Lithuanian Journal of Statistics *2014, vol. 53, No 1, pp. 34–39*
Lietuvos statistikos darbai *2014, 53 t., Nr. 1, 34–39 p.*
www.statisticsjournal.lt

MODELING OF DYNAMICS OF THE UNEMPLOYMENT INDEX AT THE REGIONAL LEVEL BY MEANS OF THE SPECTRAL ANALYSIS METHOD

Valentina Pleskach

Scientific & Technical Complex of Statistical Research
Address: Apt. 10, 35, L. Ukrainka, Kovel, Volyn region, UA-45000 Ukraine
E-mail: valya_pleskach@mail.ru

Received: July 2014 Revised: October 2014 Published: November 2014

Abstract. The model of the unemployment index at the regional level in accordance with the spectral analysis method worked out by the team of O. Morgenstein of American Princeton University is got. The unemployment indexes of Kovel town in Kovel district of Volyn region for the period of observations from January 1994 to February 2012 are the information basis of this research.

Keywords: unemployment, dynamical series, spectral analysis.

1. Introduction

The spectral analysis of time series is the most promising directions of usage of mathematical methods in statistical research. The dynamical series of the unemployment index is one of objects of this investigation.

Statisticians have been applying progressive and powerful methods of time series analysis for indicators of different areas for a long time; however, as for the time series of statistical indicators, the general statistical methods are used.

We tried to use the spectral analysis of time series for modelling the unemployment dynamics.

While studying time series fluctuations economists would come to the conclusion that these fluctuations can be superposition of several harmonic components. Unfortunately, at the initial stages of investigation, those components were called cycles, though this term assumes significant regularity in both the period and the amplitude. Many economists took part in discussing a lot of different cycles, so their final conclusions coincide only partially, but any list of the most significant economic cycles includes the following:

1. Long waves with the period of 40–60 years (they are often received during the analysis of gross national product, migration, population).
2. Construction cycles with the period of 15–20 years.
3. Principal and secondary economic cycles: principal ones – 6–11 years, secondary ones– 2–4 years.
4. Economic cycles that are a certain type of fluctuations of the countries' economic activity in the sphere of commercial entrepreneurship. The duration of economic cycles changes from 1 year (a season cycle) to 10–12 years. For example, the average duration of the economic cycles determined by the USA National Agency is about 4 years. In many cases (especially in the economic series connected with new orders, prices, equipment), the average cycle duration is 2 years.

It should be noted that in all above mentioned cases the notion *cycle* does not include full regularity. It is just to the point to cite the ideas given in the paper [1]: "There are no grounds to state that economic cycles have tendency to be repeated. Some cycles are very moderate, the others are clearly expressed, some have duration of 2, 3, 4 years the other last 8–10 years. In some of them, rise is longer than downfall, in others – vice versa. It is just the same mixture of regularity and irregularity which appears to cause the main difficulties while constructing models and statistical description of the economic series."

In our study, the cycles with such periods have been discovered: the most powerful cycle with the period of 14.2 years, then less powerful –7.9 years, and further decreasing in power come the cycle with periods of 5.4, 3.4, 2.3 years, and the season cycle – 1 year. As it can be seen, these cycles properly fit into the given classification of the economic cycles. However, in order to deeper investigate their nature it is necessary to have unemployment index analysis on other regions, but it was not the object of our study.

The main advantage of the spectral analysis is that for the study of the mixture of regularity and irregularity it provides the natural approach from mathematical point of view: fluctuations similar to economic cycles can be described as the sum of a certain number of sinusoids with their amplitudes and phases. Besides, the spectral analysis gives though theoretically approximate but practically very important method of definition of cycles reality by means of making it clear whether this or that spectrum pick gives an essential contribution in the general variance of index dynamics. Our paper is the first try to use spectral analysis for investigating unemployment index dynamics and does not open all the power of its procedures, especially cross-spectral methods.

2. Objectives of the study

Our basic task is to get an accurate model of unemployment process as the composition of the harmonious components with different amplitudes, frequencies and phases. Based on recommendations of the scientists of American Princeton University [1] the spectral analysis method can be used for this task. The spectral analysis allows modelling very different complicated processes. Such well-known methods of mathematical modelling as regression or correlation methods do not always allow modelling such complicated process as dynamics of unemployment in desired accuracy.

The principal task of modelling the unemployment index consists in the necessity of its prognosis which state administrative authorities need at the regional level. At the same time prognosis should be carried out on the base of as long time series as possible as in these cases it is more reliable.

Nowadays the most widespread method of modelling is regression analysis. However, regression analysis can model proceeding the process in short time intervals when the process can really be represented by the linear function or reduced to the linear type. However, it is known that reliability of prognosis is conditioned by the prehistory of the process studying: the longer the prehistory, the more reliable the prognosis. However, the unemployment index dynamics cannot in any way be represented by the linear function.

The regression analysis of the unemployment index at the regional level as that of the time function gave the following estimates of the regression coefficients.

$Y_i = a + bt_i$; $a = 2891.6$; $b = -8.2$; t_i – monthly index time with February 2003 as the initial date.

Thus, the unemployment index y_i can be represented as:

$y_i = Y_i + e_i$; where e_i – random errors;

$$e_i = y_i - Y_i .$$

The usage of the regression analysis gives the following variance of random errors (theory-practice).

$$\delta_{RA}{}^2 = 3702668.29$$

The usage of the spectral analysis gives the variance of random errors as follows.

$$\delta_{SA}{}^2 = 259816.3$$

Degrees of freedom for these variances are 217.

We use Fisher's F-correlation for checking up the hypothesis of essentiality of the variance divergence $\delta_{RA}{}^2$ and $\delta_{SA}{}^2$:

$$F = \frac{\delta_{RA}{}^2}{\delta_{SA}{}^2} = 14.251 \rangle F_{êð.} = 1.000 ,$$

where $F_{кр.}$ – the critical value of $F_.$ – criterion at the confidence level 1 %.

Thus, the variance of random divergences for the spectral analysis is 14.251 times less than the variance of the same divergences for the regression analysis and this makes a very essential difference. That's why I use the spectral analysis for unemployment modelling. It is the best method of modelling at this stage. No better method of unemployment index modelling at the regional level has been found so far.

3. Modelling of Dynamics of the Unemployment Index

The estimator of the spectral function in accordance with recommendations of the work [1] is as follows:

$$\hat{f}(\omega_j) = \frac{1}{2\pi}\left\{\lambda_0 C_0 + 2\sum_{k=1}^{m}\lambda_k C_k \cos\omega_j k\right\}; \tag{1}$$

$$\omega_j = \frac{\pi j}{m}, \quad j = 0,1,...,m,$$

where ω_j is the frequency of the j harmonic component, and C_k is the covariance function which is calculated according to the following formula using the data $(y_t, t = 1,2,...,n)$:

$$C_k = \frac{1}{n-k}\left\{\sum_{t=1}^{n-k} y_t y_{t+k} - \frac{1}{n-k}\sum_{t=1+k}^{n} y_t \sum_{t=1}^{n-k} y_t\right\}, \tag{2}$$

and weights λ_k are determined for the values *k* and *m*.

Many authors proposed different formulas for the calculation weights λ_k (for example, Jenkins G.M. [2], Parzen E. [3], Hennan E. [4], Grenander U. and Rosenblatt M. [5]). Difficulties connected with the usage of λ_k are in that some formulas for its calculation lead sometimes to the negative estimates of the spectral function (but the estimates have to be always really positive). In our calculations we used the Parzen's weight function, which has that useful property, that the estimates of spectral function are nonnegative. The values of weights function λ_k Parzen are calculated by means of the following formula:

$$\lambda_k = \begin{cases} 1 - \dfrac{6k^2}{m^2}\left(1 - \dfrac{k}{m}\right), & 0 \le k \le \dfrac{m}{2}, \\ 2\left(1 - \dfrac{k}{m}\right)^3, & \dfrac{m}{2} \le k \le m, \end{cases} \tag{3}$$

and respectively under usage (3) the estimator of spectral function will be the following:

$$U_j^* = \frac{C_0}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{m/2}\left[1 - \frac{6k^2}{m^2}\left(1 - \frac{k}{m}\right)\right] C_k \cos\frac{\pi k}{m} j + \frac{2}{\pi}\sum_{k=(m/2)+1}^{m}\left(1 - \frac{k}{m}\right)^3 C_k \cos\frac{\pi k}{m} j. \tag{4}$$

The estimator of covariance function is determined in this case by the formula:

$$C_k = \frac{1}{n}\sum_{t=0}^{n-k}\left(y_t - \bar{y}\right)\left(y_{t+k} - \bar{y}\right), \tag{5}$$

where

$$\bar{y} = \frac{1}{n}\sum_{t=0}^{n} y_t. \tag{6}$$

The integer m in formula (1) and in other formulas is named the cut off point [1]. The larger *m* is, the more points for which the spectrum is estimated, and the larger the variance of the estimator is at each point respectively; the smaller *m*, the better the estimate. There is no mutual opinion among statisticians about the desired value *m* at the moment, except only that it is often recommended *m* to be of the order of n/3, if the number of data is *n*, and obviously, it would be better that *m* to be of the order of *n*/5 or *n*/6 for the small number *n* [1].

According to the recommendations, presented in the work [1] it is necessary to have n >100 for estimation of spectral function. It is possible to consider n =200 as the sufficient minimum, however the rough estimates of spectral function were sometimes obtained using n =80. In our research the number values is $n = 218$ and it is sufficient amount for us.

It is necessary to note that we estimate the spectral function at $m+1$ constant distance points $\omega_j = \frac{j\pi}{m}$, $j = 0,1,...,m$. If the monthly data are analyzed, and we are interested in the separate frequency, for example with annual frequency $\frac{2\pi}{12}$, then the point ω_j for the spectral function estimate is the nearest to this number, for example,

$$\omega_j = \frac{j\pi}{m} \simeq \frac{2\pi}{12},$$

$$j \simeq \frac{m}{6}.$$

Thus, if we used m=60, the annual component will be at the 10th frequency point. By analogy, the component of the 40-monthly cycle will be at the third frequency point.

Modelling of any indicator on the basis of the spectral analysis means determination of frequencies, amplitudes and phases of periodical components (sinusoids), which can model the deterministic component of any process. Noises, differences of the model points from the real data are estimated by the mean-square error. Character of noises in the spectral analysis allows receiving the so-called “white noise” process and its spectral density is constant. In other words, the number of the harmonious components at the modelling of the unemployment index one can pick up in such a way that exactly “white noise” would remain in the residuals.

We used the computer program of the spectral analysis worked out in the Poltava Gravimetric Observatory (PGO) of S. I. Subotin Institute of Geophysics of the National Academy of Sciences of Ukraine for determination of the frequency, amplitude and phase of the harmonious components.

The technique of spectral analysis in accordance with this program is presented in the following. First, the parameters (the amplitude, initial phase and period) of the largest wave are selected. Then this wave is excluded from the initial data, and the rest of the data is processed using the spectral analysis and the parameters of the next largest wave which is in this data are selected again. The process is repeated until there is a approach to white noise.

The general theory of receiving spectrum marks of unemployment index dynamics y_i is shown by the formulas (1–2). However, the power spectrum which is a positive value but because of statistical fluctuations of the covariance function C_k, especially at the small numbers n-k in (2), can become negative. To avoid negative estimate of the spectrum the frequency picture of unemployment index dynamics is based on the formula (4) under the usage of covariance function (5) and weights function (3) which excludes appearance of negative numbers of the power spectrum. After the estimates of the spectrum are received, the stage of defining the amplitude, phase and frequency of the most powerful wave out of the estimates starts (4). The most powerful estimate wave (4) is that which has the biggest number U_j in (4), that is the biggest energy, as U_j is exactly the square of the wave amplitude: the wave that gives the biggest contribution into the variance of the index dynamics is considered to have the biggest amplitude. After the amplitude, the period and the phase of the most powerful wave is defined it is excluded.

Thus we got six harmonious components with the following amplitudes, periods and phases (see the table 1) as a result of usage of the higher mentioned software product.

Table 1. Amplitudes, Periods and Phases of Harmonious Components for the Unemployment Index in Kovel Town and Kovel District

№ of the harmonious component of the process, i	Amplitude, A_i Number of people	Period in months, T_i	Period in years, T_i	The phase of the harmonious component in degrees, λ_i
1	1780.16	170.23	14.2	274.53
2	1188.60	94.58	7.9	203.11
3	487.49	64.51	5.4	93.51
4	171.82	40.89	3.4	116.73
5	224.52	27.28	2.3	238.16
6	146.31	11.98	1.0	23.88

One can write down the statistical model of the unemployment index at usage of the spectral analysis in the following form:

$$y_i = Y_i + \Delta Y + e_i; \tag{7}$$

$$Y_i = y_0 + \sum_{i=1}^{k} A_i \sin\left(\frac{360^\circ}{T_i} t_i + \lambda_i\right) \tag{8}$$

where y_i is the actual value of the unemployment index for a *i*-month;

y_0 – is the mean value of the unemployment index for all period of observation (from January 1994 to February 2012);

ΔY – is the correction in case of displacement of the model;

e_i – are the random errors of the model;

A_i, T_i, λ_i – are the amplitude, the period and the phase of the *i* – harmonious component; t_i is the number of the month in the series of observations *i*=1, 2, 3, ….., 218.

The basic purpose of the spectral analysis is to expose the period, the amplitude and the phase of every harmonious component of the modelled process. The correction ΔY in case of displacement of the model would be absent in that case, when all interval of observations in age 18.2 was multiple to each of the six exposed periodic components. For example, the powerful harmonious component with the period in age 14.2 is not multiple to the interval of observation on the period 0.28 ((18.2–14.2):14.2=0.28); the less powerful wave in age 7.9 is not multiple to the interval of observations on the period 0.30 ((18.2–7.9*2):7.9=0.30) and etc. Therefore we have the correction ΔY in case of the vertical displacement of the spectral model for the investigated phenomenon. We can find the values of the correction ΔY with the help of the formula:

$\Delta Y = y_o - \sum Y_i / n$. In our case $y_0 = 2883$; $\sum Y_i / n = 2574$. Thus, $\Delta Y = 2883 - 2574 = 309$.

Taking into account the data of the table 1 the statistical model of the unemployment index of Kovel town and Kovel district for the value t_1 is:

$$\begin{aligned} Y_i = 3192 &+ 1780.16\sin\left(\frac{360}{170.23} t_i + 274.53^\circ\right) + 1188.6^\circ0\sin\left(\frac{360}{94.58} t_i + 203.11^\circ\right) \\ &+ 487.49\sin\left(\frac{360}{64.51} t_i + 93.51^\circ\right) + 171.82\sin\left(\frac{360}{40.89} t_i + 116.73^\circ\right) \\ &+ 224.52\sin\left(\frac{360}{27.28} t_i + 238.16^\circ\right) + 146.31\sin\left(\frac{360}{11.98} t_i + 23.88^\circ\right). \end{aligned} \tag{9}$$

The scheme 1 gives to us an idea about the character of permanent errors e_i which have the character of white noise.

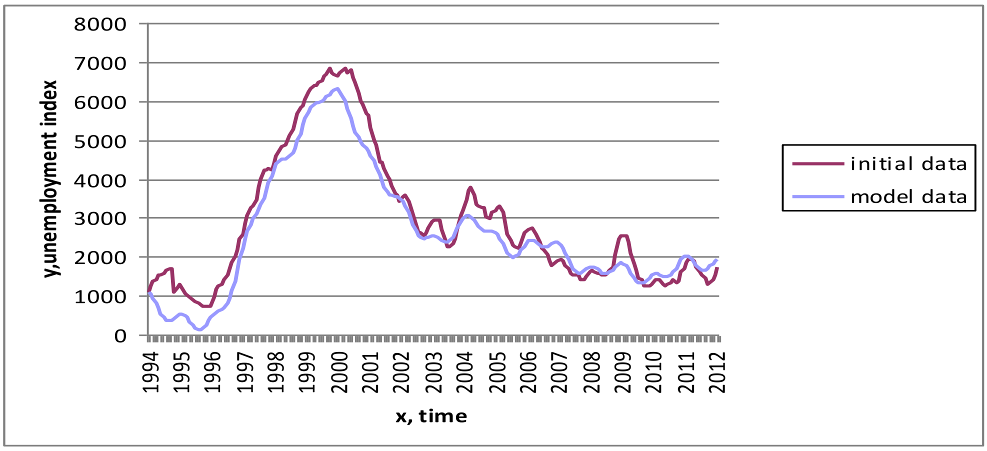

Fig. 1. The model of the unemployment index of Kovel town and Kovel district for the period of observations from January 1994 to February 2012

4. Conclusion

The basic advantage of spectral modelling of dynamic processes is that it allows modelling the process of any complexity without special difficulties. The unemployment index is the process with difficult dynamics.

Firstly, for Ukraine, we got the model of the unemployment index at the regional level by the method of the spectral analysis in cooperation with Poltava Gravimetric Observatory of the Academy of Sciences of Ukraine. This model allows professional forecasting of the unemployment index with continuous accurate definition of prognosis in the process of current activity, which is very important for the administrative authorities and realization of employment social policy.

References

1. Granger C.W. J, and Hatanaka M. 1972: *Spectral analysis of economic time series*, Moscow, Statistics.
2. Jenkins G. M. 1961: General considerations in the analysis of spectra. *Technometrics*, vol. 3, p. 133–166.
3. Parzen E. 1961: Mathematical considerations in the estimation of spectra, *Technometrics*, vol. 3, p. 167–190.
4. Hannen E.J. 1964: *Time Series Analysis*, Moscow, Science.
5. Grenander U. and Rosenblatt M. 1957: *Statistical Analysis of Stationary Time Series*, New York.

REGIONŲ NEDARBO INDEKSO DINAMIKOS MODELIAVIMAS TAIKANT SPEKTRINĖS ANALIZĖS METODĄ

Valentina Pleskach

Santrauka. Sudarytas regionų nedarbo indekso kitimo modelis, taikant O. Morgenstein iš American Princeton universiteto grupės sukurtą spektrinės analizės metodą. Volyn regiono Kovel rajono Kovel miesto 1994 m. sausio mėn. – 2012 m. vasario mėn. stebėti nedarbo rodikliai yra šio tyrimo informacinė bazė.

Reikšminiai žodžiai: nedarbas, dinamikos eilutė, spektrinė analizė.