

# A PROPOSED ALGORITHM FOR LEAST ABSOLUTE ERROR ESTIMATION

## PART II

by

Bijan BIDABAD\*

Regional Planning Bureau, Plan & Budget Ministry; and Islamic Azad  
University; Tehran, Iran.

An algorithm for the simple restricted linear model was previously given by Bidabad. This note deals with the general linear model. An algorithm is therefore proposed for this case which is highly efficient in compare to linear programming.

### 1. Introduction

In a recent paper Bidabad [1], an efficient algorithm proposed for Least Absolute Error (LAE) estimate of  $b_2$  in the simple model,

$$y_i = b_2 x_{2i} + u_i \quad (1)$$

where  $b_2$  is population parameter,  $y_i, x_{2i}$  and  $u_i$  dependent, independent and random variables respectively. This paper tries to find an algorithm for LAE estimate of  $b_j$  for  $j=1, \dots, m$ , in the general linear model,

$$y_i = \sum_{j=1}^m b_j x_{ji} + u_i \quad (2)$$

LAE estimates of  $b_j$  for  $\{j=1, \dots, m\}$  are attained by minimizing

---

\* I would like to express my sincer appreciation to Dr. M.J. Mojarrad for his guidance and encouragement at each stage of this endeavor. He has read the manuscript carefully and made interesting comments. The author is also indebted to Dr. A. Monajemi and Dr. F. Ahrabi for their valuable suggestions.

$$S = \sum_{i=1}^n |u_i| = \sum_{i=1}^n \left| y_i - \sum_{j=1}^m b_j x_{ji} \right| \quad (3)$$

The problem of LAE estimation as described by Taylor [3] and Bidabad [1] can be reduced to find those observations which have zero errors in the minimand solution of (3).

## 2. The Proposed Algorithm

To find the LAE estimates of  $\{b_j, j=1, \dots, m\}$  for (2), an algorithm is proposed to search those observations on the regression hyperplane. For  $m=2$  and  $x_{1i} = 0, i=1, \dots, n$ , (2) reduces to (1). The LAE estimate of  $b_2$  for the simple model can be obtained by the algorithm proposed by Bidabad [1]. Now, let us consider a simple unrestricted linear model in which  $m=2$  and  $x_{1i}, i=1, \dots, n$ ,

$$y_i = b_1 + b_2 x_{2i} + u_i \quad (4)$$

The objective function  $S$  to be minimized will be,

$$S = \sum_{i=1}^n |y_i - b_1 - b_2 x_{2i}| \quad (5)$$

Let  $k_1$  denotes a subscript which is in the range one to  $n$ , and assume that the  $k_1$ th observation  $(x_{2k_1}, y_{k_1})$  is a candidate to be on the regression line. If this is the case, then  $u_{k_1} = 0$  and we can transfer the origin of the  $X_2 - Y$  coordinates to the point  $(x_{2k_1}, y_{k_1})$  without any loss. For this, we should rewrite all observations as deviations from the point  $(x_{2k_1}, y_{k_1})$ ,

$$y_i^{k_1} = y_i - y_{k_1} \quad i=1, \dots, n \quad (6)$$

$$x_{2i}^{k_1} = x_{2i} - x_{2k_1} \quad i=1, \dots, n$$

Rearranging the terms,

Rewriting (6) and (7) for  $k_1 = p, q$  and substituting them into (8), yields

$$S_p = \sum_{i=1}^n |y_i^p - b_2 x_{2i}^p| \quad (11)$$

$$S_q = \sum_{i=1}^n |y_i^q - b_2 x_{2i}^q| \quad (12)$$

Using (6) and rewriting (11) and (12), it can be shown that  $S_p$  is equal to  $S_q$ ,

$$S_p = \sum_{i=1}^n |y_i - b_2 x_{2i} - (y_p - b_2 x_{2p})| \quad (13)$$

$$S_q = \sum_{i=1}^n |y_i - b_2 x_{2i} - (y_q - b_2 x_{2q})| \quad (14)$$

$S_p$  is equal to  $S_q$  if and only if the two parantheses inside the absolute value signs in (13) and (14) are equal. This can be concluded by solving the two equations in (10) for  $b_1$ , that is,

$$b_1 = y_p - b_2 x_{2p} = y_q - b_2 x_{2q} \quad (15)$$

Thus,  $S_p = S_q$ , so  $b_2^p$  and  $b_2^q$  derived from minimizing either  $S_p$  or  $S_q$  must be equal.

This gives a criterion to find the desired  $b_2$  from all  $b_2^{k_1}$ , that is when  $b_2^p = b_2^q$ . The estimated value of  $b_2$  is denoted by  $\hat{b}_2$ . Value of  $\hat{b}_1$  is simply computed from (15). Now let us summarize the whole procedure for finding the values of  $\hat{b}_1$  and  $\hat{b}_2$  for the model given in (4).

#### Crude Algorithm:

Step 0: Set  $k_1 = 1$ .

Step 1: Compute (6).

Step 2: Minimize (9) using Bidabad [1] and find  $b_2^{k_1}$ .

$$y_i = y_i^{kl} + y_{kl}$$

(7)

$$x_{2i} = x_{2i}^{kl} + x_{2kl}$$

Now substitute (7) into (5), then, the LAE minimization problem can be redefined as,

$$\min_{b_1, b_2} : S_{kl} = \sum_{i=1}^n |y_i^{kl} - b_2 x_{2i}^{kl} + (y_{kl} - b_1 - b_2 x_{2kl})| \quad (8)$$

Since we assumed that the  $k$ th observation is on the regression line, the term  $y_{kl} - b_1 - b_2 x_{2kl} = 0$ , thus (8) reduces to

$$\min_{b_2} : S_{kl} = \sum_{i=1}^n |y_i^{kl} - b_2 x_{2i}^{kl}| \quad (9)$$

The solution of the optimization problem in (9) is the same as described by Bidabad [1]. Note that when  $u_{kl} = 0$ , then the minimum value of (9) is equal to the minimum value of (5).

Let  $b_2$  derived from minimizing (9) be denoted by  $b_2^{kl}$ . By changing  $k$  from one to  $n$  and minimizing (9),  $b_2^1, \dots, b_2^n$  are attained accordingly. Now the question is this: what value of  $k$  minimizes (5)? In other words, which observations are on the regression line? Note that in two parameters LAE linear model, there exists at least two observations with zero errors. Transferring of  $X_2 - Y$  coordinates to both these points leads the minimum of (5) unchanged. Suppose  $p$ th and  $q$ th are those two observations on the regression line, thus

$$u_p = y_p - b_1 - b_2 x_{2p} = 0 \quad (10)$$

$$u_q = y_q - b_1 - b_2 x_{2q} = 0$$

Step 3: Check if  $b_2^{kl} = b_2^{kl-k}$  for  $0 < k < kl$ , then set  $\hat{b}_2 = b_2^{kl}$  and

$$\hat{b}_1 = y_{kl} - \hat{b}_2 x_{2kl} \text{ then stop.}$$

Step 4: Increase  $kl$  by one and go to step 1.

However, this crude algorithm for finding  $\hat{b}_1$  and  $\hat{b}_2$  is not computationally efficient because it usually requires testing the majority of observations. Therefore, in order to make our algorithm efficient some elaborations are necessary.

Instead of setting  $kl=1$ , let us set  $kl$  equal to an arbitrary value such as  $\alpha$ .  $\alpha$  is an integer from one to  $n$ . Suppose now,  $u_\alpha = y_\alpha - b_1 - b_2 x_{2\alpha} = 0$ , and rewrite (9)

$$\min_{b_2} : S = \sum_{i=1}^n |y_i^\alpha - b_2 x_{2i}^\alpha| \quad (16)$$

Minimizing (16) gives  $b_2^\alpha$  which is equal to:

$$b_2^\alpha = \frac{y_\beta^\alpha}{x_{2\beta}^\alpha} = \frac{y_\beta - y_\alpha}{x_{2\beta} - x_{2\alpha}} \quad (17)$$

By pivoting on  $\alpha$ th observation another point such as  $\beta$  is found which  $\beta$  refers to that observation which has zero error in the minimand solution of (16), so  $u_\beta = y_\beta - b_1 - b_2 x_{2\beta} = 0$ .

Let us denote minimum of  $S_\alpha$  in (16) as  $S_\alpha^*$ ,

$$S_\alpha^* = \sum_{i=1}^n |y_i^\alpha - b_2^\alpha x_{2i}^\alpha| \quad (18)$$

Note that

$$b_1 = y_\alpha - b_2^\alpha x_{2\alpha} = y_\beta - b_2^\alpha x_{2\alpha} \quad (19)$$

This comes from multiplying (17) by its denominator and rearranging the terms. Substitute (19) in (5),

$$S = \sum_{i=1}^n |y_i - y_\alpha - b_2^\alpha (x_{2i} - x_{2\alpha})| = \sum_{i=1}^n |y_i - y_\beta - b_2^\alpha (x_{2i} - x_{2\alpha})|$$

or,

$$S = \sum_{i=1}^n |y_i^\alpha - b_2^\alpha x_{2i}^\alpha| = \sum_{i=1}^n |y_i^\beta - b_2^\alpha x_{2i}^\beta| \quad (20)$$

Using (18) the first sum in (20) is  $S_\alpha^*$  and the second sum is  $S_\beta$  evaluated at  $b_2 = b_2^\alpha$ . Thus it can be concluded that:

$$S_\alpha^* = S_\beta \Big|_{b_2 = b_2^\alpha} \quad (21)$$

$S_\alpha^*$  is at minimum but  $S_\beta \Big|_{b_2 = b_2^\alpha}$  can be minimized yet for other values of  $b_2$ .

Therefore, an important result is derived, that is

$$S_\alpha^* > S_\beta^* \quad (22)$$

Inequality term in (22) guarantees that if we choose an arbitrary point to transfer the origin of coordinates to it and minimize the objective function (9) another point is found, then transfer of the origin of coordinates to the newly found point decreases the total sum of absolute errors. Therefore, at each transfer point we get near to the minimum of  $S$ . By a similar discussion, it can be generalized that

$$S_\alpha^* > S_\beta^* > S_\gamma^* > S_\delta^* \dots \quad (23)$$

Note that  $\alpha$  is an arbitrary starting value.  $\beta$  is derived by minimizing  $S_\alpha$ ,  $\gamma$  is derived by minimizing  $S_\beta$  and  $\delta$  is derived by minimizing  $S_\gamma$  and so on.

Now, the question is when the minimum value of  $S$  is reached? Suppose  $S^* = S_\tau^*$ , by transferring the origin of coordinates to the point  $\tau$  and minimizing  $S_\tau$  the  $v$ th observation is derived. When  $S^* = S_\tau^*$  by minimizing  $S_\tau$ ,  $v$ th observation is again found because  $S_v^* = S_\tau^* = S^*$  and the  $v$ th and  $\tau$ th observations are both on the LAE regression line. This conclusion

gives a criterion to stop the procedure. Hence

$$S_{\alpha}^* > S_{\beta}^* > S_{\gamma}^* > S_{\delta}^* > \dots > S_{\tau}^* = S_{\nu}^* = S^* \quad (24)$$

It should be noted that if the minimum solution of  $S$  is not unique, that is function  $S$  has a horizontal segment, the procedure stops when it reaches the first minimum solution.

Now, let us introduce the complete stages of the algorithm to find the LAE estimates of  $b_1$  and  $b_2$  in the simple linear model  $y_i = b_1 + b_2 x_{2i} + u_i$ ,

#### Efficient Algorithm

Step 0: Select an arbitrary observation  $\alpha$  and set  $k_1 = \alpha$ .

Step 1: Compute (6) with  $k_1 = \alpha$ .

Step 2: Minimize (9) using Bidabad [1] and find that observation which locates on the line; observation  $\beta$ .

Step 3: Compute (6) with  $k_1 = \beta$ .

Step 4: Minimize (9) and find that observation which locates on the line; observation  $\gamma$ .

Step 5: Check that if  $\gamma = \alpha$  then  $\hat{b}_2 = y_{\beta}^{\gamma} / x_{2\beta}^{\gamma}$ ,  $\hat{b}_1 = y_{\gamma} - \hat{b}_2 x_{2\gamma}$  and stop.

Step 6: Set  $\alpha = \beta$  and go to step 1.

Now we extend the above procedure for two parameters restricted model:

$$y_i = b_2 x_{2i} + b_3 x_{3i} + u_i \quad (25)$$

Let,

$$S = \sum_{i=1}^n |y_i - b_2 x_{2i} - b_3 x_{3i}| \quad (26)$$

$S$  can be written as:

$$S = \sum_{i=1}^n |x_{2i}| \left| \frac{y_i}{x_{2i}} - b_2 - b_3 \frac{x_{3i}}{x_{2i}} \right| = \sum_{i=1}^n |x_{2i}| |y_i^{sl} - b_2 - b_3 x_{3i}^{sl}| \quad (27)$$

in which,

$$y_i^{sl} = y_i / x_{2i} \quad i=1, \dots, n \quad (28)$$

$$x_{3i}^{sl} = x_{3i} / x_{2i} \quad i=1, \dots, n$$

Minimization of (27) is similar to a linear simple model which explained above. An important distinction for solving (27) compare to (5) is the expression  $|x_{2i}|$  which has been multiplied to  $|y_i^{sl} - b_2 - b_3 x_{3i}^{sl}|$ . This multiplication does not make any problem when (27) is minimized, because if

$$y_i^{slkl} = y_i^{sl} - y_{kl}^{sl} \quad i=1, \dots, n \quad (29)$$

$$x_{3i}^{slkl} = x_{3i}^{sl} - x_{3kl}^{sl} \quad i=1, \dots, n$$

then we can rewrite (27) similar to (9), so

$$S_{kl} = \sum_{i=1}^n |x_{2i}| |y_i^{slkl} - b_3 x_{3i}^{slkl}| \quad (30)$$

To minimize (30) following Bidabad's procedure [1], we should use the following expression,

$$S_{kl} = \sum_{i=1}^n |x_{2i} x_{3i}^{slkl}| |y_i^{slkl} / x_{3i}^{slkl} - b_3| \quad (31)$$

According to Bidabad [1], in applying discrete derivative to (31), the content of the first absolute value sign is used to find the subscript of that point which locates on the regression line. This is the main difference in compare to the simple unrestricted linear model (4).



For the case of including an intercept in the model given in (25), we have

$$S = \sum_{i=1}^n |y_i - b_1 - b_2 x_{2i} - b_3 x_{3i}| \quad (32)$$

Let  $k_2$  be an arbitrary subscript, then, transfer of origin of coordinates to the point related to  $k_2$  is done by deviating all observations from this point. namely,

$$\begin{aligned} y_i^{k_2} &= y_i - y_{k_2} & i=1, \dots, n \\ x_{2i}^{k_2} &= x_{2i} - x_{2k_2} & i=1, \dots, n \end{aligned} \quad (33)$$

$$x_{3i}^{k_2} = x_{3i} - x_{3k_2} \quad i=1, \dots, n$$

Rearrange terms of (33) and substitute in (32), we have

$$S = \sum_{i=1}^n |y_i^{k_2} - b_2 x_{2i}^{k_2} - b_3 x_{3i}^{k_2} + (y_{k_2} - b_1 - b_2 x_{2k_2} - b_3 x_{3k_2})| \quad (34)$$

If  $k_2$ th observation is on the regression plane, then

$$y_{k_2} - b_1 - b_2 x_{2k_2} - b_3 x_{3k_2} = 0 \quad (35)$$

So instead of minimizing (32), the followign function is to be minimized:

$$S_{k_2} = \sum_{i=1}^n |y_i^{k_2} - b_2 x_{2i}^{k_2} - b_3 x_{3i}^{k_2}| \quad (36)$$

Minimization of (36) is completely similar to that of (26) and can be

proceed as follow:

$$\begin{aligned} S_{k_2} &= \sum_{i=1}^n |x_{2i}^{k_2}| |y_i^{k_2}/x_{2i}^{k_2} - b_2 - b_3 x_{3i}^{k_2}/x_{2i}^{k_2}| \\ &= \sum_{i=1}^n |x_{2i}^{k_2}| |y_i^{s1} - b_2 - b_3 x_{3i}^{s1}| \end{aligned} \quad (37)$$

in which

$$y_i^{s1} = y_i^{k2} / x_{2i}^{k2} \quad (38)$$

$$x_{3i}^{s1} = x_{3i}^{k2} / x_{2i}^{k2}$$

Now, again, transfer the origin of the two dimensional  $Y^{s1} - X_3^{s1}$  coordinates to the arbitrary point  $k1$  by deviating  $y_i^{s1}$  and  $x_{3i}^{s1}$  from  $y_{k1}^{s1}$  and  $x_{3k1}^{s1}$  as follow:

$$y_i^{slk1} = y_i^{s1} - y_{k1}^{s1} \quad i=1, \dots, n \quad (39)$$

$$x_{3i}^{slk1} = x_{3i}^{s1} - x_{3k1}^{s1} \quad i=1, \dots, n$$

Rearranging the terms of (39) and substituting them in (37) and assuming the point  $k1$  is on the regression line, we can rewrite (37) as follow:

$$S_{k2k1} = \sum_{i=1}^n |x_{2i}^{k2}| |y_i^{slk1} - b_3 x_{3i}^{slk1}| \quad (40)$$

or,

$$S_{k2k1} = \sum_{i=1}^n |x_{2i}^{k2} x_{3i}^{slk1}| |y_i^{slk1} / x_{3i}^{slk1} - b_3| \quad (41)$$

The objective function (41) can be minimized as suggested by Bidabad [1]. Now, the procedure from (37) to (41) can be repeated with different values of  $k1$  as the Efficient Algorithm that was proposed for the simple linear model before. When the last point (M) in the process of minimizing (41) is reached, the origin of the three dimensional  $Y - X_2 - X_3$  coordinates ( $k2$ ) is transferred to this newly found point (M) and the procedure from (33) to (41) is again repeated with the exception that instead of assigning an arbitrary value to  $k1$ , we set  $k1$  equal to the previous value of  $k2$ . This procedure continues until the point found from minimizing (41) is equal to previous of  $k2$ . The values of  $\hat{b}_1, \hat{b}_2$  and  $\hat{b}_3$  can be computed according

to the followign formulas:

$$\begin{aligned} \hat{b}_3 &= y_M^{slk1} / x_{3M}^{slk1} \\ \hat{b}_2 &= y_{k1}^{sl} - \hat{b}_3 x_{3k1}^{sl} \\ \hat{b}_1 &= y_{k2} - \hat{b}_2 x_{2k2} - \hat{b}_3 x_{3k2} \end{aligned} \quad (42)$$

It should be noted that at each step it can be proved that we are getting near the minimum of S of (32). The procedure of proof is similar to that we did for the two parameters model (4). We have, therefore, omitted the proof here.

To generalize the above algorithm to the m parameters linear model (2), we should reduce the number of parameters in the same fashion as the three parameters model explained above. If the model is restricted, we can make it unrestricted by dividing all explanatory variables to one of them.

$$S = \sum_{i=1}^n |y_i - \sum_{j=2}^m b_j x_{ji}| = \sum_{i=1}^n |x_{2i} |y_i / x_{2i} - b_2 - \sum_{j=3}^m b_j x_{ji}| \quad (43)$$

If the model is unrestricted, we can make it restricted by deviating all observations from an arbitrary one.

$$S = \sum_{i=1}^n |y_i - b_1 - \sum_{j=2}^m b_j x_{ji}| = \sum_{i=1}^n |y_i^{k(m-1)} - \sum_{j=2}^m b_j x_{ji}^{k(m-1)}| \quad (44)$$

in which,

$$y_i^{k(m-1)} = y_i - y_{k(m-1)} \quad i=1, \dots, n \quad (45)$$

$$x_{ji}^{k(m-1)} = x_{ji} - x_{jk(m-1)} \quad \begin{aligned} i &= 1, \dots, n \\ j &= 2, \dots, m \end{aligned}$$

Therefore, following (43), (44) and (45) any  $m$  parameters models can be reduced to a simple restricted one parameter model and then solved according to Bidabad [1]. To do this transformation, if the model is unrestricted, we should deviate all observations from an arbitrary one and make the model restricted. Then divide all explanatory variables to one of them. This makes the model unrestricted. At this step we have reduced one of the parameters of the model. By continuing this process, we can reduce any  $m$  parameters model to a one parameter model. If the model is restricted we should start by dividing all explanatory variables by one of them which makes the model unrestricted. Now we can again reduce one of the parameters by applying the procedures explained above in order to transform unrestricted model to restricted one.

By solving one parameter model, we can then solve for the two parameters model and then three parameters model and so on.

The most delicate part of this proposed algorithm is that, at the starting point of the algorithm once  $k_1, k_2, \dots, k_{(m-1)}$  are selected arbitrarily, then the algorithm assigns the best possible values to the integers  $k_1, k_2, \dots, k_{(m-1)}$ .

To explain the procedure, let us deal with the four parameters unrestricted linear model. Once, the value of  $k_3$  is arbitrarily selected by deviating all observations from the  $k_3$ th observation. In this way, the model reduces to a three parameters restricted model. By dividing all explanatory variables to one of them, the model becomes completely similar to that of (32). By minimizing (32) according to the algorithm previously explained for the three parameters model, the subscript  $M$  corresponding to  $M$ th point is derived. This is the newly found point which its subscript  $M$  is assigned to  $k_3$ . The previous value of  $k_3$  assigned

to  $k_2$  and the previous value of  $k_2$  is assigned to  $k_1$  and whole procedure is repeated again. The procedure stops when value of  $M$  is equal to  $k_3$ .

The above important assigning technique which is essential for pivoting the origins of different size coordinates can be extended for more parameters as we did above. Again, we should note that at each succeeding step we get near the minimum of  $S$  of (2). The proof is omitted here.

### 3. Evaluations

In comparison with linear programming (LP) method, the proposed algorithm (PA) is very efficient in execution time, storage requirement and also time for swapping process. In order to compare these two algorithms 390 random experiments for different parameters models of type both restricted and unrestricted have been explored for each algorithm. In each experiment  $n$  normal random number  $u_i$  have been generated (see, M.J.Mojarrad [2]). Also,  $n \times m$ , fixed values have been selected for  $x_{ji}$  which are uniformly distributed.  $y_i$  have been computed for 10 different values of  $b_j$ 's. Three different sample size of 20, 50, 100 have been used for  $n$ . All runs have been done on machine BASF 7.68 (MVS). The software used to solve the linear programming is the package MPSX/370 V1M4 PTF7. FORTRAN VS (level 1983, optimizer 03) used to compute the proposed algorithm. The results are shown in table 1. In each cell of this table the average CPU time of 10 experiments is presented. Table 2 compares the storage requirements for both algorithms for different sample sizes, different number of parameters and variables.

Table 1.

Comparison of CPU time for LP for different sample sizes and different number of parameters and variables.

(average of 10 experiments)

par no.	var no.	n=20		n=50		n=100	
		PA	LP	PA	LP	PA	LP
1	2	0.109	9.553	0.112	9.963	0.139	11.527
2	2	0.110	9.650	0.127	10.112	0.156	11.809
2	3	0.116	9.607	0.137	9.991	0.172	11.149
3	3	0.123	9.696	0.159	10.199	0.234	12.178
3	4	0.126	9.561	0.165	10.098	0.247	11.519
4	4	0.134	9.723	0.249	10.394	0.376	12.714
4	5	0.141	9.730	0.250	10.123	0.429	11.815
5	5	0.178	9.776	0.384	10.476	0.935	13.306
5	6	0.205	9.776	0.494	10.229	1.262	11.936
6	6	0.339	9.780	1.106	10.629	2.933	13.953
6	7	0.328	9.721	1.161	10.430	2.792	12.310
7	7	0.529	9.689	2.163	10.241	6.830	14.039
7	8	0.551	9.661	2.106	10.438	5.549	12.841

In both algorithms pre-execution time of compiler and linkage-editor have been excluded.

Table 2.

Comparison of storage requirement for LP and PA for different sample size and different number of parameters and variables.

par no.	var no.	n=20		n=50		n=100	
		PA	LP	PA	LP	PA	LP
1	2	4262	59392	4982	71680	6182	86016
2	2	4894	59392	5614	71680	6814	86016
2	3	5240	59392	6200	71680	7804	86016
3	3	5822	63488	6782	71680	8386	90112
3	4	6046	63488	7126	71680	8930	90112
4	4	6840	63488	7920	71680	9724	90112
4	5	7100	63488	8304	71680	10308	90112
5	5	8084	63488	9288	75776	11292	90112
5	6	8382	63488	9706	75776	11948	90112
6	6	9690	63488	11010	75776	13258	94208
6	7	10024	63488	11462	75776	13906	94208
7	7	11508	63488	12960	75776	15416	94208
7	8	11876	63488	13448	75776	16142	94208

Byte

References

1. Bidabad B., A Proposed Algorithm For Least Absolute Error Estimation, Part I, submitted to the First International Conference on Statistical Data Analysis Based on the  $L_1$  Norm and Related Methods, 1987.
2. Mojarrad M.J., The Application of Comparative Monte Carlo Methods to Econometrics: An Efficient Monte Carlo Study of Finite Sample properties of Iterative Instrumental Variables Estimation, Ph.D. dissertation, university of Pennsylvania, 1977, pp. 125-144.
3. Taylor L.D., Estimation by Minimizing the Sum of Absolute Errors, in P. Zarembka(ed.), Frontiers in Econometrics, Academic Press, New York, 1974, pp. 169-190.