

MATRICES ESCALONADAS Y METODOS PRIMAL DUAL DE PUNTO INTERIOR

Alibeit Kakes Cruz, Departamento de Matemática Aplicada, Facultad de Matemática y Computación, Universidad de La Habana
 Darnes Vilariño Ayala, Departamento de Computación, Facultad de Matemática y Computación, Universidad de La Habana

RESUMEN

Se particionan las filas de la matriz A de un modelo general de Programación Lineal en dos subconjuntos, B₁ y B₂, de manera que al aplicar un algoritmo primal-dual de punto interior, no se utilice A tal y como es dada por datos. La ecuación normal, usual en este tipo de métodos, toma la forma B₁D²B₁^tx = b. La matriz B₂ es utilizada para obtener D². El objetivo último del trabajo es aplicar el resultado anterior a una clase particular de matrices escalonadas.

Palabras clave: Programación Lineal, Métodos de Punto Interior, algoritmos.

ABSTRACT

Given a general linear programming model, the set of rows of the matrix A is partitioned in two subsets, B₁ and B₂, which allows an interior point primal dual algorithm to deal with matrices of lower order. The normal equation associated with these methods takes the form B₁D²B₁^tx = b. Matrix B₂ is used in the calculation of D². The ultimated purpose of the paper is the application of this result to a particular class of staircase matrices.

Key words: Linear Programming, Interior Point Methods, algorihtms.

MSC: 90C05

1. INTRODUCCION

Los métodos Primal Dual de Punto Interior resuelven los problemas primal y dual (Bazaraa-Sherali, Shetty (1993), Dantzing (1963), Monteiro-Adler (1996) a Monteiro- Adler (1996b)) de la Programación Lineal dados por (1) – (6). Los resuelven de una vez, aplicando variantes del método de Newton a las tres igualdades (7) – (9) y modificando las direcciones de búsqueda y tamaño de paso, de manera que la restricción (10) se satisfaga estrictamente en cada iteración. Esa condición, x > 0, s > 0, dio origen al término “algoritmo de punto interior”.

$$\begin{aligned}
 & \min c^T x & (1) \\
 \text{s.a} & Ax = b, & (2) \\
 & x \geq 0, & (3) \\
 & \max b^T \lambda & (4) \\
 \text{s.a} & A^T \lambda + s = c, & (5) \\
 & s \geq 0, & (6) \\
 & A^T \lambda + s = c & (7) \\
 & Ax = b & (8) \\
 & x_i s_i = 0 & (9) \\
 & (x, s) \geq 0 & (10)
 \end{aligned}$$

Las expresiones (7) - (10) definen una condición necesaria y suficiente de optimalidad (Bazaraa-Sherali-Shetty (1993), Dantzing (1963)), de x y (λ, s) de los problemas Primal y Dual dados por (1)-(3) y (4)-(6), respectivamente (Dantzing. (1963), Luemberger. (1973)).

Esquema general de los Métodos Primal Dual de Punto Interior Inafctible

Dados $(x_0, \lambda_0, s_0), x_0 > 0, s_0 > 0$

Para $k = 0, 1, 2, \dots$

Resolver

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -X^k S^k e + \sigma^k \mu^k e \end{bmatrix}$$

donde $\sigma_k \in [0, 1]$ y $\mu_k = (x^k)^T s^k / n$;

hacer $(x^{k+1}, \lambda^{k+1}, s^{k+1}) \leftarrow (x^k, \lambda^k, s^k) + \alpha_k (\Delta x^k, \Delta \lambda^k, \Delta s^k)$,

eligiendo α_k de forma que $(x^{k+1}, s^{k+1}) > 0$.

Fin.

Como se observa, lo anterior es sólo un esquema, pues no contempla, por ejemplo, criterio de parada.

Los parámetros σ^k y μ^k son los parámetros de centrado y de medida de la brecha de dualidad, $c^t x - b^t \lambda$, respectivamente.

Ese esquema general conlleva a la aplicación de un método de Newton perturbado, donde $X = \text{diag}(x_1, x_2, \dots, x_n)$, $S = \text{diag}(s_1, s_2, \dots, s_n)$, $e = (1, 1, \dots, 1)^T$.

$$r_b = Ax - b, \quad r_c = A^T \lambda + s - c.$$

De ese esquema general de algoritmo se obtienen las siguientes ecuaciones:

$$A \Delta x = -r_b$$

$$A^t \Delta \lambda + \Delta s = -r_c$$

$$S \Delta x + X \Delta s = -X S e + \sigma \mu e$$

donde r_b y r_c son los residuos correspondientes.

Por despejes y sustituciones adecuadas se llega a las siguientes expresiones:

$$AD^2 A^T \Delta \lambda = -r_b + A(S^{-1} X r_c + x - \sigma \mu S^{-1} e), \quad (11)$$

$$\Delta s = -r_c - A^T \Delta \lambda,$$

$$\Delta x = -x + \sigma \mu S^{-1} e - S^{-1} X \Delta s$$

donde $D^2 = (S X^{-1})^{-1} \quad (12)$

A la ecuación (11) se le conoce como “ecuación normal” (Gonzaga (1996), Monteiro-Adler (1996a), Monteiro-Adler (1996b)). De (12) y (11) se infiere que el cálculo de $\Delta\lambda$ manipula dos matrices inversas, D^2 para construir AD^2A^T , y luego la inversa de esta última matriz.

En el caso de un modelo lineal con variables acotadas, se tiene:

$$\begin{aligned} & \min c^t x \\ \text{sa} \quad & Ax = b, \\ & x \geq 0 \\ & x \leq v \end{aligned} \tag{13}$$

El modelo dual asociado al anterior se define así;

$$\begin{aligned} & \text{Max } b^t \lambda - v^t w \\ \text{sa} \quad & A^T \lambda - w + s = c \\ & w \geq 0, s \geq 0 \end{aligned}$$

Las condiciones de KKT para ese problema quedan expresadas como:

$$\begin{aligned} Ax &= b \\ x + z &= v \\ A^T \lambda - w + s &= c \\ XSe &= \mu e \\ ZWe &= \mu e \\ (x, s, w, z) &\geq 0 \end{aligned}$$

X, Z, S y W son matrices diagonales con x_j, z_j, s_j, w_j , en la diagonal. s y z variables de holuras en P y D respectivamente.

La fórmula de Newton produce un sistema de ecuaciones que finalmente da las expresiones que siguen para D^2 y la ecuación normal:

$$\begin{aligned} \Delta\lambda &= (AD^2A^t)^{-1}[-r_b - AD^2(r_c + (W - S)e - \mu(Z^{-1} - X^{-1})e)] \\ D^2 &= (Z^{-1}W + X^{-1}S)^{-1} \end{aligned} \tag{14}$$

Las restricciones de acotación dadas por (13) pudieran ser estandarizadas y considerar $A = \begin{pmatrix} A \\ I \end{pmatrix}$. Ese tratamiento aumenta la dimensión de la matriz a considerar en la ecuación normal. En nuestro caso, preferimos trabajar con la matriz original, no aumentando la dimensión de la matriz a considerar en (14). Las restricciones tipo (13) están implícitamente involucradas en (15), como se verá a continuación.

2. ECUACION NORMAL COMPACTADA

Se considera el siguiente modelo lineal, P:

$$\begin{aligned} & \min c^t x \\ \text{s.a} \quad & Ax = b \\ & Bx \leq d \\ & x \geq 0 \end{aligned} \tag{15}$$

y su dual, D

$$\begin{aligned} & \max \quad b^t \lambda - d^t \omega \\ \text{s.a} \quad & A^t \lambda - B^t \omega \leq c \\ & \omega \geq 0 \end{aligned}$$

Las matrices A y B son de orden $(k \times n)$ y $(p \times n)$ respectivamente.

Estandarizando ambos problemas se tienen las condiciones de KKT.

$$\begin{aligned} Ax &= b \\ Bx + lz &= d \\ A^t \lambda - B^t \omega + s &= c \\ XSe &= \mu e_n \\ ZWe &= \mu e_p \\ x \geq 0, z \geq 0, w \geq 0, s \geq 0 \end{aligned}$$

X, S, Z, W matrices diagonales asociadas a los vectores x, s, z, w, respectivamente.

Se define,

$$F(x, z, \lambda, \omega, s) = \begin{pmatrix} Ax - b \\ Bx + lz - d \\ A^t \lambda - B^t \omega + s - c \\ XSe - \mu e_n \\ ZWe - \mu e_p \end{pmatrix} = 0$$

Aplicando la fórmula de Newton;

$$\begin{pmatrix} A & 0 & 0 & 0 & 0 \\ B & I & 0 & 0 & 0 \\ 0 & 0 & A^t & -B^t & I \\ S & 0 & 0 & 0 & X \\ 0 & W & 0 & Z & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta z \\ \Delta \lambda \\ \Delta \omega \\ \Delta s \end{pmatrix} = \begin{pmatrix} -r_b \\ -r_d \\ -r_c \\ -XSe + \tau \mu e_n \\ -ZWe + \tau \mu e_p \end{pmatrix}; \quad (16)$$

donde

$$\begin{aligned} r_b &= Ax - b; \\ r_d &= Bx + z - d; \\ r_c &= A^t \lambda - B^t \omega + s - c \end{aligned}$$

X, Z, S y W son matrices diagonales asociadas a los vectores x, z, s, y w respectivamente.

De (16) se obtiene:

$$A \Delta x = -r_b \quad (17)$$

$$B \Delta x + \Delta z = -r_d \quad (18)$$

$$A^t \Delta \lambda - B^t \Delta \omega + \Delta s = -r_c \quad (19)$$

$$S\Delta x + X\Delta s = -XSe + \tau\mu e_n \quad (20)$$

$$W\Delta z + Z\Delta\omega = -WZe + \tau\mu e_p \quad (21)$$

Despejando Δs en (20) y $\Delta\omega$ en (21):

$$\Delta s = -Se - \tau\mu X^{-1}e - X^{-1}S\Delta x \quad (22)$$

$$\Delta\omega = -We + \tau\mu Z^{-1}e - Z^{-1}W\Delta z \quad (23)$$

Sustituyendo (22) y (23) en (19):

$$A^t\Delta\lambda - B^t[-We + \tau\mu Z^{-1}e - Z^{-1}W\Delta z] - Se - \tau\mu X^{-1}e - X^{-1}S\Delta x = -r_c \quad (24)$$

Despejando Δz en (18), sustituyendo en (24) y agrupando Δx :

$$(X^{-1}S + B^t Z^{-1}WB)\Delta x = r_c + B^tWe - \tau\mu B^t Z^{-1}e - B^t Z^{-1}Wr_d - Se - \tau\mu X^{-1}e + A^t\Delta\lambda$$

Definamos
$$D^2 = (X^{-1}S + B^t Z^{-1}WB)^{-1} \quad (25)$$

Finalmente, despejando Δx y sustituyendo en (17) se obtiene la ecuación normal.

$$AD^2A^t\Delta\lambda = -r_b - AD^2(r_c + B^tWe - B^tZ^{-1}Wr_d - \tau\mu B^tZ^{-1}e + \tau\mu X^{-1}e - Se_n)$$

El resultado muestra que es posible resolver la ecuación normal no trabajando con la matriz $\begin{pmatrix} A \\ B \end{pmatrix}$; sino trabajando primero con B y luego con A.

2.1. Ecuación normal en una clase particular de modelos escalonados

Los modelos que consideramos a continuación son frecuentes en la economía y en la industria, Fourer (1982), Glassey (1970), Izuno-Masuzawa (1989), Propoi (1988).

$$\begin{aligned} & \min J(u) = a(T)x(T) \\ \text{s.a} & \\ & G(t)x(t) + D(t)\mu(t) \leq f(t) \\ & x(t+1) = A(t)x(t) + B(t)\mu(t) \\ & x(t) \geq 0 \\ & \mu(t) \geq 0 \\ & x(0) \\ \text{dado} & \\ & t = 0, 1, \dots, T-1 \end{aligned}$$

Las matrices $G(t)$; $D(t)$; $A(t)$ y $B(t)$ y los vectores $f(t)$, $x(t)$ y $a(T)$ tienen dimensiones $m \times n$; $m \times r$; $n \times n$; $n \times r$; m ; n ; y n respectivamente.

Tales modelos tienen un sabor dinámico, basta considerar las desigualdades como las restricciones del problema, y las otras restricciones, como las ecuaciones de ligadura.

Luego de estandarizar el modelo anterior, la matriz del sistema tendría dimensión $(m+n)T \times (n+r+m)T$. Pero no se estandariza. Se particionan las filas de esa matriz, sin estandarizar, de manera que una parte de ella sea tomada como A en la ecuación normal y la otra forme parte de D^2 . Presentaremos a continuación dos posibles filosofías de partición:

La matriz $X^{-1}S$ se particiona según la matriz del problema que se resuelve. Llamemos $\bar{U}(t)$ y $\bar{X}(t)$ a las partes de $X^{-1}S$ formadas con los vectores $u(t)$ y $x(t)$ respectivamente. Podemos entonces escribir:

$$X^{-1}S = \begin{pmatrix} \bar{U}(0) & & & & & \\ & \bar{X}(1) & & & & \\ & & \bar{U}(1) & & & \\ & & & \bar{X}(2) & & \\ & & & & \bar{U}(2) & \\ & & & & & \ddots \\ & & & & & & \bar{U}(4) \\ & & & & & & & \ddots \end{pmatrix}$$

Definamos ahora:

$$X^{-1}S(t) = \bar{U}(t); \quad t = 0, 2, 4, \dots$$

$$X^{-1}S(t) = \begin{pmatrix} \bar{X}(t) & & \\ & \bar{U}(t) & \\ & & \bar{X}(t+1) \end{pmatrix}; \quad t = 1, 3, \dots$$

Ejemplifiquemos con $T = 5$, lo que ocurre.

Para que sea compatible la suma $X^{-1}S + B^t Z^{-1}WB$; expresemos B de la siguiente manera:

$$\bar{B} = \begin{pmatrix} 0 & \bar{B}(1) & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{B}(3) & 0 \end{pmatrix}$$

Los bloques de cero corresponden a las matrices que están en $\bar{A}(t)$ pero no en $\bar{B}(t)$.

$Z^{-1}W$ se particiona según las columnas de B .

Se tiene finalmente que $X^{-1}S + B^t Z^{-1}WB$ es igual a:

$$\begin{pmatrix} X^{-1}S(0) & & & & \\ & X^{-1}S(1) & & & \\ & & \ddots & & \\ & & & X^{-1}S(4) & \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \bar{B}^t(1) & 0 \\ 0 & 0 \\ 0 & \bar{B}^t(3) \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Z^{-1}W(1) & & \\ & & \\ & & Z^{-1}W(3) \end{pmatrix} \begin{pmatrix} 0 & \bar{B}(1) & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{B}(3) & 0 \end{pmatrix}$$

$$= \begin{pmatrix} X^{-1}S(0) & & & & \\ & X^{-1}S(1) + \bar{B}^t(1)Z^{-1}W(1)\bar{B}(1) & & & \\ & & X^{-1}S(2) & & \\ & & & X^{-1}S(3) + \bar{B}^t(3)Z^{-1}W(3)\bar{B}(3) & \\ & & & & X^{-1}S(4) \end{pmatrix}$$

De la matriz diagonal por bloques anterior; concluimos que:

$$D^2(t) = XS^{-1}(t); \quad t = 0,2,\dots$$

$$D^2(t) = K^{-1}(t); \quad t = 1,3,\dots$$

donde

$$K(t) = X^{-1}S(t) + \bar{B}^t(t)Z^{-1}W(t)\bar{B}(t)$$

3. CONCLUSIONES

- 1) Con ambas particiones se logra construir la ecuación normal sin invertir la matriz del sistema original. Parte de ella se utiliza para el cálculo de la matriz D^2 , y ésta, junto a la parte restante dan lugar a la propia ecuación normal AD^2A^t .
- 2) Dos ventajas son evidentes en la segunda partición
 - Parte de $D^2(t)$ resulta inmediatamente calculable ($t = 0,2,\dots$)
 - Las inversas $K^{-1}(t)$ para t impar, pueden ser calculadas simultáneamente, es decir, en paralelo.
- 3) Como la matriz AD^2A^t resulta tridiagonal por bloques, [9], la ecuación normal aprovecha esa condición.

REFERENCIAS

- BAZARAA, M.S.; H.D.SHERALI and C.M. SHETTY (1993): **Nonlinear Programming**, John Wiley & Sons, Inc. New York
- DANTZING, G.B. (1963): **Linear Programming and Extensions**, Princeton University Press, New Jersey.
- FOURER, R. (1982): "Solving staircase linear programs by the simplex method, 1: inversión. 2: Priacing", **Mathematical Programming** 23 , 274- 313.
- FUJISAWA, K. (1997): "Exploiting Sparsity in primal dual interior algorithm for linear programming", **Mathematical Programming** 61, 79.
- GLASSEY, C.R. (1970): "Dynamic Linear Programming for Production Scheduling", **Operations Research** 18(1).
- GONDZIO, J. and J.P. VIAL (1967): "Using an interior point methods for the master problem in a descomposicion approach", **European Journal of Operational Research** 101.
- GONZAGA, C.C. (1996): "Complexity of Predictor–Corrector algorithms", **Technical Report**. Department of Mathematics University of Geneva.
- JI, J.; F. POTRA, R.A: TAPIA and Y. ZHANG (1991): "An interior Method with Polynomial Complexity and Superlinear Convergence for Linear Complementarity Problems", Department of Mathematical of Sciences, Rice University.
- KAKES, A. (2000): Tesis de Doctorado. Universidad de La Habana, Departamento de Matemáticas Aplicadas.
- LUEMBERGER, D.G.. (1973): **Introduction to linear and nonlinear programming**. Addison-Wesley Publishing Company, Inc., New York.
- MIZUNO, S. and K. MASUZAWA (1989): "Polynomial time interior point algorithms for transportation problems", **Journal of the Operation Research Society of Japan**, 32 , 371-382.

- MIZUNO, S.; M. KOJIMA and M.J. TODD (1995): "Infeasible interior point primal dual potencial reduction algorithms for linear programming", **SIAM Journal of Optimization**.
- MONTEIRO, R.D.C. and I. ADLER. (1996a): "Interior path following primal-dual algorithms", Part I : "Linear Programming". **Mathematical Programming** 44, 27- 41. North -Holland.
- _____ (1996b): "Interior path following primal-dual algorithms", Part II: "Convex Quadratic Programming", **Mathematical Programming** 44, 43-66. North-Holland.
- PROPOI, A.I. (1988): **Problems of dynamic linear programming**, Luxenburgo (Austria) IIASA. RM 76-78.
- TAPIA, R.A.; Y. ZHANG and Y. YE. (1991): **On the convergence of the iteration sequence in primal - dual interior - point methods**, Department of Mathematical of Sciences , Rice University.
- TOMLIN. J.A. (1989). "A note on comparing Simplex and Interior Methods for Linear Programming", **Mathematics of Operations Research** 15, 508-529.