Fuzzy Non-Linear Regression Analysis and Its Applications

Miroslav Pokorný[†)]

Technical University of Ostrava
Faculty of Electrical Engineering and Informatics
Dept. of Measurement and Control
17. listopadu, 708 33 Ostrava-Poruba
Czech Republic

In the real systems where the human subjective estimation is influential in the regression model or if the complex uncertain and ill-known system under study exists, we must deal with a fuzzy structure of the system. The difference between the actual values and the computed values of the dependent variable in the fuzzy regression models are due to the "indefiniteness" of the system structure. In other words, fuzziness in the system is indicated in the fuzzy system parameters - the regression coefficients in the fuzzy model.

1 Fuzzy linear regression analysis

The fuzzy parameters in the FLRM model $\underline{K}$ are defined by the normal convex fuzzy sets (fuzzy numbers). The fuzzy linear regression model, in general, has the form

$$\underline{Y}_i = \underline{K}_1 x_{i1} + \cdots + \underline{K}_n x_{in}$$

where $\underline{Y}_i$ is the computed output fuzzy form for the i-th sample, $(\underline{K}1, \cdots, \underline{K}_n)$ are the fuzzy parameters of the triangular form $\underline{K}(\alpha, c)$. The estimated val-

[†)]This paper was presented at Fuzzy Workshop, which took place in Kočovce (Slovakia), February 12–17, 1995

ues of the dependent variable $\underline{Y}_i$ in this model are computed on the fuzzy form.

The vagueness of the fuzzy linear regression model is defined by objective function J, which is a sum of all the widths of the fuzzy parameters $\underline{K}(\alpha, c)$ by formula

$$J = c_1 + c_2 + \cdots + c_n$$

This value of J should be as small as possible. The problem of linear regression is then viewed as finding of fuzzy parameters $\underline{K}$ such that J is minimized [1].

2 A new fuzzy rule based non-linear regression model

In case of above presented fuzzy linear regression model FLRM [1], the regression relation validity exists in full scale of the n dimension input variables space. If the hypothesis exists that in the various input variables intervals exists the various regression relations $y = f(x)$ and if the fitting of model is to be sufficient, it is necessary to devide the input variables space and to define appropriate partial regression models. Thus it is possible to design the production rules set which represents the rule based model of system under study with higher degree of fitting than the original regression model. The method of expressing these approximation relations by fuzzy premises gives the possibility of their smooth connection, which is not possible under classical approach.

In [2] the fuzzy rules therefore determine the division of fuzzy input space, and at the same time the input/output conventional (crisp) linear regression relation are defined in each sub-spaces and *fuzzy non-linear regresion model* (FNRM) is constructed in this way. The regression coefficients in his consequent regression relations are *crisp numbers.*

A new proposal of more general method FFRM with the consequent relations in *fuzzy form* and whose consequent parameters are *fuzzy numbers* together with methods of its construction and identification are the original contributions of presented paper. In the contrary to the actual method

FNRM [2], the fuzzy non-linear regression model FFRM is more including the vague phenomenon in its fuzzy regression relations [3].

To present a new generalized fuzzy non-linear model, we consider the following rules, in which the consequent parameters of the regression model are the *fuzzy numbers* K, i.e. for j- independent input variables and R - rules:

IF $[x_1$ is L_{11} and $\ldots$ and x_j is $L_{1j}]$
THEN $[\underline{Y}_1 = \underline{K}_{11}x_1 + \cdots + \underline{K}_{1j}x_j + \underline{K}_{10}]$
....
IF $[x_1$ is L_{R1} and $\ldots$ and x_j is $L_{Rj}]$
THEN $[\underline{Y}_R = \underline{K}_{R1}x_1 + \cdots + \underline{K}_{Rj}x_j + \underline{K}_{R0}]$

R implications divide the space of the variables x into R sub-spaces, in which there R different relations $\underline{Y}_r = f_r(x, \underline{K}_r, \underline{K}_{r0})$ are defined.

The partial results of implications $\underline{Y}_r$ are estimated by formulas of the conventional regression methods. The resulting value of conclusion $\underline{Y}$ is

$$\underline{Y} = \sum_{r=1}^{R} w_r \underline{Y}_r / \sum_{r=1}^{R} w_r$$

where R is the number of the rules and the weight value wr is given by the minimum relation

$$w_r = \min_j \mu_{rj}(x_j^0)$$

where x_j^0 are the values measured at the object and $\mu_{rj}(x_j^0)$ are grades of membership of observed values x_j^0

3 Fitting criterion

There are many criterion for measure of identified model fitting. For using in fuzzy non-linear analysis FFRM the simple Kondo criterion in form

$$FCR = (100/N) \sum_{i=1}^{N} |\underline{Y}_i^0 - \underline{Y}_i| / \underline{Y}_i^0$$

was choosen where $i = 1, 2, \cdots, N$ are number of samples.

4 Identification method

Identification of the premise structure consists of two tasks: a determination of optimal structure of independent variables and a determination of optimal partition of fuzzy space of input variables as a proper problem of fuzzy modeling [2].

The heuristic algorithm of structure optimization consists in starting the identification process with one implication (linear model) and in increasing the number of implications (rules) with simultaneous observation of the trend of the value of FCR criterion. The model has m rules in the m-th step of identification process. Individual model structures are being systematically selected from the set of input variables. The partial sub-models are identified in their selected combinations, and appropriate values of FCR criterion are calculated. The optimal premise structure is achieved in the model with the smallest FCR value. The models structure space area is hierarchically composed using the layers. The special procedure for space searching is proposed [4].

The search is stopped and identification process is finished, when the minimal values of FCR criterion in steps $(m-1)$ and m are in relation $\mathrm{FCR}_{[m]} > FCR_{[m-1]}$, and model $(m-1)$ is described as optimal model, or if the model structure space is exhausted and model m is described as optimal model.

Identification of the consequent fuzzy parameters are implemented for variables of the premise of previous phase of identification of the premise structure.

In the case of non-fuzzy consequent parameters we calculate the regression coefficients by the least squares method. In the new model FFRM we perform the generalisation to determine the consequents regression coefficients as a fuzzy parameters. Then the values of parameters are computed and the linear programming method is used: find the fuzzy parameters $\underline{K}_j(\alpha_j, c_j)$ which give the solution to the linear programming problem with the objective function min J and the constraints as in the previous classical case.

The structure identification of consequents is performed using the simple comparation method. If the value of regression coefficient $\underline{K}_j$ is lower than determined by decision-maker limit $KMIN$, then the variable x_j is implicitely eliminated from the consequent. For fuzzy numbers operation, it was necessary to create the special functions of fuzzy algebra [3].

5 Computer program module

The special computer program modul called NEFRIT (The Non-Linear Effective Fuzzy Regression Identification Tool) was developed and created [4]. Its sub-routines solve all the tasks necessary for the creation and identification of the proposed fuzzy non-linear regression model FFRM.

6 Results of numerical experiments

Some numerical experiments are performed and the identification process was tested.

To applicate the system NEFRIT-V3.0 two real tasks have been choosen. The first one presents the blast furnace herth temperature prediction. The three input independent variables (Silicium contents in pig iron, degree of direct reduction, intensity of blast furnace process) incluiding two linquistic fuzzy values (SMALL, BIG) have been determined. The fourty observations (samples) have been obtained and FFRM model have been identified with final model presented in next form:

x1S x2S x3S	y1= 0.00(0.00)x1+	8.50(6.92)x2+	2727.98(0.00)x3+	0.00(0.00)
x1S x2S x3B	y2=385.40(0.00)x1+	0.00(8.67)x2+	0.00(0.00)x3+	1597.68(0.00)
x1S x2B x3S	y3= 0.00(0.00)x1+	0.00(6.60)x2+	2663.70(0.00)x3+	31.67(0.00)
x1S x2B x3B	y4= 0.00(0.00)x1+	-18.77(6.68)x2+	3171.85(0.00)x3+	0.00(0.00)
x1B x2S x3S	y5= 0.00(0.00)x1+	0.00(4.24)x2+	0.00(0.00)x3+	1836.03(0.00)
x1B x2S x3B	y6=481.40(0.00)x1+	0.00(6.15)x2+	0.00(0.00)x3+	1242.55(0.00)
x1B x2B x3S	y7= 0.00(0.00)x1+	-8.62(0.89)x2+	0.00(0.00)x3+	1992.15(0.00)
x1B x2B x3B	y8= 0.00(0.00)x1+	-0.00(0.00)x2+	2082.36(0.00)x3+	0.00(0.00)

While the FCR value of linear fuzzy regression model is $FCR = 4.96\%$, the fitting criterion of global optimal model is improved to value $FCR = 1.31\%$.

The second identification task presents the task of an coke-oven gas cooler modelling. The three input independent variables (input gas temperature, input gas pressure, cooling water quantity) and one dependent output variable (output coke-oven gas temperature) have been determinated. To identificate this model the fourty-eight observations have been used. The final model of identification process is presented in next:

x1M x2M x3M y1=0.12(0.00)x1+3.32(0.00)x2+0.00(0.00)x3+0.00(0.00)
x1M x2M x3V y2=0.35(0.00)x1+0.00(0.00)x2+0.07(0.02)x3+0.00(0.00)
x1M x2V x3M y3=0.00(0.00)x1+2.05(0.00)x2+0.08(0.00)x3+0.00(0.00)
x1M x2V x3V y4=0.00(0.00)x1+3.74(0.00)x2+0.00(0.03)x3+0.00(0.00)
x1V x2M x3M y5=0.00(0.00)x1+4.84(0.00)x2+0.00(0.06)x3+0.00(0.00)
x1V x2M x3V y6=0.00(0.00)x1+4.31(0.00)x2+0.00(0.03)x3+0.00(0.00)
x1V x2V x3M y7=0.60(0.00)x1+0.00(0.00)x2+0.00(0.02)x3+0.00(0.00)
x1V x2V x3V y8=0.40(0.00)x1+1.61(0.00)x2+0.00(0.03)x3+0.00(0.00)

While the FCR value of linear fuzzy regression model is FCR = 6.34improved to value FCR = 2.65

7 Conclusion

The proposed fuzzy non-linear regression model FFRM including the fuzzy linear regression equations is interesting due to the vague phenomenon of system structure/parameters is better reflected.

The realization of computer program NEFRIT which solves the tasks of construction FFRM and its structure/parameters identification enables application of proposed method in real information and control systems.

References

[1] Tanaka, H., Uejima, S., Asai, K.: *Linear Regression Analysis with Fuzzy Model*, IEEE Transactions on Systems, MAC, Vol.12, 6, 1982

[2] Takagi, T., Sugeno, M.: *Fuzzy Identification of Systems and Its Application to Modeling and Control*, IEEE Transactions on Systems, MAC, Vol.15, 1, 1985

[3] Pokorný, M.: **Regression Analysis Using the Fuzzy Methods**, Doctoral Thesis, Technical University of Brno, 1994

[4] Pokorný, M., Drobisz, Z., Stonawski, I: *A Progressive Approach to Complex Ill-known Systems Modeling*, Proc.4th International DAAAM Symposium, Brno, 1993