ISSN: 2581-4230, Website: journalnx.com, June 18th & 19th, 2020

ON FACTOR ANALYSIS AND FACTORIZATION OF DYNAMIC SERIES

Safarov Javokhir Ismoilovich

Accountant in the Central election commission of the Republic of Uzbekistan

Tel.: +998(97)4484800

missing_henry@mail.ru

Abstract

The paper presents the methods of processing time series, which is of particular great interest for economic research. When processing time series, factors are used as generalized characteristics in the analysis and forecasting of complex economic processes, such as, for example, increasing the technical equipment of the industry, increasing labor productivity, demand structure, and market research, inflation in capitalist economics studies.

Keywords: methods of interpretation, factor analysis, economic problems, matrix factor loadings, simple structure.

Introduction

The result of correlation analysis is a matrix of correlation coefficients between the initial parameters. With a small number of parameters, a direct visual analysis of such a matrix is sufficient for a comparative comparison of all elements of the matrix. However, when the number of parameters approaches ten, this method of studying the correlations between the studied parameters is unsatisfactory.

Methods of factor analysis were initially developed precisely with the aim of explaining the variety of correlations arising between the initial parameters [1,2]. When analyzing the objects of observations described by a large number of parameters (several tens or more), the results of factor analysis themselves required the development of special interpretation procedures. The practical application of factor analysis has shown that the required procedures should primarily single out groups of closely related variables [3]. The development of such procedures has undergone two periods. In the first period, when these procedures were given an auxiliary role, they were developed only as heuristic procedures. In the second period, when

the procedures for grouping parameters acquired independent significance, work appeared in which these procedures were built as methods for solving precisely formulated problems.

If in a specific study it turns out that each factor has noticeably large loads on its specific group of initial parameters, then the interpretation of the factors is determined by the groups of parameters selected in this way. However, in most cases, this situation is rare.

In [4], he drew attention to the fact that sometimes in a particular study there is a priori data for an isolated study of individual groups of parameters that are interconnected in meaning. For these cases, he proposed for each such group to separately determine only one most significant factor. With this approach, the interpretation of the resulting factors is also simple: it comes down to an analysis of the load distribution for a single factor within each selected group of parameters separately.

Statement of the problem and methods of solution

The problem of factor analysis and processing of time series, which is of particular great interest for economic studies, which can rightly be considered a "growth point" for applied research with factor analysis.

Until recently, the application of factor analysis to the study of economic problems has been carried out mainly according to the traditional scheme developed in psychology and widely used in regional studies, this is the so-called *R*-analysis, the material for which is cross-sectional survey data. At the same time, some innovations in the formulation of problems appeared. This is, firstly, factorization with the subsequent construction of a regression on factors. Secondly, the intensive development of the *R*-analysis scheme, i.e. factor analysis of time series.

The processing of time series in traditional factor analysis models encounters a number of additional difficulties due to the overestimation of the correlation coefficients and factor loads arising from the trend characteristic of such series. In applied works, a number of methods are used that reduce the influence of the trend:

1. The transition from the initial statistical series defined by the values of the levels to the series characterizing the values of the growth or growth rate, i.e. from x_t to

x_t-x_{t-1} or $\frac{x_t - x_{t-1}}{x_t}$. If it is desirable to eliminate the influence of seasonal fluctuations on quarterly data, use the series conversion to x_t- x_{t-4} or to $\frac{x_t - x_{t-4}}{x_t}$ [5,6].

2. Elimination of the trend by calculating the regression line of each variable from time to time, followed by a transition from the initial values to the differences between the initial and calculated values of each series.
3. The division of the studied period into sub periods in accordance with one or another substantial hypothesis and on the basis of the assumption that for a short period of time the influence of the trend is not decisive.
4. Finally, some authors propose to consider the first factor a characteristic of the trend and exclude it from further consideration. Investigate the structure of the relationships between the variables using the residual correlation matrix.

Naturally, the choice of transforming the source data depends on the statement of the problem and the nature of the characteristics being studied. A trend is the most important component of a time series, so the exclusion of a trend should be substantiated every time, based on substantive considerations. The formal requirements of "clearing" the trend cannot be decisive. In the vast majority of cases, factor analysis is used as a tool to identify the structure of a specific information array, the conclusions are interpreted at a qualitative level. With this approach, no features of the initial series, including the presence of a trend, can serve as an obstacle to factorization. Naturally, the presence of a trend should be taken into account in the interpretation. If "uncleaned" dynamic series containing a trend are factorized, then the result may be a grouping of series according to the nature of the trend, slowly changing inertial signs should be allocated in a separate group, signs with a cyclical nature of changes in another, signs with sharply expressed translational dynamics in a third etc.

The factors identified during the processing of time series are used as generalized characteristics in the analysis and forecasting of complex economic processes, such as, for example, increasing the technical equipment of the industry, increasing labor productivity, demand structure, and market research and inflation in studies on the capitalist economy. Essentially, factor analysis appears in such cases as a special method for constructing complex

indices, and the weights of the individual components of the index (initial variables) are not set a priori by an expert, but are determined in a factor model, factor loads being such weights.

For the first time, factor analysis in such a statement of the problem was used by the English economist E. Rhodes to construct a generalized index of the economic situation, the index was based on the dynamics of 14 specific macroeconomic characteristics [7]. A similar approach to the use of factor analysis for forecasting economic conditions is now practiced in the Center for International Studies of Economic Trends according to surveys (CIRET). Factorization of these data allows reducing the influence of errors and obtaining a synthetic indicator that is sensitive to the movement of the economic cycle.

Researchers used factor analysis of time series to forecast the development of the industry. The characteristic obtained by the method of principal components of generalization describes 87% of the total variance of the four initial time series and is treated as a generalized factor in the development of the industry. Using this characteristic as an argument in the regression model made it possible to construct a satisfactory forecast of labor productivity.

The combination of time series for a number of countries for joint research within the framework of one task allows us to expand the range of observations on the relationships of the studied variables. This technique makes it possible to expand the statistical base of the study, to make factors more representative.

Interesting conclusions about the laws of price movements in the United States were obtained using factor analysis (the method of extreme grouping of parameters) by Soviet economist L.Rakhlina. To find out how the interconnections in the system of 28 indicators determining inflation are changing, factorization was carried out separately for five periods (for economic cycles). The division of indicators into interrelated groups is fairly stably reproduced from one cycle to another. The gross product deflator is most closely related to indicators of pronounced steady growth trends: the volume of government spending, the fund and hourly wage rates, the movement of bank loans, the speed of circulation of unlimited deposits. This leads to the conclusion that state fiscal policy aimed at stimulating "effective demand", as well as expanding the monetary system, is an important factor in price increases. Another group is formed by indicators, in the dynamics of which market fluctuations are

clearly expressed. These are indicators of corporate production activity: net profit, investment, new orders, unemployment, industrial production index. The factor synthesizing these indicators has a large amplitude of oscillations and is an indicator of the cycle. Differences in the grouping of indicators and the magnitude of factor loads by periods reflect the characteristics of each cycle.

The wider application of factor analysis to the study of time series is hampered by the fact that traditional factor analysis models postulate an unchanged composition of factors, i.e. the number of factors and their relationship with the set of source variables are assumed to be the same for the entire set of observations. When working with data from cross-sectional surveys, this assumption often seems reasonable enough. On the contrary, in the study of dynamics it is more logical to assume that the composition of the factors determining the development of the process under study changes significantly with the transition from one stage of development to another.

Although the factors themselves in most cases are not directly observable, a change in the determining factors is reflected in the changing structure of the interconnections of the observed time series. Therefore, the analysis of changes in the structure of these relationships would allow us to identify the most important stages in the development of the process, to give a description of these stages, i.e. justify periodization. New factor analysis algorithms make it possible to simultaneously search for periods and factors characterizing them in such a way that the description of the process under study would be as informative as possible for a given number of periods. The entire time interval for observing processes is divided into periods, within each of which the system of initial variables remains stable, while during the transition from one period to another there is a sharp change in the behavior of the system [8].

Consider a possible statement of the problem for periodization of a complex phenomenon with a changing composition of factors. Let us examine the structure of trade for a number of years. A description of the turnover structure is given by a set of n time series x_j characterizing quarterly changes in the volume of purchases of various goods. The task is to highlight the main periods in the development of trade and find the main characteristic features of each of the periods.

ISSN: 2581-4230, Website: journalnx.com, June 18th & 19th, 2020

The initial information in this problem is given in the form of a two-dimensional data matrix, where the signs x_j ($j=1,\ldots,n$) are represented by n product groups, and the number of observations corresponds to t time intervals. Initial processing of information consists in the transition from x_{jt} values characterizing the level of purchases by product groups at each of t observation points to z_{jt} values characterizing the growth rate of purchases relative to the average level of two consecutive years:

$$z_{jt} = \frac{x_{t+1}-x_t}{x_t+x_{t+1}}.$$

Further, the series go through the following three stages of processing:

1. For each time moment t, the values z_{jt} are divided by a given number of groups (L), so that close values (that is, commodity groups with approximately equal growth rates) fall into one group. The number of groups (L) is set on the basis of substantive considerations and from a comparison of calculations performed for different values of L.
2. Then, the entire studied interval of time observations is divided into a given number of periods (k) so that the partition of z_j into groups within each of k periods is kept as stable as possible. As a measure of the similarity of the two partitions, any of the coefficients of the statistical connection of nominal features, in particular, the Chuprov coefficient, can be used.
3. Finally, for each of the selected periods, such a partition of the initial characteristics into L groups is found that signs with similar changes in growth rates within the given period fall into one group.

Conclusions

The selected groups of signs are treated as factors, i.e. main characteristics of the period. For each of the factors, an average development path is calculated. Note that the composition of factors can vary significantly from period to period.

Literature

1) Lawley D., Maxwell A. Factor analysis as a statistical method. Moscow, Mir, 1967, 176p.
2) Perch I. Factor analysis. Moscow, Statistics, 1974, 254p.
3) Wilkinson, D.H., Algebraic problems of eigenvalues, Moscow, Nauka, 1970, 432p.
4) Duran B., Odedd P. Cluster analysis / Per. from English M.: URSS, 1977
5) Barret N.S., Gerardi G., Hart T.A. Factor analysis of quarterly price and wage behavior for US manufacturing // Quarterly Journal of Economics, 1975, vol. 88, No. 3, pp. 385-408.
6) Carrol J.B. An Analytical Solution for approximating simple structure in factor analysis // Psychometrics, 1983, 18, p. 23-28 /
7) Lawley D., Maxwell A. Factor analysis as a statistical method. Moscow, Mir, 1967, 176p.
8) Perch I. Factor analysis. Moscow, Statistics, 1974, 254p.