MATTERS OF BUILDING REGRESSION MODEL IN CRITERIAL RATIOS WITH REFERENCE TO ELECTROENERGY PROBLEMS

V. I. Abdullayev, A. M. Kuliyev

Azerbaijan State Oil Academy
Azadlig Avenue 20,Baku, AZ1010, Azerbaijan
Fax (994-12) 498-29-41, (994-12) 493-45-57

ABSTRACT

Analysis of stochastically determined processes in electroenergy problems has been made. Regression analysis is applied as a method at representing variables as similarity criteria, structure of which usually comprises dependent (in terms of dimensions) and one or several independent parameters. It is shown that depending on method of obtaining of random values of similarity criteria, they may become more or less correlated. In case of a strong correlation between independent criteria, the multicollinearity phenomenon occurs. In such situations use of usual least-squares method causes big errors of determination of regression equation parameters.

Keywords: stochastic, regression analysis, similarity criteria, correlation, multicollinearity, electroenergy system, least-squares method.

I. INTRODUCTION

At study of processes of stochastically determined systems, regression analysis methods are of wide application. One of the significant matters of regression analysis is authenticity of obtained results and simplicity of model itself, which can be adversely influenced by a number of significant factors.

At analysis of processes of electroenergy systems, relatively little studied methods based on use of regression analysis at criterial ratios started to be applied lately [1]. In methodological plan, regression analysis at criterial ratios does not have significant differences from classic one, where factors are usual parameters. The main difference is that in case of criterial ratios, more complicated variables – similarity criteria act as factors. As it is known, structure of different similarity criteria usually comprises one or more identical independent parameters in terms of dimensions. It causes certain difficulty at obtaining of random values of criteria on the basis of random values, which are included to their parameters' structure.

As per results of analysis, depending on form of record of similarity criteria and method of obtaining their random values, independent criteria of studied process may become more or less correlated. In case of a strong correlation between similarity criteria, the multicollinearity phenomenon occurs. In such conditions use of usual least-squares method does not give positive outcomes on assessment of regression equation's parameters, which are determined with big error [2, 3].

It should be noted that multicollinearity phenomenon occurs also in case of usual approach, when independent variables are usual parameters. There are different approaches, which allow these difficulties to be bypassed [2]. Use of criterial approach in such cases is more expedient as it allows through selection of comfortable form of record and methods of obtaining their random values to bypass these difficulties and decrease correlation connections between independent criteria.

II. RESULTS AND THEIR DISCUSSION

Let's consider different methods of building regression model in criterial ratios at the example of study of physical process.

Functional dependence reflecting connections between process factors is represented as follows:

$$F(q_1, q_2, q_3, q_4, q_5, t) = 0, \qquad (1)$$

where q – parameters of process including output parameter q_1.

Dimension of only three parameters of six included in dependence (1) can be chosen as independent [1]. Let's determine particular forms of record of similarity criteria on the basis of analysis of dimensions of -π theorem.

1) By accepting q_4, q_5 , and t as parameters with independent measurement unit, the following forms of record of similarity criteria have been obtained:

$$\pi_1 = \frac{q_1 t}{q_4 q_5}; \quad \pi_2 = \frac{q_2 q_4}{t}; \quad \pi_3 = \frac{q_3 q_4}{t^2} \qquad (2)$$

2) By accepting q_2, q_4, and q_5 as parameters with independent measurement unit, the other forms of record of similarity criteria have been obtained:

$$\pi_1 = \frac{q_1 q_2}{q_5}; \quad \pi_2 = \frac{q_3}{q_2^2 q_4}; \quad \pi_3 = \frac{t}{q_2 q_4} \qquad (3)$$

Having determined forms of record of similarity criteria, dependence (1) can be represented as criterial ratios:

$$\Phi(\pi_1, \pi_2, \pi_3)=0. \qquad (4)$$

Having accepted criteria π_1 as an output parameter of studied process, dependence π_1 on π_2 and π_3 can be represented in the form of the following polynomial:

$$\pi_1 = b_0 + \sum_{i=2}^{3} b_i \pi_i + \sum_{\substack{i=1 \\ i \neq j}}^{3} b_{ij} \pi_i \pi_j \qquad (5)$$

The following methods of obtaining random values of independent criteria π_2 and π_3 were considered to obtain polynomial (5):

1. Random values π_2 and π_3 in (2) and (3) were obtained on the basis of random values q_2, q_3, r_3 and q_4 by modeling them using normal distribution law.

2. Random values π_2 and π_3 in (2) were obtained on the basis of random values q_2 and q_4, by modeling them using normal distribution law.

3. Random values π_2 and π_3 in (3) were obtained on the basis of random values q_2 and q_4, by modeling them using normal distribution law.

Time value t during study process was kept on permanent level.

The difference of reviewed methods is that in first case random values of independent criteria π_2 and π_3 were obtained on the basis of random values of three parameters q_2, q_3 and q_4, two of which q_2 and q_4 were common. In second case random values of independent criteria π_2 and π_3 were obtained on the basis of random values q_2 and q_3, each of which is included to the structure of one similarity criteria only. In third case random values π_2 and π_3 were obtained on the basis of two common parameters q_2 and q_4.

In accordance with above methods, random values of similarity criteria were obtained by means of computer, their statistical indices were calculated and assessments of unknown coefficients of criterial dependence were determined in normalized

$$\overline{\pi}_0 = \beta_1 \overline{\pi}_1 + \beta_1 \overline{\pi}_2 \qquad (6)$$

and in real values of coefficients

$$\pi_0 = b_0 + b_1 \pi_1 + b_2 \pi_2 + b_{12} \pi_1 \pi_2, \qquad (7)$$

where β_1 and β_2 – normalized values of coefficients; b_0, b_1, b_2 and b_{12} - real values of regression equation's coefficients.

Depending on the method of obtaining random values of criteria, correlation degree between them became different. At obtaining of random values of similarity criteria on the basis of random values of parameters, which are included to the structure of one criteria only /second case/, there are no correlation connections between them $\left(\tau_{\pi_2-\pi_3} = 0{,}006\right)$.

When random values of similarity criteria are obtained on the basis of random values of three parameters /first case/, when one of them are common, independent similarity criteria become weakly correlated $\left(\tau_{\pi_2-\pi_3} = 0{,}5\right)$. At obtaining random values of similarity criteria on the basis of random values of two or more common parameters /third case/, independent similarity criteria become strongly correlated $\left(\tau_{\pi_2-\pi_3} = 0{,}94\right)$.

Having determined pair coefficients between similarity criteria, normalized and real values of coefficients of criterial-regression model (6) and (7) were further calculated. At presence of strong correlation between independent similarity criteria ($\tau \geq 0{,}75$), model's coefficients become statistically insignificant

$$\begin{aligned} \overline{\pi}_1 &= -0{,}824\overline{\pi}_2 - 0{,}156_1 \overline{\pi}_3 \\ \pi_1 &= 0{,}634 - 0{,}072\pi_2 - 0{,}022\pi_3 \end{aligned} \qquad (8)$$

Values of model's coefficients become more sensitive to limits of scattering of random values of parameters and similarity criteria. If random values of similarity criteria are scattered within ±(10÷50)% from average level, values of coefficients vary within quite large limits (β_1=0,45÷1,15; β_2=0,1÷0,51), (b_1=0,03÷0,11; b_2 =0,007÷0,08).

At weak correlation connection between similarity criteria π_2 and π_3 (τ<0,5), model's coefficients become statistically significant

$$\begin{aligned} \overline{\pi}_1 &= -0{,}657\overline{\pi}_2 - 0{,}454_1 \overline{\pi}_3 \\ \pi_1 &= 1{,}25 - 0{,}452\pi_2 - 0{,}324\pi_3 \end{aligned} \qquad (9)$$

and insensible to limits of scattering of random values of similarity criteria (β_1=0,65÷0,69; β_2=0,41÷0,45), (b_1=0,45÷0,47; b_2 =0,029÷0,33).

Results of analysis have shown that values of regression model's coefficients in criterial ratios do not depend on number of varied parameters and variation limits of their random values.

Conducted analysis allows us to conclude that at criterial-statistical analysis of physical process, random values of similarity criteria must be obtained on the basis of random values of parameters being dependent only in terms of dimensions or only dependent and one independent (common for several similarity criteria) parameters.

III.CONCLUSION

The regression equation in criterial ratios at different forms of record and methods of obtaining random values of similarity criteria has been built. It has been shown that at obtaining random values of similarity criteria on the basis of random values of two and more parameters, which are common for similarity criteria, similarity criteria become strongly correlated. The multicollinearity phenomenon distorting the value of regression equation's

coefficients occurs. To eliminate multicollinearity phenomenon, random values of similarity criteria must be obtained on the basis of random values of no more than one parameter, which is common for similarity criteria.

REFERENCES

1.*Venikov V.A., Venikov G.V.* Similarity theory and modeling (with reference to electroenergy problems). M.:"Visshaya shkola", 1984 – p. 439 (in Russian).

2. *Ayvazyan S.A. etc.* Applied statistics. Study of dependences: Reference book. /S.A.Ayvazyan, I.S.Yenyukov, L.D.Meshalkin: Under edition of S.A.Ayvazyan.- M.: Finances and statistics. 1985.- p. 487 (in Russian).

3. *Demidenko E.Z.* Linear and non-linear regression. M.: Finances and statistics, 1981- p. 302 (in Russian).