METHOD OF CONSTRUCTION OF REGRESSION EQUATION IN CONDITIONS OF SMALL SAMPLE OF INITIAL DATA.

Lev Raskin, Larysa Sukhomlyn,Dmytro Sokolov, Vitalii Vlasenko, Artur Hatunov, Serhii Andriienko
National Technical University «Kharkiv Polytechnic Institute», Kharkiv

One of the most important tasks arising in finding the type of functional dependence in conditions of small sample of initial data and large size of the factor space is the construction of regression polynomials with as few uncertain coefficients as possible. The difficulty lies in the fact that when the number of estimated coefficients of a polynomial (or a functional of another kind) is reduced, the accuracy of estimation of the desired value drops significantly.

The problem of constructing approximating polynomials of good quality with a small number of unknown coefficients is proposed as follows. It is necessary to construct a polynomial of such a kind that the number of estimated coefficients would be smaller than that of the full polynomial. A promising idea for solving this problem is as follows. We describe each coefficient of the regression polynomial by a function of some number of unknown parameters. For this purpose, we maintain an approximation polynomial of the following form:

$$L(x_1,x_2,\dots,x_m)=f_0(b_0,b_{11},\dots,b_{1m})+f_1(b_0,b_{11},\dots,b_{1m})x_1+\dots+$$
$$+f_m(b_0,b_{11},\dots,b_{1m})x_m+f_{11}(b_{21},\dots,b_{2m})x_1^2+f_{12}(b_{21},\dots,b_{2m})x_1x_2+\dots+$$
$$+f_{mm}(b_{21},\dots,b_{2m})x_m^2+\dots$$

where f_{ij}- any function $m$variables, b_{ij}- polynomial coefficient.

To construct a specific polynomial it is necessary to determine the type of functions $f_0,f_1,\dots,f_m,f_{11},f_{12},\dots,f_{\mathrm{mm}},\dots$

Two ways of specifying these polynomials were considered:

$$f_{i_1i_2\dots i_k}=\sum_{j=1}^{k}b_{ki_j}\ ,i_1,i_2,\dots,i_k=1,2,\dots,m \quad (1)$$

$$f_{i_1i_2\dots i_k}=\prod_{j=1}^{k}b_{ki_j}\,,i_1,i_2,\dots,i_k=1,2,\dots,m \quad (2)$$

$$f_0=b_0.$$

The problem of finding the optimal set $B=(b_0,b_{11}\dots b_{1m},b_{21}\dots b_{2m}\dots)^T$,of the approximating polynomial is solved by the least squares method. Of course, the above algorithm does not guarantee finding the optimal vector of estimated coefficients, but it allows us to find a good (and often better) approximation of it. Experimental evaluation of the quality of the proposed methods for obtaining approximation polynomials has revealed a noticeable advantage of truncated polynomials of type (1). The use of the proposed approximations radically reduces the number of estimated parameters, which predetermines a high degree of expediency of their practical application.

References:

1.Jeffrey S. Simonoff , Samprit ChatterjeeRegression Analysis With Applications in R, – J. Wiley 2020.