APPLICATIONS OF DUALITY TO THE LAW OF SWEEP-THE-NEGATIVES

By

Jae-Rong CHOI*

(Received Oct. 30, 1975)

1. Introduction

In Choi [5], the author described an efficient algorithm for M.L.E. and the least squares estimations under certain order restrictions.

He solved a minimizing problem

$$(1.1) \quad \min_{Ay \geq 0} (y - y_1)'(y - y_1)$$

for given vector y_1 and matrix A , where AA' is a positive definite symmetric matrix of which all off diagonal elements are not positive. The problem (1.1) is converted to

$$(1.2) \quad \min_{x \geq 0} (x - x_1)'A^{-1}(x - x_1),$$

where $x = Ay$ and $A (= AA')$ satisfies the above condition.

In this paper, we make use of the algorithm of Choi [5] for the solution of (1.1) or (1.2) in case A^{-1} is a positive definite symmetric matrix of which all off diagonal elements are not positive, considering the dualities of (1.1) or (1.2).

2. Duality

Consider the dual problems of (1.1) and (1.2). The primal problems are equivalent to

$$(2.1) \quad \min_{y \in \mathfrak{C}} \frac{1}{2} (y - y_1)'(y - y_1),$$

where $\mathfrak{C} = \{y: Ay \geq 0\}$,

* On leave of absence from the Department of Mathematics, Dong-A University, Korea.

and

$$(2.2) \quad \text{Min}_{x \geq 0} \frac{1}{2} (x - x_1)' A^{-1} (x - x_1).$$

Their dual problems are able to be obtained by Fenchel's duality theory (see Rockafellar [6], Theorem 31.4) and some discussions in Barlow and Brunk [1]. The dual cone (or polar) of $\mathfrak{C}$ is defined by

$$(2.3) \quad \mathfrak{C}^* = \{y^*: (y^*, y) \leq 0, y \in \mathfrak{C}\}.$$

For a nonsingular matrix A , we can easily get

$$(2.4) \quad \mathfrak{C}^* = \{y^*: -A'^{-1}y^* \geq 0\}$$

according to Choi [4].

The dual problem of (2.1) is to find y^* which attains the optimal solution y_0^* such that

$$(2.5) \quad -\text{Min}_{y^* \in \mathfrak{C}^*} \frac{1}{2} (y^* - y_1)' (y^* - y_1) = -\frac{1}{2} (y_0^* - y_1)' (y_0^* - y_1).$$

y_0^* is related to the optimal solution y_0 of (2.1) by the identity

$$(2.6) \quad y_0 = y_1 - y_0^*.$$

In the like manner, we have the dual of (2.1);

$$(2.7) \quad -\text{Min}_{x^* \geq 0} \frac{1}{2} (x^* + A^{-1}x_1)' A (x^* + A^{-1}x_1) + \text{const},$$

and the optimal solution x_0^* is related to the optimal solution x_0 of (2.1) by the identity

$$(2.8) \quad x_0 = x_1 + Ax_0^*.$$

The above discussions for duality give us an advantageous use of the algorithm of [5] in the case that A^{-1} is a positive definite symmetric matrix with all non-positive off diagonal elements.

3. Examples

A typical example is the case that A is an equicorrelation matrix with positive

correlation coefficients.

EXAMPLE 1. Let $y_i (i=0, 1, \dots, k)$ be distributed to a normal distribution $N(\theta_i, 1)$ respectively. Consider M. L. E. of θ_i under the restrictions

$$(3.1) \quad \lambda_i \theta_i - \lambda_0 \theta_0 \geq 0, \quad i=1, 2, \dots, k,$$

where $\lambda_0, \lambda_i > 0$.

Transform $x = Ay$

where

$$(3.2) \quad A = \begin{pmatrix} -\lambda_0 & \lambda_1 & 0 & \dots & 0 \\ -\lambda_0 & 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\lambda_0 & 0 & 0 & \dots & \lambda_k \end{pmatrix}$$

then x is distributed to $N(\mu, A)$,

where

$$(3.3) \quad A = \begin{pmatrix} \lambda_0^2 + \lambda_1^2 & \lambda_0^2 & \dots & \lambda_0^2 \\ \lambda_0^2 & \lambda_0^2 + \lambda_2^2 & \dots & \lambda_0^2 \\ \dots & \dots & \dots & \dots \\ \lambda_0^2 & \lambda_0^2 & \dots & \lambda_0^2 + \lambda_k^2 \end{pmatrix}$$

M. L. E.'s of x and y are given by the optimal solutions of (1.1) and (1.2) respectively. As shown in Appendix, A^{-1} is a positive definite matrix of which all off diagonal elements are not positive. Therefore we can get the solution of (1.2) by using the duality (2.7). Here we note that the dual problem of (1.1) can not be considered in this example because of the singularity of A . But we calculate M. L. E. of θ by making use of Proposition 2.1 in [5].

Proposition 2.1 in [5] means that M. L. E. of μ is given by the projection of a given x_1 to any hyperplane containing one supporting plane $\mathfrak{C}_M$ (see [5]), of which the index subset

$$(3.4) \quad M = \{i: x_{0i} = 0, i=1, 2, \dots, k\},$$

where x_{0i} denotes an element of M. L. E. x_0 .

The M. L. E. of θ is able to be obtained, by the Lagrangian multiplier method,

$$(3.5) \quad y_0 = y_1 - A_1'(A_1 A_1')^{-1} A_1 y_1,$$

where A_1 is a row partition of A corresponding to M .

REMARK. The above primal problem is included in Bartholomew [3], but the algorithm for determining the statistic was not discussed explicitly.

Now we will illustrate a numerical example.

EXAMPLE 2. Let $y = (y_1, \dots, y_4)$ be independently distributed with mean θ_i and unit variance and $x = Ay$ where

$$(3.6) \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.2 & 1 & 0 & 0 \\ 0.2 & 0.2 & 1 & 0 \\ 0.1 & 0.4 & 0.1 & 1 \end{pmatrix}.$$

Let us consider M. L. E. of mean μ of x under the order restriction that $\mu = A\theta \geq 0$ for given sample mean $y_1' = (-10, 1, 11.8, -4.58)$ and $x_1' = (-10, -1, 10, 4)$. (For convenience, sample size is 1.)

We can easily formulate the minimizing problems (2.1) and (2.2) taking A as (3.6) and

$$(3.7) \quad A = \begin{pmatrix} 1.0 & & & \\ 0.2 & 1.04 & & \\ 0.2 & 0.24 & 1.08 & \\ 0.1 & 0.42 & 0.20 & 1.18 \end{pmatrix}.$$

Now compute A^{-1} and $x_1^* (= -A^{-1}x_1)$, and make the matrix for sweep, we have

$$[A^{-1} \ x_1^*] = \begin{pmatrix} 1.06562 & -0.16648 & -0.15960 & -0.00400 & 12.06968 \\ -0.16648 & 1.18440 & -0.16200 & -0.38000 & -0.38040 \\ -0.15960 & -0.16200 & 1.01000 & -0.10000 & -12.25800 \\ -0.00400 & -0.38000 & -0.10000 & 1.00000 & 4.58000 \end{pmatrix}.$$

As A^{-1} is a positive definite symmetric matrix of which off diagonal elements are not positive, we can solve the duality (2.7) rather than (2.2).

According to the law of sweeping the negatives (Choi [5]), we sweep out the above matrix taking a pivot as the element (2, 2) because the second element of the last column is negative. Next we may sweep out the matrix, which is obtained by the first sweep out, taking a pivot as the element (3, 3) because the element (3, 5) is negative. Then we have as the result,

$$(3.8) \quad \begin{pmatrix} 1.00855 & 0.0 & 0.0 & -0.08547 & 9.74359 \\ -0.16581 & 1.0 & 0.0 & -0.34188 & -2.02564 \\ -0.18462 & 0.0 & 1.0 & -0.15385 & -12.46154 \\ -0.08547 & 0.0 & 0.0 & 0.85470 & 2.56410 \end{pmatrix}.$$

As the remainders of the last column in (3.8) are all positive, the solution of the dual problem, $x_0^* = (9.74359, 0.0, 0.0, 2.56410)$. By the identity (2.8), $x'_0 = (0.0, 2.02564, 12.46154, 0.0)$. Furthermore we are sure that $y'_0 = (0.0, 2.02564, 12.05641, -2.01590)$ and $y_0^* = (-10.0, -1.02564, -0.25641, -2.56410)$ from one of the identities $y'_0 = x'_0 A'$, $y_0^* = x_0^* (-A)$, and (2.6).

Appendix

Let

$$A = \begin{pmatrix} \lambda_1 + \lambda_0 & \lambda_0 & \dots & \lambda_0 \\ \lambda_0 & \lambda_2 + \lambda_0 & \dots & \lambda_0 \\ \dots & \dots & \dots & \dots \\ \lambda_0 & \lambda_0 & \dots & \lambda_k + \lambda_0 \end{pmatrix} = D + \lambda_0 E,$$

where D is a diagonal matrix of which i -th diagonal is λ_i and E is a matrix with all elements 1.

Consider $(D + \lambda_0 E)(D^{-1} + G) = I$, then $\lambda_0 E D^{-1} + DG + \lambda_0 EG = 0$, that is,

$$(1) \quad \lambda_0 E \begin{pmatrix} \lambda_1^{-1} & 0 \\ & \ddots \\ 0 & \lambda_k^{-1} \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 \\ & \ddots \\ 0 & \lambda_k \end{pmatrix} [g_{ij}] + \lambda_0 E [g_{ij}] = 0.$$

Considering the first column of (1), we get

$$(2) \quad \lambda_0 \lambda_1^{-1} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \begin{pmatrix} \lambda_1 g_{11} \\ \vdots \\ \lambda_k g_{k1} \end{pmatrix} + \lambda_0 \left(\sum_{i=1}^k g_{i1} \right) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let $\sum_{i=1}^k g_{i1} = G_1$, from (2) we have

$$(3) \quad g_{i1} = (-\lambda_0 G_1 - \lambda_0 \lambda_i^{-1}) \lambda_i^{-1}, \quad i = 1, 2, \dots, k.$$

From (3),

$$(4) \quad G_1 = -\lambda_0 \lambda_1^{-1} (\Sigma \lambda_i^{-1}) / (1 + \lambda_0 \Sigma \lambda_i^{-1}).$$

Put (4) into (3), we have

$$g_{i1} = -\{\lambda_0 / (1 + \lambda_0 \Sigma \lambda_i^{-1})\} (\lambda_i \lambda_1)^{-1}.$$

In the like manner, we get

$$G = [g_{ij}].$$

where

$$(5) \quad g_{ij} = -\{\lambda_0 / (1 + \lambda_0 \Sigma \lambda_i^{-1})\} (\lambda_i \lambda_j)^{-1}.$$

and also

$$G = -\{\lambda_0 / (1 + \lambda_0 \Sigma \lambda_i^{-1})\} \cdot D^{-1} E D^{-1}.$$

Furthermore

$$|D + \lambda_0 E| = \prod_{i=1}^k \lambda_i (1 + \lambda_0 \sum_{j=1}^k \lambda_j^{-1}) \geq 0 \quad \text{if and only if} \quad \lambda_0 \geq 1 / \Sigma \lambda_j^{-1}, \quad \text{and} \\ \lambda_0 > 1 / \Sigma \lambda_j^{-1} \text{ implies}$$

$$(6) \quad |D^{-1} + G| = \prod_{i=1}^k \lambda_j^{-1} / (1 + \lambda_0 \Sigma \lambda_i^{-1}) \geq 0.$$

From (5) and (6), A^{-1} is a positive definite symmetric matrix of which all the off diagonal elements are not positive when A is a positive definite symmetric matrix.

The author would like to thank professor A. Kudô for his valuable suggestions and guidances.

References

- [1] BARLOW, R. E. and BRUNK, H. D.: *The isotonic regression problem and its dual*. J. Amer. Statist. Assoc., **67**, 327 (1972), pp. 140-7.
- [2] BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M., and BRUNK, H. D.: *Statistical Inference under Order Restrictions*. John Wiley & Sons Inc., New York, (1972).
- [3] BARTHOLOMEW, D. J.: *A test of homogeneity of means under restricted alternatives*. J. R. Statist. Soc., **b**, 23 (1961), pp. 239-81.
- [4] CHOI, J. R.: *An equality involving orthant probabilities*. Comm. Statist., **4**(12) (1975), pp. 1167-75.
- [5] CHOI, J. R.: *The law of Sweep-the-Negatives in the estimation under order restrictions*. Mem. Fac. Sci., Kyushu Univ., Ser. A (math.), **30**, 1 (1976), pp. 135-43.
- [6] ROCKAFELLAR, R. T.: *Convex Analysis*. Princeton Univ. Press, Princeton, (1970).