

# Métodos Numéricos para la solución de Problemas de Programación Lineal: Puntos Interiores

Adrián Tame Jacobo

Santiago Córdova López

Estudiantes de Matemáticas Aplicadas del ITAM

*"Numerical Analysis is often cosidered neither beautiful nor, indeed, profound. [...] This, I belive is nonsense. A mathematical problem does not cease being mathematical just beacuse we have discretzed it. The purpose of dicretization is to render mathematical problems, often approximately, in a form accesible to efficient calculation by computers.[...]"* Arieih Iserles

Antes que nada, agradecemos a Alberto Albarran, Patricio Davila, Mariana Prud'Homme y Martín Sánchez por ayudarnos a desarrollar y permitirnros publicar lo escrito a continuación.

Un problema de optimización es cuando se busca minimizar/maximizar una función objetivo sujeta a una serie de restricciones en el dominio de la función. Además, si dicha función y restricciones son lineales entonces, el problema anterior se declara como un **Problema de Programación Lineal** (PL); debido a la cantidad de representaciones existentes de un PL es necesario estandarizar la definición, es por eso que a continuación damos la definición que usaremos junto con definiciones útiles para las soluciones.

Sea  $x \in \mathbb{R}^n$ , y sean  $c \in \mathbb{R}^n$ ;  $b \in \mathbb{R}^m$  y  $A \in \mathcal{M}_{m \times n}(\mathbb{R})$  fijos, definimos un Problema de Programación Lineal en su forma estándar (PLE)

$$\begin{array}{ll} \text{minimizar} & c^T x. \\ \text{Sujeto a} & Ax \leq b, \\ \text{y} & x \geq 0. \end{array} \quad (\text{PLE})$$

Se observa que, las restricciones lineales forman un poliedro convexo en el espacio, es por eso que: todo punto en el poliedro y su frontera se dice **solución factible**, por lo que toda la región se dice **conjunto factible**, denotado como  $C_f$ . Y todo punto que sea vértice de dicho poliedro se dice **solución básica factible**.

**Teorema 1** (Existencia de soluciones óptimas). *Toda solución óptima es una solución básica factible; geométricamente, toda solución óptima es un vértice del conjunto factible. Además, existen dos casos para que las soluciones óptimas no existan: 1) el conjunto factible es vacío y 2) es no acotado; por lo que no hay forma de encontrar una mejor solución.*

Se puede observar que la estructura de  $A$ ,  $b$ ,  $c$ , y la dimensión del problema presentan que tan "fácil" es resolver el problema analíticamente pero, como no podemos depender de esto en general entonces, la mayoría de estos problemas se resuelven numéricamente. A lo largo de la historia han existido varias formas de resolver el PLE numéricamente, que

se dividen a gran escala en: algoritmos tipo Simplex y métodos de Puntos Interiores. El propósito de nuestro artículo es introducir un método de Puntos Interiores que combina *path following* y *affine scaling*.

La idea de un método tipo Simplex, propuesto por George Dantzing en 1947, es aprovechar el teorema 1 y visitar inteligentemente los vértices del poliedro para llegar a una solución óptima o declarar que estas no existen<sup>1</sup>.

## Dualidad

A la teoría de PL se le anida el concepto de Dualidad, propuesto por John Von Neumann. La idea es poder ver el PLE de dos perspectivas: el **problema primal** (P) y el **problema dual** (D). Normalmente, el PLE es definido como el **primal** y la definición de su **dual** se presenta a continuación.

**Definición 1** (Problema Primal y Dual). <sup>2</sup> Dado el siguiente **problema primal** (P)

$$\begin{array}{ll} \text{minimizar} & c^T x. \\ \text{Sujeto a} & Ax \leq b, \\ y & x \geq 0. \end{array} \quad (P)$$

Se define su **problema dual** (D) como:

$$\begin{array}{ll} \text{maximizar} & b^T \lambda. \\ \text{Sujeto a} & A^T \lambda \geq c, \\ y & \lambda \geq 0; \lambda \in \mathbb{R}^m. \end{array} \quad (D)$$

La ventaja de tener un PL visto como un problema primal y dual radica en los teoremas conocidos como: Dualidad Fuerte, nos da una relación entre las soluciones del problema Primal y el Dual; Dualidad de Programas Lineales, nos da una extensión del teorema 1 que combina las soluciones del problema Primal y el problema Dual; y el Lema de Farkas, nos da un margen teórico para la caracterización de los problemas mencionados anteriormente y la extensión a problemas no lineales.

**Teorema 2** (Dualidad Fuerte). Sea un PL descrito como la definición 1 además, sean  $x \in C_f(\text{Primal})$ , y  $\lambda \in C_f(\text{Dual})$  tales que  $c^T x = \lambda^T b$ . Entonces,  $x$  es solución óptima del problema primal (P), y  $\lambda$  es solución óptima del problema dual (D).

**Teorema 3** (Dualidad de Programaciones Lineales). Sea un PL descrito como la definición 1, entonces sucede sólo uno de los siguientes escenarios:

1. P y D no tienen solución factible.

---

<sup>1</sup>En [1, Capítulo 13] se trata a más profundidad el método Simplex.

<sup>2</sup>El primal y dual de un PL necesitan ser definidos en dupla, ya que no tienen sentido separados, pues el concepto es relativo. Por lo anterior la definición involucra a los dos a pesar de que el PLE fue definido anteriormente como primal.

2. Si  $P$  es no acotado, entonces  $C_f(\text{Primal}) = \emptyset$ .
3. Si  $D$  es no acotado, entonces  $C_f(\text{Dual}) = \emptyset$ .
4.  $P$  tiene solución óptima, si y sólo si  $D$  tiene solución óptima. Adicionalmente, se cumple el teorema 2.

Con los teoremas anteriores, se modifica el método Simplex de tal forma que el problema se logra resolver de forma iterativa mediante la búsqueda alterna de las soluciones óptimas de  $P$  y  $D$ . A los algoritmos que emplean esta modificación se les conoce como métodos Primal-Dual, en concreto existe el método Simplex Primal-Dual<sup>3</sup>.

## Puntos Interiores

Los algoritmos de puntos interiores pertenecen a una clase diferente de algoritmos de optimización, esto ya que no son necesariamente lineales. Una diferencia importante es que los métodos de puntos interiores son mucho más costosos, relativamente, al método Simplex, pero hace un progreso mucho más significativo a cada iteración en llegar a una solución considerada óptima.

Lo anterior se debe a que el método Simplex aproxima una solución buscando sobre la frontera de la región factible, los métodos de puntos interiores pueden llegar a la solución desde dentro o fuera del conjunto factible, pero nunca sobre la frontera de la región. Al algoritmo base de puntos interiores se le llama Primal-Dual, ya que utiliza condiciones de ambos problemas de optimización para generar las condiciones necesarias y aproximar la solución óptima.

Antes de empezar con la descripción del método, transformaremos el  $D$  de tal forma que las restricciones sean igualdades. Lo anterior se logrará introduciendo una variable de holgura (*slackness variable*),  $z$ . La variable de holgura  $z$ , esta en  $\mathbb{R}^n$ , es no negativa, y es tal que  $A^T \lambda + z = b$ . Por lo tanto el  $D$ , es equivalente al siguiente PL:

$$\begin{array}{ll} \text{minimizar} & b^T \lambda. \\ \text{Sujeto a} & A^T \lambda + z = c, \\ & \lambda \geq 0, \\ \text{y} & z \geq 0, z \in \mathbb{R}^n. \end{array} \quad (\text{D con holgura})$$

**Definición 2** (Puntos Interiores). Se dice que  $(x, \lambda, z) \in \mathbb{R}^{n+m+n}$  es un punto interior, del PL, si y solo si sucede

$$x \in \Omega^0 := \{(x, \lambda, z) \in \mathbb{R}^{n+m+n} : Ax = b, A^T \lambda + z = c, (x, z) > 0\}.$$

Además, a  $\Omega^0$  se le conoce como **conjunto de puntos interiores**, asociado al PL.

<sup>3</sup>En [1, Capitulo 13] se trata a más profundidad el método Simplex Primal-Dual.

## Metodología Básica de Puntos Interiores

Primero, considérese el problema base de programación lineal,  $(P)$ , y su dual asociado, definido con holgura,  $(D \text{ con holgura})$ . De estos problemas tomamos las condiciones de Karush-Kuhn-Tucker asociadas,

$$F(x, \lambda, z) = \begin{pmatrix} A^T \lambda + z - c \\ Ax - b \\ XZe \end{pmatrix} = 0, \quad (\text{KKK})$$

$$(x, z) \geq 0.$$

Donde,  $X$  y  $Z$  son las matrices diagonales de dimensión  $n \times n$  generadas por el vector  $x$  y  $z$ , respectivamente. Es decir,  $X = (X_{(i,j)})$  con  $X_{(i,j)} := \begin{cases} x_i & \text{si } i = j, \\ 0 & \text{en otro caso.} \end{cases}$  Y de forma análoga para  $Z$ .

Los métodos de puntos interiores generan una solución por iteración de la forma  $(x^k, \lambda^k, z^k)$  tales que  $(x^k, z^k) > 0$ . Esta es la razón por la cual al método se le llama de puntos interiores, ya que *por estas desigualdades estrictas, las soluciones nunca estarán en la frontera de la región factible, y estarán en el interior del conjunto factible*.

## Método de Newton

El Teorema de Taylor nos dice que si una función es dos veces diferenciable, entonces

$$\nabla f(x + p) = \nabla f(x) + \int_{\mathcal{D}} \nabla^2 f(x + tp) p \, dt$$

y también

$$f(x + p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x + tp) p$$

para alguna  $t \in (0, 1)$ . El paso de Newton es el iterado que minimiza este modelo en términos de  $x$ . Es interesante notar que este icónico teorema es para una función, pero fácilmente se puede extender a múltiples funciones, y con este nuevo resultado, obtendremos lo necesario para presentar el algoritmo de Newton utilizado en el método de puntos interiores.

**Teorema 4.** Sea  $r : \mathbb{R}^n \rightarrow \mathbb{R}^n$  una función en un espacio convexo y diferenciable  $\mathcal{D}$  y sean  $x$  y  $x + p$  vectores en  $\mathcal{D}$ . Entonces,

$$r(x + p) = r(x) + \int_{\mathcal{D}} J(x + tp) p \, dt.$$

Tomando una aproximación de  $\int_{\mathcal{D}} J(x + tp) p \, dt$  como  $J(x)p$ , podemos definir un modelo lineal de  $r(x_k + p)$  como  $M_k(p)$ , y se obtiene

$$M_k(p) = r(x_k) + J(x_k)p.$$

En su forma más sencilla, la iteración de Newton toma un paso para el cual  $M_k(p_k) = 0$ .

### Condiciones de Optimalidad Karush Khun Tucker

Luego, las soluciones  $(x, \lambda, z)$  son caracterizadas por las condiciones de Karush–Kuhn–Tucker (**KKT**). Los métodos primal-dual para puntos interiores difieren de los métodos primales en que la dirección de Newton es calculada para  $(\lambda, z)$  así como para  $x$ . Para calcular la dirección de Newton, aplicamos las restricciones de **KKT** para el problema dual

$$F(x, \lambda, z) = \begin{pmatrix} A^T \lambda + z - c \\ Ax - b \\ XZe \end{pmatrix} = 0.$$

Así, la matriz Jacobiana  $F$  puede ser escrita como la matriz:

$$J(x, \lambda, z) = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix}.$$

Así, la solución del siguiente sistema de ecuaciones es nuestra nueva dirección de Newton  $(x_B, \lambda_B, z_B)$

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} x_B \\ \lambda_B \\ z_B \end{pmatrix} = \begin{pmatrix} -(A^T \lambda + z - c) \\ -(Ax - b) \\ -XZe \end{pmatrix}.$$

Para ver dónde está la barrera, usamos de nuevo las condiciones **KKT** asociadas a la función barrera. Vemos que son casi las mismas condiciones de **KKT** para el problema de programación lineal, sólo por el término  $XSe = \mu e$ . Ahora, resolvemos las ecuaciones de Newton:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ z & 0 & X \end{pmatrix} \begin{pmatrix} x_B \\ \lambda_B \\ z_B \end{pmatrix} = \begin{pmatrix} -(A^T \lambda + z - c) \\ -(Ax - b) \\ -(XZe - \mu e) \end{pmatrix}.$$

De lo cual obtenemos

$$(x_{k+1}, \lambda_{k+1}, z_{k+1}) = (x_k, \lambda_k, z_k) + \alpha(x_B, \lambda_B, z_B),$$

donde  $(x_B, \lambda_B, z_B)$  es nuestra nueva dirección de Newton.

### Path Following

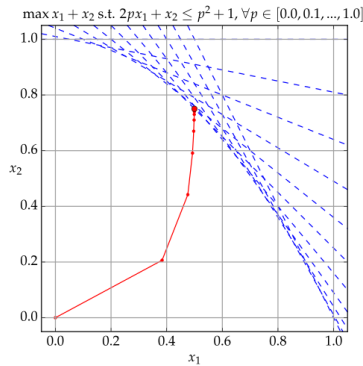
La trayectoria que sigue el algoritmo general es la **trayectoria central**. Para poder describirla, y mejor explicar los conceptos ya presentados, definimos la región factible y la región estrictamente factible.

$$\mathcal{F} = \{(x, \lambda, z) | Ax = b, A^T \lambda - z = c, (x, z) \geq 0\}$$

$$\mathcal{F}^0 = \{(x, \lambda, z) | Ax = b, A^T \lambda - z = c, (x, z) > 0\}$$

Ahora, podemos definir la dirección central. La dirección central  $\mathcal{C}$  es una colección de puntos  $(x_\tau, \lambda_\tau, z_\tau)$  que siguen un arco a través de la región factible estricta  $\mathcal{F}^0$ . Todos estos puntos satisfacen unas condiciones que están parametrizadas por un escalar  $\tau > 0$ .

Figura 1: Ejemplo del método. Las líneas azules son las restricciones y la línea roja es la trayectoria que siguen las iteraciones del método de puntos interiores para llegar a aproximar la solución óptima. [4]



## Affine Scaling

Cualquier algoritmo básico de optimización sigue el mismo esbozo de como funciona: primero, se determina el paso que se da, y luego se determina una medida de optimalidad de la iteración. La dirección afin (*affine scaling direction*) es la que se obtiene de resolver el sistema de ecuaciones asociado a las condiciones Karush-Kuhn-Tucker con el método de Newton. Este método se presenta a detalle en la siguiente sección. Generalmente no se utiliza el punto generado por el método de Newton exactamente, ya que esta puede rápidamente violar la condición necesaria de  $(x, z) < 0$ . Por esto, generalmente se toma un paso de Newton hacia un punto tal que  $x_i z_i = \sigma \mu$ , donde  $\mu = \frac{x^T z}{n}$  es la medida de dualidad, con  $n$  la dimensión del vector  $x$ , y  $\sigma \in [0, 1]$  un elemento reductor.

Notemos que entre los métodos de puntos interiores y el simplex, los primeros requieren un número polinomial de iteraciones, pero cada iteración puede ser costoso en términos de programación, por esto, calculando la dirección de Newton y una  $\alpha$  óptima para la longitud de los pasos son operaciones no triviales. Por otro lado, cada iteración del método simplex es relativamente fácil de calcular, pero esto necesita un número exponencial de iteraciones en el peor de los casos.

## Primal-Dual Predictor-Corrector Interior Point

Este método aprovecha el concepto de trayectoria central, resolviendo dos sistemas de direcciones de Newton el predictor y el corrector. Así las direcciones se mantienen en la trayectoria central, es decir cumpliendo las condiciones KKT, y dan pasos óptimos creados por el Affine Scaling.

La idea del método es generar una primera dirección afin, una predicción y dar un corrector

que centraliza la trayectoria primal,  $\tau_p$ , y otro corrector que centraliza la trayectoria dual,  $\tau_d$ , y termina en el cálculo de la  $\alpha$  para dar un nuevo punto.<sup>4</sup>

**Definición 3** (Notaciones del Algoritmo). Se define  $a$ ,  $r_p$  como el residuo primal dado por  $Ax - b$ , y  $r_d$  como el residuo dual dado por  $A^t\lambda + z - c$ .

Y se define  $diag(x)$ , con  $x$  un vector de dimensión  $k$ , se define como la matriz cuadrada con entradas  $(diag(x))_{i,j} = \begin{cases} x_i & \text{si } i = j, \quad \forall i, j = 1, \dots, k. \\ 0 & \text{e.o.c.} \end{cases}$

## Simplex vs PDPC

A continuación presentamos una comparación entre el método Simplex, programado con dos fases, contra PDC. Usamos el problema de Simple Klee-Minty definido en [3]. Además, debemos de ser sinceros y confesar que el problema anterior esta diseñado para que el Simplex trabaje en su peor caso.

El problema de Simple Klee-Minty es

$$\begin{aligned} &\text{minimizar} && c^T x. \\ &\text{Sujeto a} && Ax = b, \\ &&& y \quad x \geq 0. \end{aligned}$$

Con

$$\begin{aligned} c^T &= -(1, 1, \dots, 1) \\ A &= \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 2 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 2 & 2 & \dots & 2 & 1 \end{pmatrix} \\ b^T &= (2^1 - 1, 2^2 - 1, \dots, 2^n - 1) \end{aligned}$$

donde  $n$  es la dimensión del problema.

Este problema nos resulta interesante ya que la solución se conoce de [3] y es

$$x^* = (0, 0, \dots, 0, 2^n - 1)^T,$$

Los resultados son generados por rutinas hechas en MatLab, en una computadora con 16 Gb de memoria Ram y un procesador Intel64 Family 6 Model 158 Stepping 9 GenuineIntel~2808 Mhz, y son los siguientes<sup>5</sup>

<sup>4</sup>Una implementación del algoritmo está al final del artículo.

<sup>5</sup>Todos los tiempos están en segundos.

n = 10	Simplex	PDPC
Iteraciones	177	13
Tiempo Maquina	0.003320	0.020507

n = 20	Simplex	PDPC
Iteraciones	21891	18
Tiempo Maquina	0.196782	0.062952

n = 30	Simplex	PDPC
Iteraciones	2.69254e+06	22
Tiempo Maquina	24.187374	0.144416

Se omite el valor objetivo debido a que ambos métodos lograron darlo, y por eso no se consideró como una criterio de optimalidad en este ejemplo.

## Conclusiones

Se presentaron las ideas generales detrás de la teoría de Programación Lineal, se introdujo una metodología para resolver dichos problemas distinta del tradicional Simplex.

Debemos de hacer notar que este método tiene mayor importancia en problemas de optimización no lineales. Por lo anterior, podemos pensar que resolver PL's mediante este método es usar fuerza bruta. Sin embargo, el ejemplo de Klee-Minty, elegido de forma mal intencionada, nos muestra que existen PL's que no pueden ser resueltos de forma óptima, en tiempo, por el Simplex, por lo que tener otra herramienta es útil de forma practica y teórica.

Además, creemos en resultados recreables. Es decir, si el lector desea los códigos implementados y las rutinas en MatLab, o conoce alguna forma de optimizarlo, nos puede contactar a través de [santiago.cordova@itam.mx](mailto:santiago.cordova@itam.mx) o [adrian.tame.jacobo@gmail.com](mailto:adrian.tame.jacobo@gmail.com).

## Referencias

- [1] George Nocedal and S. Wright (2016). *Numerical Optimization*, 2<sup>nd</sup> ed. United States: Springer-Verlag New York.
- [2] S. Mehrotra (1992). On the Implementation of a Primal-Dual Interior Point Method. *SIAM Journal on Optimization* # 4, vol 2: 575-601. doi: 10.1137/0802028. <https://epubs.siam.org/doi/abs/10.1137/0802028>
- [3] Kitahara, Tomonari et al Mizuno, Shinji *Klee-Minty's LP and Upper Bounds for Dantzig's Simplex Method*
- [4] Karmarkar - Wikipedia - Interior Point Method [https://en.wikipedia.org/wiki/Interior-point\\_method#/media/File:Karmarkar.svg](https://en.wikipedia.org/wiki/Interior-point_method#/media/File:Karmarkar.svg)



**Algoritmo 1:** PDPC

**input** :  $A, c, b$ , y un punto de inicio  $(x_0, \lambda_0, z_0)$

**output**:  $(x, \lambda, z)$ , solución del problema

1 **while**  $\|Ax_k + -b\| > 0, \|A^T \lambda_k + z_k - c\| > 0$  y  $x_{k+1}^T z_{k+1} > 0$ . **do**

2     Encontrar las direcciones mediante un paso de Newton.

$$\begin{pmatrix} 0 & A^T & I & 0 \\ A & I & 0 & 0 \\ Z & 0 & X & 0 \end{pmatrix} \begin{pmatrix} d_x^{aff} \\ d_\lambda^{aff} \\ d_z^{aff} \end{pmatrix} = \begin{pmatrix} -r_d \\ -r_p \\ -XZe \end{pmatrix}$$

3     Se calculan los siguientes parámetros

$$D_x \leftarrow \text{diag}(d_x^{aff}).$$

$$D_z \leftarrow \text{diag}(d_z^{aff}).$$

$$D_\lambda \leftarrow \text{diag}(d_\lambda^{aff}).$$

$$\mu \leftarrow \frac{x^T z}{n}, \text{ conocida como la medida dual.}$$

$$\tau \leftarrow \left( \frac{(x + d_x)^T (z + d_z)}{n} / \frac{x^T z}{n} \right)^3, \text{ conocido como el corrector.}$$

4     Se resuelve el sistema con correcciones

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} d_x \\ d_\lambda \\ d_z \end{pmatrix} = \begin{pmatrix} -r_d \\ -r_p \\ -XZe + D_x D_z e + \tau \mu e \end{pmatrix}$$

5     Se determinan las  $\alpha$ 's

$$\alpha_x \leftarrow \min \left\{ 1, \eta \min \left\{ 1, \min_{i: (d_x)_i < 0} \left\{ -\frac{x_i}{(d_x)_i} \right\} \right\} \right\}$$

$$\alpha_z \leftarrow \min \left\{ 1, \eta \min \left\{ 1, \min_{i: (d_z)_i < 0} \left\{ -\frac{z_i}{(d_z)_i} \right\} \right\} \right\}$$

$$\alpha_\lambda \leftarrow \min \left\{ 1, \eta \min \left\{ 1, \min_{i: (d_\lambda)_i < 0} \left\{ -\frac{\lambda_i}{(d_\lambda)_i} \right\} \right\} \right\}$$

6      $(x_{k+1}, \lambda_{k+1}, z_{k+1}) \leftarrow (x_k, \lambda_k, z_k) + (\alpha_x d_x, \alpha_\lambda d_\lambda, \alpha_z d_z)$ .

7      $k \leftarrow k + 1$

8 **Regresar**  $(x_{k+1}, \lambda_{k+1}, z_{k+1})$