6th St.Petersburg Workshop on Simulation (2009) 1-6

About the matter of forecasting by using simultaneous equations models[1]

Boris Yu. Lemeshko[2], Alexey E. Tsheglov[3]

Abstract

The procedure of building the interval predictions of endogenous variables was analyzed in case of violation of supposition about noise normality and in case of samples with finite sizes. The comparison of accuracy of predictions with using different methods of estimating reduced form of SEM.

1 The problem definition

Simultaneous equations models (SEM) are widely used for solving applied problems, especially when economical processes are defined.

The structural form of SEM looks like [1]:

$$GY_t = HX_t + \Delta_t, t = 1, 2, ..., n, \tag{1}$$

where $G = (\gamma_{ij})_{m,m}$ - coefficients matrix with m endogenous variables $Y_t = (y_t^{(1)}, y_t^{(2)}, ..., y_t^{(m)})^T$, $H = (h_{ij})_{m,p}$ - coefficients matrix with p exogenous variables $X_t = (x_t^{(1)}, x_t^{(2)}, ..., x_t^{(p)})^T$, and $\Delta_t = (\delta_t^{(1)}, \delta_t^{(2)}, ..., \delta_t^{(m)})^T$ - vector of structural disturbances. It is supposed that coefficients γ_{ij} are normalized by condition of $\gamma_{ii} = 1$.

The reduced form of SEM looks like [1]:

$$Y_t = \Pi X_t + E_t, t = 1, 2, ..., n, \tag{2}$$

where $\Pi = G^{-1}H$, $E_t = G^{-1}\Delta_t = (\varepsilon_t^{(1)}, \varepsilon_t^{(2)}, ..., \varepsilon_t^{(m)})^T$.

The main aim of constructing simultaneous equations models consists of forecasting values of variables interesting for researcher.

Point prediction of endogenous variables $\tilde{Y}_{n+\tau}$ at τ temporal times forward is constructed by the formula [1]:

$$\tilde{Y}_{n+\tau} = \tilde{\Pi} X_{n+\tau}, \tag{3}$$

[1]This work was partly supported by the Federal Agency for Education Ministry of Russian Federation in the Analysis of departmental target program "Development of the scientific potential of high school" (project 2.1.2/3970)

[2]Novosibirsk state technical university, E-mail: `Lemeshko@fpm.ami.nstu.ru`

[3]Novosibirsk state technical university, E-mail: `tleha@ngs.ru`

where $\tilde{\Pi}$ - estimation of Π matrix of reduced form of SEM (2), $X_{n+\tau}$ - values of exogenous variables in time $n+\tau$.

True values of endogenous variables $Y_{n+\tau}$ with adjusted confidence probability $1-\alpha$ are concentrated inside ellipsoid of dispersion with center in point $\tilde{Y}_{n+\tau}$, which derivation is based on the fact, that Hotelling's statistics has F-distribution of Fisher with degrees of freedom m and $n-p-m+1$ in case of normal distribution of noise.

Hotelling's statistics looks like following [1, 2]:

$$H = \frac{n-p-m+1}{(n-p)m} T, \tag{4}$$

where $T = (\tilde{Y}_{n+\tau} - Y_{n+\tau})^T \tilde{\Sigma}^{-1}_{\tilde{\varepsilon}(\tau)} (\tilde{Y}_{n+\tau} - Y_{n+\tau})$, and $\tilde{\Sigma}^{-1}_{\tilde{\varepsilon}(\tau)}$ - estimation of covariance matrix of forecast errors:

$$\tilde{\varepsilon}(\tau) = \tilde{Y}_{n+\tau} - Y_{n+\tau}, \tag{5}$$

where $Y_{n+\tau}$ - true values of endogenous variables at forecasting period.

Estimation $\tilde{\Pi}$ of matrix Π for reduced form of SEM (2) can be found by different ways.

The first way consists in applying ordinary least-squares method for reduced form of SEM. Further, this estimation will be named as Reduce OLS.

The second way consists in finding firstly estimations of matrixes G and H in structural form (1) by using something special methods of unknown parameters estimating (for example, two-stage least squares (TSLS), LIML or SDUNB [3]), and in finding secondly estimation of matrix in reduced form (2) by using formula $\Pi = G^{-1}H$.

When the system is just-identified (all unknown coefficients of structural form are unambiguously restored by coefficients of reduced form), elements of matrix Π of reduced form do not have any limitations. Therefore estimations found by the first and the second ways will be coincident.

If the system is overidentified (unknown coefficients of any equation in structural form are ambiguously restored by coefficients of reduced form), elements of matrix Π of reduced form have limitations, which are not considered by Reduce OLS estimations [4].

Methods of estimating separate equation, such as TSLS, LIML, SDUNB, partly consider these limitations. Therefore, it is naturally supposed, that estimations of matrix Π of reduced form, which are found against estimations of structural form matrixes by these methods, have better properties than by using Reduce OLS. And it must provide more accurate forecast (3).

Thus, the verification of this supposition is the first problem solving in this article.

The second problem consists in researching distribution of Hotelling's statistics (4) in case of violation of noise normality supposition under small sample size.

Efforts were conducted by using the computer modeling method, which provides to research statistical regularities simpler and faster than analytical procedures [5, 6]. Size N of samples of analyzed objects values (statistics, forecast errors, estimations) was equal to 10000.

During researching structural disturbances were modeled with having two-sided exponential distribution which density look like:

$$f_{Dexp}(x;\theta_0,\theta_1,\lambda) = \frac{\lambda}{2\theta_1\Gamma\left(\frac{1}{\lambda}\right)} exp\left(-\left(\frac{|x-\theta_0|}{\theta_1}\right)^{\lambda}\right), \tag{6}$$

where λ - form parameter. Particulars of two-sided exponential distribution are normal distribution ($\lambda = 2$) and Laplace distribution ($\lambda = 1$).

2 Properties of estimations of reduced form and point prediction

Efforts were conducted under following overidentified system:

$$\begin{cases} y_t^{(1)} = -\gamma_{12}y_t^{(2)} + h_{11}x_t^{(1)} + h_{12}x_t^{(2)} + \delta_t^{(1)} \\ y_t^{(2)} = -\gamma_{21}y_t^{(1)} + h_{23}x_t^{(3)} + h_{24}x_t^{(4)} + \delta_t^{(2)} \end{cases} , t = 1, ..., n. \tag{7}$$

At that, it was supposed that true values of parameters were equal to $\gamma_{12} = 1, h_{11} = 3, h_{12} = 2, \gamma_{21} = -2, h_{23} = 4, h_{24} = 1$, and exogenous variables values were within the range of $[-5, 5]$.

The matrix of reduced form in this case looks like:

$$\Pi = \begin{bmatrix} 1 & 0.66666 & -1.33333 & -0.33333 \\ 2 & 1.33333 & 1.33333 & 0.33333 \end{bmatrix}. \tag{8}$$

As results for estimations by TSLS, LIML, SDUNB methods are very similar, as results only for SDUNB estimations are cited (Reduce SDUNB in following text).

On figure 1 empiric distributions of Reduce OLS and Reduce SDUNB estimations of p_{11} parameter are cited in case of sample size $n = 15$ and normal noise distribution with dispersion 1.

As shown on figure 1, Reduce SDUNB estimations have smaller dispersion than Reduce OLS estimations, and therefore they are more exactly [4], that is confirmation of made supposition.

For this case study of point predictions quality led to following results: mean values of forecast errors (5) are near zero value (for Reduce OLS 0.00189, for Reduce SDUNB -0.00256), it means unbiasedness of forecasts, and standard deviation for Reduce SDUNB estimation (0.51165) is smaller than for Reduce OLS estimation (0.59540). Thus, forecast, based on Reduce OLS estimations, for overidentified system with small sample size is less exactly than forecast, based on Reduce SDUNB estimations [4], that is also confirmation of made supposition.

When rising sample size, difference between estimations of reduced form matrix becomes less essential, and therefore, difference between forecast errors distributions also becomes less essential.

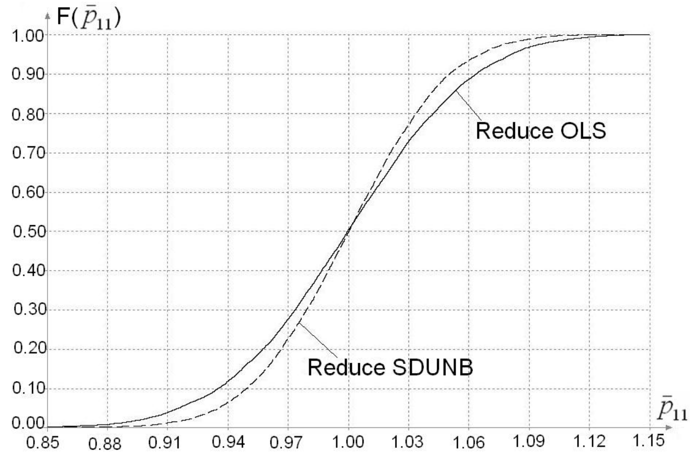

Figure 1: Distributions of Reduce OLS and Reduce SDUNB estimations of p_{11} parameter with $n = 15$, and normal noise distribution.

3 Hotelling's statistics

Fisher distribution is particular of beta-distribution II:

$$F_{m,n-p-m+1} = Be_{II}\left(0, \frac{n-p-m+1}{m}, \frac{m}{2}, \frac{n-p-m+1}{2}\right), \tag{9}$$

where density of beta-distribution II looks like:

$$f(x) = Be_{II}(\theta_0, \theta_1, \alpha, \beta) = \frac{1}{\theta_1 B(\alpha, \beta)} \left(\frac{x-\theta_0}{\theta_1}\right)^{\alpha-1} \left(1 + \frac{x-\theta_0}{\theta_1}\right)^{-\alpha-\beta}. \tag{10}$$

Efforts show that in case of normal distributing structural disturbances the distribution of Hotelling's statistics, calculated with using Reduce OLS estimation of reduced form matrix, is in well accord with Fisher distribution with m and $n-p-m+1$ degrees of freedom in case of small sample sizes as well as big sample sizes [4]. For example, when checking goodness of empiric distribution of Hotelling's statistics for system (7) in case of normal noise distribution and sample size 15 with becoming Fisher distribution $F_{2,10}$ $(Be_{II}(0,5,1,5))$ achieved significance levels were obtained as sufficiently high (chi-square criterion: 0.89267, Kolmogorov criterion: 0.48774, ω^2 criterion: 0.5023, Ω^2 criterion: 0.58103). Graphs are not shown because they are visually identical.

If structural disturbances have distribution different from normal, the distribution of Hotelling's statistics with using Reduce OLS estimations is already different from Fisher distribution with m and $n-p-m+1$ degrees of freedom [4]. For example, distribution of statistics (4) for system (7) is shown on figure 2 with sample size 15 and structural disturbances distributed by normal distribution law ($\lambda = 2$), two-sided exponential distribution law with form parameter 4 ($\lambda = 4$) and Laplace distribution law ($\lambda = 1$).

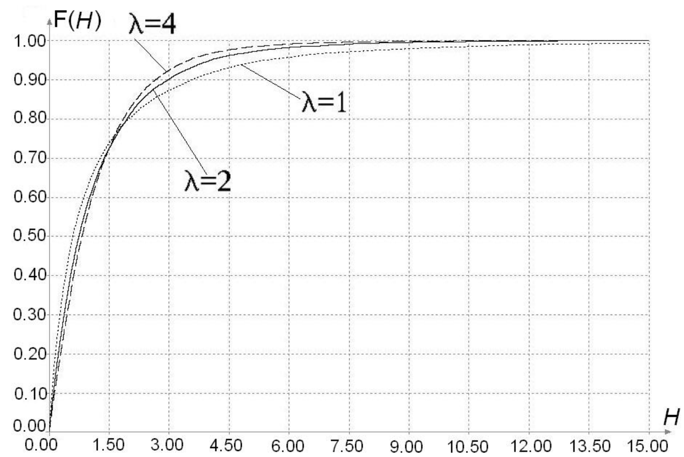

Figure 2: Distribution of Hotelling's statistics for system (7) with $n = 15$ and different noise distribution.

As shown on figure 2, using of upper 100α-percent points of "classic" Fisher distribution, in case of two-sided exponential distribution with form parameter 4, will conduct to intervals covering true values of endogenous variables with probability more than $P = 1-\alpha$, and, in case of Laplace distribution, will conduct to intervals covering true values of endogenous variables with probability less than $P = 1-\alpha$. As the modeling shows, for case shown on figure 2, with using 10-percent point of $F_{2,10}$-distribution the probability of covering true values of endogenous variables equals to 0.8979 for normal distribution law, equals to 0.9151 for two-sided exponential distribution with form parameter 4, and equals to 0.8811 for Laplace distribution.

In order to build correctly the interval prediction in case of violation of noise normality, for finding upper percentile it is possible to use the empiric distribution function of statistics (4), obtained by modeling process, or rough analytic model well approximating empiric distribution function. Efforts show, that distribution of Hotelling's statistics is described in the best way by beta-distribution II or III and Γ-distribution in case of violation of noise normality.

Density of beta-distribution III looks like:

$$f(x;\theta_0,\theta_1,\alpha,\beta,\delta) = \frac{\delta^\alpha}{\theta_1 B(\alpha,\beta)} \frac{\left(\frac{x-\theta_0}{\theta_1}\right)^{\alpha-1}\left(1-\frac{x-\theta_0}{\theta_1}\right)^{\beta-1}}{\left[1+(\delta-1)\frac{x-\theta_0}{\theta_1}\right]^{\alpha+\beta}}, \tag{11}$$

and density of Γ-distribution is shown as:

$$f(x;\theta_0,\theta_1,\alpha,\beta) = \frac{\alpha\left(\frac{x-\theta_0}{\theta_1}\right)^{\alpha\beta-1}}{\theta_1\Gamma(\beta)} exp\left\{-\left(\frac{x-\theta_0}{\theta_1}\right)^\alpha\right\}. \tag{12}$$

Using upper percentiles of true distribution of Hotelling's statistics, calculated with Reduce SDUNB estimations, for building interval predictions for overidentified systems with small sample sizes will conduct to narrower intervals than Reduce OLS estimations with similar probability $1 - \alpha$.

4 Conclusions

Method of computer modeling is accessible and effective instrument for researching statistical regularities. Naturally, the resume obtained by using computer modeling has less degree of generalization than analytical methods, because the modeling is released for concrete system with concrete values of variables and parameters. However, experiments, conducted many times and over wide range of values of parameters set into computer, are able to give the researcher valuable information [1] about statistical regularities, which is impossible to obtain resting only upon asymptotic deducing based on analytical results.

In this paper the method of computer modeling was successfully applied for researching methods of forecasting with using SEM.

References

[1] Ayvazyan S.A., Mkhitaryan V.S. Applied statistics and the foundation of econometrics. Textbook for college. - Moscow: UNITY, 1998. - 1022 p. (in Russian)

[2] Hooper J.W., Zellner A. The error of forecast for multivariate regression model / Econometrica, Vol. 29, pp. 544-555, 1961.

[3] Oberhelman D., Kadiyala K.R. An empirical investigation of the distribution of modified LIML estimators for equations with more than two endogenous vari-ables. Sankhya: The Indian Journal of Statistics, Vol. 61, Series B, Pt. 2, pp. 518-532, 1999.

[4] Tsheglov A.E., Lemeshko B.Yu. The testing hypotheses and construction forecasts for models in the form of systems of simultaneous equations // 9^{th} international conference on actual problems of electronic instrument engineering proceedings. In 7 Volumes. Novosibirsk: NSTU, 2008. Volume 6, pp. 70-74. (in Russian)

[5] Denisov V.I., Lemeshko B.Yu., Postovalov S.N. Applied statistics. Rules checking experienced distribution agreement with the theoretical. Part 1. The criteria for the type of chi-square. - Novosibirsk: NSTU, 1998. - 126 p. (in Russian)

[6] Lemeshko B.Yu., Postovalov S.N. Applied statistics. Rules checking experienced distribution agreement with the theoretical. Part 2. Nonparametric criteria. - Novosibirsk: NSTU, 1999. - 85 p. (in Russian)