EuroEconomica

Issue 2(32)/2013 *ISSN: 1582-8859*

The LSM adjustment method of time series

Cătălin Angelo IOAN[1]
Gina IOAN[2]

[1]*Danubius University of Galati, Department of Economics,*
catalin_angelo_ioan@univ-danubius.ro
[2]*Danubius University of Galati, Department of Economics,*
gina_ioan@univ-danubius.ro

Abstract. This paper presents a method for adjusting the time series in the direction of the minimum deviation of the transition process from one period to another. The formulas so obtained satisfy the condition of invariance of aggregate data, the stability on successive re-adjustment action and the minimum deviation of the linear forecast compared to the original.

Keywords: adjustment, regression

1. Introduction

The statistical adjustment is of fundamental importance in the analysis of phenomena showing a pronounced seasonal character. Removing occasional disturbances allows to analyze the general tendency regardless of any gaps that may temporarily affect the phenomenon.

On the other hand, from the point of view of the new data, the authors believe that they need to satisfy a set of prerequisites.

On the one hand, the amount of data traveling over a reference period should be zero, because, otherwise, the aggregate of the two sets of data - the original and modified, lead to different overall results, and therefore different conclusions.

On the other hand, in a regression analysis, the difference between the regression functions corresponding to the two data sets could be minimal.

The study of several statistical adjustment methods ([1]-[11]) reveals an aspect which, from the point of view of the authors, leads to a change in the data more or less justified. Thus, for example, quarterly adjustments are made in conditions of relative movement only for the data in each quarter individually. In this case, however, it is neglect the phenomenon of continuous variation from one period to another.

The adjustment method that we propose, in what following, consist in the determining the best shift of data so that the total variance, using the method of least squares, from one period to the next, being minimal.

2. The LSM adjustment method of time series

Let the date set $(i, y_i)_{i=\overline{1,kp}}$ where n=kp is the number of statistical data and, also, the sets $B_j=\{p(j-1)+s \mid s=\overline{1,p}\}$, $j=\overline{1,k}$, card B_j=p.

Considering a partition of the data sets B_j, $j=\overline{1,k}$, that is a group of them in k consecutive groups of length equal to p, we aim to determine the parameters $(\beta_j)_{j=\overline{1,p}}$ such that considering the new data set: $(\tilde{y}_i)_{i=\overline{1,kp}}$ with $\tilde{y}_{i+ps} = y_{i+ps} + \beta_i$, $i=\overline{1,p}$, $s=\overline{0,k-1}$, the sum of squared deviations of the data $\tilde{y}_i$ from $\tilde{y}_{i-1}$ be minimum.

So we have the condition:

$$(1)\quad \min\left(\sum_{i=2}^{n}(\tilde{y}_i - \tilde{y}_{i-1})^2\right)$$

Let note, for the beginning:

$$(2)\quad \begin{cases} S_1 = \sum_{s=1}^{k-1}(y_{1+ps} - y_{ps}) \\ S_j = \sum_{s=0}^{k-1}(y_{j+ps} - y_{j+ps-1}),\ j=\overline{2,p} \end{cases}$$

and also:

$$(3)\quad T_j = \frac{\sum_{s=0}^{k-1} y_{j+ps}}{k},\ j=\overline{1,p}$$

From (2), (3) follows:

$$(4)\quad \begin{cases} S_1 = kT_1 - kT_p - y_1 + y_{kp} \\ S_j = kT_j - kT_{j-1},\ j=\overline{2,p} \end{cases}$$

From the definitions of $\tilde{y}_{i+ps}$ we have:

$$(5)\quad \sum_{i=2}^{n}(\tilde{y}_i - \tilde{y}_{i-1})^2 = \sum_{i=2}^{p}(\tilde{y}_i - \tilde{y}_{i-1})^2 + \sum_{s=1}^{k-1}(\tilde{y}_{1+ps} - \tilde{y}_{ps})^2 + \sum_{s=1}^{k-1}\sum_{i=2}^{p}(\tilde{y}_{i+ps} - \tilde{y}_{i+ps-1})^2 =$$

$$k\sum_{i=2}^{p}(\beta_i - \beta_{i-1})^2 + (k-1)(\beta_1 - \beta_p)^2 + 2\sum_{i=2}^{p}(\beta_i - \beta_{i-1})(y_i - y_{i-1}) + 2(\beta_1 - \beta_p)\sum_{s=1}^{k-1}(y_{1+ps} - y_{ps}) +$$

$$2\sum_{i=2}^{p}(\beta_i - \beta_{i-1})\sum_{s=1}^{k-1}(y_{i+ps} - y_{i+ps-1}) + \sum_{i=2}^{p}(y_i - y_{i-1})^2 + \sum_{s=1}^{k-1}(y_{1+ps} - y_{ps})^2 + \sum_{s=1}^{k-1}\sum_{i=2}^{p}(y_{i+ps} - y_{i+ps-1})^2$$

therefore (1) becomes:

$$(6)\quad F(\beta_1,...,\beta_p) = k\sum_{i=2}^{p}(\beta_i - \beta_{i-1})^2 + (k-1)(\beta_1 - \beta_p)^2 + 2(\beta_1 - \beta_p)S_1 + 2\sum_{i=2}^{p}(\beta_i - \beta_{i-1})S_i = \text{minimum}$$

The necessary condition for minimum of the function F is, after (6):

$$(7)\quad \begin{cases} \frac{\partial F}{\partial \beta_1} = 2(k-1)(\beta_1 - \beta_p) - 2k(\beta_2 - \beta_1) + 2S_1 - 2S_2 = 0 \\ \frac{\partial F}{\partial \beta_j} = 2k(\beta_j - \beta_{j-1}) - 2k(\beta_{j+1} - \beta_j) + 2S_j - 2S_{j+1} = 0,\ j=\overline{2,p-1} \\ \frac{\partial F}{\partial \beta_p} = 2k(\beta_p - \beta_{p-1}) - 2(k-1)(\beta_1 - \beta_p) - 2S_1 + 2S_p = 0 \end{cases}$$

or, after a rearrangement, of the unknowns:

(8) $$\begin{cases}(2k-1)\beta_1 - k\beta_2 - (k-1)\beta_p = -S_1 + S_2 \\ -k\beta_{j-1} + 2k\beta_j - k\beta_{j+1} = -S_j + S_{j+1},\ j=\overline{2,p-1} \\ -(k-1)\beta_1 - k\beta_{p-1} + (2k-1)\beta_p = S_1 - S_p\end{cases}$$

The system solutions are:

(9) $$\begin{cases}\beta_j = \frac{1}{k(p(k-1)+1)}\left(k(j-p)S_1 + (j-p)(k-1)\sum_{i=2}^{p}S_i + (p(k-1)+1)\sum_{i=j+1}^{p}S_i\right) + \lambda,\ j=\overline{1,p-1} \\ \beta_p = \lambda, \lambda \in \mathbf{R}\end{cases}$$

We will proceed to a change of parameter, in the sense of uniformity for the displacements:

(10) $$\frac{1}{k(p(k-1)+1)}\left(-k\frac{p(p-1)}{2}S_1 - \frac{p(p-1)}{2}(k-1)\sum_{i=2}^{p}S_i + (p(k-1)+1)\sum_{i=2}^{p}(i-1)S_i\right) + p\lambda = p\mu,\ \mu \in \mathbf{R}$$

After (9), (10) we can write, finally:

(11) $$\beta_j = \frac{1}{2kp(p(k-1)+1)}\left(p(2j-p-1)\left(kS_1 + \sum_{i=2}^{p}(k-1)S_i\right) + 2(p(k-1)+1)\left(p\sum_{i=j+1}^{p}S_i - \sum_{i=2}^{p}(i-1)S_i\right)\right) + \mu,\ j=\overline{1,p}$$

where for j=p, the penultimate sum is ignored, or, using (3) in terms of average periods:

(12) $$\beta_j = \frac{1}{2p(p(k-1)+1)}\left(p(2j-p-1)(T_1 - T_p - y_1 + y_{kp}) + 2(p(k-1)+1)\sum_{i=1}^{p}T_i\right) - T_j + \mu,\ j=\overline{1,p}$$

Noting now:

(13) $$\alpha = \frac{T_1 - T_p - y_1 + y_{kp}}{p(k-1)+1}$$

(14) $$\gamma = \frac{\sum_{i=1}^{p}T_i}{p} - \frac{(p+1)(T_1 - T_p - y_1 + y_{kp})}{2(p(k-1)+1)}$$

we have:

(15) $$\beta_j = \alpha j + \gamma - T_j + \mu,\ j=\overline{1,p}$$

From the fact that $\sum_{j=1}^{p}(\alpha j + \gamma) = \sum_{j=1}^{p}T_j$ we obtain, immediately:

(16) $$\sum_{j=1}^{p}\beta_j = p\mu$$

(17) $$\sum_{j=1}^{p}j\beta_j = \alpha\frac{p(p+1)(2p+1)}{6} + (\gamma+\mu)\frac{p(p+1)}{2} - \sum_{j=1}^{p}jT_j$$

3. The condition of invariance of aggregated values

We now ask a question concerning the movements of β_j such that $\sum_{s\in B_j}\tilde{y}_s = \sum_{s\in B_j} y_s,\ j=\overline{1,k}$ or, in other words, the aggregate of the adjusted date in any of the periods k equaling that of the original data. The condition is therefore $\sum_{j=1}^{p}\beta_j=0$ that is, after (16) returns to μ=0. Therefore, we obtain:

(18) $$\beta_j = \frac{1}{2p(p(k-1)+1)}\left(p(2j-p-1)(T_1-T_p-y_1+y_{kp})+2(p(k-1)+1)\sum_{i=1}^{p}T_i\right)-T_j,\ j=\overline{1,p}$$

4. The condition of constancy displacements in successive adjustments

Next, we will put the following question: if after a readjustment of data, we will apply again the above algorithm, what happens with the new data?

Let therefore:

(19) $$\begin{cases}\tilde{S}_1=\sum_{s=1}^{k-1}(\tilde{y}_{1+ps}-\tilde{y}_{ps})=\sum_{s=1}^{k-1}(y_{1+ps}-y_{ps})+\sum_{s=1}^{k-1}(\beta_1-\beta_p)=S_1+(k-1)(\beta_1-\beta_p)\\ \tilde{S}_j=\sum_{s=0}^{k-1}(\tilde{y}_{j+ps}-\tilde{y}_{j+ps-1})\sum_{s=0}^{k-1}(y_{j+ps}-y_{j+ps-1})+\sum_{s=0}^{k-1}(\beta_j-\beta_{j-1})=S_j+k(\beta_j-\beta_{j-1}),\ j=\overline{2,p}\end{cases}$$

that is:

(20) $$\begin{cases}\tilde{S}_1=S_1+(k-1)(\beta_1-\beta_p)\\ \tilde{S}_j=S_j+k(\beta_j-\beta_{j-1}),\ j=\overline{2,p}\end{cases}$$

From the formulas (11) we obtain:

(21) $$\tilde{S}_j=\frac{k}{p(k-1)+1}S_1+\frac{k-1}{p(k-1)+1}\sum_{i=2}^{p}S_i,\ j=\overline{1,p}$$

From (21) it follows therefore:

(22) $$\tilde{S}_1=\tilde{S}_2=...=\tilde{S}_p$$

Using again (11), we find that for any j=$\overline{1,p}$: $\tilde{\beta}_j=\mu$.

Therefore, as to an adjustment of the adjusted data to be zero, the displacement will have to be null, that is μ=0, therefore again the relations (18).

5. The influence of the adjusted data in linear regression

We now ask, naturally, the problem of the behavior of linear regression coefficients at these changes.

For the data set $(j,y_j)_{j=\overline{1,kp}}$, the linear regression equation is y=ax+b+ε where:

(23) $$a=6\frac{2\sum_{j=1}^{kp}jy_j-(kp+1)\sum_{j=1}^{kp}y_j}{kp(k^2p^2-1)}$$

(24) $$b=2\frac{(2kp+1)\sum_{j=1}^{kp}y_j-3\sum_{j=1}^{kp}jy_j}{kp(kp-1)}$$

ε being the residual variable.

For the new data set $\left(j,\tilde{y}_j\right)_{j=\overline{1,kp}}$ and the regression equation y= $\tilde{a}$ x+ $\tilde{b}$ + $\tilde{\varepsilon}$, we obtain first:

(25) $$\sum_{j=1}^{kp}\tilde{y}_j=\sum_{i=1}^{p}\sum_{s=0}^{k-1}\left(y_{i+ps}+\beta_i\right)=\sum_{j=1}^{kp}y_j+k\sum_{j=1}^{p}\beta_j$$

(26) $$\sum_{j=1}^{kp}j\tilde{y}_j=\sum_{i=1}^{p}\sum_{s=0}^{k-1}\left((i+ps)y_{i+ps}+(i+ps)\beta_i\right)=\sum_{j=1}^{kp}jy_j+k\sum_{j=1}^{p}j\beta_j+p\frac{(k-1)k}{2}\sum_{j=1}^{p}\beta_j$$

and after (16), (17), (23), (24):

(27) $$\tilde{a}=a+6\frac{2\sum_{j=1}^{p}j\beta_j-(p+1)\sum_{j=1}^{p}\beta_j}{p\left(k^2p^2-1\right)}=a+\frac{2}{kp+1}\frac{\alpha p(p+1)(2p+1)+3\gamma p(p+1)-6\sum_{j=1}^{p}jT_j}{p(kp-1)}$$

(28) $$\tilde{b}=b+\frac{-6\sum_{j=1}^{p}j\beta_j+(pk+3p+2)\sum_{j=1}^{p}\beta_j}{p(kp-1)}=b-\frac{\alpha p(p+1)(2p+1)+3\gamma p(p+1)-6\sum_{j=1}^{p}jT_j}{p(kp-1)}-\mu$$

From this, it follows:

(29) $$\left|\tilde{a}-a\right|=\frac{2}{p\left(k^2p^2-1\right)}\left|\alpha p(p+1)(2p+1)+3\gamma p(p+1)-6\sum_{j=1}^{p}jT_j\right|$$

(30) $$\left|\tilde{b}-b\right|=\frac{2}{p(kp-1)}\left|\alpha p(p+1)(2p+1)+3\gamma p(p+1)-6\sum_{j=1}^{p}jT_j+\mu\right|\le$$

$$\frac{2}{p(kp-1)}\left|\alpha p(p+1)(2p+1)+3\gamma p(p+1)-6\sum_{j=1}^{p}jT_j\right|+\frac{2}{p(kp-1)}|\mu|$$

From (29) and (30) results:

(31) $$\left|\tilde{a}x+\tilde{b}-ax-b\right|\le\left|\tilde{a}-a\right|x+\left|\tilde{b}-b\right|\le$$

$$\frac{2}{p\left(k^2p^2-1\right)}\left|\alpha p(p+1)(2p+1)+3\gamma p(p+1)-6\sum_{j=1}^{p}jT_j\right|\left(x+\frac{1}{kp+1}\right)+\frac{2}{p(kp-1)}|\mu|$$

From (31) we obtain that for an exogenous variable value x, the minimum difference of the forecast will be for μ=0 so again the expression (18).

6. The Quarterly data adjusting

In particular, in the case of quarterly adjustments for p=4, the formula (11) for μ=0 becomes:

(32) $$\beta_1=\frac{1}{4k(4k-3)}\left(-6kS_1+(6k-3)S_2+2kS_3-(2k-3)S_4\right)$$

$$\beta_2=\frac{1}{4k(4k-3)}\left(-2kS_1-(6k-5)S_2+(6k-4)S_3+(2k-1)S_4\right)$$

$$\beta_3=\frac{1}{4k(4k-3)}\left(2kS_1-(2k-1)S_2-(6k-4)S_3+(6k-5)S_4\right)$$

$$\beta_4 = \frac{1}{4k(4k-3)}\left(6kS_1 + (2k-3)S_2 - 2kS_3 - (6k-3)S_4\right)$$

and in terms of averages, form (18):

(33) $$\beta_1 = \frac{1}{4(4k-3)}\left(-3(4k-1)T_1 + (4k-3)T_2 + (4k-3)T_3 + (4k+3)T_4 + 6y_1 - 6y_{4k}\right)$$

$$\beta_2 = \frac{1}{4(4k-3)}\left((4k-5)T_1 - 3(4k-3)T_2 + (4k-3)T_3 + (4k-1)T_4 + 2y_1 - 2y_{4k}\right)$$

$$\beta_3 = \frac{1}{4(4k-3)}\left((4k-1)T_1 + (4k-3)T_2 - 3(4k-3)T_3 + (4k-5)T_4 - 2y_1 + 2y_{4k}\right)$$

$$\beta_4 = \frac{1}{4(4k-3)}\left((4k+3)T_1 + (4k-3)T_2 + (4k-3)T_3 - 3(4k-1)T_4 - 6y_1 + 6y_{4k}\right)$$

7. Conclusions

In view of the above theory, it can be concluded that the method described satisfy a number of requirements. On the one hand, the aggregation was not affected, the overall conclusions drawn on the basis of the adjusted data being identical with respect to the original data. On the other hand, the re-adjustment is stationary in the sense that a new adjustment of the already-adjusted data, does not produce effects. Not in the last, the effect of forecast leads to minimal errors relative to the original data.

8. References

1. Box, G.E.P., Hillmer, S.C., Tiao, G.C. (1978), *Analysis and Modeling of Seasonal Time Series*, Seasonal Analysis of Economic Time Series, Washington, 309-334
2. Brockwell, P., Davis, R. (1987), *Time Series: Theory and Methods*, Springer-Verlag, Berlin
3. Burridge, P., Wallis, K.F. (1988), *Prediction Theory for Autoregressive Moving Average Processes*, Econometric Reviews, 7, 65-95
4. Engle, R.F. (1978), *Estimating Structural Models of Seasonality*, Seasonal Analysis of Economic Time Series, Washington, 281-297
5. Fisher, B. (1995), *Decomposition of Time Series: Comparing Different Methods in Theory and Practice*, Eurostat, April 1995
6. Gersh, W., Kitagawa, G. (1983), *The Prediction of Time Series with Trends and Seasonalities*, Journal of Business and Economic Statistics, 1, 253-264
7. Ghysels, E. (1993), *Seasonality and Econometric models*, Journal of Econometrics, Special issue, 55, No. 1-2
8. Harvey, A.C. (1993), *Time Series Models*, Deddington
9. Nerlove, M., Grether, D.M., Carvalho, J.L. (1979), *Analysis of Economic Time Series: A Synthesis*, Academic Press, New York
10. Stephenson, J.A., Farr, H.T. (1972), *Seasonal Adjustment of Economic Data by Application of the General Linear Statistical Model*, Journal of the American Statistical Association, 67, 37-45
11. Wallis, K.F. (1974), *Seasonal Adjustment and Relations Between Variables*, Journal of the American Statistical Association, 69, 18-31