No. 213 (84-9)

On a Variable Selection Procedure for Morimune's Modified Limited Information Maximum Likelihood Method

by

Haruo ONISHI

February 1984

On a Variable Selection Procedure for Morimune's Modified Limited Information Maximum Likelihood Method

University of Tsukuba
Haruo ONISHI

Abstract: W.A. Fuller [5] proposed a modification of the limited information maximum likelihood (LIML) method to have it possess exact moments. Furthermore, the Fuller's modified LIML method is asymptotically unbiased to $O(T^{-1})$. K. Morimune [8] further modified the LIML method along the line of the Fuller's modified LIML method. The Morimune's modified LIML (MF-LIML) method also possesses exact moments and is asymptotically unbiased to $O(T^{-1})$. Since the MF-LIML method reduces the thick tails of the distribution in the LIML method, the MF-LIML method is appropriate to be used for the estimation of equations in a simultaneous equation model. The j-th best subset problem is formulated for the variable selection problem of the MF-LIML method. The solution to the first best subset problem (j=1) or the ultimately best subset among the J solutions obtained by solving the first to the J-th best subset problems (e.g. J=10) in one computer-run is proposed as a pragmatical solution to the variable selection problem for the MF-LIML method. Such a subset possesses the following characteristics in addition to the above statistical properties: (i) meaningful from the viewpoint of a research field in question, (ii) just- or over-identifiable, (iii) satisfying the magnitude conditions for (values of linear functions of) estimated coefficients (including the sign conditions), the Durbin-Watson test ([2], [3], [4]), the relative absolute error test and the turning point test [9], if applied, and (iv) showing the highest (or suffi-

ciently high) adjusted coefficient of determination. The proposed variable selection procedure can make the laborious work of estimation by the MF-LIML method less time-consuming, labor-consuming and costly.

1. Introduction

Variable selection procedures of estimation methods for equations in a simultaneous equation model (called a model from here on) have not been proposed so far in the literature. As a result, researchers have wasted much time, brain labor and other resources like paper and electricity in the work of estimation. When the variable selection problem is solved for the MF-LIML method, applied researchers will use this method more often. The author would like to formulate the j-th best subset problem for the MF-LIML method and introduce the computer package OEPP to solve it. The (ultimately) best subset obtained by solving the first (to the J-th) best subset problem(s) in one computer-run can be used as a pragmatical solution to the variable selection problem for the MF-LIML method. The OEPP has been developed for socio-economic analysis and forecasting by the author [10].

2. Derivation of All Possible Meaningful Subsets Estimable With the MF-LIML Method

The variable selection problem for the MF-LIML method is defined as finding efficiently the best subset, if it exists, that is (i) derivable from a given set of all possible included predetermined variable candidates (called all possible included candidates from here on), all possible explanatory endogenous variable candidates (called all possible endogenous candidates from here on) and all possible excluded predetermined variable candidates (called all possible excluded candidates from here on) specified for an explained endogenous variable (called an explained variable from here on), (ii) meaningful from the viewpoint of the information (e.g. theories, field surveys, experiments, empirical studies, etc.) from a research field in question, e.g. economics, (iii) estimable with the MF-LIML method and (iv) satisfactory with respect to estimated coefficients and observations. Condition (i) implies that if N possible included, endogenous and excluded candidates are specified for an explained variable, the best subset is one of 2^N-1 possible subsets when it exists. For instance, when N=20, the best subset is one of $2^{20}-1=1048575$ possible subsets. Condition (ii) implies that the best subset must have only included and endogenous candidates necessary for an equation which is reasonable for the research but never has any unnecessary included and endogenous ones. It must be kept in mind that condition (ii) does not refer to excluded candidates at all. Condition (iii) implies that the best subset must be just- or over-identifiable (called identifiable from here on). Condition (iv) is checked through information from the research field. Conditions (ii) and (iii) must be satisfied before estimation.

Thus, it is of great importance to derive only meaningful and identifiable subsets from a given set of all possible included, endogenous and excluded candidates specified for an explained variable. The author [9] pointed out that when the ordinary least squares method is

Table 1. Variable Classifications for the MF-LIML Method by Functional Format $y=F(X^1:Y:X^2)$ in Case of Three Candidates A, B and C

Names	Classifications	Selection of Candidates
Absolutely Important	/A,B,C/	(1) A,B,C
Optionally Important	<A,B,C>	(1) A,B,C; (2) A,B; (3) A,C; (4) B,C; (5) A; (6) B; or or (7) C
Exclusively Important	</A,B,C/>	(1) A; (2) B; or (3) C
Gradually Important	<+A,B,C+>	(1) A,B,C; (2) A,B; or (3) A
Exclusively Optional	<*A,B,C*>	(1) A; (2) B; (3) C; or (4) Empty (no selection)
Gradually Optional	<-A,B,C->	(1) A,B,C; (2) A,B; (3) A; or (4) Empty
Completely Optional	A,B,C	(1) A,B,C; (2) A,B; (3) A,C; (4) B,C; (5) A; (6) B; (7) C; or (8) Empty
Fixed	(D,E)	For instance, if B=(D,E) above, all B's must be replaced with D,E.
NAI & NU Included Predetermined	'A,B,C'	If all or some of A, B and C are not selected as included candidates, they must be selected as excluded ones.
Constant Term	$C	Always selected, if any.

Footnotes: (1) NAI & NU stands for non-absolutely-important and non-uniquely. (2) y=explained variable, X^1=set of included candidates, Y=set of endogenous candidates, X^2=set of excluded or NAI & NU included candidates. (3) X^1, Y and X^2 are classified by the above rules. (4) Candidates A, B and C can be expressed with at most 8 alphanumeric symbols (and minus time lag numbers in parentheses like HLWK(-2)). (5) Candidates in a functional format are separated from each other by a blank, a comma (,), /, <, >, </, />, <+, +>, <*, *>, <-, ->, (,), : or '. (6) For example, AA=F($C/BC/D:<+E E(-1)+>FG</I(I1,I2)/>:<J,K>'D'/L,M,N,PQ/). (7) OEPP can handle nested variable classifications.

utilized, explanatory candidates can be classified on the basis of information from a research field into 8 groups. These are (i) absolutely important (or forced or core), (ii) optionally important, (iii) exclusively important, (iv) gradually important, (v) exclusively optional, (vi) gradually optional, (vii) fixed and (viii) completely optional groups. These variable classifications can be used for included, endogenous and excluded candidates for the MF-LIML method. We assume that an explained variable y and a set X of all possible included, endogenous and excluded candidates are loaded into a computer by a functional format like y=F(X) and the X is classified by the following rules: we enclose (i) absolutely important candidates within /.../, (ii) optionally important candidates within <...>, (iii) exclusively important candidates within </.../>, (iv) gradually important candidates within <+...+>, (v) exclusively optional candidates within <*...*>, (vi) gradually optional candidates within <-...->, (vii) fixed candidates within (...) and (viii) completely optional candidates freely in a functional format. Table 1 summarizes the meanings and characteristics of these basic classifications and how to select classified candidates.

The MF-LIML method requires that candidates be classified into 3 sets of included, endogenous and excluded ones. To distinguish these sets from each other, we postulate that a set X^1 of all possible included candidates x_k^1's ($0 \leqq k \leqq K$), a set Y of all possible endogenous candidates y_ℓ's ($1 \leqq \ell \leqq L$) and a set X^2 of all possible excluded candidates x_m^2's ($1 \leqq m \leqq M$) specified for an explained variable y are separated with a colon (:) in a functional format like $y=F(X^1:Y:X^2)$, where x_0^1 stands for a constant term, X^1 is entered first in a functional format, followed by Y which is followed by X^2, and sets X^1, Y and X^2 are classified on the basis of information from a research field and econometrics. In addition to the above variable classifications, we need one more variable classification from the technical viewpoint of the derivation of all possible identifiable subsets. We define a uniquely included candidate as the one which could appear in the equation at hand but never appears in the other equations and identities in a model. On the other hand, a non-uniquely included candidate is defined as the one which could appear in the equation at hand and does appear in at least one of the other equations and identities in a model. Only when non-uniquely included candidates are not absolutely important, must special attention be paid to these included ones. If a non-uniquely-included and non-absolutely-important candidate is not selected as an included one in a subset, such a candidate can be used as an excluded one in the same subset. To guarantee this case, we add non-uniquely-included and non-absolutely-important candidates to a group of excluded candidates X^2 and enclose them within '...'. Accordingly, non-uniquely-included and non-absolutely-important candidates are entered into not only X^1 but also X^2 in functional format $y=F(X^1:Y:X^2)$. When we derive only meaningful and identifiable subsets, we ignore a subset which (i) possesses the same candidate not only as an included one but also as an excluded one or (ii) does not possess at least one of the candidates enclosed within '...'. Now we can derive only meaningful and identifiable subsets from a given set of all possible included, endogenous and excluded candidates specified for an explained variable in any case.

Let us give an example. Suppose that a researcher enters the following functional format into a computer:

```
ABC=F($C/DD,E/<*F,GG*>:</HI(HJ1,HJ2)/>/KLM/:'GG'
     /PA,PB,PC/)                                          (1)
```

where $C stands for a constant term.

Then, he has the following 6 equations:

$$ABC=a_0+a_1DD+a_2E+a_3F+a_4HI+a_5KLM \quad (2)$$

$$ABC=b_0+b_1DD+b_2E+b_3GG+b_4HI+b_5KLM \quad (3)$$

$$ABC=c_0+c_1DD+c_2E+c_3F+c_4HJ1+c_5HJ2+c_6KLM \quad (4)$$

$$ABC=d_0+d_1DD+d_2E+d_3GG+d_4HJ1+d_5HJ2+d_6KLM \quad (5)$$

$$ABC=e_0+e_1DD+e_2E+e_3HI+e_4KLM \quad (6)$$

$$ABC=f_0+f_1DD+f_2E+f_3HJ1+f_4HJ2+f_5KLM \quad (7)$$

It is known from format (1) that F is a uniquely included candidate, while GG is a non-uniquely included candidate. Whether DD and E are uniquely or non-uniquely included candidates does not matter, because they are absolutely important. Equations (2), (4), (6) and (7) are estimated with the MF-LIML method by using excluded candidates GG, PA, PB and PC. However, equations (3) and (5) are estimated by using excluded candidates PA, PB and PC without GG. Equations (2), (3), (4), (6) and (7) are over-identified, whereas equation (5) is just-identified.

If 'GG' is replaced with GG in format (1), 4 more equations are estimated in addition to the above 6 equations. The additional equations have exactly the same forms as equations (2), (4), (6) and (7), but they are estimated with the MF-LIML method by using excluded candidates PA, PB and PC without GG. GG can be treated as a completely optional excluded candidate, when GG is not selected as an included candidate.

3. The j-th Best Subset Problem for the MF-LIML Method

We assume that there are N (cross-sectional) units like industries, classes, plots, households and T observation times to allow for time series, cross-sectional and pooled data and $y_n(t)$ stands for the explained variable's datum in unit n at observation t, where $1 \leq n \leq N$ and $1 \leq t \leq T$.

Let (X_i^1, Y_i, X_i^2) stand for the i-th meaningful and identifiable subset which is derived from $y=F(X^1:Y:X^2)$ and satisfies $X_i^1 \cap X_i^2 = \phi$, K_i, L_i and M_i for the numbers of candidates in X_i^1, Y_i and X_i^2, respectively, (A_i, B_i) for a row coefficient vector of (X_i^1, Y_i) and y as a vector for $(y_1(1), y_2(1), \ldots, y_N(1), \ldots, y_1(T), y_2(T), \ldots, y_N(T))'$. We can express the i-th meaningful and identifiable subset as follows:

$$y = X_i^1 A_i' + Y_i B_i' + u \tag{8}$$

where $u \sim N(0, \sigma^2 I)$ stands for a disturbance term and X_i^2 is used for the calculation of coefficients and the asymptotic variance-covariance matrix of coefficients.

We rewrite equation (8) as follows:

$$-(y, Y_i)(-b_0, B_i)' - X_i^1 A_i' = u \quad \text{with } b_0 = 1 \tag{9}$$

We would like to formulate the j-th best subset problem of equation (9) estimated with the MF-LIML method as follows:

Find subset (X_i^1, Y_i, X_i^2) derivable from a given set (X^1, Y, X^2) of all possible included, endogenous, and excluded candidates specified for an explained variable y and estimate coefficient vector $(\bar{A}_i, \bar{B}_i)$ such that

(I) subset (X_i^1, Y_i) is meaningful from the viewpoint of a research field in question,

(II) subset (X_i^1, Y_i, X_i^2) is just- or over-identifiable, namely, $1 \leqq L_i \leqq M_i$,

(III) subsets X_i^1 and X_i^2 have no common candidates, namely, $X_i^1 \cap X_i^2 = \phi$,

(IV) $(\bar{A}_i, \bar{b}_0, \bar{B}_i)$ and its variance-covariance matrix must be calculated as follows:

$$(W_i^1 - \bar{r}_i W_i)(-\bar{b}_0, \bar{B}_i)' = 0 \text{ with } \bar{b}_0 = 1 \tag{10}$$

and

$$\bar{A}_i' = (X_i^{1\prime} X_i^1)^{-1} X_i^{1\prime} (y, Y_i)(-\bar{b}_0, \bar{B}_i)' \tag{11}$$

with the estimated asymptotic variance-covariance matrix [6] of $(\bar{A}_i, \bar{B}_i)$

$$\bar{s}_i^2 \begin{bmatrix} G(X_i^1, X_i^1) & -G(X_i^1, Y_i) \\ -G(Y_i, X_i^1) & G(Y_i, Y_i) \end{bmatrix}$$

for the $\bar{r}_i$ defined as

$$\bar{r}_i = \tilde{q}_i - \bar{p}_i / (NT - K_i - M_i) \tag{12}$$

for $\bar{p}_i = 1 + \{(M_i - L_i)/(NT - K_i - M_i)\}\{(1 + \tilde{q}_i) / \sum_{j=1}^{L_i+1} (\tilde{q}_i - \tilde{q}_{ij})\}$

with respect to $\tilde{q}_i = \text{Min}\{\tilde{q}_{i1}, \tilde{q}_{i2}, \ldots, \tilde{q}_{i(L_i+1)}\}$ of (L_i+1) eigen values $\tilde{q}_{i1}, \tilde{q}_{i2}, \ldots, \tilde{q}_{i(L_i+1)}$ of

$$|W_i^1 - q_i W_i| = 0 \tag{13}$$

where

$$W_i^1 = (y, Y_i)' Z_i^1 (y, Y_i) \text{ for } Z_i^1 = I - X_i^1 (X_i^{1\prime} X_i^1)^{-1} X_i^{1\prime} \tag{14}$$

$$W_i = (y, Y_i)' Z_i (y, Y_i) \text{ for } Z_i = I - X_i (X_i' X_i)^{-1} X_i' \tag{15}$$

$$X_i = (X_i^1, X_i^2) \tag{16}$$

$$G(Y_i, Y_i) = \{Y_i' (\bar{r}_i Z_i - Z_i^1) Y_i\}^{-1} \tag{17}$$

$$G(Y_i, X_i^1) = G(Y_i, Y_i) Y_i' X_i^1 (X_i^{1\prime} X_i^1)^{-1} \tag{18}$$

$$G(X_i^1, Y_i) = (X_i^{1\prime} X_i^1)^{-1} X_i^{1\prime} Y_i G(Y_i, Y_i) \tag{19}$$

$$G(X_i^1, X_i^1) = (X_i^{1\prime} X_i^1)^{-1} \{I + X_i^{1\prime} Y_i G(Y_i, Y_i) Y_i' X_i^1 (X_i^{1\prime} X_i^1)^{-1}\} \tag{20}$$

$$\tilde{y}_i = X_i^1 \tilde{A}_i' + Y_i \tilde{B}_i' \quad (21)$$

$$\tilde{s}_i^2 = (y-\tilde{y}_i)'(y-\tilde{y}_i)/(NT-K_i-L_i) \quad (22)$$

(V) $\tilde{C}_i=(\tilde{A}_i,\tilde{B}_i)$ satisfies the following magnitude condition (including the sign condition), if necessary:

$$D_{hi}^1\tilde{C}_i \pm |D_{hi}^2\tilde{C}_i| \pm |D_{hi}^3\tilde{C}_i| \leq d_h^1,\ D_{hi}^1\tilde{C}_i \pm |D_{hi}^2\tilde{C}_i| \pm |D_{hi}^3\tilde{C}_i| \geq d_h^2,$$

$$\text{and/or } d_h^1 \leq D_{hi}^1\tilde{C}_i \pm |D_{hi}^2\tilde{C}_i| \pm |D_{hi}^3\tilde{C}_i| \leq d_h^2 \text{ for } 1 \leqq h \leqq H \quad (23)$$

where D_{hi}^k for k=1,2,3, d_h^1 and d_h^2 stand for a row vector of coefficients, a lower bound and an upper bound of the h-th magnitude condition, respectively, $|D_{hi}^k\tilde{C}_i|$ for k=2,3 stands for the absolute value of $D_{hi}^k\tilde{C}_i$, "$\geq$" and "$\leq$" indicate ">" or "$\geqq$" and "<" or "$\leqq$", respectively, and "$\pm$" indicates "+" or "-",

(VI) the Durbin-Watson statistic DW_i defined below is significant at a level specified by a researcher, when N=1, $T \geqq 15$ and $K_i+L_i \leqq 5$:

$$DW_i = \sum_{t=2}^{T} \{\tilde{e}_{i1}(t)-\tilde{e}_{i1}(t-1)\}^2 / \sum_{t=1}^{T} \tilde{e}_{i1}(t)^2 \quad (24)$$

for $\tilde{e}_{i1}(t)=y_1(t)-\tilde{y}_{i1}(t)$ for $1 \leqq t \leqq T$,

(VII) $\tilde{y}_i$ whose elements are denoted by $\tilde{y}_{in}(t)$'s satisfies the following relative absolute error test, if necessary:

$$100 \times |\{y_n(t)-\tilde{y}_{in}(t)\}/y_n(t)| \leqq w \text{ for } 1 \leqq n \leqq N \text{ and } 1 \leqq t \leqq T \quad (25)$$

where w (%) is specified by a researcher,

(VIII) $\tilde{y}_i$ satisfies the turning point test, if necessary:

if

$$\{y_n(t)-y_n(t-1)\}\{y_n(t+1)-y_n(t)\}<0 \quad (26)$$

and

$$100 \times \mathrm{Min}[|\{y_n(t)-y_n(t-1)\}/y_n(t)|,|\{y_n(t)-y_n(t+1)\}/y_n(t)|] \geqq v$$

then (27)

$$\{y_n(t)-y_n(t-1)\}\{\tilde{y}_{in}(t)-\tilde{y}_{in}(t-1)\}>0 \quad (28)$$

and

$$\{y_n(t+1)-y_n(t)\}\{\tilde{y}_{in}(t+1)-\tilde{y}_{in}(t)\}>0 \tag{29}$$

for $1\leqq n\leqq N$, $2\leqq t\leqq T-1$ and $T\geqq 3$

where v (%) is specified by a researcher,

and

(IX) $(\tilde{A}_i,\tilde{B}_i)$ shows the j-th highest adjusted coefficient of determination measured by

$$RR_i=\mathrm{Max}\{0,1-(1-R_i)(NT-1)/(NT-K_i-L_i)\} \tag{30}$$

where

$R_i=1-(y-\tilde{y}_i)'(y-\tilde{y}_i)/(y-\bar{y}E)'(y-\bar{y}E)$, $\bar{y}=\sum_{n=1}^{N}\sum_{t=1}^{T}y_n(t)/NT$,

and $E=(1,1,1,\ldots,1)'$ with dimension (NTx1).

Let us briefly explain the above problem. Conditions (I) and (V) are necessary from the viewpoint of information from a research field in question. Conditions (II), (III), (IV) and (VI) are related to statistics and econometrics. Conditions (VII), (VIII) and (IX) are important criteria to evaluate meaningful and identifiable subsets through the comparison of estimated observations with actual ones. When an estimated model which contains lagged endogenous candidates is used for the final test or forecasting, condition (VIII) becomes quite useful. Although various criteria have been proposed to measure a goodness of fit, we adopted an adjusted coefficient of determination to measure a goodness of fit, because it has been used so often in the literature of applied research and can assume a value between 1 (for the perfect fitting) and 0 (for the worst fitting).

If a researcher does not have any additional criterion, the best subset is defined as the one which (i) is meaningful, identifiable and estimable with the MF-LIML method, (ii) satisfies all discrete (or pass-or-fail) criteria applied from conditions (V), (VI), (VII) and (VIII), and (iii) has the highest adjusted coefficient of determination. However, when a researcher has to use a new criterion, the ultimately best subset is defined as the one which (i) is one of the best J subsets obtained by solving the first to the J-th best subset problems in one computer-run and (ii) satisfies the new criterion. Of course, he has to find by himself the ultimately best subset among the best J subsets through the new criterion. Finally, we regard the (ultimately) best subset as a pragmatical solution, which can be regarded as a pragmatically best subset, to the variable selection problem for the MF-LIML method.

4. An Example

By estimating a macro agricultural production function of Cobb-Douglas type with the data of Japanese agriculture from 1965 to 1979, we would like to demonstrate how to use the proposed procedure. Let us introduce the following variable notations: LY=log(agricultural outputs in real monetary terms), LL=log(labor), LKA=log(KA)=log(animal capital), LKP=log(KP)=log(plant capital), LKM=log(KM)=log (machine capital), LK=log(KA+KP+KM), LKR=log(KA+KP+KM*R), R=an estimated use rate of machine capital where $0 \leqq R \leqq 1$, LAX=log(A-X)=log(cultivated acreage minus damaged and abandoned acreage), LCAX=log((A-X)xMin(C,1)), C=a cropping index of rice where C=1, C>1 and C<1 for average, rich and poor harvest, respectively, LQ=log(intermediate goods and services), LWIQ=log(wheat import quantity), LRFI=log(real farm income), LRRPP=log(real producer price of rice determined by the government), LRWRF=log(real wage rate of farming), DVCS=dummy variable which assumes 1 for cold summer and 0 for normal or hot summer, T=time trend, LT=log(T) and \$C=constant term.

We assume that (i) LY is an explained variable, (ii) DVCS and LT are uniquely included candidates, (iii) T is a non-uniquely included candidate, (iv) DVCS is completely optional, (v) LT and T are exclusively optional, (vi) LL, LK, LKR, LAX, LCAX and LQ are endogenous candidates, (vii) LL is absolutely important, (viii) LK and LKR are exclusively important, (ix) LAX and LCAX are exclusively important, (x) LQ is completely optional, (xi) LWIQ, LRFI, LRRPP, LRWRF, LRFI(-1), LKA(-1) and LKM(-1) are excluded candidates, where, for example, LRFI(-1) stands for LRFI with time lag number 1, and (xii) each meaningful and identifiable subset must use all excluded candidates available to that subset as instrumental ones. The last assumption (xii) implies that all excluded candidates LWIQ to LKP(-1) must be treated as absolutely important and non-uniquely included candidate T which is not absolutely important must be added to the group of excluded candidates as 'T' in a functional format. Then, we can load

```
LY=F($C,DVCS<*LT,T*>:/LL/</LK,LKR/></LAX,LCAX/>LQ:'T'
/LWIQ,LRFI,LRRPP,LRWRF,LRFI(-1),LKA(-1),LKP(-1)/)   (31)
```

where y, X^1, Y and X^2 in the j-th best subset problem are as follows:

```
y={LY}, X^1={$C,DVCS<*LT,Y*>},
Y={/LL/</LK,LKR/></LAX,LCAX/>LQ}, and
X^2={'T'/LWIQ,LRFI,LRRPP,LRWRF,LRFI(-1),LKA(-1),LKP(-1)/}
```

17 non-constant candidates appear in format (31) so that we

have 2^{17}-1=131071 possible subsets derivable from format (31). However, by the information on candidates, we know that only 48 subsets are meaningful and over-identifiable. The remaining 131,023 subsets are meaningless or cannot be estimated with the MF-LIML method. Since the number of excluded candidates exceeds that of the maximum meaningful combination of endogenous candidates, there are no meaningful and just-identifiable subsets derivable from format (31).

In order to check for and avoid unrealistic production functions, we would like to introduce the following conditions: (i) a free sign (implying that a sign is not determined before estimation) for $C, a negative sign for DVCS and positive signs for LT, T, LL, LK, LKR, LAX, LCAX and LQ, (ii) 0.1<LL(0.5, (iii) 0.1<LK+LKR(0.5, (iv) 0.1<LAX+LCAX(0.6, (v) 0.1<LQ(0.3, (vi) 0.85<LL+LK+LKR+LAX +LCAX+LQ(1.15, (vii) 5 % Durbin-Watson statistic test, (viii) 5 % relative absolute error test, (ix) 10 % turning point test and (x) the minimum adjusted coefficient of determination is 0.7, where the variable notations in (i) to (vi) imply their coefficients and "(" stands for "≦". Condition (vi) is related to increasing, constant or decreasing returns to scale in the agricultural production.

When we solved the first best subset problem in one computer-run, we had the following equation at about 2 minutes 36 seconds CPU time by the FACOM-M200 (about 13 MIPS/CPU):

```
LY=1.03882+0.017300*T+0.238848*LL+0.596795*LCAX+0.148143*LK
  (2.3429)(0.008871) (0.170560)  (0.596795)    (0.059841)

RR=0.9016, SD=0.02031, K=2.020, FA=0.0123, DW=1.934,

EV=2.19151, 3.32492, 18.3337, 57.4175                    (32)
```

where numbers in parentheses, RR, SD, K, FA, DW and EV stand for standard deviations of asymptotic variances of coefficients, adjusted coefficient of determination, standard deviation of asymptotic variance of a disturbance term, k of k-class estimator, first-order autocorrelation coefficient, Durbin-Watson statistic and eigen values, respectively. The excluded candidates used for equation (32) are LWIQ, LRFI, LRRPP, LRWRF, LRFI(-1), LKA(-1) and LKP(-1). The non-uniquely included candidate T is selected in the equation, so that candidate T was not selected as an excluded candidate, as assumed above. The sum of the coefficients of candidates LL, LCAX and LK is 0.983786 which indicates slightly decreasing returns to scale in the agricultural production.

5. Summary

The MF-LIML method is useful for the esimation of equations in a simultaneous equation model. However, procedures to solve the variable selection problem for this estimation method have not been discussed so far in the literature. The author formulated the j-th best subset problem and proposed to regard the (ultimately) best subset of the first (to the J-th) best subset problem(s) as a pragmatical solution to the variable selection problem for the MF-LIML method. The proposed variable selection procedure installed in the package OEPP may be able to save much time, (brain) labor and other resources like paper and electricity which would have been wasted during a trial and error process of estimation.

References

[1] Anderson, T.W. and Rubin, H. (1950) The asymptotic properties of estimators of the parameters of a single equation in a complete system of stochastic equations, Annals of Mathematical Statistics 21 570-582.

[2] Durbin, J. and Watson, G.S. (1950) Testing for serial correlation in least squares regression I, Biometrica 37 409-428.

[3] Durbin, J. and Watson, G.S. (1951) Testing for serial correlation in least squares regression II, Biometrica 38 159-178.

[4] Durbin, J. and Watson, G.S. (1971) Testing for serial correlation in least squares regression III, Biometrica 58 1-19.

[5] Fuller, W.A. (1977) Some properties of a modification of the limited information estimator, Econometrica 45 (4) 939-953.

[6] Goldberger, A.S. (1966) Econometric Theory, 3rd ed., Wiley, NY.

[7] Liu, T.C. and Breen, W.J. (1969) The covariance matrix of the limited information estimator and the identification test, Econometrica 37 (2) 222-227.

[8] Morimune, K. (1983) Approximate distribution of the k-class estimators when the degree of overidentifiability is large compared to the sample size, Econometrica 51 (3) 821-841.

[9] Onishi, H. (1983) A variable selection procedure for econometric models, Computational Statistics and Data Analysis 1 (2) 85-95.

[10] Onishi, H. (1983) Computer Package OEPP for Socio-Economic Analysis and Forecasting, Institute of Socio-Economic Planning, University of Tsukuba.