ISSN 1064-5624, Doklady Mathematics, 2019, Vol. 100, No. 3, pp. 514–518.

MATHEMATICS

Factor Model for the Study of Complex Processes

Academician B. N. Chetverushkin[a] and V. A. Sudakov[a,b,*]

Received August 4, 2019

Abstract—An original concept of building a factor model based on frames has been proposed. A method has been developed for calculating the values of factors by searching for the eigenvalues of matrices of pairwise comparisons of the degree of influence of slots. Possible applications of the factor model in the study of complex processes and systems are discussed.

DOI: 10.1134/S1064562419060036

1. INTRODUCTION

An analysis of complex systems, including engineering, economic, and social ones, requires that a large number of diverse characteristics be simultaneously taken into account. Since the vector of characteristics is of high dimension, the construction of a quantitative model is a difficult task.

One approach to modeling such systems is to construct a factor model [1]. It represents a directed graph in which the vertices are the characteristics (factors) determining the state of the system under study and the arcs are the relations between the various factors.

The degree of the influence exerted by one factor on another is reflected differently. In the simplest case, it is characterized by a plus or minus sign, which suggests that an increase in the value of one factor leads to an increase or decrease in the value of the other factor. Such models are called cognitive. As a rule, they are based on expert judgments.

In a more complicated version, each arc is assigned an influence coefficient characterizing the degree of the influence exerted by one factor on the other. The values of these coefficients can be obtained using expert judgments or can be determined from experience, including probabilistic estimates [2]. More complicated cases when the coefficients depend on factors (i.e., there is a nonlinear dependence) are also possible.

In this work, we mainly discuss a problem with statistical influence coefficients.

[a] *Federal Research Center Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, 125047 Russia*
[b] *Plekhanov Russian University of Economics, Moscow, 117997 Russia*
**e-mail: sudakov@ws-dss.com*

2. FORMALIZATION OF FACTOR MODEL

A factor model can be formalized in the matrix form

$$X = AX + F. \tag{1}$$

Here, X is the n-dimensional vector of factor values, A is the matrix of influence coefficients between the corresponding factors, and F is the vector of values characterizing an external influence.

The values of the factors can be found using the iterative procedure

$$X^{(k+1)} = AX^{(k)} + F. \tag{2}$$

The rate of convergence of this procedure (and the possibility of convergence) is determined by the eigenvalues of the matrix A.

In many cases, the problem under study has a high dimension, i.e., the number of factors is large. Then the amount of computations in procedure (2) increases sharply. A visual representation of the graph becomes difficult to perceive. To cope with this situation, we propose modeling factors on the basis of the concept of frames.

A frame is a method for knowledge representation in artificial intelligence (AI) [3] and corresponds to a certain concept of the real world. A frame consists of slots containing structured knowledge. The value of a slot is the value of a factor at a given moment of time. Accordingly, a frame is a subject-area concept aggregating a number of slots according to a semantic feature. An analogy of the transition from slots to a frame is graph condensation [4].

The frame structure makes it possible to reduce the complexity of the system description. Without considering the frame structure, it is possible to examine how the slots of one frame influence the slots of another frame. By analogy with graph condensation, the concentration of several vertices (factors) at a single frame vertex yields a graph of lower dimension. The internal influence of some slots on others within a frame is specified by loops at corresponding vertices.

We introduce the following notation: $i = 1, 2, \ldots, m$ is the frame index; $I = \{i\}$ is the set of frame indices; J_i is the tuple of factor indices of the ith frame; q_i is the cardinality of J_i; and (i_1, i_2) is the arc describing the influence exerted by the i_1th frame on the i_2th frame.

Consider the directed graph of frame relations $G = (I, \{(i_1, i_2)\})$. In G, for each arc (i_1, i_2) and each factor $j^* \in J_{i2}$, we construct an inverse symmetric matrix of pairwise comparisons of the degree of influence of $j \in J_{i1}$ on j^* [5]:

$$U_{i_1 j^*} = \begin{Vmatrix} 1 & u_{12} & u_{13} & u_{14} & \ldots & u_{1q_{i_1}} \\ \frac{1}{u_{12}} & 1 & u_{23} & u_{24} & \ldots & u_{2q_{i_1}} \\ \frac{1}{u_{13}} & \frac{1}{u_{23}} & 1 & u_{34} & \ldots & u_{3q_{i_1}} \\ \frac{1}{u_{14}} & \frac{1}{u_{24}} & \frac{1}{u_{34}} & 1 & \ldots & u_{4q_{i_1}} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{1}{u_{1q_{i_1}}} & \frac{1}{u_{2q_{i_1}}} & \frac{1}{u_{3q_{i_1}}} & \frac{1}{u_{4q_{i_1}}} & \ldots & 1 \end{Vmatrix}. \tag{3}$$

From the equation

$$(\lambda_{i_1 j^*} I - U_{i_1 j^*}) w_{i_1 j^*} = 0, \tag{4}$$

we find the maximum eigenvalue $\lambda_{i_1 j^*}$ and the corresponding vector of influence factors $w_{i_1 j^*}$.

Both criteria characterizing the efficiency of the system performance and other characteristics of the system, for example, resource limitations can be used as factors. The concept of frames admits a fairly broad interpretation. For example, we can introduce a hierarchy of inheritance, which makes it possible to apply identical types of slots to frames with a single ancestor.

A generalized influence matrix A over all factors can be defined as

$$A = \begin{Vmatrix} w_{i_1 1}^{1} & 0 & \ldots & 0 \\ w_{i_1 1}^{2} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ w_{i_1 1}^{q_{i_1}} & 0 & \ldots & 0 \\ 0 & w_{i_2 2}^{1} & \ldots & 0 \\ 0 & w_{i_2 2}^{2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & w_{i_2 2}^{q_{i_2}} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & w_{i_n n}^{1} \\ 0 & 0 & \ldots & w_{i_n n}^{2} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & w_{i_n n}^{q_{i_n}} \end{Vmatrix}. \tag{5}$$

The matrix elements are denoted as $A = \|a_{ij}\|$. They are given by the formulas

$$a_{ij} = \begin{cases} w_{i_1 j}^{i^*} \ \text{ if } \ \exists i_1 \exists i_2 : ((i_1, i_2) \in G) \\ \wedge (i \in J_{i_1}) \wedge (j \in J_{i_2}), \\ 0 \quad \text{otherwise}, \end{cases} \tag{6}$$

where i^* is the index of the factor i in the tuple J_{i1}.

The columns j where there is an influence of the frame i_k contains the vectors $w_{i_k j}$. The other elements are set to zero.

The initial factor influence coefficients are normalized as follows:

$$A^* = \text{norm}\langle A \rangle = A \cdot \begin{Vmatrix} \sum_{k=1}^{n} a_{k1} & 0 & \ldots & 0 \\ 0 & \sum_{k=1}^{n} a_{k2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sum_{k=1}^{n} a_{kn} \end{Vmatrix}. \tag{7}$$

Thus, $\sum_{k=1}^{n} a_{kj}^* = 1 \ \ \forall j$, which is the condition for A^* to be a stochastic matrix.

Consider the sequence of stochastic matrices defined by the relation

$$A^1 = A^*, \quad A^h = \text{norm}\langle A^{(h-1)} \cdot A^* \rangle \quad \text{for} \quad h = 2, 3, \ldots. \tag{8}$$

This procedure demonstrates the propagation of influence of factors along all possible routes in the influence graph associated with the matrix A. Its elements show only the direct influence of each factor of the system on the other factors. The slots of a frame can include other frames, which leads to a hierarchy of factors influencing each other. Factors can influence each other indirectly through transit factors. All possible routes of influence through transit factors have to be considered. For this purpose, we construct a sequence of influence matrices: the matrix A, its square, cube, fourth power, etc.

Consider the limit

$$\lim_{k \to \infty} \sum_{h=1}^{k} \frac{A^h}{k}. \tag{9}$$

If a sequence converges to a limit, then its Cesàro sum converges to the same limit [6]. It may happen that sequence (8) does not converge to a unique limit, but the average of Cesàro sums (9) has a unique limit point. Thus, the following two versions are possible:

(i) Starting at some h, the matrix A^h changes by at most a given small error (according to the Cayley–

Table 1. Initial values of the factor influence coefficients

	P	I	R	S	N	Market 1	Market 2	Q	Pr	C	Commodity 1
P	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0
S	0.83	0.5	0.13	0	0	0	0	0	0	0	0
N	0.17	0.5	0.88	0	0	0	0	0	0	0	0
Market 1	0	0	0	0.7	0.7	1.0	0	0	0	0	0
Market 2	0	0	0	0.3	0	0	1.0	0	0	0	0
Q	0.14	0.29	0.33	0.33	0.6	0	0	0	0	0	0
Pr	0.5	0.57	0.26	0.33	0.28	0	0	0	0	0	0
C	0.36	0.14	0.41	0.33	0.13	0	0	0	0	0	0
Commodity 1	0	0	0	0	0	0	0	1.0	0.5	0.3	1.0

Table 2. Limiting values of the factor influence coefficients

	P	I	R	S	N	Market 1	Market 2	Q	Pr	C	Commodity 1
P	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0
Market 1	**0.18**	**0.19**	**0.2**	0.35	0.41	1.0	0	0	0	0	0
Market 2	**0.06**	**0.04**	**0.01**	0.15	0	0	1.0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0
Pr	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0
Commodity 1	0.76	0.77	0.79	0.5	0.59	0.0	0.0	1.0	1.0	1.0	1.0

Hamilton theorem); hence, the solution has been found.

(ii) The matrix A^h changes with a given period. Then the period and the average over all versions of A^h in it have to be found.

3. TESTING OF THE MODEL

The mathematical model was tested as applied to modeling an innovation-active enterprise. Of course, choosing strategic or tactical approaches used by the enterprise is a task of decision support. It is reasonable to define the aims and criteria for evaluating the quality of the decisions made and to use methods for multiobjective analysis of alternatives. However, before estimating decisions, we need to compute the expected values of criteria resulting from making certain decisions. Criteria can be computed using subject area models. For semi-structured problems, an effective approach, in our view, relies in the construction of frame-based factor models proposed above.

Figure 1 presents a factor model of an innovation-active enterprise. The frames are denoted by rounded rectangles. The name of a frame is given in the upper section of the corresponding rectangle. The factors to be modeled are structured according to frame slots and are shown in the lower section of the frames. The influence of the frames on each other is shown by solid arrows (shaded in black).

The enterprise can work at several markets and can manufacture products of several types. If they are added to the model in the form of independent frames with corresponding slots and influence relations, the visual perception of the model is significantly complicated. All commodities and all markets have identical features, so they are modeled using the concept of generalization, which is shown by unshaded arrows. The resulting scheme allows us to simplify variations in the structure of the model, for example, in the case

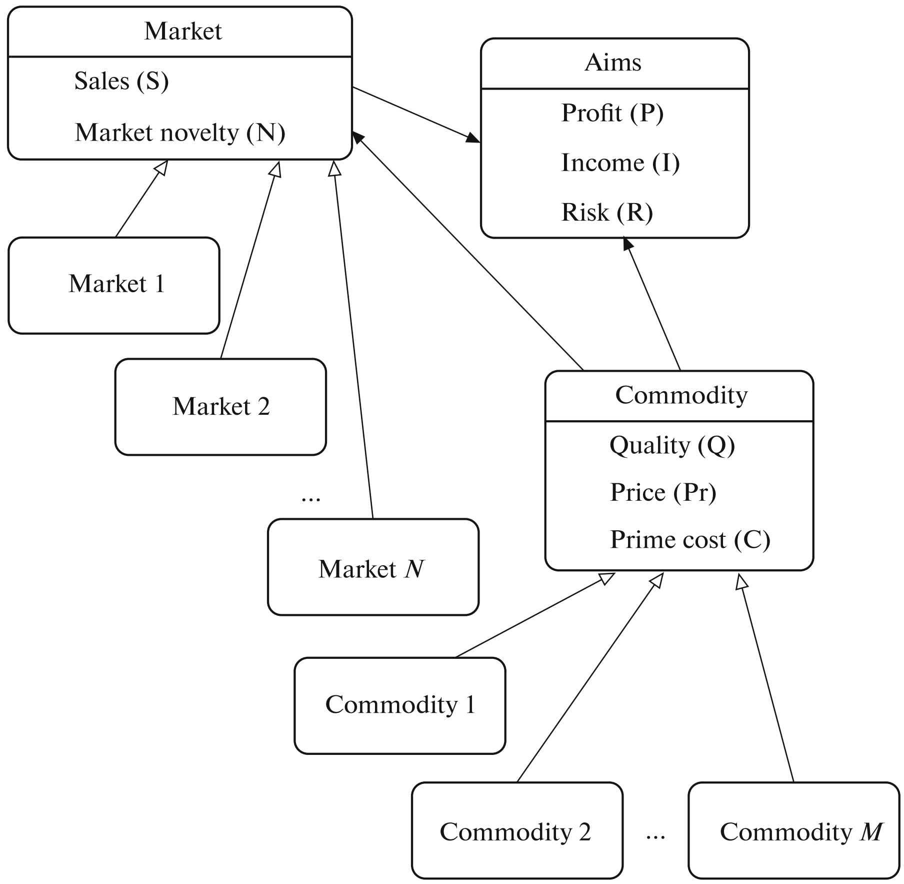

Fig. 1. Model of an innovation-active enterprise.

of adding new types of commodities and in analyzing the expediency of developing new markets. Within the model, we can perform a 'what if' analysis, for example, determine what the profit will be if a commodity with certain characteristics appears on the market.

Table 1 shows the elements of the matrix A calculated in the problem of choosing a strategy for entering one of two alternative sales markets (Market 1 or Market 2) in the case of a single commodity (Commodity 1). These coefficients were obtained by applying expert pair comparisons and computing eigenvalues with the use of formulas (3) and (4). The following notation is used for the factors: P is profit, I is income, R is risk, S is sales, N is market novelty, Q is quality, Pr is price, and C is the prime cost.

The degree of influence of the slots of a frame on a single slot of another frame is normalized; for example, for the influence on P, we have S + N = 1 and Q + Pr + C = 1.

The computations based on formulas (7) and (8) produced a matrix A^2 that remained unchanged up to the error $\varepsilon = 2.886 \times 10^{-3}$. This matrix is shown in Table 2. Zero values in A^2 say that the influence of one factor on another decays over time and passes to other factors.

Table 2 shows that it is more reasonable to enter Market 1, since this decision is Pareto optimal for all three target factors. If a Pareto optimal decision is not obtained, then we need to introduce an integral factor of the quality of the enterprise performance or to use a multiobjective analysis method for examining the alternatives [7].

4. CONCLUSIONS

The developed model can be used to solve the following problems:

determining factors having the largest effect on the efficiency of a system;

predicting the development of complicated dynamical systems with feedback;

assessing and making administrative decisions by analyzing factors that are criteria in the problem under consideration;

modeling innovative activities of enterprises in the case of incomplete or fuzzy information.

Since the example considered in this paper is of low dimension, it can be viewed as educational. However, this approach can be successfully applied to various high-dimensional factor models in scientific-engineering and social-economic problems of current interest and in modeling economy development of corporations, large regions, and countries.

REFERENCES

1. Jae-On Kim and C. W. Mueller, *Factor Analysis: Statistical Methods and Practical Issues* (Sage, Newbury Park, 1989).
2. B. N. Chetverushkin, V. P. Osipov, and V. I. Baluta, Preprint No. 43, IPM RAN (Keldysh Inst. of Applied Mathematics, Russian Academy of Sciences, Moscow, 2018). https://doi.org/10.20948/prepr-2018-43
3. M. Minsky, "A framework for representing knowledge," in *The Psychology of Computer Vision*, Ed. by P. H. Winston (McGraw-Hill, New York, 1975). https://doi.org/10.1016/B978-1-4832-1446-7.50018-2
4. R. Sedgewick, *Algorithms in C++*, Parts 1–4: *Fundamentals, Data Structures, Sorting, Searching*, 3rd ed. (Addison-Wesley Professional, Reading, Mass., 1998).
5. T. L. Saaty, *Decision Making with Dependence and Feedback: The Analytic Network Process* (Rws, New York, 2001).
6. I. I. Volkov, "Cesàro summation methods," in *Encyclopedia of Mathematics*, Ed. by M. Hazewinkel (Springer Science+Business Media B.V./Kluwer Academic, 2001).
7. V. P. Osipov and V. A. Sudakov, Preprint No. 30, IPM RAN (Keldysh Inst. of Applied Mathematics, Russian Academy of Sciences, Moscow, 2015).

Translated by I. Ruzanova