A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S

FOLIA OECONOMICA 68, 1987

Bogusław Guzik *

SAMPLE GENERATION FOR SIMULATION STUDIES OF NON-LINEAR ECONOMETRIC MODELS

The practically used methods for estimation the parameters of a non-linear econometric model usually produce approximate solutions (estimates). So, the problem arises, how much the results obtained by means of approximate methods differ from the exact ones. The answer may be searched for either by formal analysis of convergency of algorithms or by simulational experiments. The second approach is dealt with in the paper.

Let us consider an approximate method M, for estimation the parameters of a non-linear econometric model[1]. On the basis of a numerical experiment the results produced by method M are to be compared to the results produced by an exact method[2], which is the ordinary least square method.

One can propose the following standard procedure:

1) fix a vector of observational results on a dependent variable, and a matrix of observational results on explanatory variables,

---

*Assistant Professor, Economic Academy in Poznań.

[1] "Model" will be understood as a function adjusted to the observational results.

[2] The comparisons can concern e.g. a degree of adjustment, undetermination parameters, stochastic characteristics etc. We assume that the method M is not identical to the standard method. This non-identity of methods can be the result of the difference of approximation (or estimation) criteria, or - if the method M criterion is (at least intetionally) identical to the criterion in the standard method - it may be the result of a different procedure of finding out an extreme point (or rather a point close to the extreme point, because the methods of searching for the minimum of the non-linear approximation criteria are usually iterative methods).

2) derive the vector of parameters of a given non-linear function by means of the ordinary l.s.m. and by means of the method M,

3) compare the results derived by means of both methods.

However, many difficulties arise there.

Firstly, the estimation of the parameters of a non-linear function is numerically difficult because it needs solving many (more or less) complicated systems of non-linear equations. For example, even in a very simple case where the function is a power one $f = a_1 x^{a_2}$ the system of equations:

$$a_1 \sum_l x_l^{2a_2} - \sum_l y_l x_l^{a_2} = 0$$

$$a_1^2 \sum_l x_l^{2a_2} \ln x_l - a_1 \sum_l y_l x_l^{a_2} \ln x_l = 0$$

have to be solved with regard to a_1 and a_2. The symbol l denotes a variable numbering the observational results ($l = 1, \ldots, L$), and x_l, y_l denote the l-th observational result on the explanatory variable and on the dependent variable, respectively.

Secondly, this type of systems can usually be solved only approximately. Consequently, the results of the method M would be compared with approximations of the results of the ordinary l.s.m.

For these reasons another approach to the type of simulation in question is suggested. The general scheme of this approach can be formulated as follows:

1) fix the vector a^u of the parameters of given non-linear function;

2) having the vector a^u and the matrix X of observational results on the explanatory variables find such a vector y^g of the dependent variable values, that application of the ordinary l.s.m. directly to the numerical data $\{y^g \vdots X\}$, provides a vector a^g identical to a^u;

3) approximate method M is employed to statistical data $\{y^g \vdots X\}$, vector of model parameters a^M is derived, and vector a^M is compared to vector a^u.

This approach, therefore, is of an indirect character. The problem is not in determining a vector of parameters a^u by means of the ordinary l.s.m. in terms of matrix X and vector y, but in generation of a vector y of the values of the dependent variable in terms of matrix X and vector of the parameters a^u.

The approach suggested here possesses some important advantages. The given vector a^u is an exact one; difficulties connected with solution of non-linear systems of equations are avoided, the problem - as we shall see later - consists in solving a system of homogenous linear equations; besides the values of some characteristics of special interest related to the standard method can be easily planned (making the analytical experiment something like a controlled experiment).

The aim of this paper is to present a method of generating the vector y^g. In order to avoid misunderstandings it should be emphasized that we are not interested in any special method M of estimation of the parameters of a non-linear model, or in any procedure of comparing results of this method with results of the ordinary l.s.m. Instead, we shall focus on a way of generation of the vector y^g such as to satisfy certain conditions.

1

Before formulation of a way of generation of y^g, it might be useful to mention the basic ideas of the ordinary l.s.m.

Let f(a, x) be the examined non-linear function with the parameters $a = (a_i)_{i=1}^{K}$ and explanatory variables $x = (x_j)_{j=1}^{P}$. Taking for granted that the vector $y^e = (y_l)$, $l = 1, \ldots, L$ of the observational results on the dependent variable and the matrix $X = (x_{lj})$ $l = 1, \ldots, L$; $j = 1, \ldots, P$ of the observational results on the explanatory variables are given, let us assume for the time being that we are going to estimate the vector a by means of the ordinary l.s.m.

We assume that $L > K$. Let x_l stands for the l-th row of the matrix X and e_l for the residual

$$(1) \qquad e_l = y_l^e - f_l$$

where

(2) $f_l = f(a, x_l), \quad l = 1, \ldots, L.$

The residual sum of squares

(3) $S = \sum_l e_l^2 ,$

reaches the minimum value if the relation

(4) $\frac{\partial S}{\partial a_i} = 0 \quad \text{for} \quad i = 1, \ldots, K$

holds.
Since $\partial S/\partial a_i = \sum_l \partial e_l^2/\partial a_i = -\sum_l 2e_l \frac{\partial f_l}{\partial a_i}$ relation (4) can be represented as

(5) $\sum_l a_{il} e_l = 0 \quad (i = 1, \ldots, K)$

where

(6) $a_{il} = \frac{\partial f_l}{\partial a_i} \quad (l = 1, \ldots, L).$

Let us now assume that the vector $a = a^u$ and the matrix X are given. We are going to generate such a vector $y^g = (y_l^g)$ $l = 1, \ldots, L$ that if the ordinary l.s.m. was used to the data one would obtain the vector $a^g = a^u$.

Since (5) holds for the vector a derived by means of the ordinary l.s.m., so for fixed a^u the vector y^g should be generated in such a way as to

(7) $\sum_l a_{il}^u (y_l^g - f_l^u) = 0 \quad (i = 1, \ldots, K)$

where

(8) $f_l^u = f(a^u, x_l)$

$$(9) \qquad \partial^u_{il} = \frac{\partial f^u_l}{\partial a^u_i} .$$

In these formulae f^u_l, ∂^u_{il} $(l = 1, \ldots, L;\ i = 1, \ldots, K)$ are given, because the vector a^u and the matrix X are given. Let us also denote

$$(10) \qquad e^g_l = y^g_l - f^u_l .$$

One can conclude from (7) that a generation of y^g is actually reduced to solving a system of homogenous linear equations

$$(11) \qquad \sum_l \partial_{il} e^g_l = 0 \qquad (i = 1, \ldots, K)$$

with regard to e^g_l $(l = 1, \ldots, L)$.
Then

$$(12) \qquad y^g_l = f^u_l + e^g_l \qquad (l = 1, \ldots, L).$$

If (e^j_l) $l = 1, \ldots, L$ stands for the j-th $(j = 1, \ldots, L-K)$ basic solution of system (11) with the fixed basic system, then the solutions of system (11) can be represented as

$$(13) \qquad e^g_l = \sum_{j=1}^{L-K} p_j e^j_l \qquad (l = 1, \ldots, L)$$

where p_j are arbitrary coefficients[3].

2

In the above procedure the coefficients p_j are arbitrary. However, some additional postulates imposed on the vector y^* usually calls for an appropriate definition of these coefficients. The following postulates can be mentioned as examples:

1) the vector $y^* > 0$;

---

[3] The coefficient p_j can be determined e.g. by means of the random numbers generator.

2) the coefficient of undetermination

$$(14) \qquad w^* = \frac{\sum_l (y_l^* - f_l^u)^2}{\sum_l (y_l^* - \bar{y}^*)^2} \quad \text{where } \bar{y}^* = \frac{1}{L}\sum_l y_l^*$$

is not greater than 1;

3) the coefficient of undetermination is equal to a given value.

Some examples will illustrate the allowance for these postulates. Let us assume that certain vectors y^g, e^g obtained by means of the procedure do not satisfy the respective postulate. We will provide the satisfaction of this postulate by means of an appropriate modification of the vectors y^g, e^g. The vectors a^u, $f^u = (f_l^u)$, $l = 1, \ldots, L$ as well as the matrix X will remain unchanged.

The modified (new) quantities will be marked with an asterisk. We shall assume that

$$(15) \qquad f_l^u > 0 \qquad (l = 1, \ldots, L)$$

and that the function $f(x, a)$ is such that with the vector a satisfying (5), the relation

$$(16) \qquad \sum_l f_l e_l = 0 \qquad \text{where} \quad e_l = y_l - f_l$$

holds. It should be noticed that non-linear functions $f(x, a)$ which are the most popular ones in econometric analysis satisfy this postulate if - as it has been assumed here - the inequality (5) holds. For instance for the power function $f_l = a_1 \prod_{i=2}^{I} x_{li}^{a_i}$ and for the exponential function $f_l = a_1 \exp(\sum_{i=2}^{I} a_i x_{li})$ the following equality holds $\partial_{1l} = f_l/a_1$. So, if $\sum_l \partial_{1l} e_l = 0$, as it is in (5), then $\sum_l f_l e_l = \sum_l a_1 \partial_{1l} e_l = a_1 \sum_l \partial_{1l} e_l = 0$. In turn for the Törnqvist II function $f_l = \frac{a_1 x_l - a_2}{a_3 + x_l}$, we have $\partial_{1l} = x_l/(a_3 + x_l)$, $\partial_{2l} = -1/(a_3 + x_l)$. Taking $\sum_l f_l e_l$ we get

$$\sum_l f_l e_l = \sum_l \frac{a_1 x_l}{a_3 + x_l} e_l - \sum_l \frac{a_2}{a_3 + x_l} e_l =$$

$$= a_1 \sum_l \partial_{1l} e_l - a_2 \sum_l \partial_{2l} e_l = 0$$

because postulate (5) says that $\sum_l \partial_{1l} e_l, \sum_l \partial_{2} e_l = 0$.

Postulate A ($y^* > 0$). Let us assume that some elements of the vector y^g are negative. The residuals e_l^g should be modified in such a way as to get the new residuals e^* being solution of system of homogenous linear equations (11) and providing the new vector $y^* = (y_l^*)$ $l = 1, \dots, L$ to be positive.

Notice that by assumption (15) ($f_l^u > 0$) there is $y_l^* > 0$ if all negative modified residuals satisfy $|e_l^*| < f_l^u$. It can achieved if

(17) $e_l^* = e_l^g m \qquad (l = 1, \dots, L)$

where

(18) $m = n \min_{l \in L^-} \left| \frac{f_l^u}{e_l^g} \right|$

(19) $L^- = \{l : e_l^g < 0\}$.

Note that by definition of the vector e^g (see (12)) the modification of residuals can be represented as

$$(20)\quad e^*_l = \sum_{j=1}^{L-K} p^*_j e^j_l \qquad (l = 1, \ldots, L)$$

where

$$(21)\quad p^*_j = p_j m \qquad (j = 1, \ldots, L-K).$$

Because (e^*_l), $l = 1, \ldots, L$ is a linear combination of basic solutions of system (11), the "new" residulas are solutions of system (11), as well.

Postulate B ($w^* \leqslant 1$). First of all let us find out when the coefficient of undetermination

$$(22)\quad w = \frac{\sum_l e^2_l}{\sum_l (y_l - \bar{y})^2}, \quad \bar{y} = \frac{1}{L}\sum_l y_l, \; e_l = y_l - f_l$$

for the function f(y, a) whose parameter vector satisfies system (5), is not greater than 1, i.e. let us find out when

$$(23)^4\quad \sum_l e^2_l \leqslant \sum_l (y_l - \bar{y})^2.$$

Let us transform the expression $\sum_l (y_l - \bar{y})^2$. Since $y_l = f_l - e_l$, therefore

$$(24)\quad \sum_l (y_l - \bar{y})^2 = \sum_l \{(f_l - \bar{f}) + (e_l - \bar{e})\}^2 =$$

$$= \sum_l (f_l - \bar{f})^2 + \sum_l (e_l - \bar{e})^2 +$$

[4] Moreover, the vector a derived by means of the ordinary l.s.m. and a^u for $y_l = y^g_l$ ($l = 1, \ldots, L$) satisfy the postulate (5).

$$+ 2 \sum_{l} (f_l - \bar{f})(e_l - \bar{e})$$

where $\bar{e} = \frac{1}{L} \sum_{l} e_l$, $\bar{f} = \frac{1}{L} \sum_{l} f_l$.

Moreover,

$$\sum_{l} (e_l - \bar{e})^2 = \sum_{l} e_l^2 - L\,\bar{e}^2$$

$$(25) \quad \sum_{l} (f_l - \bar{f})(e_l - \bar{e}) = \sum_{l} f_l e_l - L\bar{e}\bar{f} = - L\bar{e}\bar{f}.$$

The latter equality is derived under the assumption $\sum_{l} f_l e_l = 0$.

After appropriate substitutions to (23) and reduction of the component $\sum_{l} e_l^2$, one can see that inequality (23) is equivalent to the inequality

$$(26) \quad - L\bar{e}^2 - 2\,L\bar{f}\bar{e} + \sum_{l} (f_l - \bar{f})^2 > 0.$$

The inequality holds if $\bar{e} = 0$[5]. Further, we shall ignore this case and focus on $\bar{e} \neq 0$.

The left-hand side if inequality (26) is a quadratic trinomial with regard to $\bar{e}$. As the coefficient at $\bar{e}^2$ is negative, and the intercept term of this trinomial is non-negative, so the discriminant of this trinomial is non-negative and there are two real roots $\bar{e}_I$, $\bar{e}_{II}$, ($\bar{e}_I \leq \bar{e}_{II}$) of the trinomial. Since $-L < 0$ the inequality $w \leq 1$ holds[6] for $\bar{e} \in [\bar{e}_I, \bar{e}_{II}]$.

---

[5] Notice, that if $f(a, x) = g(a, x) + a_0$, then $\bar{e} = 0$. Thus $\partial f_l / \partial a_0 = 1$ and the equation of system (5) related to the parameter a_0 says that $\sum_{l} 1\, e_l = 0$.

[6] It follows from the presented conditions for $w \leq 1$ that:

Let certain vectors e^g, y^g be already derived and the coefficient of undetermination

(27) $$w^g = \frac{\sum_l (e_l^g)^2}{\sum_l (y_l^g - \bar{y}^g)^2}, \quad \bar{y}^g = \frac{1}{L}\sum_l y_l^g$$

be greater than 1. On the basis of e_l^g we are going to derive such "new" residuals e_l^* (as well as "new" values of y_l^*), that the coefficient of undetermination

(28) $$w^* = \frac{\sum_l (y_l^* - f_l^u)^2}{\sum_l (y_l^* - \bar{y}^*)^2} \quad \text{where} \quad \bar{y}^* = \frac{1}{L}\sum_l y_l^*$$

is not greater than 1 under the vector $e^* = (e_l^*)$ $l = 1, \ldots, L$ satisfying system (11).

If "new" residuals are proportional to e_l^g:

(29) $$e_l^* = k\, e_l^g$$

they really are solution of system (11).

Let $\bar{e}^*$ stand for the mean of the residuals e_l^* ($l = 1, \ldots, L$). From the above condition, $w \leqslant 1$, w^* will be not greater than 1 if

(30) $$\bar{e}^* \in [\bar{e}_1, \bar{e}_2]$$

where $\bar{e}_1$, $\bar{e}_2$ are the roots of the trinomial

(31) $$b_2 z^2 + b_1 z + b_0$$

1) if $\bar{e} = 0$, then $w = 1$ when $\sum_l (f_l - \bar{f})^2 = 0$, i.e. $f_l = \text{const}$ ($l = 1, \ldots, l$); 2) if $\bar{e} \neq 0$, then $w = 1$ when $\bar{e} = \bar{e}_I$ or when $\bar{e} = \bar{e}_{II}$.

with coefficients

$$(32)\quad b_2 = -L, \quad b_1 = -2L\bar{f}^u, \quad b_o = \sum_l (f_l^u - \bar{f}^u)^2$$

where $\bar{f}^u = \frac{1}{L}\sum_l f_l^u$.

On the basis of e^g vector e^* can be generated as follows.

Let $\bar{e}^g = \frac{1}{L}\sum_l e_l^g$, $\bar{e}^g \neq 0$, and

$$(33)\quad k_1 = \bar{e}_1/\bar{e}_g, \qquad k_2 = \bar{e}_2/\bar{e}^g.$$

The "new" residuals can be determined from

$$(34)\quad e_l^* = k\, e_l^g \qquad \text{where} \quad k \in [k_1, k_2].$$

Notice that in this case

$$\bar{e}^* = \frac{1}{L}\sum_l k\, e_l^g = k\, \bar{e}^g \in [\bar{e}_1, \bar{e}_2],$$

what means that the inequality $w^* \leqslant 1$ holds.

Postulate C ($w^* = w_o$, $0 < w_o \leqslant 1$). Let the coefficient of undetermination w^g for the vectors e^g, y^g be different from w_o, and let $0 < w^g \leqslant 1$. On the basis of e^g we are going to derive such a vector e^* being a solution of (11) and providing the equality $w^* = w_o$[7].

If the "new" residuals are proportional to e_l^g:

$$(35)\quad e_l^* = r\, e_l^g \quad (l = 1, \ldots, L), \qquad r \neq 0$$

the vector e^* is a solution of (11). The problem is how to determine the coefficient r. The coefficient w^* is defined by (28). Let us transform this formula into

[7] The postulate $w^* = w^o$ may appear e.g. when we are going to check up the quality of the method M with respect to the function adjusted by means of the ordinary l.s.m.

$$(36)\qquad w^* = \frac{\sum_l (e_l^*)^2}{\sum_l (y_l^* - \bar{y}^*)^2}, \quad e_l^* = y_l^* - f_l^u.$$

If the residuals are defined by (35), then the postulate (5) also holds. Therefore, by virtue of (24) and (25)

$$(37)\qquad \sum_l (y_l^* - \bar{y}^*)^2 = \sum_l (f_l^u - \bar{f}^u)^2 + \sum_l (e_l^* - \bar{e}^*)^2 - - 2L\,\bar{e}^*\,\bar{f}^u.$$

Taking $e_l^* = re_l^g$, it results from (36) and (37) that

$$(38)\qquad w^* = \frac{r^2 \sum_l (e_l^g)^2}{\sum_l (f_l^u - \bar{f}^u)^2 + r^2 \sum_l (e_l^g - \bar{e}^g)^2 - 2L\,r\,\bar{e}^g\,\bar{f}^u}.$$

Let us consider the quotient w^*/w^g. Postulating that $w^* = w_o$ we also postulate that $\frac{w^*}{w^g} = q$, where

$$(39)\qquad q = \frac{w_o}{w^g}$$

The equality $\frac{w^*}{w^g} = q$ holds if

$$(40)\qquad r^2 \left[q \sum_l (e_l^g - \bar{e}^g)^2 - \sum_l (y_l^g - \bar{y}^g)^2 \right] - r\left[2q\,L\bar{e}^g\bar{f}^u\right] + \sum_l (f_l^u - \bar{f}^u)^2 = 0.$$

The left-hand side of this equality is a quadratic trinomial with respect to r.

It follows that the value of r determining the new residuals e_l^* by means of the residuals e_l^g (see (35)) can be taken as a real root (if any) of the quadratic trinomial

$$(41) \qquad c_2 z^2 + c_1 z + c_o$$

where

$$(42) \qquad \begin{cases} c_2 = q \sum_l (e_l^g - \bar{e}^g)^2 - \sum_l (y_l^g - \bar{y}^g)^2 \\ c_1 = -2q\, L\, \bar{e}^g\, \bar{f}^u, \qquad c_o = \sum_l (f_l^u - \bar{f}^u)^2 . \end{cases}$$

Let us now consider whether the real roots of trinomial (41) exist.

1. The case $\bar{e}^g = 0$. Transforming the expression $\sum_l (f_l - f)^2$ and letting $f_l^u = y_l^g - e_l^g$ by assumption, $\sum_l f_l^u e_l^g = 0$, we get

$$\sum_l (f_l^u - \bar{f}^u)^2 = \sum_l \left\{(y_l^g - \bar{y}^g) - (e_l^g - \bar{e}^g)\right\}^2 =$$

$$= \sum_l (y_l^g \bar{y}^g)^2 + \sum_l (e_l^g - \bar{e}^g)^2 + 2\, L \bar{y}^g \bar{e}^g - 2 \sum_l (e_l^g)^2 .$$

Since $\bar{e}^g = 0$, the postulate (40) will now take form

$$r^2 \left[q \sum_l (e_l^g)^2 - \sum_l (y_l^g - \bar{y}^g)^2 + q \sum_l (y_l^g - \bar{y}^g)^2 - \right.$$

$$\left. - q \sum_l (e_l^g)^2 = 0. \right.$$

Consequently,

$$(43) \qquad r^2 = \frac{q\left[\sum_l (y_l^g - \bar{y}^g)^2 - \sum_l (e_l^g)^2\right]}{\sum_l (y_l^g - \bar{y}^g)^2 - q\sum_l (e_l^g)^2}.$$

The numerator of this formula is non-negative, as for $\bar{e}^g = 0$ the coefficient of undetermination $w^g \leqslant 1$ (cf. (26)), and the expression in the square brackets is the difference of the numerator and the denominator of the formula defining w^g.

The expression $q\sum_l (e_l^g)^2$ in the denominator of (43) can be transformed into the form

$$q\sum_l (e_l^g)^2 = \frac{w^*}{w^g}\sum_l (e_l^g)^2 = w^*\frac{\sum_l (y_l^g - \bar{y}^g)^2}{\sum_l (e_l^g)^2}\sum_l (e_l^g)^2 =$$

$$= w^*\sum_l (y_l^g - \bar{y}^g)^2$$

so

$$\sum_l (y_l^g - \bar{y}^g)^2 - q\sum_l (e_l^g)^2 = (1 - w^*)\sum_l (y_l^g - \bar{y}^g)^2 \geqslant 0$$

because $w^* \leqslant 1$.

This means that for $\bar{e}^g = 0$ the value of the right-hand side of formula (43) is non-negative and consequently the trinomial (41) has a real root.

2. The case $\bar{e}^g \neq 0$. In light of (42), the discriminant of trinomial (41) is certainly non-negative if

$$(44) \qquad c_2 = q\sum_l (e_l^g - \bar{e}^g)^2 - \sum_l (y_l^g - \bar{y}^g)^2 \leqslant 0.$$

This inequality is equivalent to the inequality

$$\frac{\sum_{l} (e_l^g - \bar{e}^g)^2}{\sum_{l} (y_l^g - \bar{y}^g)^2} \leq \frac{1}{q}$$

which means that it is equivalent to

$$(45) \quad w^g - \frac{L(\bar{e}^g)^2}{\sum_{l} (y_l^g - \bar{y}^g)^2} \leq \frac{1}{q}$$

because

$$\sum_{l} (e_l^g - \bar{e}^g)^2 = \sum_{l} (e_l^g)^2 - L(\bar{e}^g)^2, \quad w^g = \frac{\sum_{l} (e_l^g)^2}{\sum_{l} (y_l^g - \bar{y}^g)^2}.$$

Inequality (45) holds if $\frac{1}{q} \geq 1$, because

$$w^g - \frac{L(\bar{e}^g)^2}{\sum_{l} (y_l^g - \bar{y}^g)^2} \leq w^g$$

and because, as it was assumed, $w^g \leq 1$.
If we postulate the coefficient w^* not to be greater than w^g the coefficient q is not greater than 1 (i.e. $\frac{1}{q} \geq 1$).

Notice that we can always fix such a vector e^g (satisfying postulate (11)) that $w^g \geq w^*$. In the extreme case it can be such a vector e for which the coefficient of undetermination $w = 1$. The conditions for $w = 1$ are given in footnote 6.

Thus the problem of determination such a coefficient r that $w^* = w_o$ for the residuals e_l^* ($l = 1, \ldots, L$), is solvable.

3. Final remark is only loosely connected with the basic problem of the paper, but it illustrates an application of the presented approach to sample generation.

In our didactic practice we are often trying to solve approximation or estimation problems in such a way so as solutions were expressed by "simple" numbers. Let us assume as an example that we are going to work out a problem dealing with estimation of the parameters of a model $f = \sum_{i=1}^{K} a_i x_i$ by means of the ordinary l.s.m.[8] We want to fix such values of the explanatory variables so as the vector of parameters is expressed by "simple" numbers. The elaboration of the problem can be as follows:

1) we determine a matrix of the values of the explanatory variables $X = (x_{li})$, $l = 1, \ldots, L$; $i = 1, \ldots, K$, and a vector of parameters a^u which is numerically simple, i.e. we assume that $f_l = f_l^u$, where $f_l^u = \sum_{i=1}^{K} a_i^u x_{li}$;

2) such a vector $y^g = (y_l^g)$ that the application of the ordinary l.s.m. to the numerical data $\{y^g \vdots X\}$ results in a vector $a^g = a^u$, is derived as follows:

$$y_l^g = f_l^u + e_l^g \qquad (l = 1, \ldots, L),$$

where e_l^g are solutions to the system of homogenous linear equations

$$(*) \sum_l x_{li}\, e_l^g = 0 \qquad (i = 1, \ldots, K).$$

System (*) is a particular case of the system $\sum_l \partial_{il}^u e_l^g = 0$ (see (11)). Now the derivatives $\partial_{il}^u = \partial f_l^u / \partial a_i$ are equal to x_{li} ($l = 1, \ldots, L$; $i = 1, \ldots, K$). Any linear combination of the vectors $e^g = (e_l^g)$, $l = 1, \ldots, L$ which are solutions of (*), is also a solution of (*)[9].

---

8 This model can be a linearized version of a non-linear model.

9 Postulating the simplicity of calculations we should obviously choose only vectors of residuals e^g with elements expressed by "simple" numbers.

Example. Let $f = a_1x_1 + x_2a_2$, where $x_2 = 1$. Let us assume that

$$X' = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad a^u = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Hence $(f_1^u) = [5 \quad 8 \quad 11 \quad 14 \quad 17 \quad 20 \quad 23 \quad 26]$. The procedure of generation of the vector $y^g = (y_1^g)$ can be represented as checking whether the equations

$$(**) \sum_l x_{l1}e_l^p = 0, \quad \sum_l e_l^p = 0$$

hold for different vectors of residuals e^p.

This procedure is illustrated in Tab. 1. $\sum_i$ denotes the value $\sum_l x_{li}e_l^p$. Let us assume that $e^g = e_l^p$. A text of the problem

Table 1

Calculations concerning the procedure of generation of the vector y^g

x_{l1}	x_{21}	e_l^p I	$\sum_1$	$\sum_2$	e_l^p II	$\sum_1$	$\sum_2$	e_l^p III	$\sum_1$	$\sum_2$
1	1	1	1	1	0	0	0	-1	-1	-1
2	1	1	2	1	0	0	0	-1	-2	-1
3	1	0	0	0	2	6	2	2	6	2
4	1	-4	-16	-4	-1	-4	-1	0	0	0
5	1	1	5	1	-1	-5	-1	2	10	2
6	1	0	0	0	-1	-6	-1	-1	-6	-1
7	1	0	0	0	-1	-7	-1	-1	-7	-1
8	1	1	8	1	2	16	2	0	0	0
Sum	total		0	0		0	0		0	0

could be as follows: "using the ordinary l.s.m. adjust linear function $f = a_1x_1 + a_2$ to the observation results

x_{1l}	1	2	3	4	5	6	7	8
y_l	6	9	11	10	18	20	23	27

We shall get $f = 3x_1 + 2$.

Some other postulates can be imposed on the vector e^g. For instance, we could demand that the estimate s of the standard deviation of the random component ($s^2 = \frac{S}{L - K}$, where S is the residual sum of squares), should be expressed by a "simple" number. In this case we have to select the residuals in such a way that they satisfy the system of equations (*) and that S has a fixed value.

In our example if we postulated s = 1, we would have to find such residuals that equalities (**) hold and that S = 6. It is so e.g. when $e^g = [-1\ 0\ 0\ 2\ 0\ 0\ -1\ 0]$, or when $e^g = [-1\ 1\ 1\ 0\ -1\ -1\ 1\ 0]$.

Bogusław Guzik

GENEROWANIE PRÓBEK DLA BADAŃ SYMULACYJNYCH

W artykule jest rozważany następujący rodzaj badań symulacyjnych. Za pomocą eksperymentu numerycznego, wyniki oszacowania parametrów strukturalnych nieliniowego modelu ekonometrycznego f(a,X) dla zmiennej y, uzyskane za pomocą danej metody estymacji M, są porównywane z wynikami uzyskanymi za pomocą zwykłej metody najmniejszych kwadratów (l.s.m.).

W tym przypadku jest proponowana następująca standardowa procedura:

1° ustal wektor y i macierz X wyników obserwacji na zmiennych,

2° wyprowadź wektor a^o za pomocą l.s.m. i wektor a^M za pomocą metody M.

Pojawiają się różne trudności. Najważniejszą jest aproksymacja wektora a^o. Z tego względu jest proponowane inne podejście do badania:

1° ustal wektor a^u parametrów modelu

2° na bazie a^u i X wyprowadź taki wektor y^G wartości y, że $a^o = a^u$ przy $\{y^G; X\}$.

Problem sformułowany w ten sposób można sprowadzić do rozwiązania układu jednorodnych równań liniowych. W artykule jest opisany sposób określania wektorów y^G i wektorów spełniających pewne dodatkowe warunki.