MATHEMATICAL SCIENCES AND APPLICATIONS E-NOTES
4 (1) 118-123 (2016) ©MSAEN

Parameter Estimation Based Type-II Fuzzy Logic and Comparison with Robust Methods

*K. Ş. Kula and T. E. Dalkılıç

(Communicated by Murat TOSUN)

Abstract
Parameter estimation is one of the important stages of regression analysis. In the regression analysis, while parameter estimation by classical methods there are a number of assumptions need to be satisfied. One of them is error are normally distributed. In the case that the data sets have outliers, providing of this assumption becomes more difficult. When a data set has outliers, robust methods such as the M method (Huber, Hampel, Andrews and Tukey) are used for estimating parameters. In this paper the Adaptive Network Based Fuzzy Inference System (ANFIS) is used to parameter estimation which is the neural network architecture based type-II fuzzy logic. The proposed method has the properties of a robust method, because the process does not give permission to the intuitional and is not affected by the outliers. Consequently, another aim of this study is, to compare the proposed method with the robust methods that are mentioned above.

Keywords: *Type-II fuzzy logic, outliers, robust regression, membership function.*

AMS Subject Classification (2010): *Primary: 62J05, 62K25.*

***Corresponding author*

1. Introduction

The first serious step for the fuzzy set theory has been taken in 1965 by Lotfi A. Zadeh. The last 35 years, the theory and application of fuzzy sets have been developed rapidly. If there is uncertainty in the components of the membership function, fuzzy set is converted in to Type-II fuzzy set. It can be say that Type-II fuzzy logic is a gener-alization of Type-I fuzzy logic in the sense that fuzziness is not only limited to the linguistic variables but also is present in definition of the component of membership function. Some of the studies on Type-II fuzzy logic given as follows: Türkşen (1999), described the fuzzy rules for both the Type-I and Type-II fuzzy logic theory. Karnik and Mendel (1999) defined a Type-II fuzzy inference system whit uncertainty rules. They also provided a practical algorithm for performing union, intersection and complement to Type-II fuzzy sets (Karnik and Mendel, 2001). Mendel (2007), described the advances for general and interval Type-II fuzzy logic system in the study titled Advance in Type-II fuzzy sets and systems. Definitions of robust M methods which are commonly used in the literature (for details, see [4, 6, 7, 8, 11, 16]).

The remainder of the paper is organized as follows. In the Section 2 the general information about Type-II fuzzy logic are given. In the Section 3 which is the main part of this article, special algorithm for parameter estimation by ANFIS in case where the component of membership function is fuzzy is given. Finally numerical example is given in Section 4.

Received : *29–August–2015,* ***Accepted :*** *03–December–2015*

This work was supported by the Scientific Research Projects Council of Ahi Evran University, Krsehir Turkey under Grant FBA-11-20.

2. Type-II Fuzzy Logic

Type-II fuzzy set, originally introduced by Zadeh as a generalization of concept of ordinary fuzzy set (Zadeh, 1975). Type-II fuzzy sets are characterized by suitable membership function. The degree of each element of Type-II fuzzy set is takes values in interval [0,1]. Since the membership degree in the Type-I fuzzy set is a crisp number in [0,1], Type-II is different from Type-I. If there is uncertainty in membership degree, in the shape of membership function or in some of parameters, Type-II fuzzy sets can be used . When the membership of an element cannot determine in a set as 0 or 1, Type-I fuzzy sets are used [12-15]. Similarly, when the situation is so fuzzy that trouble determining the membership degree even as a crisp number in [0,1], Type-II fuzzy sets are used. There are many real-world problems the exact form of the membership functions are indeterminable. Consider the fuzzy set characterized by normal membership function with standard deviation and mean can take values in $[m_1, m_2]$, the membership function is defined as;

$$\mu(x) = \exp\left\{-\left[\frac{x-m}{\sigma}\right]^2\right\};\ \ m \in [\,m_1,\ m_2] \tag{2.1}$$

and in both cases $\mu(x)$ is a fuzzy set. In this study, the unknown parameters of regression model will be obtained in the event of the independent variables are fuzzy sets that characterized by normal membership function and mean of the membership function is a fuzzy number like as $m \in [m_1, m_2]$. Fuzzy adaptive network based fuzzy inference system will be used in order to obtain the unknown parameters of regression model [1,2]. In this study, the unknown parameters of regression model are determined by used the algorithm for parameter estimation that given in Section three. In this regression model independent variables are fuzzy sets and characterized by Gaussian membership function. The mean that one of the components of the membership function is a fuzzy number like as $m \in [m_1, m_2]$. Fuzzy adaptive network will be used in obtain the unknown parameters of regression model that based fuzzy logic and Type-II fuzzy set.

3. An Algorithm for Parameter Estimation

The determination process of unknown regression parameters is begins with definition class numbers of independent variables. Than the priori parameters of each class are obtained which are characterized the distribution. In this work independent variables are come from Gaussian distribution than we are interested in center $(m)$and spread (σ). The algorithm to determine the unknown parameters of regression model is defined as follows:

STEP 1: Class numbers of independent variables are determinate heuristically.

STEP 2: A priori parameter set include center and spread is determined. These parameters values are depend on class number of independent variables and its range.

STEP 3: $\overline{w}^L$ weights are calculated using the membership function of Gaussian distribution. These weights are outputs from the third layer of the adaptive network and obtained by the following ways; The h^{th} node in the first layer of the adaptive network is defined as

$$f_{1,h} = \mu_{F_h}(x_i)\,. \tag{3.1}$$

Where fuzzy cluster related to fuzzy rules are indicated with $F_1, F_2, ..., F_h$ and μ_{F_h} is the membership function related to F_h. Membership functions for F_h are defined as

$$\mu_{F_h}(x_i) = \exp\left[-\left(\frac{x_i - m_h}{\sigma_h}\right)^2\right]. \tag{3.2}$$

Here, $\{m_h, \sigma_h\}$ is priori parameter set suitable for Gaussin distribution and m is a fuzzy parameter and takes values in the range of $m \in [m_1, m_2]$. w^L weights are obtained from the multiplication of these membership degrees and defined as $w^L = \mu_{F_L}(x_i)\,\mu_{F_L}(x_j)$.

$\overline{w}^L$ weights are normalization of the w^L and determinated by $\overline{w}^L = \frac{w^L}{\sum_{L=1}^{m} w^L}$.

STEP 4: If m is uncertain that one of the priori parameter, unknown coefficients of regression model that called posterior parameter obtained as a fuzzy number shape of $c_i^L = \left(a_i^L,\ b_i^L\right)\ (i = l, ..., p)$. In this case, equality $Z = \left(B^T B\right)^{-1} B^T Y$ is used for determining the c_i^L.

STEP 5: By using the posteriori parameter set, the regression model indicated with $Y^L = c_0^L + c_1^L x_1 + c_2^L x_2 + ... + c_p^L x_p$.

Setting out from the models and weights specified in Step 3, the prediction values are obtained using $\hat{Y} = \sum_{L=1}^{m} \overline{w}^L Y^L$.

STEP 6: The amount of the error of the model is obtained by $\varepsilon = \frac{\sum_{k=1}^{n}(Y_k - \hat{Y}_k)^2}{n}$.

If $\varepsilon < \phi$, the process is completed. Then the posteriori parameter has been determinate as parameters of regression model. If $\varepsilon \geq \phi$, then Step 7 is begins. Here, ϕ is a small fixed value that determinate by decision makers.

STEP7: Central priori parameters specified in Step 2 are updated with $m_h^{'} = m_h \pm t$. Here, t is size of step; $t = \frac{\max(x_{ji}) - \min(x_{ji})}{a}$; $j = 1, 2, ..., n$; $i = 1, 2, ..., p$ and a is stable value which is determinant of size of step and therefore iteration number.

STEP 8: Predictions for each priori parameter obtained by change and error criterion related to these predictions are counted with $\varepsilon_k = Y_k - \widehat{Y_k}$. Here; Y_k is k. predicted outcome and $\widehat{Y_k}$ is k. network output of input vector. The lowest of error criterion is defined. Priori parameters giving the lowest error specified, and prediction obtained via the models related to these parameters is taken as output.

4. Application

Data set used in numerical application is selected from the literature and consists of one dependent variable which is indicated by Y and two independent variables which are indicated by X_1 and X_2. The data set have 30 observations and two of them are outliers. Outlier observation numbers are 22 and 25. This data set is given in Table 1.

Table 1. Data set having one dependent and two independent variables

No	X_1	X_2	Y	No	X_1	X_2	Y
1	24.7000	15.0000	2.6500	**16**	26.2000	22.6000	5.0200
2	24.8000	17.0000	2.6300	**17**	23.9000	22.6000	4.7700
3	26.5000	19.4000	4.9500	**18**	28.1000	23.4000	5.3600
4	29.6000	20.1000	4.4900	**19**	23.0000	18.5000	3.8500
5	25.7000	19.5000	3.1700	**20**	26.0000	16.4000	2.9400
6	25.0000	20.1000	3.8800	**21**	23.6000	21.0000	4.2800
7	21.6000	16.3000	3.9000	**22**	22.4000	15.0000	5.5000
8	24.7000	18.3000	3.5100	**23**	22.6000	19.4000	4.3100
9	25.9000	18.3000	3.9000	**24**	23.4000	20.3000	4.1300
10	25.6000	18.7000	3.4700	**25**	27.5000	22.0000	3.6400
11	27.9000	21.9000	5.5300	**26**	36.0000	19.4000	3.4200
12	25.8000	20.0000	3.4800	**27**	25.2000	20.2000	3.2100
13	26.2000	20.2000	4.3500	**28**	24.7000	14.6000	2.6200
14	24.8000	21.5000	4.3800	**29**	23.4000	21.7000	4.8600
15	27.7000	20.6000	4.3900	**30**	26.2000	21.8000	4.9700

The algorithm proposed in Section three and the defined methods M (Huber, Hampel, Tukey, Andrews) were executed with programs written in MATLAB. From the program, the regression models based fuzzy inference systems are as follows;

$$
\begin{aligned}
\widehat{Y}_1 &= (17500; 1642) + (1040; 103)X_1 + (140; 12)X_2 \\
\widehat{Y}_2 &= (9550; 932) - (810; 79)X_1 + (150; 13)X_2 \\
\widehat{Y}_3 &= (-14150; 1544) + (860; 91)X_1 + (20; 3)X_2 \\
\widehat{Y}_4 &= (-152420; 14251) + (6600; 590)X_1 + (430; 39)X_2 \\
\widehat{Y}_5 &= (112640; 12213) - (4250; 421)X_1 - (390; 40)X_2 \\
\widehat{Y}_6 &= (-103450; 9894) + (4210; 398)X_1 - (40; 5)X_2 \\
\widehat{Y}_7 &= (-870560; 86121) + (25170; 2316)X_1 + (2910; 289)X_2 \\
\widehat{Y}_8 &= (389640; 37865) - (13030; 1298)X_1 + (2360; 241)X_2 \\
\widehat{Y}_9 &= (-40086; 3945) + (11600; 1063)X_1 + (140; 12)X_2.
\end{aligned} \tag{4.1}
$$

Regression model estimates, which are obtained from robust regression methods and the least squares method (LSM), are located in Table 2.

Table 2. The estimation of regression parameters

	Constant	$\widehat{\beta_1}$	$\widehat{\beta_2}$
LSM	1.0904	-0.0576	0.2278
Huber	-0.4174	-0.0426	0.2828
Hampel	-0.8071	-0.0376	0.2950
Tukey	-1.2326	-0.0328	0.3103
Andrews	-1.1091	-0.0328	0.3036

The residuals,which belong to estimates for LSM, belong to estimates for models from robust regression methods and belong to estimates from regression models for network, are located in Table 3.

Table 3. The residuals belong to observations for all methods

No	**LSM Residual**	**Huber Residual**	**Hampel Residual**	**Tukey Residual**	**Andrews Residual**	**Network Residual**
1	-0.4353	-0.1207	-0.0391	0.0372	0.0156	-0.1341
2	-0.9052	-0.7020	-0.6453	-0.6002	-0.6084	0.1306
3	0.9660	1.0119	1.0307	1.0308	1.0387	0.0215
4	0.5250	0.4862	0.4807	0.4551	0.4680	-0.0022
5	-0.8829	-0.8305	-0.8089	-0.8065	-0.7979	0.2052
6	-0.3499	-0.3200	-0.3022	-0.3056	-0.2931	-0.3379
7	0.3400	0.6295	0.7109	0.7822	0.7691	0.0004
8	-0.2588	-0.1090	-0.0640	-0.0338	-0.0353	-0.0698
9	0.1320	0.2473	0.2826	0.3025	0.3030	0.0533
10	-0.4064	-0.3086	-0.2767	-0.2615	-0.2583	-0.2244
11	1.0570	0.9447	0.9259	0.8808	0.9056	0.0018
12	-0.6811	-0.6576	-0.6426	-0.6484	-0.6364	0.1585
13	0.1664	0.1729	0.1834	0.1727	0.1860	-0.2849
14	-0.1804	-0.2244	-0.2227	-0.2466	-0.2247	-0.0838
15	0.2017	0.1638	0.1618	0.1377	0.1538	0.0252
16	0.4263	0.3339	0.3225	0.2841	0.3095	-0.0220
17	0.0439	-0.0142	-0.0140	-0.0413	-0.0161	0.0565
18	0.5568	0.3591	0.3209	0.2519	0.2868	0.0007
19	-0.1306	0.0171	0.0646	0.0954	0.0971	0.0003
20	-0.3894	-0.1711	-0.1132	-0.0647	-0.0768	-0.0032
21	-0.2356	-0.2342	-0.2203	-0.2308	-0.2123	-0.0090
22	2.2823	2.6312	2.7245	2.8118	2.7901	0.0005
23	0.1014	0.2056	0.2441	0.2630	0.2707	-0.0101
24	-0.2376	-0.1948	-0.1713	-0.1701	-0.1563	0.0553
25	-0.8788	-0.9906	-1.0087	-1.0533	-1.0279	-0.0120
26	-0.0170	-0.1130	-0.1422	-0.1879	-0.1794	-0.0013
27	-1.0312	-1.0098	-0.9942	-1.0001	-0.9869	0.3667
28	-0.3742	-0.0376	0.0489	0.1313	0.1070	0.0910
29	0.1734	0.1393	0.1457	0.1254	0.1486	-0.0377
30	0.4219	0.3405	0.3315	0.2962	0.3202	0.0652
ERROR	0.4448	0.4678	0.4815	0.4979	0.4932	0.0173

The weights related to the observations that are used in estimation methods for regression models, are located in Table 4. The weights for robust methods are expression of that observation's effect on one model for each of the outlier observations of the robust method. On the other hand, weight obtained from the network is an expression of that observation's effect on more than one model, which are expressed in Equation (4.1). For this reason, eight different weights, which are called membership degrees of observation, are located in Table 4. The residuals, which belong to estimates from regression models in Equation (4.1) and belong to estimates for models from robust regression methods, are located in Table 4. The proposed algorithm was executed with a program written in MATLAB. In the stage of step operating, data sets have one dependent variables and this variable has an outlier observation.

The defined methods M (Huber, Hampel, Tukey, Andrews) were executed with programs written in MATLAB. The residuals of from the robust methods and LSM are large, but the residuals from the proposed algorithm based network are small. This is because, this method depend on fuzzy clustering.

As it can be seen in a numerical example, error related to estimations obtained via the network according to error criterion is lower than errors obtained via all the other methods.

Table 4. The weight related to observation for all methods

*No	*LSM	*Huber	*Hampel	*Tukey	*Andrews	The membership degrees of the observation to belong to the models in Equation (4.1)								
						w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
1	1	1	1	0.9992	0.4762	0.3397	0.3374	0.3351	0.2347	0.2331	0.2315	0.1101	0.1094	0.1087
2	1	0.8805	1	0.7940	0.4321	0.7464	0.7439	0.7413	0.5227	0.5209	0.5192	0.2487	0.2479	0.2471
3	1	0.6109	0.6838	0.4606	0.3546	0.9826	0.9833	0.9839	0.8691	0.8697	0.8703	0.5224	0.5227	0.5231
4	1	1	1	0.8786	0.4498	0.6705	0.6718	0.6730	0.9079	0.9096	0.9113	0.8353	0.8369	0.8385
5	1	0.7443	0.8713	0.6453	0.4019	0.9810	0.9818	0.9826	0.7774	0.7781	0.7787	0.4186	0.4190	0.4193
6	1	1	1	0.9443	0.4657	0.9005	0.9022	0.9039	0.6482	0.6494	0.6507	0.3171	0.3177	0.3183
7	1	0.9819	0.9913	0.6641	0.4069	0.3899	0.3881	0.3863	0.1759	0.1751	0.1743	0.0539	0.0537	0.0535
8	1	1	1	0.9993	0.4760	0.9011	0.8996	0.8981	0.6225	0.6214	0.6203	0.2922	0.2917	0.2912
9	1	1	1	0.9454	0.4650	0.9694	0.9683	0.9671	0.7896	0.7887	0.7878	0.4371	0.4365	0.4360
10	1	1	1	0.9591	0.4681	0.9910	0.9905	0.9900	0.7746	0.7742	0.7738	0.4114	0.4112	0.4110
11	1	0.6543	0.7612	0.5857	0.3819	0.5251	0.5277	0.5304	0.5630	0.5658	0.5686	0.4101	0.4121	0.4142
12	1	0.9399	1	0.7619	0.4281	0.9343	0.9359	0.9375	0.7507	0.7532	0.7532	0.4098	0.4105	0.4112
13	1	1	1	0.9820	0.4720	0.9073	0.9092	0.9110	0.7701	0.7717	0.7733	0.4442	0.4451	0.4460
14	1	1	1	0.9635	0.4700	0.6394	0.6421	0.6448	0.4478	0.4497	0.4516	0.2131	0.2140	0.2149
15	1	1	1	0.9886	0.4733	0.7853	0.7874	0.7896	0.8190	0.8213	0.8235	0.5805	0.5821	0.5836
16	1	1	1	0.9518	0.4645	0.5518	0.5546	0.5575	0.4684	0.4708	0.4732	0.2701	0.2715	0.2729
17	1	1	1	0.9990	0.4762	0.4967	0.4992	0.5018	0.3074	0.3090	0.3105	0.1293	0.1299	0.1306
18	1	1	1	0.9620	0.4662	0.2510	0.2529	0.2548	0.2766	0.2787	0.2808	0.2071	0.2087	0.2103
19	1	1	1	0.9945	0.4750	0.7919	0.7912	0.7905	0.4331	0.4327	0.4324	0.1610	0.1608	0.1607
20	1	1	1	0.9975	0.4755	0.6456	0.6428	0.6400	0.5332	0.5308	0.5285	0.2992	0.2979	0.2966
21	1	1	1	0.9680	0.4707	0.6676	0.6699	0.6722	0.3965	0.3979	0.3992	0.1600	0.1606	0.1611
22	1	0.2349	0.0978	0	0	0.2584	0.2566	0.2549	0.1301	0.1293	0.1284	0.0445	0.0442	0.0439
23	1	1	1	0.9586	0.4673	0.7473	0.7479	0.7484	0.3869	0.3872	0.3874	0.1361	0.1362	0.1363
24	1	1	1	0.9826	0.4732	0.7587	0.7604	0.7621	0.4384	0.4394	0.4403	0.1721	0.1725	0.1729
25	1	0.6240	0.6987	0.4415	0.3569	0.5224	0.5251	0.5277	0.5301	0.5328	0.5355	0.3655	0.3674	0.3693
26	1	1	1	0.9788	0.4723	0.0855	0.0856	0.0856	0.2789	0.2791	0.2793	0.6182	0.6186	0.6190
27	1	0.6122	0.7089	0.4865	0.3655	0.8947	0.8965	0.8983	0.6619	0.6633	0.6646	0.3328	0.3335	0.3341
28	1	1	1	0.9896	0.4748	0.2729	0.2709	0.2689	0.1885	0.1871	0.1857	0.0885	0.0878	0.0872
29	1	1	1	0.9905	0.4735	0.5249	0.5273	0.5297	0.3033	0.3047	0.3061	0.1191	0.1196	0.1202
30	1	1	1	0.9476	0.4637	0.5955	0.5983	0.6012	0.5054	0.5079	0.5103	0.2915	0.2929	0.2943

5. Conclusion

In this study, independent variables are Gaussian distributed and regression models are formed by adaptive network using membership functions that are produced for Gaussian distribution. Since the central parameter m that one of the posteriori parameters is uncertain and takes values in the range $m \in [m_1, m_2]$ the unknown parameters of regression model are obtained as fuzzy number. To demonstrate the validity of the regression model that obtained from adaptive network, the predicted values from this model are compared whit predicted values from least square estimates and predicted values from robust methods. In case of the data set have outliers, according to the indicated error criterion, the error related to the predictions that are obtained from the adaptive network are less than errors obtained from the other methods.

6. References

[1] Aisbett, J.,Rickart, J. T., Morgenthaler, D., Multivariate modeling and type-2 fuzzy sets, Fuzzy Sets and Systems 163 (2011) 78-95.

[2] Castillo, O., Melin, P. Type-2 fuzzylogic: theory and applications, Sipringer, 2008.

[3] Cheng, C.B., Lee, E.S., Switching Regression Analysis by Fuzzy Adaptive Network, Europen Journal of Operational Research 128, (2001) 647-668.

[3] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., and Stahel, W.A., Robust Statistics, John- Willey and Sons, New-York (1986).

[5] Hisao, I., Manabu, N., Fuzzy Regression Usin Asymmetric Fuzzy Coefficients and Fuzzied Neural Networks, Fuzzy Sets and Systems 119 (2001) 273-290.

[6] Hogg, R.V. Statistician robustness: One View of Its Use in Applications Today. The American Statistician, 33, (1979) 108-115.

[7] Huber, P.J., Robust statistics. John Willey and Son (1981).

[8] Huynh, H., A Comparison of For Approaches to Robust Regression, Psychological Bulletin, 92, (1982) 505-512.

[9] Karnik N.K.,Mendel, J.M.,Type-2 Fuzzy logic systems, IEEE Transaction on Fuzzy Systems 7 (1999) 643-658.

[10] Karnik N.K.,Mendel, J.M., Centroid of a type-2 fuzzy set, Information Sciences 132 (2001) 195-220.

[11] Kula, K.S.and Apaydn, A., Fuzzy robust regression analaysis based on the ranking of fuzzy sets, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16,(2008) 663-681.

[12] Kula, K. S., Dalkilic, T. E., Parameter Estimation Based Type-II Fuzzy Logic, TWMS J. Pure Appl. Math, 4-2, (2013) 187-193.

[13] Mendel, J.M., John, R.I.B., Type-2 fuzzy sets made simple, IEEE Transaction on Fuzzy Systems 10 (2002) 117-127.

[14] Mendel, J. M.,Type-2 Fuzzy Sets and Systems: An Overview, IEEE Computational Intelligence Magazine Fabuary (2007) 21-29.

[15] Mendel, J. M.,Advances in type-2 fuzzy sets and systems, Information Sciences 177 (2007) 84-110.

[16] Rousseeuw, P.J, Leroy, A.M., Robust regression and outlier detection. John Willey and Son, (1987).

[17] Türkşen, I.B., Type I and Type II fuzzy systems modelling, Fuzzy Sets and Systems 106 (1999) 11-34.

[18] Zadeh, L.A., The concept of a linguistic variable and its application to approximate reasoning-1,Information Science 8 (1975) 199-249.

Affiliations

Kamile Şanlı KULA
Address: Ahi Evran University, Faculty of Sciences and Arts
Department of Mathematics, Kırşehir, Turkey.
e-mail: sanli2004@hotmail.com

Türkan Erbay Dalkiliç
Address: Karadeniz Technical University, Faculty of Sciences
Department of Statistics and Computer Sciences, Trabzon, Turkey.
e-mail: tedalkilic@gmail.com