# UN ALGORITMO NO LINEAL PARA PROGRAMACION LINEAL

*Hector Jairo Martinez R.*
*Departamento de Matemáticas*
*Universidad del Valle*

**Abstract**

Morshedi and Tapia (1987) argue that Karmarkar algorithm results from a clever steepest descent formulation of the nonlinear program that arises when squared-slack variables are used. By replacing the steepest descent component with a succesive quadratic programming (SQP) component, we present a new algorithm and demostrate that it is locally q-quadratically convergent, and the variables that converge to zero do so q-superlinearly.

**Resumen**

En 1987, Morshedi y Tapia demostraron que el algoritmo de Karmarkar para programación lineal se puede deducir de una formulación especial del método del gradiente para programación no lineal. Reemplazando el método del gradiente con el método de programación cuadrática sucesiva (PCS), presentamos un nuevo algoritmo para programación lineal y demostramos que es localmente convergente con una rata de convergencia q-cuadrática, además, que las variables que convergen a cero lo hacen con una rata q-superlineal.

## INTRODUCCION

En 1987, Morshedi y Tapia demostraron que el algorítmo de Karmarkar (1984) para programación lineal se puede deducir de una formulación especial del método del gradiente para programación no-lineal.

En efecto, como se presenta en Martínez (1990), el algorítmo de Karmarkar es equivalente al método del gradiente aplicado al problema de programación no-lineal que resulta cuando el problema de programación lineal

$$
\begin{aligned}
&\text{Minimizar } c^T x && (1.1a)\\
&\text{Sujeto a } \quad Ax = 0 && (1.1b)\\
&\qquad\qquad e^T x = 1 && (1.1c)\\
&\qquad\qquad x \geq 0, && (1.1d)
\end{aligned}
$$

donde $e^T = (1, 1, \ldots, 1)$, es transformado en un problema de programación no-lineal con restricciones de igualdad usando variables de holgura al cuadrado. Es decir,

$$\begin{aligned} &\text{Minimizar } c^T x && (1.2a)\\ &\text{Sujeto a} \quad Ax = 0 && (1.2b)\\ &\qquad\qquad e^T x = 1 && (1.2c)\\ &\qquad\qquad x = y^2 && (1.2d) \end{aligned}$$

o lo que es igual,

$$\begin{aligned} &\text{Minimizar } c^T y^2 && (1.3a)\\ &\text{Sujeto a} \quad Ay^2 = 0 && (1.3b)\\ &\qquad\qquad e^T y^2 = 1 && (1.3c)\\ &\qquad\qquad x = y^2 && (1.3d) \end{aligned}$$

A partir de este importante resultado, es natural pensar en un algorítmo alterno que use el método de Programación Cuadrática Sucesiva (**PCS**) en lugar del método del gradiente para resolver el problema (1.3) .

El método **PCS** es una herramienta conocida y muy efectiva en optimización no-lineal. La principal diferencia con el método del gradiente es que en cada iteración , en lugar de usar una aproximación lineal de la función objetivo, se usa una aproximación cuadrática de la función de Lagrange asociada al problema de programación no-lineal como la función objetivo del subproblema. Es de esperarse que el método **PCS** tenga una convergencia local más rápida que el método del gradiente.

En este artículo presentamos el algorítmo básico del método **PCS** con variables de holgura al cuadrado y demostramos la convergencia local de este nuevo método para programación lineal. También, demostraremos que este método tiene la propiedad deseable de que las variables que convergen a cero lo hacen a una rata q-superlineal. La implementación numérica, al igual que el uso práctico de este algorítmo, es en este momento el tema de tesis de grado de uno de mis estudiantes del Programa de Magister en nuestro Departamento.

## EL METODO PCS CON VARIABLES DE HOLGURA AL CUADRADO

Una presentación general del método **PCS** está dada en la sección 6.5.3 de Gill, Murray y Wright (1981), en Tapia (1978) y en Tapia (1980). A continuación derivaremos el método **PCS** para el problema (1.3), el cual será a su vez el método **PCS** con variables de holgura al cuadrado para el problema (1.1).

Dado el vector no-negativo $x_0 \in R^n$ como una aproximación a $x^*$, una solución del problema (1.1) y usando la ecuación (1.3d), definimos $y_0 = \sqrt{x_0}$ como una aproximación a $y^*$, una solución del problema (1.3). Sea $E_c = diag(y_c)$, y $\mu_c$ y $\eta_c$ estimadores de los multiplicadores de Lagrange asociados a la solución $y^*$ correspondientes a las restricciones (1.3b) y (1.3c) respectivamente. La filosofía de la Programación Cuadrática Sucesiva sugiere obtener una aproximación mejorada de $y^*$ de la forma $y_c + s$, donde $\acute{s}$ es la solución del subproblema de programación cuadrática asociado al método **PCS**. Finalmente, la aproximación mejorada de $x^*$ se obtendrá de la ecuación (1.3d), ó de una aproximación a dicha ecuación.

La función objetivo del subproblema de programación cuadrática es

$$c^T E_c s + \tfrac{1}{2} s^T \Lambda_c s, \tag{2.1}$$

donde $\Lambda_c = diag(\lambda_c)$ y

$$\lambda_c = c + A^T \mu_c + \eta_c e. \tag{2.2}$$

Si escribimos $y = y_c + s$, entonces la linealización dada por la serie de Taylor de la función $y^2$ en el punto $y_c$ está dada por

$$y_c^2 + 2E_c s \tag{2.3}$$

Por lo tanto, la linealización de Taylor para la restricción (1.3b) en el punto $y_c$ es

$$A(y_c^2 + 2E_c s) = 0 \tag{2.4a}$$

y si $x_c$ es un punto factible del problema (1.1) (2.4a) se reduce a

$$AE_c s = 0 \tag{2.4b}$$

Similarmente, la linealización de Taylor para la restricción (1.3c) en el punto $y_c$ es

$$e^T(y_c^2 + 2E_c s) = 1 \tag{2.5a}$$

y si $x_c$ es factible, (2.5a) se reduce a

$$e^T E_c s = 0 \tag{2.5b}$$

Ahora tenemos todos los ingredientes para presentar el método **PCS** con variables de holgura al cuadrado para el problema (1.1).

**Algorítmo PCS con variables de holgura al cuadrado:**

Dado $x_0 > 0$ , un punto inicial factible, y $\lambda_0 > 0$
Tome $x = x_0$
$\lambda = \lambda_0$
$y = \sqrt{x}$
Repetir hasta obtener convergencia
$E = diag(y)$
$\Lambda = diag(\lambda)$
Calcular $(\acute{s}, \mu, \eta)$ como un punto de Kuhn Tucker del problema

$$\text{Minimizar } c^T Es + \tfrac{1}{2} s^T \Lambda s \tag{2.6a}$$
$$\text{sujeto a } \quad AEs = 0 \tag{2.6b}$$
$$e^T Es = 0 \tag{2.6c}$$
$$\| s \| \leq \delta \tag{2.6d}$$

$$\text{Calcular } \lambda = c + A^T \hat{\mu} + \hat{\eta} e \tag{2.7a}$$
$$y = y + \hat{s} \tag{2.7b}$$

$x = y^2$
**FIN**

Dos comentarios son importantes a este nivel del trabajo. El primero es sobre (2.5) , la aproximación lineal de la ecuación (1.3c), y el segundo es sobre una propiedad indeseable que tienen $\hat{\mu}$ y $\hat{\eta}$ , los multiplicadores de Lagrange en el subproblema (2.6).

1. La linealización (2.5) de la ecuación (1.3c), llamada linealización de escalamiento afin en Morshedi y Tapia (1987), no es del mismo tipo que la usada por Karmarkar (1984) en su algorítmo original. Esta aproximación también pudo usarse aquí y es posible que dé al algorítmo un mejor comportamiento global; sinembargo, no nos fué posible incluirla en nuestro análisis de convergencia local. Este hecho amerita más investigación.
2. Es bien sabido que cuando se usan variables de holgura al cuadrado, si un multiplicador de Lagrange correspondiente a una restricción donde una variable de holgura al cuadrado haya sido usada toma el valor cero, éste será cero en todas las siguientes iteraciones. Por lo tanto, la fórmula (2.7a) será incapaz de predecir correctamente aquellos multiplicadores que equivocadamente tomen el valor cero en alguna iteración siempre que $\hat{\mu}$ y $\hat{\eta}$ hayan sido obtenidos del subproblema (2.6). Una buena alternativa para salvar este inconveniente podría ser el estimar $\hat{\mu}$ y $\hat{\eta}$ usando otra forma de estimar multiplicadores diferente a la aquí propuesta. (Tapia (1977) presenta y analiza algunas de estas formas alternas).

## ANALISIS DE CONVERGENCIA

Se sabe que el método **PCS** es equivalente a aplicar el Método de Newton a las condiciones de Kuhn-Tucker de primer orden para el problema de optimización en cuestión (ver Tapia (1978)). Analizaremos la convergencia del método presentado en la sección anterior usando las condiciones de primer orden para los problemas (1.2) y (1.3) .

Sean $\mu$ , $\eta$ y $\lambda$ los multiplicadores de Lagrange correspondientes a las restricciones (1.2b), (1.2c) y (1.2d) respectivamente. Las condiciones de primer orden para el problema (1.2) las podemos escribir de la siguiente forma:

$$\begin{aligned} c + A^T\mu + \eta e - \lambda &= 0 \\ \Lambda y &= 0 \\ Ax &= 0 \\ e^T x &= 1 \\ x &= y^2 \end{aligned} \tag{3.1}$$

donde $\Lambda = diag(\lambda)$ . Sea $(x_c, y_c, \mu_c, \eta_c, \lambda_c)$ una aproximación a $(x^*, y^*, \mu^*, \eta^*, \lambda^*)$, punto de Kuhn-Tucker para el poblema (1.2) y solución de (3.1) . Cálculos simples muestran que $(\Delta x, \Delta y, \Delta \mu, \Delta \eta, \Delta \lambda)$ , la corrección dada por el método de Newton, es la solución del siguiente sistema de ecuaciones lineales

$$\begin{pmatrix} 0 & 0 & A^T & e & -I \\ 0 & \Lambda_c & 0 & 0 & E_c \\ A & 0 & 0 & 0 & 0 \\ e^T & 0 & 0 & 0 & 0 \\ -I & 2E_c & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta \mu \\ \Delta \eta \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} c + A^T \mu_c + \eta_c e - \lambda_c \\ \Lambda_c y_c \\ A x_c \\ e^T x_c - 1 \\ y_c^2 - x_c \end{pmatrix} \tag{3.2}$$

donde $E_c = diag(y_c)$ y $\Lambda_c = diag(\lambda_c)$ .

En forma similar, sea $\mu$ y $\eta$ los multiplicadores de Lagrange correspondientes a las restricciones (1.3b) y (1.3c) respectivamente. Las condiciones de primer orden para el problema (1.3) se pueden escribir asi:

$$\begin{aligned} diag(y)(c + A^T \mu + \eta e) &= 0 \\ Ay^2 &= 0 \\ e^T y^2 &= 1 \\ x &= y^2 \end{aligned} \tag{3.3}$$

Sea $(x_c, y_c, \mu_c, \eta_c)$ una aproximación dada a $(x^*, y^*, \mu^*, \eta^*)$ , un punto de Kuhn-Tucker del problema (1.3) y solución de (3.3) . Algunos calculos muestran que $(\Delta x, \Delta y, \Delta \mu, \Delta \eta)$, la corrección dada por el método de Newton, es la solución del sistema de ecuaciones lineales

$$\begin{pmatrix} 0 & diag(c + A^T \mu_c + \eta_c e) & A^T E_c & y_c \\ 0 & 2E_c A & 0 & 0 \\ 0 & 2y_c^T & 0 & 0 \\ -I & 2E_c & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta \mu \\ \Delta \eta \end{pmatrix} = - \begin{pmatrix} E_c(c + A^T \mu_c + \eta_c e) \\ Ay_c^2 \\ y_c^T y_c - 1 \\ y_c^2 - x_c \end{pmatrix} \tag{3.4}$$

donde $E_c = diag(y_c)$.

Además, si $\mu, \eta$ y $\lambda$ son los multiplicadores de Lagrange correspondientes a las restricciones (1.1b), (1.1d) y (1.1c) respectivamente, entonces las condiciones de primer orden para el problema (1.1) se pueden escribir asi

$$\begin{aligned} c + A^T \mu + \eta e - \lambda &= 0 \\ \Lambda x &= 0 \\ Ax &= 0 \\ e^T x &= 1 \\ x &\geq 0 \\ \lambda &\geq 0 \end{aligned} \tag{3.5}$$

donde $\Lambda = diag(\lambda)$ .

Nuestro primer resultado es sobre la equivalencia de los problemas hasta ahora propuestos y lo expresamos en el siguiente lema.

**Lema 1**

(i) *Si* $(x^*, \mu^*, \eta^*, \lambda^*)$ *es un punto de Kuhn-Tucker para el problema (1.1), entonces* $(x^*, \sqrt{x^*}, \mu^*, \eta^*, \lambda^*)$ *es un punto de Kuhn-Tucker para el problema (1.2) y* $(x^*, \sqrt{x^*}, \mu^*, \eta^*)$ *es un punto de Kuhn-Tucker para el problema (1.3).*

(ii) *Si* $(x^*, y^*, \mu^*, \eta^*)$ *es un punto de Kuhn-Tucker para el problema (1.3), entonces* $(x^*, y^*, \mu^*, \eta^*, \lambda^*)$ *, donde* $\lambda^* = c + A^T\mu^* + \eta^*$ *es un punto de Kuhn-Tucker para el problema (1.2). Si además,* $\lambda^* \geq 0$ *, entonces* $(x^*, \mu^*, \eta^*, \lambda^*)$ *es también un punto de Kuhn-Tucker para el problema (1.1).*

(iii) *Si* $(x_+, y_+, \mu_+, \eta_+, \lambda_+)$*, donde por notación* $a_+ = a_c + \Delta a$*, es la nueva aproximación a la solución del problema (1.2) dada por el método de Newton a partir de la aproximación* $(x_c, y_c, \mu_c, \eta_c, \lambda_c)$ *entonces* $(x_+, y_+, \mu_+, \eta_+)$ *es la nueva aproximación a la solución del problema (1.3) dada por el método de Newton a partir de la aproximación* $(x_c, y_c, \mu_c, \eta_c)$ *.*

(iv) *Si* $(x_+, y_+, \mu_+, \eta_+)$ *es la nueva aproximación a la solución del problema (1.3) dada por el método de Newton a partir de la aproximación* $(x_c, y_c, \mu_c, \eta_c)$*, entonces* $(x_+, y_+, \mu_+, \eta_+, \lambda_+)$ *con* $\lambda_+ = c + A^T\mu_+ + \eta_+ e$ *es la nueva aproximación a la solución del problema (1.2) dada por el método de Newton a partir de la aproximación* $(x_c, y_c, \mu_c, \eta_c, \lambda_c)$ *.*

**Demostración**
Partes (i) y (ii) resultan al comparar los sistemas de ecuaciones lineales (3.1) , (3.2) y (3.5) . Similarmente, partes (iii) y (iv) resultan de comparar los sistemas de ecuaciones lineales (3.2) y (3.4) . □

El siguiente teorema es una adaptacion del Teorema 3 dado en Tapia (1980), recordando que la propiedad de complementariedad estricta se cumple para $x^*$, una solución del problema (1.1), si $(x^*)_i$ y $(\lambda^*)_i$ no son ambos cero para todo $i = 1, 2, \ldots, n$ donde $\lambda^*$ es el vector de multiplicadores de Lagrange asociado con las restricciones de no negatividad (1.1d) .

**Teorema 1** *Suponga que la secuencia* $\{(y_k, \lambda_k)\}$ *ha sido generada por el algoritmo* **PCS** *con variables de holgura al cuadrado y que converge a* $(y^*, \lambda^*)$ *donde* $x^* = y_*^2$ *es una solución del problema (1.1) y* $\lambda^*$ *es el vector de multiplicadores de Lagrange asociado a las restricciones de no negatividad (1.1d) . Ademas, suponga que la propiedad de complementariedad estricta se cumple para* $x^*$ *.*

(i) *Si* $\lim_k (y_k)_i = 0$ *, entonces la rata de convergencia es q-superlineal .*

(ii) *Si* $\lim_k (\lambda_k)_i = 0$ *, entonces la rata de convergencia es q-superlineal .*

**Demostración**
Por el Lema (3.1), podemos asumir que cada iteracion es calculada usando (3.2). Sin perdida de generalidad, supongamos que ninguna de las variables converge en un numero finito de pasos . La segunda ecuacion en (3.2) se puede escribir como

$$\Lambda_c(y_c + \Delta y) + E_c \Delta\lambda = 0 \tag{3.6}$$

Como $\Delta_c$ y $E_c$ son matrices diagonales, no singulares y además $\Delta_c^{-1}\lambda_c = e$ , podemos escribir (3.6) asi

$$E_c^{-1}(y_c + \Delta y) + \Lambda_c^{-1}(\lambda + \Delta\lambda) = e. \tag{3.7}$$

Ambas partes (i) y (ii) , son consecuencia directa de (3.7) y la propiedad de complementariedad estricta . □

**Teorema 2** *Sea $(x^*, \mu^*, \eta^*, \lambda^*)$ un punto de Kuhn-Tucker para el problema (1.1) (solución de (3.5)). Suponga que*

**(i)** *La propiedad de complementariedad estricta se cumple para $x^*$ .*

**(ii)** *La matriz $[A^T \quad e]$ es de rango completo,*

*Entonces existe $\epsilon > 0$ tal que si $(x_0, \lambda_0) > 0$ y $\| (x_0, \lambda_0) - (x^*, \lambda^*) \| < \epsilon$, la secuencia $\{(x_k, y_k, \mu_k, \eta_k, \lambda_k)\}$ generada por el algorítmo* PCS *con variables de holgura al cuadrado converge q-cuadraticamente a $(x^*, \sqrt{x}^*, \mu^*, \eta^*, \lambda^*)$ .*

**Demostración**

Condiciones (i) y (ii) garantizan que la matriz del sistema (3.2) es no-singular para todo punto en una vecindad de $(x^*, \sqrt{x}^*, \mu^*, \eta^*, \lambda^*)$ . Es claro que nuestro algorítmo, bajo los supuestos del presente teorema, se reduce al método de Newton definido por (1.2), seguido por algunas modificaciones. Es posible demostrar que dichas modificaciones no destruyen la rata de convergencia del método de Newton la cual es q-cuadratica . □

## Referencias


1. Dennis, J.E. Jr. and Schnabel R.B. (1983), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall, Englewood Cliffs, New Jersey.
2. Gill, P.E., Murray W. and Wright M.H. (1981). *Practical Optimization*, Academic Press, New York.
3. Karmarkar, N. (1984). *A new polynomial time algorithm for linear programming*, Combinatoria, vol. 4, pp. 373–395.
4. Martínez, H.J. (1990). *Aplicación de una derivación del método de Karmarkar*, Lecturas Matemáticas, Vol.XI, Nos.1–2–3, 1990.
5. Morshedi, A.M. and Tapia R.A. (1987). *Karmarkar as a classical method, TR 87–7.* Mathematical Sciences Dept., Rice University, Houston, TX.
6. Tapia, R.A. (1974). *A stable approach to Newton's method for general mathematical programming problems in $R^n$*, JOTA, 22, pp. 135–194.
7. Tapia, R.A. (1977). *Diagonalized multiplier methods and a new implementation*, in Nonlinear Programming 3, (O., Mangasarian, R. Meyer and S. Robinson, eds. ), pp. 125–164, Academic Press, New York.

8. Tapia, R.A. (1978). *Quasi-Newton methods for equality constrained optimization: equivalence of existing methods and a new implementacion*, in Nonlinear Programming 3, O Mangasarian, R. Meyer and S. Robinson, eds, Academic Press, New York.

9. Tapia, R.A. (1980). *On the role of slack variables in quasi-Newton methods for constrained optimization*, in Numerical Optimization of Dynamical Systems, (L.C.W. Dixon and G.P. Szegö, eds.), North-Holland Publishing Company, pp. 235-246.