Journal of the Korean
Statistical Society
Vol. 21, No. 2, 1992

"Pool-the-Maximum-Violators" Algorithm

Kikuo Yanagi,[1] Akio Kudô[2] and Yong-Beom Park[3]

ABSTRACT

The algorithm for obtaining the isotonic regression in simple tree order, the most basic and simplest model next to the simple order, is considered. We propose to call it *"Pool-the-Maximum-Violators"* algorithm (PMVA) in conjunction with the *"Pool-Adjacent-Violators"* algorithm (PAVA) in the simple order. The dual problem of obtaining the isotonic regression in simple tree order is our main concern. An intuitively appealing relation between the primal and the dual problems is demonstrated. The interesting difference is that in simple order the required number of pooling is at least the number of initial violating pairs and any path leads to the solution, whereas in the simple tree order it is at most the number of initial violators and there is only one advisable path although there may be some others leading to the same solution.

1. INTRODUCTION

This paper is a continuation of Kim et al.(1990), with the same kind of pedagogical motivation, where a kind of parallelism between the PAVA and the "Sweep-the-Negatives" algorithm was discussed.

[1] Graduate School of Natural Science and Technology, Okayama University, Okayama, 700, Japan.
[2] Department of Mathematics, Kyushu University 33, Fukuoka, 812, Japan.
[3] Department of Natural Sciences, Pusan National University of Technology, Pusan, 608-739, Korea.

We assume that the book by Robertson et. al. is at the disposal of the readers, and we make free use of the terminologies therein. In case of the simple order. PAVA is available, whereas for the simple tree order an algorithm was explained in their book as Example 1.3.2. In short the algorithm is finding out the observation which stays far most from the simple tree order restriction and pooling it with the control observation. We, therefore, propose to call this algorithm "Pool-the-Maximum-Violators" algorithm.

The PAVA and PMVA are of the straight forward nature in the sense that each step of the process is surely approaching the final value, and that steps needed are always not more than the number of observations. In PAVA there are a number of different routes to the final solution and we can arrange the routes in some order and enumerate all of them. On the other hand, in case of the algorithm for the simple tree order set of the route is merely conceptional. This paper is motivated by this difference, and we will show that the a kind of parallelism in the algorithms between the dual problem of the simple tree order and the simple order.

The algorithm of Sweep-the-Negatives was first introduced by Choi(1976). The detailed demonstration about this can be found in Kim et al.(1990). The sweep out operation is applied on an $n \times (n+1)$ matrix, and the algorithm has also straight forward nature in the same sense as stated above (cf. Kennedy and Gentle 1980). The approach in Robertson et al.(1988) is based on minimum lower set algorithm. We treat this problem along the line of Kudô(1963), Kudô and Choi(1975), and Kim et al.(1990).

2. SIMPLE TREE ORDER

Instead of the notations in the Example 1.3.2 cited above, we use different notations. Let $X = \{0, 1, \ldots, n\}$, $g(x_i) = x_i$, $w(x_i) = w_i$, and $g^*(x_i) = \widehat{\theta}_i$.

Let the observation vector, the underlying vector and the corresponding weight vector be denoted by $(x_0, \mathbf{x}')' = (x_0, x_1, \ldots, x_n)'$, $(\theta_0, \theta')' = (\theta_0, \theta_1, \ldots, \theta_n)'$, and $(w_0, \mathbf{w}')' = (w_0, w_1, \ldots, w_n)'$.

Corresponding to the simple tree order, we have the convex polyhedral cone:

$$\Omega = \{(\theta_0, \theta')'; \theta_0 \in (-\infty, +\infty), \theta_0 \geq \theta_1, \ldots, \theta_0 \geq \theta_n\}.$$

Let the matrix corresponding weight be denoted by $W = \text{diag}(w_1, \ldots, w_n)$.

The problem is to determine the scalar $\widehat{\theta}_0$ and the vector $\widehat{\theta}'$ whose combination is the solution of the minimization problem:

$$\min_{\theta\in\Omega}\{w_0(\theta_0-x_0)^2+(\theta-\mathbf{x})'W(\theta-\mathbf{x})\} = w_0(\hat{\theta}_0-x_0)^2+(\hat{\theta}-\mathbf{x})'W(\hat{\theta}-\mathbf{x}). \quad (2.1)$$

Let $\hat{\theta}(\theta_0)$ be the solution of the intermediate minimization problem:

$$\text{Minimize} \quad (\theta-\mathbf{x})'W(\theta-\mathbf{x}) = (\hat{\theta}(\theta_0)-\mathbf{x})'W(\hat{\theta}(\theta_0)-\mathbf{x}),$$

$$\text{subject to the inequalities } \theta_0 \geq \theta_i, \quad i=1,\ldots,n, \quad (2.2)$$

that is minimizing the second term of (2.1) where θ_0 is fixed.

Let us introduce the variables: $\mathbf{y}=(y_1,\ldots,y_n)', y_i = x_0 - x_i, \mu = (\mu_1,\ldots,\mu_n)'$, $\mu_i = \theta_0 - \theta_i$, for $i=1,\ldots,n$.

The new variable can be rewritten by introducing the $n\times(n+1)$ matrix : $\mathbf{A} = (\mathbf{1}_n, -\mathbf{E}_n)$ where $\mathbf{1}_n = (1,\ldots,1)'$: the vector of size n with all elements 1 and $\mathbf{E}_n$ is the unit matrix of order n, and $\mathbf{y} = \mathbf{A}(x_0,\mathbf{x}')'$ and $\mu = \mathbf{A}(\theta_0,\theta')'$.

Now we introduce the auxiliary quantities ;

$$w^{(0)} = w_0^{-1}, w^{(k)} = w_k^{-1}, u_{i,k} = \sum_{t=i}^{k} w_t, \bar{y}^{(k)} = \frac{1}{u_{1,k}}\sum_{t=1}^{k} w_t y_t,$$

where $i=0$ or $1, k=1,2,\ldots,n$.

Note the relation, $\mathbf{E}_n = \mathbf{1}_n\mathbf{1}_n'$ and let $\Lambda = A\mathrm{diag}(w_0,W)^{-1}A' = W^{-1}+w^{(0)}\mathbf{E_n}$. From the Binomial inverse theorem of Siotani et al.(1985) and Problem 2.9 of Rao(1973), we have $\Lambda^{-1} = W - (1/u_{0,n})(w_iw_j)$.

In order to solve (2.2), we first solve the minimizing problem

$$\min_{\mu\geq 0}(\mu-\mathbf{y})'\Lambda^{-1}(\mu-\mathbf{y}) = (\hat{\mu}-\mathbf{y})'\Lambda^{-1}(\hat{\mu}-\mathbf{y}) \quad (2.3)$$

to get the solution $\hat{\theta}(\theta_0) = (\hat{\theta}_1(\theta_0),\ldots,\hat{\theta}_n(\theta_0))'$. In case the original observation vector $(x_0,\mathbf{x}')$ is in the simple tree order, the relation $y_1 \geq 0,\ldots,y_n\geq 0$ holds true, where we have $\hat{\mu}_1 = y_1,\ldots,\hat{\mu}_n = y_n$; hence $\hat{\theta}_1(\theta_0) = x_1,\ldots,\hat{\theta}_n(\theta_0) = x_n$ is the solution to the problem (2.2). This implies that the observation vector itself is the solution.

The problem arises when the observation vector $(x_0,\mathbf{x}')$ is not in the simple tree order. Let the minimizing problem (2.3) be the primal problem, and then its dual problem is the minimizing problem given in the following form:

$$\min_{\mu^*\geq 0}(\mu^*+\Lambda^{-1}\mathbf{y})'\Lambda(\mu^*+\Lambda^{-1}\mathbf{y}) = (\hat{\mu}^*+\Lambda^{-1}\mathbf{y})'\Lambda(\hat{\mu}^*+\Lambda^{-1}\mathbf{y}). \quad (2.4)$$

By $\mu^* = W\nu^*$, (2.4) becomes

$$\min_{\nu^*\geq 0}(\nu^* + W^{-1}\Lambda^{-1}\mathbf{y})'W'\Lambda W(\nu^* + W^{-1}\Lambda^{-1}\mathbf{y})$$
$$= (\hat{\nu}^* + W^{-1}\Lambda^{-1}\mathbf{y})'W'\Lambda W(\hat{\nu}^* + W^{-1}\Lambda^{-1}\mathbf{y}). \tag{2.5}$$

We now apply Proposition 3.1 of Choi(1976). The problem can be solved by applying the Sweep-the-Negatives-algorithm on the matrix of size $(n, \overline{n+1})$:

$$\left[W^{-1}\Lambda^{-1}W^{-1\prime} - W^{-1}\Lambda^{-1}\mathbf{y}\right] = \begin{pmatrix} w^{(1)} - u_{0,n}^{-1} & & -u_{0,n}^{-1} & u_{0,n}^{-1}u_{1,n}\overline{y}^{(n)} - y_1 \\ & \ddots & & \vdots \\ -u_{0,n}^{-1} & & w^{(n)} - u_{0,n}^{-1} & u_{0,n}^{-1}u_{1,n}\overline{y}^{(n)} - y_n \end{pmatrix} \tag{2.6}$$

In order to see what the solution will be like, we assume, without loss of generality, the original observation obeys the ordering of the form $x_1 \geq x_2 \geq \cdots \geq x_n$. This implies the inequalities

$$u_{0,n}^{-1}u_{1,n}\overline{y}^{(n)} - y_1 \geq u_{0,n}^{-1}u_{1,n}\overline{y}^{(n)} - y_2 \geq \cdots \geq u_{0,n}^{-1}u_{1,n}\overline{y}^{(n)} - y_n.$$

This means that in the above $(n, \overline{n+1})$ matrix the following property holds true. All of the off-diagonal elements in the left (n, n) square matrix are non-positive while those of the last column are arranged in non-increasing order from the top to the bottom of it. If the observation vector $(x_0, \mathbf{x}')$ is not in simple tree order, there will be a cluster of negative elements in the bottom of the last column, at least one cluster, and all the remainings are positive. By sweeping out this matrix taking the (n, n) element as the pivot, the resulting matrix is of the form:

$$\begin{pmatrix} w^{(1)} - u_{0,n-1}^{-1} & & -u_{0,n-1}^{-1} & 0 & u_{0,n-1}^{-1}u_{1,n-1}\overline{y}^{(n-1)} - y_1 \\ & \ddots & & \vdots & \vdots \\ -u_{0,n-1}^{-1} & & w^{(n-1)} - u_{0,n-1}^{-1} & 0 & u_{0,n-1}^{-1}u_{1,n-1}\overline{y}^{(n-1)} - y_{n-1} \\ -w_n u_{0,n-1}^{-1} & \cdots & -w_n u_{0,n-1}^{-1} & 1 & w_n(u_{0,n-1}^{-1}u_{1,n-1}\overline{y}^{(n-1)} - y_n) \end{pmatrix}.$$

The upper left square matrix is order $n-1$, while the last column are of the same property as in the original matrix. About sweep out operation, we can take submatrix as a pivot, but in this paper, for convenience of explanation, we regard the terminology "sweep out operation" as the one taking a diagonal element as the pivot. If the element next to the last in the far right column is negative, we can proceed the same manner until the upper k elements become positive in the last column, the lower $m = n - k$ become negative, and lower right square matrix of order m becomes the largest unit submatrix in the left square matrix of order n. The forms are:

$$\begin{pmatrix} w^{(1)} - u_{0,k}^{-1} & & -u_{0,k}^{-1} & 0 & \cdots & 0 & u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_1 \\ & \ddots & & \vdots & & \vdots & \vdots \\ -u_{0,k}^{-1} & & w^{(k)} - u_{0,k}^{-1} & 0 & \cdots & 0 & u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_k \\ -w_{k+1}u_{0,k}^{-1} & \cdots & -w_{k+1}u_{0,k}^{-1} & 1 & & 0 & w_{k+1}(u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_{k+1}) \\ \vdots & & \vdots & & \ddots & & \vdots \\ -w_n u_{0,k}^{-1} & \cdots & -w_n u_{0,k}^{-1} & 0 & & 1 & w_n(u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_n) \end{pmatrix}.$$

Note the same properties remain true in the original and the final matrices, and also, if any, those in the intermediate stages. The necessary and sufficient condition for the termination of the Sweep-the-Negatives algorithm after the m-th operation is $u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_k \geq 0$.

This condition can be rewritten in the following simple form in terms of $(x_0, \mathbf{x})$ $(x_1 \geq \cdots \geq x_n)$:

$$\widetilde{x}^{(k)} - x_k \leq 0. \tag{2.7}$$

where $\widetilde{x}^{(k)} = \dfrac{1}{u_{0,k}}\sum_{i=0}^{k} w_i x_i = u_{0,k}^{-1}(w_0 x_0 + u_{1,k}\overline{x}^{(k)}),\ \overline{x}^{(k)} = \dfrac{1}{u_{1,k}}\sum_{i=1}^{k} w_i x_i \quad (k = 1, 2, \cdots, n).$

The solution of (2.5) is $\widehat{\nu}^* = (u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_1, \ldots, u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_k, 0, \ldots, 0)'$.

The solution of the dual minimization problem (2.4) is $\widehat{\mu}^* = (w_1(u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_1), \ldots, w_k(u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)} - y_k), 0, \ldots, 0)'$.

Therefore the solution of the primal problem (2.3) is

$$\widehat{\mu} = \mathbf{y} + \Lambda\widehat{\mu}^* = (0, \ldots, 0, y_{k+1} - u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)}, \ldots, y_n - u_{0,k}^{-1}u_{1,k}\overline{y}^{(k)})', \tag{2.8}$$

which yields the intermediate solution $\widehat{\theta}(\theta_0) = \theta_0 - \widehat{\mu}$. Putting the above results into the minimization problem (2.1), we can easily see that the solution is expressible in terms of the original observation

$$(\widehat{\theta}_0, \widehat{\theta}')' = (\widetilde{x}^{(k)}, \ldots, \widetilde{x}^{(k)}, x_{k+1}, \ldots, x_n)', \tag{2.9}$$

where k is determined by the condition (2.7). This is identical with the solution obtained by the PMVA.

3. CONCLUDING REMARKS

It is interesting to note the following relation. Let k be the required times of applying the PMVA, and the same solution can be obtained by applying the sweep out operations $l = (n - k)$ times on the dual minimization problem. This is natural and quite appealing because it reflects the geometric nature of the primal and its dual problem.

In Kim et al.(1990), close parallelism was explored , in simple order case, between the PAVA and the Sweep-the-Negatives algorithm. In them, the operation of PAVA or Sweep-the-Negatives can be done once all. After this operation, new adjacent violators or negatives may come out again, requiring the further step although eventual termination of the process is guaranteed.

Let the initial number of violators be s in simple tree order case, and then there will be $t = n - s$ negative components in the last vector of the dual problem. Hence the minimum possible number of necessary sweep out is t but possibly more than t. This means that the initial s violators should not be pooled once in altogether.

It is interesting to see the difference between simple and simple tree order. In the former the required number of pooling is at least the number of initial violating pairs and in the latter it is at most the number of initial violators. Another aspect is that in the former there are a number of paths leading to the solution whereas in the latter there is only one advisable path although there may be some others leading to the same solution.

REFERENCES

(1) Choi,J.R.(1976). The law of Sweep-the-Negatives in the estimation under order restrictions. *Memoirs of the Faculty of Science, Kyushu University*, A, Vol. 30, 135-143.

(2) Kennedy,W.J. and Gentle,J.E.(1980). *Statistical Computing*, New York, Marcel Dekker, Inc..

(3) Kudô,A.(1963). A multivariate analogue of the one-sided test. *Biometrika*, Vol. 50, 403-418.

(4) Kim,B.D., Choi,J.R. and Kudô,A.(1990). The Algorithm of Sweep-the-Negative and Its Applications to Order Restricted Inference. *Journal of The Korean Statistical Society*, Vol. 19, 54-70.

(5) Kudô,A. and Choi,J.R.(1975). A generalized multivariate analogue of the one-sided test. *Memoirs of the Faculty of Science, Kyushu University*, A, Vol. 29, 303-328.

(6) Rao,C.R.(1973). *Linear Statistical Inference and Its Applications.* 2nd ed., New York, John Wiley and Sons.

(7) Robertson,T.,Wright,F.T. and Dykstra,R.L.(1988). *Order restricted statistical inference*, New York, John Wiley and Sons.

(8) Siotani,M.,Hayakawa,T. and Fujikoshi,Y.(1985). *Modern Multivariate Statistical Analysis, A Graduate Course and Handbook*, American Sciences Press, Inc..