga que $f : \mathbb{R} \rightarrow V$ es una función de una variable (con V un espacio normado) derivable en un punto a . Entonces

$$\begin{aligned} \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a) &\Rightarrow \lim_{x \rightarrow a} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = 0 \\ &\Rightarrow \lim_{x \rightarrow a} \frac{f(x) - f(a) - f'(a)(x - a)}{|x - a|} = 0 \end{aligned}$$

Como $L(h) = f'(a)h$ es una función lineal de \mathbb{R} en V , se deduce que f es diferenciable en el sentido de Fréchet. Recíprocamente, como las funciones lineales de \mathbb{R} en V son de la forma

$$h \mapsto m \cdot h$$

la condición de diferenciabilidad dice que

$$\lim_{x \rightarrow a} \frac{f(x) - [f(a) + m(x - a)]}{|x - a|} = 0$$

lo cual equivale, invirtiendo la implicaciones anteriores, a que f es derivable en a y $m = f'(a)$.

Afirmación 4.1. Sea $f : \text{int}(S_n^+) \rightarrow \mathbb{R}$ dada por

$$f(X) = \ln \det X$$

Denotando

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f(X)}{\partial x_{11}} & \cdots & \frac{\partial f(X)}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(X)}{\partial x_{n1}} & \cdots & \frac{\partial f(X)}{\partial x_{nn}} \end{bmatrix}$$

Se tiene entonces que $\nabla f(X) = X^{-1}$

Demostración. Sea $X \in \text{int}(S_n^+)$ y sea $H \in S_n$ tal que $X + H \in \text{int}(S_n^+)$. Se tiene que

$$\begin{aligned} f(X + H) - f(X) &= \ln \det(X + H) - \ln \det(X) \\ &= \ln \det(X^{-1}(X + H)) \\ &= \ln \det(I + X^{\frac{1}{2}}HX^{\frac{1}{2}}) \end{aligned}$$

Usando en los valores propios de $X^{\frac{1}{2}}HX^{\frac{1}{2}}$ que la media geométrica es menor o igual a la media aritmética se tiene que

$$\begin{aligned} \ln \det(I + X^{\frac{1}{2}}HX^{\frac{1}{2}}) &\leq \ln \left(\frac{1}{n} \text{Tr}(I + X^{\frac{1}{2}}HX^{\frac{1}{2}}) \right)^n \\ &= n \ln \left(\frac{1}{n} \text{Tr}(I + X^{\frac{1}{2}}HX^{\frac{1}{2}}) \right) \\ &= n \ln \left(1 + \frac{1}{n} \text{Tr}(X^{\frac{1}{2}}HX^{\frac{1}{2}}) \right) \end{aligned}$$

Usando que la función logaritmo es creciente, se tiene que

$$f(X + H) - f(X) \leq \text{Tr}(X^{\frac{1}{2}}HX^{\frac{1}{2}})$$

Esto prueba que X^{-1} es un subgradiente de f en X . Como por hipótesis f es diferenciable, el subgradiente es único e igual a $\nabla f(X)$ \square

Tomando en cuenta la Definición 4.2, se tiene la siguiente

Afirmación 4.2. Sea $f : \text{int}(S_n^+) \rightarrow \mathbb{R}$, dada por

$$f(X) = \ln \det X$$

Si $\nabla^2 f$ denota la derivada de ∇f con respecto a X , entonces $\nabla^2 f(X)$ es el operador lineal que satisface

$$\nabla^2 f(X)H = -X^{-1}HX^{-1}, \quad \forall H \in S_n \quad (4.1)$$

para una X invertible dada.

Demostración. Se probará que la función 4.1 cumple con las condiciones de la definición anterior, es decir, cumple que:

$$\lim_{\|H\| \rightarrow 0} \frac{\|\nabla f(X + H) - \nabla f(X) - \nabla^2 f(X)H\|}{\|H\|} = 0$$

Para ello, sea $H \in S_n$ tal que la matriz $X + H$ sea invertible. Entonces se tiene que

$$\begin{aligned}
& \|\nabla f(X + H) - \nabla f(X) - \nabla^2 f(X)H\| \\
&= \|(X + H)^{-1} - X^{-1} + X^{-1}HX^{-1}\| \\
&= \|(X + H)^{-1} (I - (X + H)X^{-1} + (X + H)X^{-1}HX^{-1})\| \\
&= \|(X + H)^{-1} (I - I - HX^{-1} + HX^{-1} + HX^{-1}HX^{-1})\| \\
&= \|(X + H)^{-1}(HX^{-1}HX^{-1})\| \\
&\leq \|(X + H)^{-1}\| \|H\| \|X^{-1}HX^{-1}\|
\end{aligned}$$

Lo cual prueba que la definición se cumple. \square

Por último, se dará un resultado que será de utilidad en la siguiente sección:

Afirmación 4.3. Sean $A_i \in S_n, i = 1, \dots, m$ tales que el conjunto $\{A_i, i = 1, \dots, m\}$ es linealmente independiente; $Y, Z \succ 0$. Entonces la matriz $M \in S_m$ con entradas

$$M_{ij} = \text{Tr}(A_i Z A_j Y), \quad i, j = 1, \dots, m$$

es positiva definida

Demostración. Sea $x \in \mathbf{R}^n, x \neq \vec{0}$. Se probará que la forma cuadrática

$$q(x) = \sum_{i,j} M_{ij} x_i x_j$$

es no negativa para la x dada. Para ello, note que

$$q(x) = \text{Tr} \left(\left(\sum_i x_i A_i \right) Z \left(\sum_j x_j A_j \right) Y \right)$$

Tomando $A(x) = \sum_i x_i A_i$ (que es positiva semidefinida por la independencia lineal) se puede reescribir $q(x)$ como

$$q(x) = \text{Tr} (A(x) Z A(x) Y) > 0$$

donde la desigualdad se sigue de que $0 \neq A(x) Z A(x) \succ 0$ y $Y \succ 0$ \square

4.2. La búsqueda primal

Recuerde del Problema 3.11 que las condiciones de optimalidad para el problema perturbado son

$$\begin{aligned} \text{Tr}(A_i X) &= b_i \quad i = 1, \dots, m \\ \sum_{i=1}^m y_i A_i + S &= C \\ X S &= \mu I \\ X \succeq 0, S \succeq 0, \mu &> 0 \end{aligned}$$

y que la solución del mismo es denotada por

$$\{X(\mu), S(\mu), y(\mu)\}$$

Esta solución da una representación paramétrica del camino central. La existencia y unicidad de la solución se sigue de que es idéntica al mínimo de la función estrictamente convexa

$$f(X, S, \mu) = \frac{1}{\mu} \text{Tr}(XS) - \ln \det(XS)$$

definida claramente en la región $\mathcal{P} \times \mathcal{D}$. Se le llama *función primal-dual de la barrera*. Esta función resulta ser la suma de las funciones primal y dual de la barrera, definidas en \mathcal{P} y \mathcal{D} respectivamente, dadas por

$$f_p(X, \mu) = \frac{1}{\mu} \text{Tr}(CX) - \ln \det X \quad \text{y} \quad f_d(y, S, \mu) = -\frac{1}{\mu} b^t y - \ln \det S$$

En efecto:

$$\begin{aligned} f_p(X, \mu) + f_d(y, S, \mu) &= \\ &= \frac{1}{\mu} \text{Tr}(CX) - \ln \det X - \frac{1}{\mu} b^t y - \ln \det S \\ &= \frac{1}{\mu} [\text{Tr}(CX) - b^t y] - [\ln \det X + \ln \det S] \\ &= \frac{1}{\mu} [\text{Tr}(CX) - b^t y] - [\ln(\det X \det S)] \\ &= \frac{1}{\mu} [\text{Tr}(CX) - b^t y] - \ln \det(XS) \\ &= \frac{1}{\mu} \text{Tr}(XS) - \ln \det(XS) \end{aligned}$$

El camino central primal corresponde a los minimizadores $X(\mu)$ de $f_p(X, \mu)$. Por esta razón, a μ se le conoce como el *parámetro centrador* o *parámetro de la barrera*. Para una X y una μ dadas, se define

$$(S(X, \mu), y(X, \mu)) = \arg \min_{S \in S_n, y \in \mathbb{R}} \left\{ \left\| \frac{1}{\mu} X^{\frac{1}{2}} S X^{\frac{1}{2}} - I \right\| : \sum_{i=1}^n y_i A_i + S = C \right\} \quad (4.2)$$

Es decir, $S(X, \mu)$ satisface las restricciones duales con la restricción perturbada de ser semidefinida y minimiza la desviación del par $(X, S(X, \mu))$ del camino central, donde la desviación está definida como

$$\delta(X, \mu) = \left\| \frac{X^{\frac{1}{2}} S X^{\frac{1}{2}}}{\mu} - I \right\|$$

Se dirá que X está suficientemente centrada si $\delta(X, \mu)$ es menor a alguna tolerancia dada. En los siguientes párrafos, se verá que se pueden expresar los puntos de cada iteración en términos de $S(X, \mu)$.

La dirección para la función de la barrera primal en un punto (X, μ) está definido como

$$\Delta X = \arg \min_{\Delta X} \langle \nabla f_p, \Delta X \rangle + \frac{1}{2} \langle \nabla^2 f_p(X, \mu) \delta X, \Delta X \rangle \quad (4.3)$$

$$= \arg \min_{\Delta X} Tr(\nabla f_p, \Delta X) + \frac{1}{2} Tr(\nabla^2 f_p(X, \mu) \delta X, \Delta X) \quad (4.4)$$

sujeto a las condiciones de factibilidad:

$$Tr(A_i \Delta X) = 0, \quad i = 1, 2, \dots, m$$

donde ∇f_p y $\nabla^2 f_p$ denotan el gradiente y el Hessiano de f_p respectivamente. Cabe aclarar, que hay un pequeño abuso de notación en la definición de ΔX , sin embargo, los términos ΔX que aparecen del lado derecho de la igualdad hacen referencia a la iteración anterior. A continuación se da la derivación de una expresión explícita de ΔX . Para ello, usando las Afirmaciones 4.1 y 4.2, se tiene que

$$\begin{aligned} \nabla f_p(X, \mu) &= \frac{1}{\mu} C - X^{-1} \\ \nabla^2 f_p(X, \mu) \Delta X &= X^{-1} \Delta X X^{-1} \quad \forall \Delta X \in S_n \end{aligned}$$

Substituyendo el gradiente y el Hessiano en la Ecuación 4.3:

$$\Delta X = \arg \min_{\Delta X} \left\{ Tr \left(\frac{C \Delta X}{\mu} \right) - Tr(X^{-1} \Delta X) + \frac{1}{2} Tr((X^{-1} \Delta X)^2) \right\}$$

sujeto a

$$\text{Tr}(A_i \Delta X) = 0, \quad i = 1, 2, \dots, m$$

Las condiciones de optimalidad para este problema son

$$\frac{1}{\mu} C - X^{-1} + X^{-1} \Delta X X^{-1} + \sum_{i=1}^m y_i A_i = 0 \quad (4.5)$$

$$\text{Tr}(A_i \Delta X) = 0, \quad i = 1, 2, \dots, m$$

Una simple manipulación en las condiciones de optimalidad dan como resultado

$$\text{vec} \left(X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \right) = - [I - \mathcal{A}_X^t (\mathcal{A}_X \mathcal{A}_X^t)^{-1}] \left(\text{vec} \left(\frac{1}{\mu} X^{\frac{1}{2}} C X^{-\frac{1}{2}} - I \right) \right) \quad (4.6)$$

donde \mathcal{A}_X es una matriz de $m \times n^2$ con vectores renglón $\left[\text{vec} \left(X^{\frac{1}{2}} A_j X^{\frac{1}{2}} \right) \right]$, para $j = 1, 2, \dots, m$. La expresión 4.6 es la proyección ortogonal del vector $\text{vec} \left(\frac{1}{\mu} X^{\frac{1}{2}} C X^{\frac{1}{2}} - I \right)$ en el espacio nulo de \mathcal{A}_X . Observe que el rango de \mathcal{A}_X está dado por el espacio generado de

$$\left\{ \text{vec} \left(X^{\frac{1}{2}} A_i X^{\frac{1}{2}} \right), i = 1, 2, \dots, m \right\}$$

y el espacio nulo es el complemento ortogonal. Retomando el espacio de matrices simétricas S_n , es claro que la dirección ΔX se obtiene de la proyección ortogonal de la matriz $\left[\frac{1}{\mu} X^{\frac{1}{2}} C X^{\frac{1}{2}} - I \right]$ en el complemento ortogonal del generado por

$$\{ X^{\frac{1}{2}} A_1 X^{\frac{1}{2}}, \dots, X^{\frac{1}{2}} A_m X^{\frac{1}{2}} \}$$

Considere la siguiente proyección

$$P_{\mathcal{A}_X} : S_n \rightarrow S_n$$

$$P_{\mathcal{A}_X}(M) = \arg \min_{W \in S_n} \{ \|W - M\| : \text{Tr}(X^{\frac{1}{2}} A_i X^{\frac{1}{2}} W) = 0, i = 1, 2, \dots, m \} \quad (4.7)$$

Se está en condiciones de expresar ΔX en términos de $S(X, \mu)$. Para ello, se tiene la siguiente

Afirmación 4.4. La dirección primal tiene las siguientes dos representaciones

$$\Delta X = -X^{\frac{1}{2}} \left(P_{\mathcal{A}_X} \left(\frac{X^{\frac{1}{2}} C X^{\frac{1}{2}}}{\mu} - I \right) \right) X^{\frac{1}{2}} = - \left(\frac{X S(X, \mu) X}{\mu} - X \right)$$

Demostración. Note que las condiciones de optimalidad para

$$\min_{W \in S_n} \left\{ \left\| W - \left(\frac{X^{\frac{1}{2}} C X^{\frac{1}{2}}}{\mu} - I \right) \right\| : Tr(X^{\frac{1}{2}} A_i X^{\frac{1}{2}} W) = 0, i = 1, 2, \dots, m \right\} \quad (4.8)$$

son

$$\left. \begin{aligned} W - \left(\frac{X^{\frac{1}{2}} C X^{\frac{1}{2}}}{\mu} - I \right) + \sum_{i=1}^m \xi_i X^{\frac{1}{2}} A_i X^{\frac{1}{2}} &= 0 \\ Tr(A_i X^{\frac{1}{2}} W X^{\frac{1}{2}}) &= 0, i = 1, 2, \dots, m \end{aligned} \right\} \quad (4.9)$$

Las condiciones de optimalidad para

$$\min_{y \in \mathbb{R}^m, S \in S_n} \left\{ \left\| \frac{X^{\frac{1}{2}} S X^{\frac{1}{2}}}{\mu} - I \right\| : \sum_{i=1}^m y_i A_i + S = C \right\} \quad (4.10)$$

pueden ser escritas como

$$\left. \begin{aligned} \frac{X S X}{\mu^2} - Q &= \frac{X}{\mu}, \\ Tr(A_i Q) &= 0, \quad i = 1, \dots, m, \\ \sum_{i=1}^m y_i A_i + S &= C \end{aligned} \right\} \quad (4.11)$$

donde $Q \in S_n$. Tomando la solución del Sistema 4.11 como $(y(X, \mu), S(X, \mu), Q(X, \mu))$, entonces

$$\xi(X, \mu) = \frac{1}{\mu} y(X, \mu) \quad \text{y} \quad W(X, \mu) = \mu X^{-\frac{1}{2}} Q(X, \mu) X^{-\frac{1}{2}}$$

satisface la primera ecuación del Sistema 4.9. La segunda ecuación del Sistema 4.11 muestra que

$$Tr(A_i X^{\frac{1}{2}} W(X, \mu) X^{\frac{1}{2}}) = 0, \quad i = 1, \dots, m$$

Así una solución óptima al Problema 4.8 puede ser construida a partir de una solución óptima del Problema 4.10. Como ambos problemas se resuelven a través de proyecciones ortogonales en conjuntos convexos, y tienen soluciones únicas, se sigue la equivalencia de las dos definiciones de ΔX . \square

El Lema 4.4 muestra la dirección ΔX puede ser separada en dos términos:

$$\Delta X = \frac{1}{\mu} \Delta X^a + \Delta X^c$$

donde

$$\Delta X^a = -X^{\frac{1}{2}} \left(P_{\mathcal{A}_X} \left(X^{\frac{1}{2}} C X^{\frac{1}{2}} \right) \right) X^{\frac{1}{2}}$$

y

$$\Delta X^c = X^{\frac{1}{2}} (P_{\mathcal{A}_X}(I)) X^{\frac{1}{2}}$$

Los términos ΔX^a y ΔX^c son llamados componentes *escalador afín* y *centrador* de la dirección. En la práctica, el cálculo de ΔX se lleva a cabo como a continuación se describe. Primero note que las condiciones de optimalidad descritas por las ecuaciones 4.11 pueden reescribirse como:

$$\sum_{i=1}^m \text{Tr}(X A_i X A_i) = \text{Tr}(X A_j X C) - \mu \text{Tr}(A_j X), \quad j = 1, \dots, m$$

Al solucionar este sistema de $m \times m$ se obtiene $y(X, \mu)$. La matriz de coeficientes $[\text{Tr}(X A_i X A_j)]$ del sistema lineal anterior es simétrica y positiva definida pues las matrices $A_i, i = 1, \dots, m$, forman un conjunto linealmente independiente (Afirmación 4.3). Tomando $S(X, \mu) = \sum_{i=1}^m y_i(X, \mu) A_i - C$, ΔX toma la siguiente expresión

$$\Delta X = -\frac{1}{\mu} X S(X, \mu) X + X$$

Considere ahora la actualización primal

$$X^+ = X + \Delta X = 2X - \frac{1}{\mu} X S(X, \mu) X \quad (4.12)$$

La pareja $(X^+, S(X, \mu))$ satisface la las igualdades primales y duales para el problema perturbado. Los siguientes resultados muestran que también se cumplen las condiciones del problema perturbado si X está suficientemente centrada.

Afirmación 4.5. Si $X \succ 0$ y $\delta(X, \mu) < 1$, entonces $S(X, \mu) \succ 0$

Demostración. Por definición

$$\begin{aligned}\delta(X, \mu)^2 &= \left\| \frac{X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}}}{\mu} - I \right\|^2 \\ &= \text{Tr} \left(\left(\frac{1}{\mu} X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} - I \right)^2 \right) \\ &= \sum_{i=1}^n \left(\frac{1}{\mu} \lambda_i (X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}}) - 1 \right)^2\end{aligned}$$

Usando $\delta(X, \mu) < 1$, se tiene

$$\begin{aligned}\sum_{i=1}^n \left(\frac{1}{\mu} \lambda_i (X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}}) - 1 \right)^2 < 1 &\Rightarrow \\ \lambda_i \left(X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} \right) > 0 \quad \forall i &\Rightarrow \\ S(X, \mu) \succ 0 &\end{aligned}$$

□

Con la siguiente afirmación quedará demostrado que X^+ es factible si X está suficientemente centrada.

Afirmación 4.6. Sea X^+ como la Ecuación en 4.12. Si $X \succ 0$ y $\delta(X, \mu) < 1$, entonces $X^+ \succ 0$

Demostración. Observe que X^+ puede reescribirse como

$$X^+ = X^{\frac{1}{2}} \left(2I - X^{\frac{1}{2}} \frac{S(X, \mu)}{\mu} X^{\frac{1}{2}} \right) X^{\frac{1}{2}} \quad (4.13)$$

Como

$$\delta(X, \mu) < 1$$

es decir

$$\left\| \frac{1}{\mu} X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} - I \right\| < 1$$

se sigue que

$$\sum_{i=1}^n \lambda_i^2 \left(\frac{X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}}}{\mu} - I \right) < 1$$

Por ello se tiene que

$$\lambda_i \left(\frac{1}{\mu} X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} \right) \in (0, 2) \quad \forall i$$

lo cual implica que

$$\lambda_i \left(2I - X^{\frac{1}{2}} \frac{S(X, \mu)}{\mu} X^{\frac{1}{2}} \right) \in (0, 2) \quad \forall i$$

Y por lo tanto, $X^+ \succ 0$ por la expresión 4.13 □

Así queda demostrado para una solución factible primal suficientemente centrada, se garantiza que en la siguiente iteración, el punto encontrado será factible. Con ello queda descrito el algoritmo. En la siguiente sección se discutirá qué tan rápido se converge al óptimo y su desempeño en el peor escenario.

4.3. Convergencia y eficiencia

El objetivo de este apartado es demostrar que el algoritmo converge a un ϵ -óptimo a lo más en $O(\sqrt{n} \ln(\frac{1}{\epsilon}))$ iteraciones para un punto inicial suficientemente centrado. Esta idea de "suficientemente centrado" quedará clara al final de esta sección. Antes de dar la primera afirmación, recuerde la desigualdad de Cauchy-Schwartz: Sean V un espacio vectorial sobre \mathbb{R} con un producto escalar (denotado por $\langle \cdot, \cdot \rangle$) y cuya norma es inducida por ese producto escalar; $v_1, v_2 \in V$; entonces

$$|\langle v_1, v_2 \rangle| \leq \|v_1\| \|v_2\|$$

Afirmación 4.7. Sean $X \succ 0$ y $\delta(X, \mu) < 1$. Entonces, con cada actualización X^+ se satisface que $\delta(X^+, \mu) \leq \delta(X, \mu)$.

Demostración. Por definición

$$\begin{aligned}\delta(X^+, \mu)^2 &= \left\| \frac{X^{+\frac{1}{2}} S(X^+, \mu) X^{+\frac{1}{2}}}{\mu} - I \right\|^2 \\ &\leq \left\| \frac{X^{+\frac{1}{2}} S(X, \mu) X^{+\frac{1}{2}}}{\mu} - I \right\|^2 \\ &= \text{Tr} \left(\left(\frac{1}{\mu} S(X, \mu) X^+ - I \right)^2 \right)\end{aligned}$$

Substituyendo el valor de X^+ se tiene

$$\begin{aligned}\delta(X^+, \mu)^2 &\leq \text{Tr} \left(\left(\frac{1}{\mu} S(X, \mu) \left[2X - \frac{1}{\mu} X S(X, \mu) X \right] - I \right)^2 \right) \\ &= \text{Tr} \left(\left(\frac{1}{\mu} S(X, \mu) X - I \right)^4 \right) \\ &\leq \text{Tr} \left(\left(\frac{1}{\mu} S(X, \mu) X - I \right)^2 \right)^2 \\ &= \left\| \frac{1}{\mu} X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}} - I \right\|^4 = \delta(X, \mu)^4\end{aligned}$$

De donde la última desigualdad se sigue de la desigualdad de Cauchy-Schwartz \square

Una vez que se determina una X suficientemente centrada, es decir, que $\delta(X, \mu) < \tau$ para alguna τ , se puede reducir el parámetro μ . Todo ello con el objetivo de disminuir el valor de $\delta(X, \mu)$ con cada iteración. Formalmente, se tienen los siguientes dos resultados.

Afirmación 4.8. Sean $X \in S_n$ y tome $\mu^+ = (1 - \theta)\mu$, con $0 < \theta < 1$. Entonces

$$\delta(X, \mu^+) \leq \frac{1}{1 - \theta} (\delta(X, \mu) + \theta\sqrt{n})$$

Demostración. Usando la definición de $S(X, \mu^+)$ se tiene

$$\begin{aligned} \delta(X, \mu^+) &= \left\| \frac{X^{\frac{1}{2}} S(X, \mu^+) X^{\frac{1}{2}}}{(1-\theta)\mu} - I \right\| \\ &\leq \left\| \frac{X^{\frac{1}{2}} S(X, \mu) X^{\frac{1}{2}}}{(1-\theta)\mu} - I \right\| \\ &\leq \frac{1}{1-\theta} \left(\left\| \frac{X^{\frac{1}{2}} S(X, \mu^+) X^{\frac{1}{2}}}{\mu} - I \right\| + \theta \|I\| \right) \\ &= \frac{1}{1-\theta} (\delta(X, \mu) + \theta\sqrt{n}) \end{aligned}$$

□

El resultado anterior permite tener un parámetro θ que garantiza que con cada iteración el nuevo punto permanezca suficientemente centrado con respecto al nuevo parámetro μ^+ . El siguiente resultado garantiza que, con una iteración en μ y en X , el nuevo punto también queda suficientemente centrado.

Afirmación 4.9. Sea $\delta(X, \mu) \leq \frac{1}{2}$ y $\theta = \frac{1}{4\sqrt{n+2}}$. Después de una iteración X^+, μ^+ , se tiene $\delta(X^+, \mu^+) \leq \frac{1}{2}$

Demostración. Usando las afirmaciones 4.7 y 4.8 se tiene

$$\begin{aligned} \delta(X^+, \mu^+) &\leq \frac{1}{1-\theta} (\delta(X^+, \mu) + \theta\sqrt{n}) \\ &\leq \frac{1}{1-\theta} (\delta(X^+, \mu)^2 + \theta\sqrt{n}) \end{aligned}$$

substituyendo del valor de θ

$$\delta(X^+, \mu^+) \leq \frac{4\sqrt{n} + 2}{4\sqrt{n} + 1} \left(\frac{1}{4} + \frac{\sqrt{n}}{4\sqrt{n} + 2} \right) = \frac{1}{2}$$

Lo cual completa la prueba. □

Para poder probar definitivamente la complejidad polinomial del algoritmo, se requiere del siguiente resultado que acota el intervalo dual en términos de δ .

Afirmación 4.10. Si $\delta(X, \mu) \leq 1$, entonces

$$\mu (n - \delta(X, \mu)\sqrt{n}) \leq \text{Tr}(CX) - b^t y(X, \mu) \leq \mu (n + \delta(X, \mu)\sqrt{n})$$

Demostración. Observe que

$$\text{Tr}(CX) - b^t y(X, \mu) = \text{Tr}(XS(X, \mu)) = \text{Tr}(X^{\frac{1}{2}}S(X, \mu)X^{\frac{1}{2}})$$

Usando la desigualdad de Cauchy-Schwartz se tiene

$$\delta(X, \mu)\sqrt{n} = \left\| \frac{X^{\frac{1}{2}}S(X, \mu)X^{\frac{1}{2}}}{\mu} - I \right\| \|I\| \geq \left| \frac{\text{Tr}(XS(X, \mu))}{\mu} - n \right|$$

Lo cual implica

$$n - \delta(X, \mu)\sqrt{n} \leq \frac{\text{Tr}(XS(X, \mu))}{\mu} \leq n + \delta(X, \mu)\sqrt{n}$$

Multiplicando todo por μ se completa la prueba. \square

Se está en condiciones de dar la cota para el peor caso del algoritmo a través del siguiente

Teorema 4.1. Sea $\epsilon > 0$, $\theta = \frac{1}{4\sqrt{n+2}}$, $\mu_0 > 0$, $X_0 \succ 0$ un punto inicial estrictamente factible tal que $\delta(X_0, \mu_0) \leq \frac{1}{2}$. Entonces el algoritmo terminará en a lo más $\lceil 6\sqrt{n} \ln \frac{n\mu_0}{\epsilon} \rceil$ iteraciones, el último punto generado X y $S(X, \mu)$ serán estrictamente factibles y el intervalo dual estará acotado por $\text{Tr}(XS(X, \mu)) \leq \frac{3}{2}\epsilon$.

Demostración. Después de cada iteración el nuevo punto encontrado será estrictamente factible y $\delta(X, \mu) \leq \frac{1}{2}$, por la Afirmación 4.9. En k iteraciones se tendrá $\mu = (1 - \theta)^k \mu_0$. Las iteraciones se detienen si k es tal que

$$n\mu_0(1 - \theta)^k < \epsilon$$

Tomando logaritmo natural de ambos lados de la desigualdad, la expresión se convierte en

$$-k \ln(1 - \theta) > \ln \frac{n\mu_0}{\epsilon}$$

Como $-\ln(1 - \theta) > \theta$, la desigualdad se mantiene si

$$k\theta > \ln \frac{n\mu_0}{\epsilon}$$

lo cual implica

$$k > 6\sqrt{n} \ln \frac{n\mu_0}{\epsilon}$$

para θ como en las hipótesis. Esto prueba la primera parte del teorema. Ahora, sea X el último punto generado por el algoritmo. Por la Afirmación 4.5 $S(X, \mu) \succ 0$. Más aún, el intervalo dual está acotado por

$$\begin{aligned} Tr(XS(X, \mu)) &\leq n\mu \left(1 + \frac{\delta(X, \mu)}{\sqrt{n}} \right) \\ &\leq \epsilon \left(1 + \frac{\delta(X, \mu)}{\sqrt{n}} \right) \leq \frac{3}{2}\epsilon \end{aligned}$$

Donde la primera desigualdad se sigue de la Afirmación 4.10. □

Con este último resultado, queda sustentado que este algoritmo es más rápido que el Algoritmo Simplex o que el algoritmo de punto interior de Karmarkar. Es importante notar, que el alcance de la PSD no se limita a problemas lineales, si no que se pueden resolver más generales. Así, la PSD es una herramienta poderosa tanto por su alcance, como por su eficiencia algorítmica.

Ejemplo

Recuerde el problema 1.6 del primer capítulo:

$$\begin{aligned} \max \quad & 2^{n-1}x_1 + 2^{n-2}x_2 + \dots + 2x_{n-1} + x_n \\ \text{sujeto a:} \quad & \\ & x_1 \leq 5 \\ & 4x_1 + x_2 \leq 25 \\ & 8x_1 + 4x_2 + x_3 \leq 125 \\ & \vdots \\ & 2^n x_1 + 2^{n-1}x_2 + \dots + 4x_{n-1} + x_n \leq 5^n \\ & x_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned}$$

A continuación se resolverá tanto con el Método Simplex, como con el método expuesto en este capítulo. Se resolverá primero con el Método Simplex. Esta es la secuencia de pivotes para $n = 3$, que pasa por los 8 puntos extremos del poliedro, empezando en el origen. Las variables de holgura serán x_4, x_5 y x_6 .

Tabla inicial

	x_1	x_2	x_3	
x_4	1*			5
x_5	4	1		25
x_6	8	4	1	125
$-z$	4	2	1	0

Tabla 1

	x_4	x_2	x_3	
x_1	1			5
x_5	-4	1*		5
x_6	-8	4	1	85
$-z$	-4	2	1	-20

Tabla 2

	x_4	x_5	x_3	
x_1	1*			5
x_2	-4	1		5
x_6	8	-4	1	65
$-z$	4	-2	1	-30

Tabla 3

	x_1	x_5	x_3	
x_4	1			5
x_2	4	1		25
x_6	-8	-4	1*	25
$-z$	-4	-2	1	-50

Tabla 4

	x_1	x_5	x_6	
x_4	1*			5
x_2	4	1		25
x_3	-8	-4	1	25
$-z$	4	2	-1	-75

Tabla 5

	x_4	x_5	x_6	
x_1	1			5
x_2	-4	1*		5
x_3	8	-4	1	65
$-z$	-4	2	-1	-95

Tabla 6

	x_4	x_2	x_6	
x_1	1*			5
x_5	-4	1		5
x_3	-8	4	1	85
$-z$	4	-2	-1	-105

Tabla 7

	x_1	x_5	x_3	
x_4	1			5
x_5	4	1		25
x_3	8	4	1	125
$-z$	-4	-2	-1	-125

Esta es la última tabla con el valor óptimo $5^3 = 125$. Para resolver este mismo problema usando las herramientas de la programación semidefinida, considere el problema primal escrito en su forma estándar:

$$\begin{aligned}
 \min \quad & -4x_1 - 2x_2 - x_3 \\
 \text{sujeto a:} \quad & \\
 & x_1 + x_4 = 5 \\
 & 4x_1 + x_2 + x_5 = 25 \\
 & 8x_1 + 4x_2 + x_3 + x_6 = 125 \\
 & x_i \geq 0 \quad i = 1, 2, \dots, 6
 \end{aligned}$$

El cual tiene su correspondiente dual asociado:

$$\begin{aligned}
 \max \quad & 5y_1 + 25y_2 + 125y_3 \\
 \text{sujeto a:} \quad & \\
 & y_1 + 4y_2 + 8y_3 + s_1 = -4 \\
 & \quad y_2 + 4y_3 + s_2 = -2 \\
 & \quad \quad y_3 + s_3 = -1 \\
 & y_1 + s_4 = 0 \\
 & \quad y_2 + s_5 = 0 \\
 & \quad \quad y_3 + s_6 = 0
 \end{aligned}$$

$$y \in \mathbb{R}^3 \quad s_j \geq 0 \quad j = 1, 2, \dots, 6$$

Observe que los problemas Primal y Dual están escritos de las formas:

$$(P) \quad \min \text{Tr}(CX)$$

sujeto a: $AX = b$
 $X \succeq 0$

Y su dual

$$(D) \quad \max b^t y$$

sujeto a: $A^t \text{diag}(y) + S = C$
 $S \succeq 0, y \in \mathbb{R}^n$

Con:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 & 1 & 0 \\ 8 & 4 & 1 & 0 & 0 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 5 \\ 25 \\ 125 \end{bmatrix}$$

$$S = \text{diag}\{(s_1, s_2, s_3, s_4, s_5, s_6)\} \quad X = \text{diag}\{(x_1, x_2, x_3, x_4, x_5, x_6)\}$$

$$C = \text{diag}\{(-4, -2, -1, 0, 0, 0)\}$$

Tome ahora los problemas perturbados como el Problema 3.11:

$$\begin{aligned} \min \quad & -4x_1 - 2x_2 - x_3 - \mu \ln(\det(X)) \\ \text{sujeto a:} \quad & \\ & x_1 + + + + + = 5 \\ & 4x_1 + x_2 + + + + = 25 \\ & 8x_1 + 4x_2 + x_3 + + + = 125 \\ & x_i \geq 0 \quad i = 1, 2, \dots, 6 \end{aligned}$$

Y el dual:

$$\begin{aligned} \max \quad & 5y_1 + 25y_2 + 125y_3 + \mu \ln(\det(S)) \\ \text{sujeto a:} \quad & \\ & y_1 + 4y_2 + 8y_3 + s_1 = -4 \\ & + y_2 + 4y_3 + + s_2 = -2 \\ & + + y_3 + + + s_3 = -1 \\ & y_1 + + + + + + s_4 = 0 \\ & + y_2 + + + + + + s_5 = 0 \\ & + + y_3 + + + + + + s_6 = 0 \end{aligned}$$

$$y \in \mathbb{R}^3 \quad s_j \geq 0 \quad j = 1, 2, \dots, 6$$

Aprovechando que tanto X como S son matrices diagonales, se pueden reescribir los dos problemas anteriores como:

$$\begin{aligned} \min \quad & -4x_1 - 2x_2 - x_3 - \mu \sum_{i=1}^6 \ln x_i \\ \text{sujeto a:} \quad & x_1 + + + + + = 5 \\ & 4x_1 + x_2 + + + + = 25 \\ & 8x_1 + 4x_2 + x_3 + + + = 125 \\ & x_i \geq 0 \quad i = 1, 2, \dots, 6 \end{aligned}$$

Y el dual:

$$\begin{aligned} \max \quad & 5y_1 + 25y_2 + 125y_3 + \mu \sum_{i=1}^6 \ln s_i \\ \text{sujeto a:} \quad & y_1 + 4y_2 + 8y_3 + s_1 = -4 \\ & + y_2 + 4y_3 + + s_2 = -2 \\ & + + y_3 + + + s_3 = -1 \\ & y_1 + + + + + + s_4 = 0 \\ & + y_2 + + + + + + s_5 = 0 \\ & + + y_3 + + + + + + s_6 = 0 \\ & y \in \mathbb{R}^3 \quad s_j \geq 0 \quad j = 1, 2, \dots, 6 \end{aligned}$$

■ Primera iteración

Ahora, de acuerdo con el Teorema 4.1 se tiene que:

$$\theta = \frac{1}{4\sqrt{6} + 2} \approx 0.8476042360$$

Y sean

$$\begin{aligned} X_0 &= \text{diag}\{(2, 9, 9, 3, 8, 64)\} & y_0 &= (-1, -1, -2)^t \\ S_0 &= \text{diag}\{(20, 7, 1, 1, 1, 2)\} & \epsilon &= 4 \end{aligned}$$

un punto interior factible primal, y (S_0, y) un punto interior factible dual. Tomando $\mu_0 = 39$ se cumple que $\delta(X_0, \mu_0) \leq \frac{1}{2}$. Observe que por tratarse de un problema de programación lineal, las Condiciones 4.5 pueden escribirse como

$$\begin{aligned} c - \mu_0 X_0^{-1} \vec{1} + \mu_0 X^{-2} \Delta x - A^t y &= 0 \\ A \Delta x &= 0 \end{aligned}$$

Con Δx la dirección que se busca. Este sistema tiene la siguiente solución:

$$\Delta x = (I - X_0^2 A^t (A X_0^2 A^t)^{-1} A) \left(X_0 \vec{1} - \frac{1}{\mu_0} X_0^2 c \right) \approx \begin{bmatrix} -0.632769 \\ 0.311099 \\ 9.69507 \\ 0.632769 \\ 2.21998 \\ -5.87732 \end{bmatrix}$$

$$y = (A X_0^2 A^t)^{-1} A (X_0^2 c - \mu_0 X_0 \vec{1}) \approx \begin{bmatrix} -10.258 \\ -3.5222 \\ -0.665336 \end{bmatrix}$$

Se hace las actualizaciones de las variables primales y duales:

$$X_1 = X_0 + \text{diag}(\Delta x) \approx \begin{bmatrix} 1.36723 & 0 & 0 & 0 & 0 & 0 \\ 0 & 9.3111 & 0 & 0 & 0 & 0 \\ 0 & 0 & 18.6951 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.63277 & 0 & 0 \\ 0 & 0 & 0 & 0 & 10.22 & 0 \\ 0 & 0 & 0 & 0 & 0 & 58.1227 \end{bmatrix}$$

$$y_1 = y; \quad S_1 = \text{diag}(c - A^t y) \approx \text{diag}(25.6695, 4.18355, -0.334664, 10.258, 3.5222, 0.665336)$$

$$\mu_1 = (1 - \theta)^1 \mu_0 \approx 35.6943$$

■ Segunda Iteración

$$\Delta x = (I - X_1^2 A^t (A X_1^2 A^t)^{-1} A) \left(X_1 \vec{1} - \frac{1}{\mu_1} X_1^2 c \right) \approx \begin{bmatrix} 0.59633 \\ -1.18689 \\ 20.7668 \\ -0.059633 \\ 0.948359 \\ -16.4963 \end{bmatrix}$$

$$y = (A X_1^2 A^t)^{-1} A (X_1^2 c - \mu_1 X_1 \vec{1}) \approx \begin{bmatrix} -9.98695 \\ -3.16851 \\ -0.788419 \end{bmatrix}$$

$$X_2 = X_1 + \text{diag}(\Delta x) \approx \begin{bmatrix} 1.42686 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8.12421 & 0 & 0 & 0 & 0 \\ 0 & 0 & 39.4619 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.57314 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11.1683 & 0 \\ 0 & 0 & 0 & 0 & 0 & 41.6264 \end{bmatrix}$$

$$y_2 = y; \quad S_2 \approx \text{diag}(24.9683, 4.32219, -0.211581, 9.98695, 3.16851, 0.788419)$$

$$\mu_2 = (1 - \theta)^2 \mu_1 \approx 29.8998$$

- Tercera Iteración

$$\Delta x = (I - X_2^2 A^t (A X_2^2 A^t)^{-1} A) \left(X_2 \vec{1} - \frac{1}{\mu_2} X_2^2 c \right) \approx \begin{bmatrix} -0.054314 \\ -2.11189 \\ 32.715 \\ 0.054314 \\ 2.32915 \\ -23.8329 \end{bmatrix}$$

$$y = (A X_2^2 A^t)^{-1} A (X_2^2 c - \mu_2 X_2 \vec{1}) \approx \begin{bmatrix} -8.24076 \\ -2.11887 \\ -1.12954 \end{bmatrix}$$

$$X_3 = X_2 + \text{diag}(\Delta x) \approx \begin{bmatrix} 1.37255 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6.01232 & 0 & 0 & 0 & 0 \\ 0 & 0 & 72.1769 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.62745 & 0 & 0 \\ 0 & 0 & 0 & 0 & 13.4975 & 0 \\ 0 & 0 & 0 & 0 & 0 & 17.7935 \end{bmatrix}$$

$$y_3 = y; \quad S_3 \approx \text{diag}(21.7526, 4.63705, 0.129543, 8.24076, 2.11887, 1.12954)$$

$$\mu_3 = (1 - \theta)^3 \mu_2 \approx 29.9231$$

■ Cuarta Iteración

$$\Delta x = (I - X_3^2 A^t (A X_3^2 A^t)^{-1} A) \left(X_3 \vec{1} - \frac{1}{\mu_3} X_3^2 c \right) \approx \begin{bmatrix} -0.130851 \\ -1.29012 \\ 6.23325 \\ 0.130851 \\ 1.81352 \\ -0.025972 \end{bmatrix}$$

$$y = (A X_3^2 A^t)^{-1} A (X_3^2 c - \mu_3 X_3 \vec{1}) \approx \begin{bmatrix} -6.09139 \\ -1.47014 \\ -1.29017 \end{bmatrix}$$

$$X_4 = X_3 + \text{diag}(\Delta x) \approx \begin{bmatrix} 1.2417 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4.7222 & 0 & 0 & 0 & 0 \\ 0 & 0 & 78.4101 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.7583 & 0 & 0 \\ 0 & 0 & 0 & 0 & 15.311 & 0 \\ 0 & 0 & 0 & 0 & 0 & 17.7675 \end{bmatrix}$$

$$y_4 = y; \quad S_4 \approx \text{diag}(18.2933, 4.63081, 0.290168, 6.09139, 1.47014, 1.29017)$$

$$\mu_4 = (1 - \theta)^4 \mu_3 \approx 16.0847$$

Continuando las iteraciones como hasta ahora, en la décima iteración se tiene que:

$$X_{10} = X_9 + \text{diag}(\Delta x) \approx \begin{bmatrix} 0.116835 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.223724 & 0 & 0 & 0 & 0 \\ 0 & 0 & 122.515 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.88317 & 0 & 0 \\ 0 & 0 & 0 & 0 & 24.3089 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.655616 \end{bmatrix}$$

$$S_{10} = \begin{bmatrix} 4.314610 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2.05341 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.005912 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.148262 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.029764 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.00591 \end{bmatrix}$$

Así, $\|X_{10}S_{10}\| = 3.79476$. Así, se ha alcanzado el ϵ -óptimo. Observe que la tercera entrada de la diagonal de X_{10} se acerca al valor óptimo buscado de 125. El hecho de que el número de iteraciones sea parecido al número de iteraciones requeridas por el Método Simplex, se explica por el tamaño del problema. Se ha mostrado que para problemas grandes, la diferencia en iteraciones es observable.

Conclusiones

A pesar de existir herramientas más generales y eficientes que el Simplex, hoy en día sigue siendo el método más usado y enseñado. Algunas razones para seguir utilizando el Simplex son que:

1. Es relativamente fácil de entender.
2. Es fácil de programar en una computadora.
3. En la práctica, casi no se encuentran problemas que el Simplex resuelva en tiempo exponencial.

Sin embargo, los problemas de optimización requieren de algoritmos eficientes, debido al tamaño de los mismos. Con el paso del tiempo se han ido desarrollando otros métodos iterativos para encontrar óptimos a problemas con restricciones, y éstos han resultado ser más rápidos que el Simplex. En particular, la PSD abre nuevas posibilidades para resolver problemas muy complejos de manera rápida.

Bibliografía

- [1] Cárdenas, Humberto; Lluís, Emilio; Raggi, Francisco; Tomás, Francisco. *Álgebra Superior: conjuntos y combinatoria, introducción al álgebra lineal, estructuras numéricas, polinomios y ecuaciones*. 2a. Edición. Trillas. México 1990.
- [2] Friedberg H., Stephen; Insel J., Arnold; Spence E., Lawrence. *Algebra Lineal*. Primera edición. Publicaciones Cultural. Universidad Estatal de Illinois.
- [3] Montalvo Duran, Francisco. *Cálculo diferencial e integral en varias variables*. Departamento de Matemáticas de la Universidad de Extremadura 2003.
- [4] Haaser B., Norman; La Salle P., Joseph; Sullivan A., Joseph. *Análisis Matemático. Curso de introducción. Vol 2*. 2a. edición. Trillas 1990.
- [5] De Klerk, Etienne. *Interior Point Methods for Semidefinite Programming*. Tesis Doctoral, Universidad Técnica de Delf, Holanda 1997.
- [6] Padberg, Manfred. *Linear Optimization and Extensions*. 2a. Edición. Springer 1999.
- [7] Bazaraa, S. Mokhtar; Jarvis, J. John; Sherali, D. Hanif. *Linear Programming and Network Flows*. 2a. Edición. Wiley, Estados Unidos 2005.
- [8] Vanderbei, J. Robert. *Linear Programming: Foundations and Extensions*. 2a. Edición. Department of Operations Research and Financial Engineering, Princeton University. Princeton 2001.
- [9] Dasgupta, S; Papadimitriou, C.H.; Vazirani, U.V. *Algorithms* 2006