Application of Goal Programming in Fuzzy Linear Regression

H. Hassanpur[1] and H. R. Maleki[2]

Abstract. Most of previous works on fuzzy linear regression concentrated on minimizing a function of spreads of fuzzy numbers, and did not take to account the centers of them, which may be important for decision maker. In this paper a goal programming approach is proposed in which both spreads and centers of fuzzy data are considered in the model. In contrast to the most of previous methods, handling both symmetric and asymmetric trapezoidal and triangular fuzzy data is another feature of proposed approach.
Key words: Goal programming; Fuzzy linear regression

1. Introduction

Fuzzy linear regression (FLR) introduced by Tanaka et al. [4] in 1982, and developed rapidly through past three decades. Some authors (e.g. [2,4]) used mathematical programming (include linear and non-linear programming), and some others (e.g. [1,3]) used least-squares method to calculate regression coefficients.

In this paper, a goal programming (GP) approach is presented to calculate the coefficients of FLR model with crisp input and trapezoidal (triangular) fuzzy output. A trapezoidal fuzzy number, say $\tilde{A}$ in denoted by $\tilde{A}=(a_L,a_U,\alpha,\beta)$ where a_L (a_U) is lower (upper) middle point and α (β) is left(right) spread of $\tilde{A}$. If $a_L=a_U=a$, $\tilde{A}$ is a triangular fuzzy number and denoted by $\tilde{A}=(a\ ,\alpha,\beta)$.

2. Proposed approach

Suppose the given inputs are positive crisp numbers x_{ij} and the observed responses are trapezoidal fuzzy numbers $\tilde{y}_i=(\ y_{iL},y_{iU},l_i,r_i)$, i=1,2,…,n, j=1,2,…,p (n is the number of observations and p is the number of independent variables). Assume that the FLR coefficients are $\tilde{A}_j=(\ a_{jL},a_{jU},\alpha_j,\beta_j)$, j=0,1,…,p. Consider the FLR model as:

$$\tilde{Y}=\tilde{A}_0+\tilde{A}_1x_1+\cdots+\tilde{A}_px_p=\sum_{j=o}^{p}\tilde{A}_jx_j \tag{1}$$

where $x_0=1$. To estimate $\tilde{A}_j$'s, we try to close the lower middle points, upper middle points, left spreads, and right spreads of observed response $\tilde{y}_i$ to those of estimated response $\tilde{Y}_i=\sum_{j=o}^{p}\tilde{A}_jx_{ij}$, respectively, $\forall i$, as much as possible. This can be done by solving the following GP model:

$$Min\ \ z=\sum_{i-1}^{n}(n_{iL}+p_{iL}+n_{iU}+p_{iU}+n_{il}+p_{il}+n_{ir}+p_{ir})$$

[1] Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, I.R. of Iran
E-mail address: hassanpur@graduate.uk.ac.ir
[2] Faculty of Basic Sciences, Shiraz University of Technology, Shiraz, I.R. of Iran
E-mail address: maleki@sutech.ac.ir
This research is supported in part by Fuzzy Systems and Applications Center of Excellence, Shahid Bahonar University of Kerman, Kerman, I.R. of Iran.

$$\text{s.t.} \quad \sum_{j=0}^{p} a_{jL} x_{ij} + n_{iL} - p_{iL} = y_{iL} \qquad i=1,2,\ldots,n, \tag{2}$$

$$\sum_{j=0}^{p} a_{jU} x_{ij} + n_{iU} - p_{iU} = y_{iU} \qquad i=1,2,\ldots,n, \tag{3}$$

$$\sum_{j=0}^{p} \alpha_{j} x_{ij} + n_{il} - p_{il} = l_{i} \qquad i=1,2,\ldots,n, \tag{4}$$

$$\sum_{j=0}^{p} \beta_{j} x_{ij} + n_{ir} - p_{ir} = r_{i} \qquad i=1,2,\ldots,n, \tag{5}$$

$$n_{ik} . p_{ik} = 0 \qquad i=1,2,\ldots,n, \quad k \in \{L,U,l,r\} \tag{6}$$

$$a_{jL}, a_{jU} \text{ unrestricted}, \quad \alpha_j, \beta_j \geq 0, \qquad j=0,1,\ldots,p. \tag{7}$$

In above model, for each i, n_{iL} and p_{iL} are the negative and positive deviations between the lower middle points of estimated and observed response, respectively. Also, n_{iU} and p_{iU}, n_{il} and p_{il}, and n_{ir} and p_{ir} are similar deviations between the upper middle points, left spreads, and right spreads of them, respectively.
The above GP model has some interesting feature: I) The constraints (6) can be removed and solved the obtained LP model by available LP solvers. II) For symmetric data, the constraints (4) and (5) are equivalent. So one of them can be removed and solved a smaller model. III) For triangular data, $a_{jL} = a_{jU}$ and $y_{iL} = y_{iU}$. So the constraints (2) and (3) are equivalent. Therefore one of them can be removed. IV) The constraints (2), (3), (4), and (5) are independent. Hence, the above GP model (and its relative LP model) can be decomposed to four separated GP (LP) models. This reduces the number of constraints of GP (LP) model by a noteworthy amount.
The results of some numerical examples and simulation studies showed that the proposed approach is better than several previous methods, according to a criterion of goodness.

3. Conclusion

A goal programming approach for evaluation of fuzzy linear regression model is proposed in this paper. The proposed approach introduces a goal programming model for calculating the regression coefficients. The proposed model is separable. So it can be decomposed to some smaller models, which is very useful especially when the number of observations is large. Furthermore, the proposed approach can handle both trapezoidal and triangular fuzzy numbers easily. Numerical results show better fitness of our model to the given data, in comparison with several previous ones.

References

[1] A.R. Arabpour and M. Tata, Estimating the parameters of a fuzzy linear regression model, Iranian Journal of Fuzzy Systems, To appear.
[2] M. Hojati, C.R. Bector and K. Smimou, A simple method for computation of fuzzy linear regression, European Journal of Operational Research 166 (2005) 172- 184.
[3] B. Kim and R.R. Bishu, Evaluation of fuzzy linear regression models by comparing membership functions, Fuzzy Sets and Systems 100 (1998) 343-352.
[4] H. Tanaka, S. Uejima and K. Asai, Linear regression analysis with fuzzy model, IEEE Transactions on Systems, Man, and Cybernetics 12 (1982) 903-907.