Štefan Varga

Robust estimations in classical regression models versus robust estimations in fuzzy regression models

Kybernetika, Vol. 43 (2007), No. 4, 503--508

Persistent URL: http://dml.cz/dmlcz/135792

KYBERNETIKA — VOLUME 43 (2007), NUMBER 4, PAGES 503 – 508

ROBUST ESTIMATIONS IN CLASSICAL REGRESSION MODELS VERSUS ROBUST ESTIMATIONS IN FUZZY REGRESSION MODELS

Štefan Varga

In this paper are presented two robust estimators of unknown fuzzy parameters in the fuzzy regression model and investigated the relationship between these robust estimators in the classical regression model and in the fuzzy regression model.

Keywords: fuzzy regression model, estimations, robust estimations, predictions

AMS Subject Classification: 62J12

1. ROBUST ESTIMATIONS IN CLASSICAL REGRESSION MODELS

The classical regression model (linear in parameters) is studied in the form

$$y = a_1 f_1(x) + a_2 f_2(x) + \cdots + a_m f_m(x)$$

where $f_i(x)$ are known functions of the input variable x (predictor), y is an output variable (response) and $a = (a_1, a_2, \ldots, a_m)^T$ is the vector of unknown parameters. The observed value

$$y_i = a_1 f_1(x_i) + a_2 f_2(x_i) + \cdots + a_m f_m(x_i) + e_i$$

measured in the point x_i with the error e_i $(i = 1, 2, \ldots, n)$ is a random variable with some probability distribution (the most frequently normal distribution). The uncertainty of the value y_i $(i = 1, 2, \ldots, n)$ is expressed by a probability distribution or at least by the expectation

$$\mathrm{E}(y_i) = a_1 f_1(x_i) + a_2 f_2(x_i) + \cdots + a_m f_m(x_i)$$

$(\mathrm{E}(e_i) = 0)$ and the variance

$$\mathrm{Var}\,(y_i) = \mathrm{Var}\,(e_i) = \sigma_i^2.$$

Practically all types of estimators of the vector of unknown parameters $a = (a_1, a_2, \ldots, a_m)^T$ in the classical regression model are functions of residuals r_i (distances between observed values y_i and estimated values est y_i; $i = 1, 2, \ldots, n$)

$$r_i = y_i - \mathrm{est}\, y_i$$

The most known robust estimators of the vector of unknown parameters $a = (a_1, a_2, \ldots, a_m)^T$ in the studied model are M estimator and LTS estimator.

M estimator (generalization of Maximum likelihood estimator)

$$\mathrm{est}_M a = \arg \min_{a \in \mathbb{R}^m} \sum_{i=1}^{n} \rho(r_i)$$

minimizes the sum of the function values of the residuals $r_i = y_i - \mathrm{est}\, y_i$ of the even function ρ. If the function $\rho(x) = x^2$, the M estimator defined above, is equal to the classical least square estimator. There are a lot of possibilities to choose the function ρ. This function is usually chosen like that it increased to infinity (minus infinity) more slowly than the function $\rho(x) = x^2$ (see [2, 3]).

Least trimmed squares (LTS) estimator

$$\mathrm{est}_{\mathrm{LTS}} a = \arg \min_{a \in \mathbb{R}^m} \sum_{i=1}^{k} r_{(i)}^2$$

minimizes the sum of squares of k smallest residuals $r_{(i)}$. The residuals in the sum are ordered $r_{(1)}^2 \leqq r_{(2)}^2 \leqq r_{(i)}^2 \cdots r_{(k)}^2 \cdots \leqq r_{(n)}^2$ and the number of the residuals $k \in [n/2, n]$.

Application of these robust estimators is suitable, when the assumptions of the classical regression model are not fully satisfied (e. g. independence of observations, normality of measurement errors, outliers, etc.).

Both these robust estimators are asymptotically unbiased and normally distributed under the conditions of regularity ([2, 3]). Other statistical properties of these estimators can be expressed, for example, by the breakdown point $\varepsilon^\star$ ([5, 7]). If $k = [n/2] + [(m+1)/2]$ in the least trimmed square estimator, then the breakdown point $\varepsilon^\star = 0.5$.

2. ROBUST ESTIMATIONS IN FUZZY REGRESSION MODELS

Very natural generalization of the classical regression model is the fuzzy regression model studied in the form

$$Y = A_1 f_1(x) + A_2 f_2(x) + \cdots + A_m f_m(x)$$

where the input variable x (predictor) is a crisp (real) variable, $f_i(x)$ $(i = 1, 2, \ldots, m)$ are known real functions of the variable x, Y is an output fuzzy variable (response) and $A = (A_1, A_2, \ldots, A_m)^T$ is the vector of unknown fuzzy parameters ([8, 9]). It is easy to see that if the fuzzy numbers Y, A_i $(i = 1, 2, \ldots, m)$ are crisp (real number is a special case of fuzzy number), then the fuzzy regression model is equal to the classical regression model.

The uncertainty of an observation Y_i in the point x_i $(i = 1, 2, \ldots, n)$ is expressed by a membership function μ_{Y_i} of the fuzzy number Y_i. We do not have

any probability distribution, any expectation and any variance of the observed value Y_i $(i = 1, 2, \ldots, n)$.

The principle question is how to estimate the vector of unknown fuzzy parameters in the fuzzy regression model and how to define a quality of the estimator. One eventuality could be to generalize not only model, but to generalize the estimators defined in the classical regression model to the estimators in the fuzzy regression model too. What does it mean? It means that, for example,

$$\mathrm{est}_{\mathrm{LTS}} A = \mathrm{est}_{\mathrm{LTS}}(A_1, A_2, \ldots, A_m)^T$$

is the least trimmed squares estimator of the vector of unknown fuzzy parameters in the fuzzy regression model, if it is equal to the least trimmed squares estimator in the classical regression model

$$\mathrm{est}_{\mathrm{LTS}} a = \mathrm{est}_{\mathrm{LTS}}(a_1, a_2, \ldots, a_m)^T = \arg \min_{a \in \mathbb{R}^m} \sum_{i=1}^{k} r_{(i)}^2$$

in the case that the observation Y_i $(i = 1, 2, \ldots, n)$ in the fuzzy regression model is a crisp (real) number, a special case of fuzzy number with the membership function

$$\mu_{Y_i}(x) = \begin{cases} 1, & x = Y_i \\ 0, & x \neq Y_i. \end{cases}$$

Because the difference of two fuzzy numbers is a fuzzy number, we will not minimize a sum of squares of differences $Y_i - \mathrm{est}\, Y_i$ between observed and estimated fuzzy values, but distances between them that can be defined as crisp numbers.

The most commonly used in practice are symmetric triangular fuzzy numbers

$$A = \langle a, s \rangle$$

where a is a center and s a spread of the fuzzy number A. This fuzzy number is about a. Its membership function is

$$\mu_A(x) = \begin{cases} 1 - \frac{|x-a|}{s}, & a - s \leq x \leq a + s \\ 0, & \text{otherwise.} \end{cases}$$

For addition of two fuzzy numbers $A = \langle a, s_1 \rangle$, $B = \langle b, s_2 \rangle$, we can use

$$A + B = \left\langle a + b, \sqrt[w]{s_1^w + s_2^w} \right\rangle$$

and for multiplication of the fuzzy number $A = \langle a, s_1 \rangle$ with the real number k

$$kA = \left\langle ka, \sqrt[w]{|k|} s_1 \right\rangle$$

where the parameter $w \in [1, \infty]$. We have the set of arithmetic, but the most interesting are the limit situations. For $w = 1$

$$A + B = \langle a + b, s_1 + s_2 \rangle, \qquad kA = \langle ka, |k| s_1 \rangle,$$

and for $w = \infty$

$$A + B = \langle a + b, \max\{s_1, s_2\}\rangle, \qquad kA = \langle ka, s_1\rangle.$$

The distance of two fuzzy numbers that is a real number and that is a generalization of the Euclidean distance of two real numbers is the Diamond distance (see [1]) defined for two fuzzy numbers $A = \langle a, s_1\rangle,\ B = \langle b, s_2\rangle$, by the formula

$$d^2(A, B) = (a - b)^2 + \frac{2}{3}(s_1 - s_2)^2.$$

Now when we have defined arithmetic and distance for fuzzy numbers we can specify the studied fuzzy regression model and define robust estimators for unknown fuzzy parameters.

The studied fuzzy regression model is

$$Y = A_1 f_1(x) + A_2 f_2(x) + \cdots + A_m f_m(x)$$

where the input variable x (predictor) is a crisp variable, $f_i(x)$ $(i = 1, 2, \ldots, m)$ are known real functions of the variable x, Y is an output fuzzy variable (response), the observation Y_i $(i = 1, 2, \ldots, n)$ is a symmetric triangular fuzzy number (y_i is a center and z_i is a spread)

$$Y_i = \langle y_i, z_i\rangle$$

$(y_i \in \mathbb{R},\ z_i \in \mathbb{R}^+)$ and $A = (A_1, A_2, \ldots, A_m)^T$ is the vector of unknown symmetric triangular fuzzy parameters

$$A_i = \langle a_i, s_i\rangle$$

$(a_i \in \mathbb{R},\ s_i \in \mathbb{R}^+)$.

To estimate the vector of unknown fuzzy parameters $A = (A_1, A_2, \ldots, A_m)^T$ means, to estimate the vector of all centers $a = (a_1, a_2, \ldots, a_m)^T$ and the vector of all spreads $s = (s_1, s_2, \ldots, s_m)^T$ of the parameters.

Definition 1. *M estimator* of the vector of unknown fuzzy parameters $A = (A_1, A_2, \ldots, A_m)^T$ in the fuzzy regression model is

$$\operatorname{est}_M A = \operatorname{est}_M(A_1, A_2, \ldots, A_m)^T = \operatorname*{arg\,min}_{a\in\mathbb{R}^m,\ s\in\mathbb{R}^{m+}} \sum_{i=1}^{k} \rho(d_i)$$

where d_i is the Diamond distance of the fuzzy numbers Y_i and $\operatorname{est} Y_i$ $(i = 1, 2, \ldots, n)$

$$d_i^2 = d^2(Y_i, \operatorname{est} Y_i).$$

The choice of the function ρ is the same as in the classical regression model ([2, 3]).

Definition 2. *Least trimmed squares estimator* of the vector of unknown fuzzy parameters $A = (A_1, A_2, \ldots, A_m)^T$ in the fuzzy regression model is

$$\operatorname{est}_{\mathrm{LTS}} A = \operatorname{est}_{\mathrm{LTS}}(A_1, A_2, \ldots, A_m)^T = \operatorname*{arg\,min}_{a\in\mathbb{R}^m,\ s\in\mathbb{R}^{m+}} \sum_{i=1}^{k} d_{(i)}^2$$

where the distances between observed and estimated values are ordered

$$d^2_{(1)} \leqq d^2_{(2)} \leqq \cdots \leqq d^2_{(k)} \leqq \cdots \leqq d^2_{(n)}$$

and $k \in [n/2, n]$.

Theorem 1. The M estimator of the vector of the centers a and the vector of the spreads s of the unknown fuzzy parameters in the fuzzy regression model is

$$\begin{aligned} \operatorname{est}_M(a,s) &= \operatorname{est}_M(a_1,\ldots,a_m,s_1,\ldots,s_m)^T \\ &= \underset{a\in\mathbb{R}^m,\, s\in\mathbb{R}^{m+}}{\arg\min} \sum_{i=1}^{n} \rho\left(\sqrt{(y_i - a^T f_i)^2 + \frac{2}{3}\left(z_i - \sqrt[w]{s^w |f_i|}\right)^2}\right) \end{aligned}$$

where new elements in the formula are two column vectors $f_i = (f_1(x_i),\ldots,f_m(x_i))^T$, $|f_i| = (|f_1(x_i)|,\ldots,|f_m(x_i)|)^T$ and one row vector $s^w = (s_1^w,\ldots,s_m^w)$.

P r o o f. It is enough to prove that the Diamond distance of the observed value Y_i and the estimated value $\operatorname{est} Y_i$ in the definition 1 is

$$d_i = d(Y_i, \operatorname{est} Y_i) = \sqrt{(y_i - a^T f_i)^2 + \frac{2}{3}\left(z_i - \sqrt[w]{s^w |f_i|}\right)^2}.$$

The observation $Y_i = \langle y_i, z_i\rangle$ but what is the fuzzy number $\operatorname{est} Y_i$? Using arithmetic presented in this paper we have

$$\begin{aligned} \operatorname{est} Y_i &= \langle a_1, s_1\rangle \cdot f_1(x_i) + \cdots + \langle a_m, s_m\rangle \cdot f_m(x_i) \\ \operatorname{est} Y_i &= \left\langle a_1 f_1(x_i) + \cdots + a_m f_m(x_i),\ \sqrt[w]{s_1^w |f_1(x_i)| + \cdots + s_m^w |f_m(x_i)|}\right\rangle \\ \operatorname{est} Y_i &= \langle a^T f_i,\ \sqrt[w]{s^w |f_i|}\rangle \end{aligned}$$

and the square of the distance of the fuzzy numbers $Y_i = \langle y_i, z_i\rangle$ and $\operatorname{est} Y_i$ is

$$d_i^2 = d^2(Y_i, \operatorname{est} Y_i) = (y_i - a^T f_i)^2 + \frac{2}{3}\left(z_i - \sqrt[w]{s^w |f_i|}\right)^2.$$

Theorem 2. The least trimmed squares estimator of the vector of the centers a and the vector of the spreads s of the unknown fuzzy parameters $A = (A_1, A_2, \ldots, A_m)^T$ in the fuzzy regression model is

$$\begin{aligned} \operatorname{est}_{\mathrm{LTS}}(a,s) &= \operatorname{est}_{\mathrm{LTS}}(a_1,\ldots,a_m,s_1,\ldots,s_m)^T \\ &= \underset{a\in\mathbb{R}^m,\, s\in\mathbb{R}^{m+}}{\arg\min} \sum_{i=1}^{k} \left((y_i - a^T f_i)^2 + \frac{2}{3}\left(z_i - \sqrt[w]{s^w |f_i|}\right)^2\right)_{(i)} \end{aligned}$$

where $k \in [n/2, n]$ is a number of the least distances between observed and fitted values of the fuzzy variable Y that their sum is minimized.

The p r o o f of Theorem 2 is a simple modification of the proof of Theorem 1.

Theorem 3. If the observations $Y_i = \langle y_i, z_i \rangle$; $(i = 1, 2, \ldots, n)$ in the fuzzy regression model are crisp $(z_i = 0;\ i = 1, 2, \ldots, n)$ then two estimators of the unknown fuzzy parameters presented in Theorems 1, 2 are crisp too and equal to the analogous estimators in the classical regression model.

P r o o f . The spreads of all observations $z_i = 0;\ (i = 1, 2, \ldots, n)$ and therefore

$$\sum_{i=1}^{k} d_{(i)}^2 = \min$$

if all elements of the vector $s^w = (s_1^w, \ldots, s_m^w)$ are zero. It means that $s_i = 0$ $(i = 1, 2, \ldots, m)$ thus the fuzzy parameters $A_i = \langle a_i, 0 \rangle$ are crisp $(i = 1, 2, \ldots, m)$ and

$$\sum_{i=1}^{k} d_{(i)}^2 = \sum_{i=1}^{k} \left(y_i - a^T f_i\right)_{(i)}^2 = \sum_{i=1}^{k} \left[y_i - (a_1 f_1(x_i) + \cdots + a_m f_m(x_i))\right]_{(i)}^2 = \sum_{i=1}^{k} r_{(i)}^2$$

that is the formula for the least trimmed square estimator in the classical regression model.

The p r o o f for the M estimator is analogous.

ACKNOWLEDGEMENT

This paper was supported by the grant VEGA 1/2005/05 and by the project APVV 0375-06.

(Received February 1, 2006.)

REFERENCES

[1] A. Bárdossy and L. Duckstein: Fuzzy Rule? Based Modeling with Applications to Geophysical, Biological and Engrg. Systems. CRC Press, Boca Raton 1995.
[2] P. J. Huber: Robust estimation of a location parameter. Ann. Math. Statist. *35* (1964), 73–101.
[3] P. J. Huber: Robust Statistics. Wiley, New York 1981.
[4] G. J. Klir and B. Yuan: Fuzzy Sets and Fuzzy Logic – Theory and Applications. Prentice Hall PTR, Upper Saddle River, NJ 1995.
[5] P. J. Rousseeuw: Least median of squares regression. J. Amer. Statist. Assoc. *79* (1984), 871–880.
[6] M. Šabo: On T-reverse of T-norms. Tatra Mt. Math. Pub. *12* (1997), 35–40.
[7] Š. Varga: Robust estimations in statistics. In: PRASTAN 1998, pp. 86–114.
[8] Š. Varga: Robust estimations in fuzzy linear regression models. In: Quo Vadis Computational Intelligence, Physica–Verlag, Heidelberg 2000, pp. 239–246.
[9] Š. Varga and M. Šabo: Linear regression with fuzzy variables. In: The State of the Art in Computational Intelligence. Physica–Verlag, Heidelberg 2000, pp. 99–103.

Štefan Varga, Slovak University of Technology, Department of Mathematics, Faculty of Chemical and Food Technology, Radlinského 9, 812 37 Bratislava. Slovak Republic.
e-mail: stefan.varga@stuba.sk