SOI: 1.1/TAS DOI: 10.15863/TAS
International Scientific Journal
Theoretical & Applied Science

p-ISSN: 2308-4944 (print) **e-ISSN:** 2409-0085 (online)

Year: 2020 **Issue:** 05 **Volume:** 85

Published: 16.05.2020 **http://T-Science.org**

QR – Issue QR – Article

Javokhir Ismoilovich Safarov
Central election commission of the Republic of Uzbekistan
reseacher

FACTOR ANALYSIS AS PART OF GENERAL MATHEMATICAL METHODS OF PROCESSING

***Abstract**: The paper presents the mathematical basis for the application of factor analysis to solve economic problems. When describing the method for finding factor weights, the data used is specified as a square matrix (in the particular case). In the analysis of economic problems, incorrect equations were obtained, and the method of principal components is used to determine factor weights. The paper also substantiates the importance of the time factor for factor analysis methods.*

***Key words**: factorial weights, factorial analysis, economic problems, incorrect equations, object.*
***Language**: English*
***Citation**: Safarov, J. I. (2020). Factor analysis as part of general mathematical methods of processing. *ISJ Theoretical & Applied Science, 05 (85)*, 113-116.
Soi: http://s-o-i.org/1.1/TAS-05-85-23 ***Doi***: crossref https://dx.doi.org/10.15863/TAS.2020.05.85.23
***Scopus ASCC**: 2200.*

Introduction

Assessment of the values of general factors at the objects of observation. The task of interpreting the results of factor analysis at the substantive level is usually understood as the task of explaining the resulting factors and determining the name of factors on the basis of such an explanation. For this, as was described in detail above, the unevenness in the distribution of factor loads by common factors is analyzed [1,2].

New possibilities for interpreting the results of factor analysis were proposed in geography, where the main data for this were data on the distribution of the values of estimates of the factors themselves over real objects of observation. Moreover, the so-called factor weights were usually used as estimates of factors [3].

The possibility of using this data for interpretation is associated with two circumstances. First, it was possible to easily classify the studied objects by factors and thereby widely used for interpretation the general qualitative ideas about these objects (to compare the resulting classification with any known classification). Secondly, the obtained values of the factors could be superimposed on the map and thereby reveal the relationship of factors, for example, with climate, mineral distribution and other general geographical characteristics. Apparently, something similar can be done when analyzing data of a different nature. It is important that the wasps of the calculated factors can be compared with the axes of some external parameters that could explain the results. The present work is partially devoted to the solution of these problems [4].

Mathematical substantiation of the application of the method of factor analysis.

In the general case, the basic model of factor analysis does not allow expressing factors through the initial parameters precisely, since there are more factors in the model than the initial parameters. For this reason, the exact formulation of the problem of measuring factors is given as the problem of finding the best estimate in one sense or another for factors under the condition of already known factor loads and commonalities [8-11].

When describing the method of finding factor weights, we will use the following notation introduced in the first chapter: F - for the matrix of values of general factors on the objects under consideration and Y - for the matrix of values of the common parts of all parameters on these objects. In this notation we have:

$$Y' = AF', \tag{1}$$

whence

$$A'Y' = A'AF' \quad \text{or} \quad F' = (A'A)^{-1}A'Y'. \tag{2}$$

(Hereinafter, as usual, the prime denotes the transposition of the matrix).

Let matrix A be obtained by the method of principal factors. Then

$$\sum_{j=1}^{n} a_{jp}^2 = \lambda_p; \quad \sum_{j=1}^{n} a_{jp} \cdot a_{jq} = 0 \quad (p \neq q), \qquad (3)$$

where λ_p - eigenvalues of the reduced matrix (R-D^2).

In the matrix form, relations (3) can be written in the following form:

$$A'A = \Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_m). \qquad (4)$$

Substitute (4) in equation (2):

$$F' = \lambda^{-1}A'Y'. \qquad (5)$$

In practice, the incorrect assumption is often made that in (5) hypothetical vectors Y can be replaced by empirical vectors Z. As a result of the replacement, we obtain:

$$F' = \lambda^{-1}A'Z'. \qquad (6)$$

These equations make it possible to express F in terms of Z, and the more accurate they are, the less characteristic the parameters. In the case when the characteristics are equal to zero (the method of principal factors is transformed into the method of principal components), equations (6) are absolutely accurate. Estimates of factors determined using equations (6) are called factor weights.

Before the advent of computers, instead of equations (6), even more “simple” equations were used to determine factor weights:

$$F' = A'Z'. \qquad (7)$$

The factor weights obtained from (7) are auxiliary quantities that play the role of estimates of the contributions of objects to each of the common factors. Geometrically, the factor weight f_{jp} is defined as the projection of the object Z defined in the space Z onto the auxiliary vectors a_p constructed in the same space and made up of the factor loads of the factor F_p.

At present, incorrect equations (6) and (7) are used extremely rarely, and to determine factor weights, they usually resort to the method of principal components, i.e. use the same equation (6), but in which A' are calculated using the principal component method.

Factors in this case are expressed through the initial parameters exactly using the following formula:

$$F_p = \frac{\sum_{i=1}^{n} \sigma_i^{(p)} z_i}{\sqrt{\sum_{i,j=1}^{n} \sigma_i^{(p)} \sigma_j^{(p)} r_{ij}^{(p)}}} \qquad (8)$$

where $\left\| r_{ij}^{(p)} \right\|$ is the matrix coinciding for $i \neq j$ with the matrix of residual correlation coefficients obtained at the r-th step of applying the centroid method, but as $r_{ii}^{(p)}$, the units ($r_{ii}^{(p)} = 1$); $\left(\sigma^{(p)} = \left\{\sigma_i^{(p)}, \ i = \overline{1,n}\right\}\right.$ is a vector with components modulo equal to unity, which is calculated to determine the factor loads on the factor F_p.

About the application of automatic classification methods.

The data matrix, as noted above, can be viewed either as a set of parameters, or as a set of objects of observation. In the first case, the structure of parameter relationships is studied, the basis of which is usually the matrix of scalar products of normalized columns of the data matrix (matrix of correlation coefficients).

In the second case, the structure of the relationships between the objects of observations is studied. The matrix of Euclidean distances between the row vectors of the data matrix often acts as its basis (the element R_{ij} of such a matrix acts as a characteristic of the degree of similarity between the i-th and j-th objects).

The purpose of automatic classification methods with respect to the matrix of distances between objects is quite similar to the purpose of factor analysis methods with respect to the matrix of correlation coefficients between parameters. It comes down to the division of objects of observation into classes in one sense or another of close objects. Moreover, just as in a factor analysis a group of strongly interconnected parameters corresponds to a generalized factor in the methods of automatic classification, the class of close objects corresponds to the standard of a typical representative of the class.

It should also be noted that the methods of automatic classification are as diverse as the methods of factor analysis.

The differences between automatic classification methods and factor analysis methods are associated, firstly, with the differences between parameter vectors and object vectors, which are usually made from meaningful judgments. The parameter vector z_j can be replaced without prejudice to the opposite - z_j, while replacing the object z_i with - z_i changes the meaning of the object to the opposite.

Secondly, in the data matrix, as a rule, the number of parameters is significantly less than the number of objects. The number of the latter in practical cases is hundreds, thousands, and sometimes tens of thousands. This circumstance necessitated the development of special algorithms for solving the automatic classification problem that allow one to classify objects without resorting to explicit calculation of the distance matrix between all possible pairs of objects. Finally, thirdly, in contrast to the methods of factor analysis, in which the matrix of correlation coefficients takes an exceptional place as a form of describing the relationships between parameters, in the methods of automatic classification, the matrix of Euclidean distances is just a particular example of numerous matrices of coefficients of similarity between objects. As a result, it turned out that the methods of factor analysis for the most part are specifically intended for the analysis of Gram matrices, which is the matrix of correlation coefficients, and the methods of automatic classification are adapted for processing similarity matrices of various kinds.

In order to give a more concrete idea of the problem of automatic classification of objects, we describe for example how it is formally formulated in the framework of the method of potential functions.

We introduce the elementary proximity function between two objects *K(x,y)*, called the potential function. For a pair of identical objects, it takes a maximum value, and as the difference between the compared objects x and y increases, it gradually decreases. An example of a family of such functions is given by the formula

$$K(x,y) = \frac{a}{b + cR^d(x,y)},$$

where {*a*, *b*, *c*, *d*} are arbitrary constants; *R*(*x*, *y*) - Euclidean distance between objects.

Let all the objects in some way already be divided into classes. Using the introduced function, the proximity (potential) of the *j*-th object of the *r*-th selected group of objects (*r*-th class) is determined by the sum

$$K_p(x_j) = \sum_{\substack{x_i \in A_p \\ i \neq j}} K(x_j, y_j),$$

where the summation is over the entire object of class A_p except for the *j*-th object.

In these notations, the automatic classification problem is formulated as the problem of finding such a separation of the entire set of objects $\{x_s,\ s = 1,2,\dots,N\}$ into k classes, in which the weighted average density of all classes is maximized, given by the sum

$$y = \sum_{p=1}^{h} \frac{1}{N_p(N_p - 1)} \qquad \sum_{i=1}^{N_p} K_p(x_{ip}),$$

where N_p - is the number of objects of the *r*-th class; x_{ip} - *i*-the class object.

The idea of an algorithm that solves this problem is as follows. Consider the arbitrary separation of the entire set of objects. For some arbitrary division of the set of objects into classes, we calculate the value of the weighted average density, which we denote by I_{st}. We select some object from the *q*-th class and place its *p*-th class. We calculate the weighted average density of the new separation I_n. If it turns out that $I_{st} > I_n$, then we keep the old separation of objects into classes; if it turns out that $I_n > I_{st}$, then the object under consideration is transferred from the q- the class to the *r*-th class. After that, we take some other object and again check whether it is possible to rearrange the selected object from one class to another to increase the average density of classification, etc. until there is a situation in which there is not a single object whose rearrangement from one class to another it would be possible to increase the average weighted classification density. The classification fixed by such a situation is chosen as the final one.

About data processing.

For data processing, a linguistic approach is proposed.

In most studies, until recently, factor analysis methods and automatic classification methods were used, as a rule, separately: in some, only factor analysis methods, in others, only automatic classification methods. However, the structural information that these methods distinguish differs significantly. When applying factor analysis, as a rule, information on the composition of groups of strongly related parameters is studied, and when applying automatic classification, information on the composition of close clusters in the distribution of objects is studied.

The combined use of these methods for processing the same data has only recently been applied. A new synthetic approach to processing a data matrix, in which the results of applying the methods of factor analysis and automatic classification were interconnected.

The main idea of this approach is to develop a standard scheme for describing all structural information that is contained in the matrix in the data as a whole. The proposed scheme is extremely simple: the data matrix is divided into bands representing groups of strongly related parameters, and each strip is divided into blocks of objects that are close in one sense or another. A conditional image of this circuit is given in fig. 1.

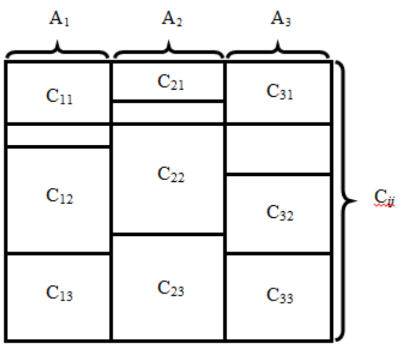

Fig. 1. Data matrix analysis circuit.
A_1, A_2, A_3 - groups of parameters, C_{ij} - groups of objects.

The implementation of such a data matrix analysis scheme can be carried out in many ways. The final classification of objects that this scheme gives is obtained by the intersection of all classifications built for individual bands. The code of each class of such a final classification can be considered as a complex phrase that serves for objects of this class. This complex phrase is composed as an enumeration of elementary phrases; an elementary phrase is the answer to the question: by which group of parameters and to which class the given object belongs. Since such answers are usually easily interpreted in meaningful terms, the resulting codes make it possible to present in a clear form the variety of information in the original data matrix.

The set of received phrase codes can be viewed as a kind of data description language. The dictionary of this language contains two types of words: the words of the first type are the names of the selected groups of strongly related parameters (or the corresponding factors that often replace these groups in such cases), the words of another type are the names of blocks of objects that are close (homogeneous) in terms of the parameters of a separate group parameter. The main role in the framework of this language is played by the already mentioned elementary phrases that highlight one block of objects from one of the group of parameters. With the help of these elementary phrases, it is possible to characterize not only the intersection of all classifications in all groups of parameters (final classification), but also consider any partial classification that is the intersection of classifications in specially selected groups of parameters.

It is important to note that most automatic classification algorithms are constructed in such a way that, along with finding the desired separation of a given set of objects, they build a decisive rule by which any new object that does not participate in the constructed classification can be attributed to one of the classes that belong to that class, to which he is closest. Using these rules for new objects, you can find out their belonging to classes for each selected group of parameters.

Conclusions.

Thus, the generated language is adapted not only to characterize the data matrix being studied, but also to describe all objects of the type whose representatives made up the matrix under consideration.

The indicated approach to processing a data matrix was called linguistic precisely because of the ability to consider the processing result as an automatically generated language for describing the information being studied.

References:

1. Lowley, D., & Maxwell, A. (1967). *Factor analysis as a statistical method*. (p.176). Moskva: Mir.
2. Perch, I. (1974). *Factor analysis*. (p.254). Moscow: Statistics.
3. Wilkinson, D.H. (1970). *Algebraic problems of eigenvalues*. (p.432). Moscow: Nauka.
4. Akobir, S. (2017). *Decision trees - general principles of work*. [Electronic resource] Retrieved from http://www.basegroup.ru/trees/
5. (n.d.). Retrieved from http://www.dmitry-kazakov.de/postgra/chapter2.htm
6. Pozdnyakova, A. (2016). *CIU Discriminant analysis and other multidimensional marketing research methods* (Consumer Insights Ukraine). Kiev.
7. Alekseychuk, N.E., & Kovnir, V.N. (2012). *Ekonomicheskaya teoriya*. – Moscow: Yurayt.
8. Ashirov, F.X. (2005). *Osnovi ekonomicheskoy teorii i praktiki*. (p.407). Moscow: RUDN.
9. Balikoev, V.Z. (2010). *Obshaya ekonomicheskaya teoriya: uchebnik*. (p.685). Moskva: Omega. - 5-e izd..
10. Butirnova, T.S. (2004). *Ekonomika*. Uchebnoe posobie. (p.364). Moscow: Ekonomika.
11. Kiseleva, Ye.A., & Sofronchuk, M.V. (2011). Ekonomicheskaya teoriya. /Pod red. M.N. Chepurina. (p.551). Moscow: Delovaya literatura.