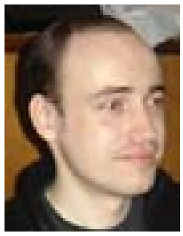
AN APPLICATION OF MINSADBESD REGRESSION



Ciprian Costin POPESCU
PhD Candidate, University Assistant
Department of Mathematics, Academy of Economic Studies, Bucharest, Romania
Calea Dorobantilor 15-17 Street, Sector 1, Bucharest, Romania

E-mail: cippx@yahoo.com

Sudradjat SUPIAN[1]
PhD Candidate

Department of Mathematics, Faculty of Mathematics and Natural Sciences
Padjadjaran University, Jl. Raya Bandung-Sumedang Jatinangor 40600, Bandung, Indonesia

E-mail: adjat03@yahoo.com

Abstract: *In this work, an application of the modified minsadbed (minimizing sum of absolute differences between deviations) approach for a fuzzy environment is given. This type of regression was used for a statistical model with two real parameters and experimental observations which implies real numbers (see Arthanary and Dodge[2]). We develop minsadbed to minsadbesd (minimizing sum of absolute differences between squared deviations) which is more suitable for our model on vague sets. The models on fuzzy sets are described by Ming, Friedman and Kandel[3]; these authors estimate the parameters pre-eminently using least squares. We make an attempt for another method, as in the following writing.*

Key words: *mindsadbesd regression, application, fuzzy models*

1. The minsadbesd approach

Consider the model composed by n observations X_i, Y_i which are put in the forms $[\underline{X_i}(r), \overline{X_i}(r)]$, $[\underline{Y_i}(r), \overline{Y_i}(r)]$ where $\underline{X_i}(r)$, $\overline{X_i}(r)$, $\underline{Y_i}(r)$, $\overline{Y_i}(r)$ are real functions defined on closed interval $[0,1]$ (see Goetschel&Voxman[4], Ming, Friedman and Kandel [5]). The model is approximately described by a regression line given by the equation $Y = a + bX$, $(a,b) \in R \times R^*$. We put the additional conditions that the line pass through the point M of form $(M_X(r), M_Y(r))$, where $\underline{M_X}(r) = \overline{M_X}(r) =$const.$\in R$, $\underline{M_Y}(r) = \overline{M_Y}(r) =$const.$\in R$. Thus we have the initial relation $\underline{M_Y} = a + b\underline{M_X}$. For the inputs $X_i, i = 1\dots,n$ the distance between an observed value Y_i and the corresponding theoretical value $Y_i = a + bX_i$ is[6]:

$$D_i = \sqrt{\left[\int_0^1 \left(a + b\underline{X}_i(r) - \underline{Y}_i(r)\right)^2 dr + \int_0^1 \left(a + b\overline{X}_i(r) - \overline{Y}_i(r)\right)^2 dr\right]} \text{ if } b > 0$$

and

$$d_i = \sqrt{\left[\int_0^1 \left(a + b\overline{X}_i(r) - \underline{Y}_i(r)\right)^2 dr + \int_0^1 \left(a + b\underline{X}_i(r) - \overline{Y}_i(r)\right)^2 dr\right]} \text{ if } b < 0 .$$

Case 1: $b > 0$.

In this case we solve the problem under the assumption that $b > 0$.

The minsadbesd algorithm lead us to solve the problem

$$\min_{(a,b)\in R\times R_+^*} \sum_{i<j} \left|D_i^2 - D_j^2\right| \quad (1.1)$$

or

$$\min_{(a,b)\in R\times R_+^*} \sum_{i<j} \left|\int_0^1 \left(a + b\underline{X}_i(r) - \underline{Y}_i(r)\right)^2 dr + \int_0^1 \left(a + b\overline{X}_i(r) - \overline{Y}_i(r)\right)^2 dr - \right.$$

$$\left. - \int_0^1 \left(a + b\underline{X}_j(r) - \underline{Y}_j(r)\right)^2 dr - \int_0^1 \left(a + b\overline{X}_j - \overline{Y}_j(r)\right)^2 dr\right| \quad (1.2)$$

For all $i < j$, $i, j = 1,\ldots,n$, we make the substitutions:

$$\begin{cases} p_i(r) = \underline{X}_i(r) - \underline{M}_X(r) \\ q_i(r) = \underline{Y}_i(r) - \underline{M}_Y(r) \end{cases}, \quad \begin{cases} P_i(r) = \overline{X}_i(r) - \overline{M}_X(r) \\ Q_i(r) = \overline{Y}_i(r) - \overline{M}_Y(r) \end{cases},$$

$$\begin{cases} p_j(r) = \underline{X}_j(r) - \underline{M}_X(r) \\ q_j(r) = \underline{Y}_j(r) - \underline{M}_Y(r) \end{cases}, \quad \begin{cases} P_j(r) = \overline{X}_j(r) - \overline{M}_X(r) \\ Q_j(r) = \overline{Y}_j(r) - \overline{M}_Y(r) \end{cases}.$$

Thus (1.2) is equivalent to

$$\min_{b\in R_+^*} \sum_{i<j} \left| b^2 \int_0^1 \left(p_i^2(r) + P_i^2(r) - p_j^2(r) - P_j^2(r)\right) dr - \right.$$

$$- 2b\int_0^1 \left(p_i(r)q_i(r) + P_i(r)Q_i(r) - p_j(r)q_j(r) - P_j(r)Q_j(r)\right) dr +$$

$$\left. + \int_0^1 \left(q_i^2(r) + Q_i^2(r) - q_j^2(r) - Q_j^2(r)\right) dr \right| \quad (1.3)$$

or

$$\min_{b\in R_+^*} \sum_{i<j} \left|a_{ij}b^2 + b_{ij}b + c_{ij}\right| \quad (1.4)$$

where

$$a_{ij} = \int_0^1 \left(p_i^2(r) + P_i^2(r) - p_j^2(r) - P_j^2(r)\right) dr, \quad c_{ij} = \int_0^1 \left(q_i^2(r) + Q_i^2(r) - q_j^2(r) - Q_j^2(r)\right) dr,$$

$$b_{ij} = -2\int_0^1 \left(p_i(r)q_i(r) + P_i(r)Q_i(r) - p_j(r)q_j(r) - P_j(r)Q_j(r)\right) dr .$$

Let $f_{ij}(b) = \left|a_{ij}b^2 + b_{ij}b + c_{ij}\right|$. For function $a_{ij}b^2 + b_{ij}b + c_{ij}$, we have $\Delta_{ij} = b_{ij}^2 - 4a_{ij}c_{ij}$.

The sign of the discriminant is unknown. We have four cases which depends on signs of a_{ij}, Δ_{ij}; consequently, the graph of $f_{ij}(b)$ has one of the forms shown in Fig. 1-4.

Case 1.1.

The "easy" case appears when all the discriminants are negative, namely $\Delta_{ij} \le 0, \forall i = 1,\dots,n$. In this situation the functions have the forms shown in Fig. 3 or Fig. 4.

The problem (1.4) is equivalent to

$$\min_{b\in R_+^*}\left[Ab^2 + Bb + C\right] = \min_{b\in R_+^*}\left[\left(\sum_{i<j}\left|a_{ij}\right|\right)b^2 + Bb + C\right] \quad (1.5)$$

where $B = \sum_{i<j} \pm \left|b_{ij}\right|$ (this writing means that some of the terms are positively and the others are negatively, depending on the concrete signs of a_{ij}) and $C = \sum_{i<j} \pm \left|c_{ij}\right|$.

The unique minimizing point for the function $u(b) = Ab^2 + Bb + C$ (see also Fig. 5) is

$$b^* = -\frac{B}{2\sum_{i<j}\left|a_{ij}\right|}.$$

Case 1.2.

Δ_{ij} have random signs.

The graph of the continuous function $\sum_{\substack{i,j=\overline{1,n}\\ i<j}} f_{ij}(b)$ is composed from small pieces which are parts from the functions given by the equations $Db^2 + Eb + F$ where D, E, F are real numbers with general form $D = \sum_{i<j} \pm \left|a_{ij}\right|$, $E = \sum_{i<j} \pm \left|b_{ij}\right|$, $F = \sum_{i<j} \pm \left|c_{ij}\right|$ (see Fig. 6).

We consider the following sets: $S_1 = \left\{-\frac{c_{ij}}{b_{ij}} / \textit{ for all } i < j \textit{ which gives } a_{ij} = 0, b_{ij} \ne 0\right\}$,

$$S_2 = \left\{\frac{-b_{ij} \pm \sqrt{b_{ij}^2 - 4a_{ij}c_{ij}}}{2a_{ij}} / \textit{ for all } i < j \textit{ which gives } a_{ij} \ne 0, b_{ij}^2 - 4a_{ij}c_{ij} > 0\right\},$$

$$S_3 = \left\{-\frac{E}{2D} / \textit{ for all } D \ne 0, D > 0\right\}.$$

Thus the feasible set is $S = \bigcup_{i=1}^{3} S_i$ and the problem (1.4) becomes $\min_{b\in S}\sum_{i<j} f_{ij}(b)$ which is relatively easy to settle, as in Section 2.

Case 2: $b < 0$.

$$\min_{(a,b)\in R\times R_{-}^{*}} \sum_{i<j} \left| d_i^2 - d_j^2 \right| \quad (1.6)$$

gives

$$\min_{(a,b)\in R\times R_{-}^{*}} \sum_{i<j} \left| \int_0^1 \left(a + b\overline{X_i}(r) - \underline{Y_i}(r)\right)^2 dr + \int_0^1 \left(a + b\underline{X_i}(r) - \overline{Y_i}(r)\right)^2 dr - \right.$$

$$\left. - \int_0^1 \left(a + b\overline{X_j}(r) - \underline{Y_j}(r)\right)^2 dr - \int_0^1 \left(a + b\underline{X_j} - \overline{Y_j}(r)\right)^2 dr \right| \quad (1.7)$$

The problem (1.7) is equivalent with

$$\min_{b\in R_{-}^{*}} \sum_{i<j} \left| b^2 \int_0^1 \left(p_i^2(r) + P_i^2(r) - p_j^2(r) - P_j^2(r)\right) dr - \right.$$

$$- 2b \int_0^1 \left(p_i(r)Q_i(r) + P_i(r)q_i(r) - p_j(r)Q_j(r) - P_j(r)q_j(r)\right) dr +$$

$$\left. + \int_0^1 \left(q_i^2(r) + Q_i^2(r) - q_j^2(r) - Q_j^2(r)\right) dr \right| \quad (1.8)$$

which becomes

$$\min_{b\in R_{-}^{*}} \sum_{i<j} \left| a_{ij} b^2 + b'_{ij} b + c_{ij} \right| \quad (1.9)$$

if

$$b'_{ij} = -2\int_0^1 \left(p_i(r)Q_i(r) + P_i(r)q_i(r) - p_j(r)Q_j(r) - P_j(r)q_j(r)\right) dr .$$

We denote $g_{ij}(b) = \left| a_{ij} b^2 + b'_{ij}\, b + c_{ij} \right|$. For function $a_{ij} b^2 + b'_{ij}\, b + c_{ij}$, we have $\Delta'_{ij} = b'^2_{ij} - 4a'_{ij}\, c'_{ij}$.

Case 2.1.

First, we consider the case $\Delta'_{ij} \le 0, \forall i = 1,\dots,n$.

Thus the graph of $g_{ij}(b)$ has one of the two forms shown in Fig. 3 and Fig. 4.

Then the problem (1.9) is equivalent with

$$\min_{b\in R_{-}^{*}} \left[Ab^2 + B'b + C' \right] = \min_{b\in R_{-}^{*}} \left[\left(\sum_{i<j} \left| a_{ij} \right| \right) b^2 + B'b + C' \right] \quad (1.10)$$

where $B' = \sum_{i<j} \pm \left| b'_{ij} \right|$, $C' = \sum_{i<j} \pm \left| c_{ij} \right|$.

The unique minimizing point for function $\left[Ab^2 + B'b + C' \right]$ is $b^{**} = -\dfrac{B'}{2\sum_{i<j} \left| a_{ij} \right|}$.

The approach is the same as in first case but with other coefficients.

In both situation, a is obtained from the condition that the line pass through the initial fixed point.

Case 2.2.

All the comments stored in case 1.2 keeps their validity.

For $D' = \sum_{i<j} \pm |a_{ij}|$, $E' = \sum_{i<j} \pm |b'_{ij}|$, $F' = \sum_{i<j} \pm |c_{ij}|$ we have

$$S'_1 = \left\{ -\frac{c_{ij}}{b'_{ij}} \,/\, for \;\; all \;\; i < j \;\; which \;\; gives \;\; a_{ij} = 0, b'_{ij} \neq 0 \right\},$$

$$S'_2 = \left\{ \frac{-b'_{ij} \pm \sqrt{b'^2_{ij} - 4a_{ij}c_{ij}}}{2a_{ij}} \,/\, for \;\; all \;\; i < j \;\; which \;\; gives \;\; a_{ij} \neq 0, b'^2_{ij} - 4a_{ij}c_{ij} > 0 \right\},$$

$S'_3 = \left\{ -\frac{E'}{2D'} \,/\, for \;\; all \;\; D' \neq 0, D' > 0 \right\}$ and (1.4) is equivalent with

$$\min_{b \in S'_1 \cup S'_2 \cup S'_3} \sum_{i<j} g_{ij}(b) = \min_{b \in S'} \sum_{i<j} g_{ij}(b) \quad (1.11)$$

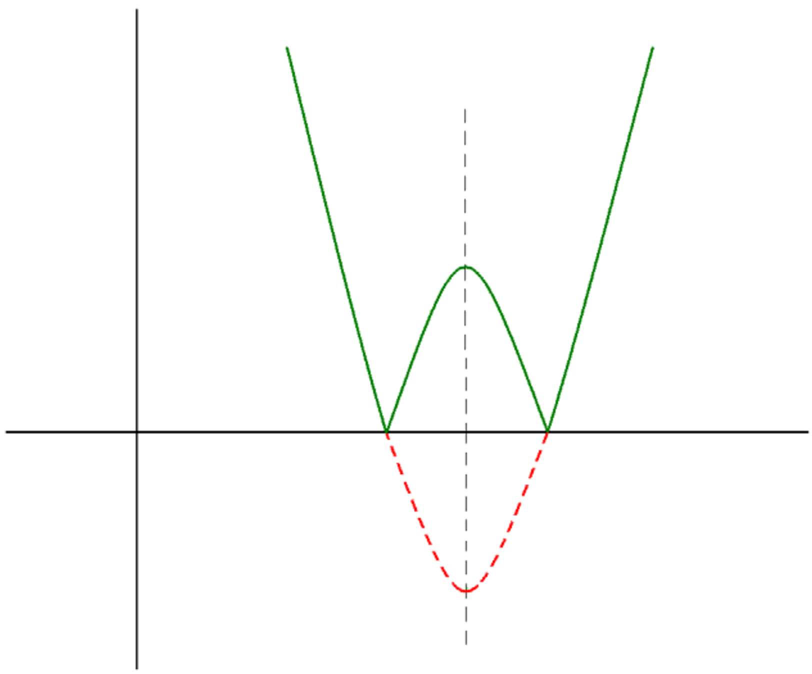


Figure 1. The graph of $f_{ij}(b)$ (green) if $\Delta_{ij} > 0$ and $a_{ij} > 0$

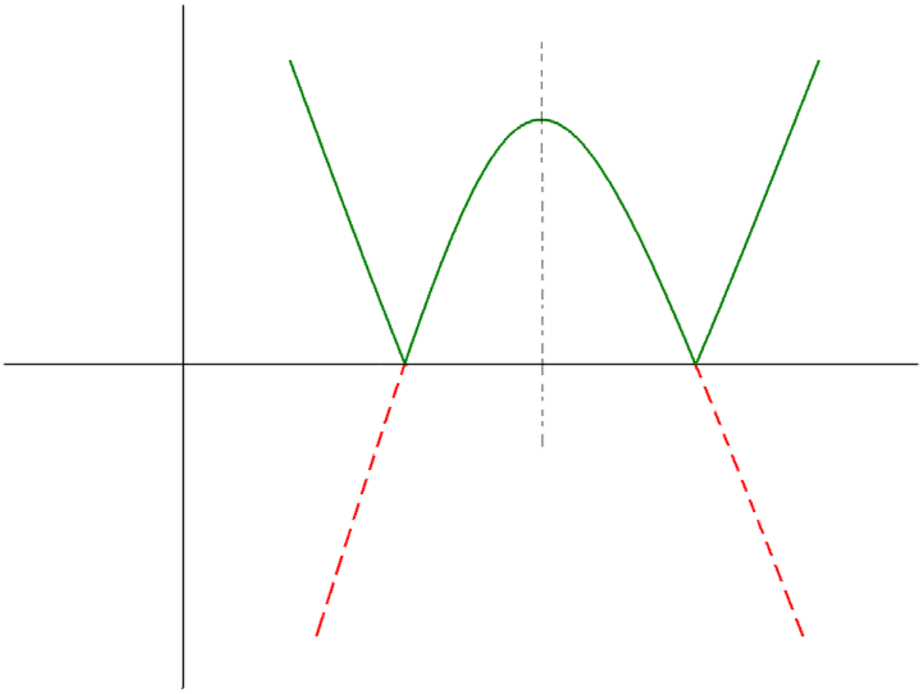


Figure 2. *The graph of* $f_{ij}(b)$ *(green) if* $\Delta_{ij} > 0$ *and* $a_{ij} < 0$

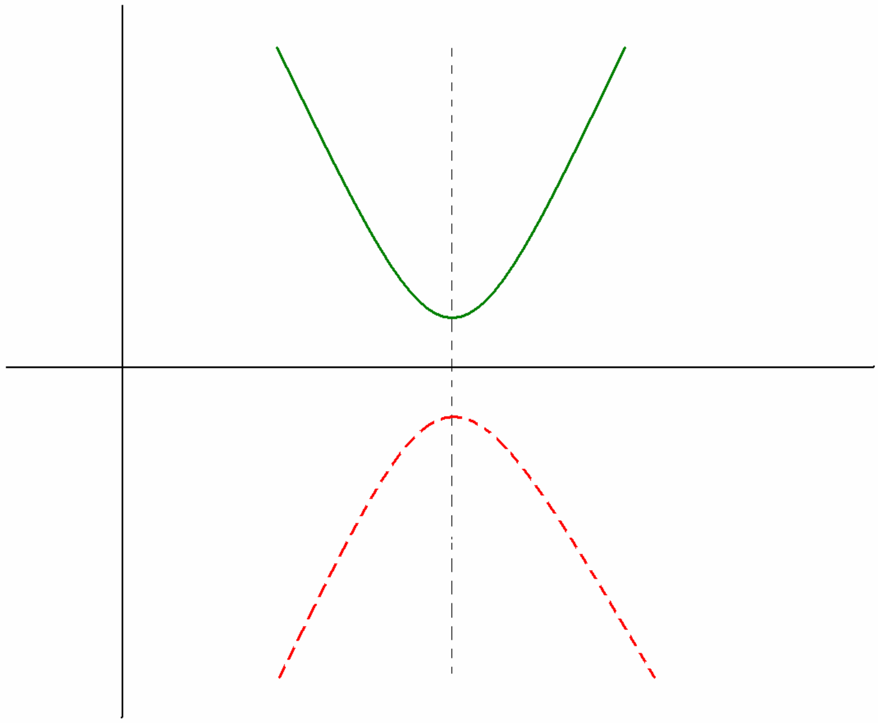


Figure 3. The graph of $f_{ij}(b)$ if $\Delta_{ij} < 0$ and $a_{ij} < 0$

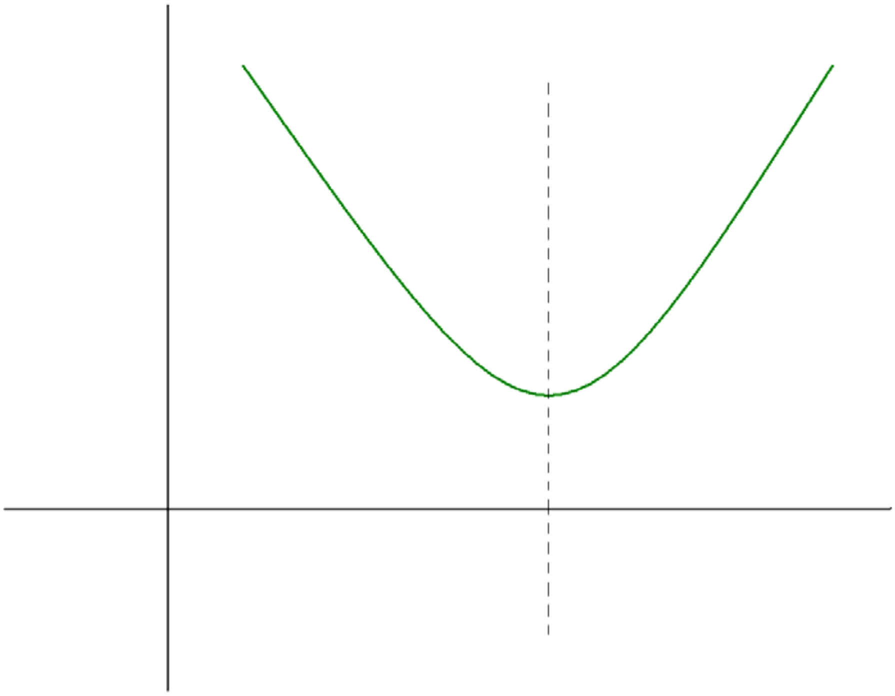


Figure 4. The graph of $f_{ij}(b)$ if $\Delta_{ij} < 0$ and $a_{ij} > 0$

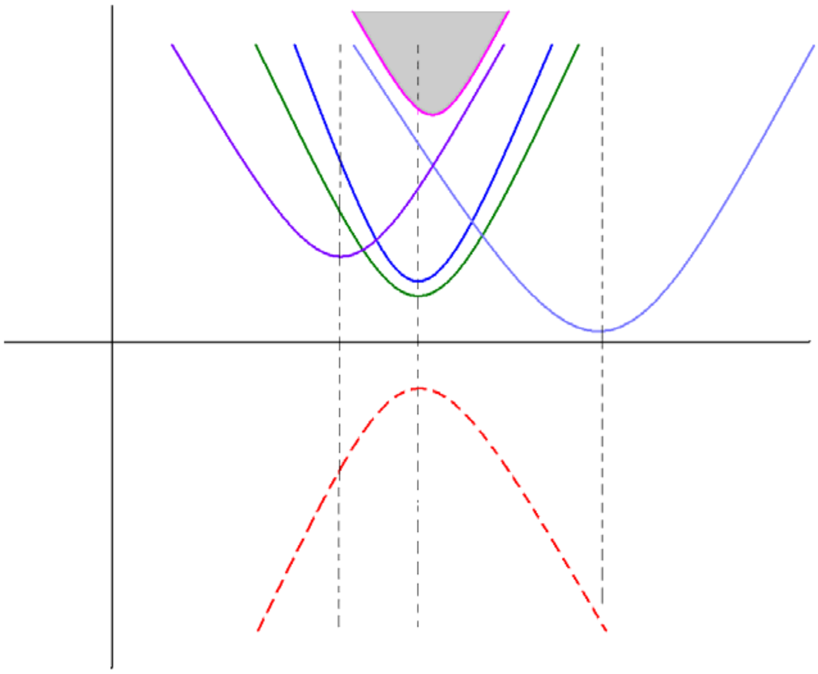


Figure 5. The graphs of the functions $f_{ij}(b)$, $\sum_{i<j} f_{ij}(b)$ when $\Delta_{ij} \leq 0, \Delta'_{ij} \leq 0, \forall i, j = \overline{1,n}, i < j$; the surface bounded by the graph of $\sum_{i<j} f_{ij}(b)$ is colored in gray

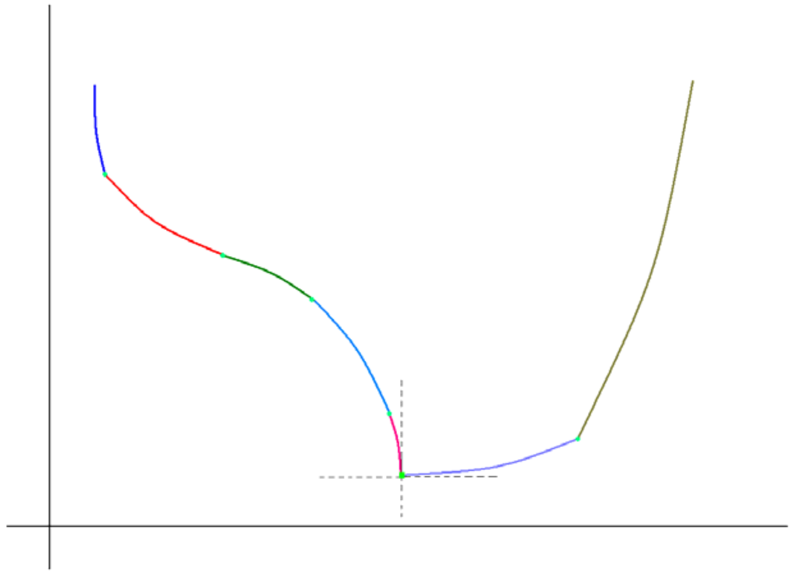


Figure 6. The graph of the function $\sum_{i<j} f_{ij}(b)$ when $\Delta_{ij}, \Delta'_{ij}\ (i, j = 1,\ldots,n; i < j)$ have random signs

2. Example

We test the method for fixed point $M = (1,2)$ and the fuzzy data:

$$X_1 = [3 - r, 3 + r];\ Y_1 = [5 + r, 6 + r]$$
$$X_2 = [4r, 5r];\ Y_2 = [4 + 7r, 10 + r]$$
$$X_3 = [3 + r, 7 - r];\ Y_3 = [9r, 8 + r]$$

Then

i	p_i	P_i	q_i	Q_i
1	$2 - r$	$2 + r$	$3 + r$	$4 + r$
2	$-1 + 4r$	$-1 + 5r$	$2 + 7r$	$8 + r$
3	$2 + 3r$	$6 - r$	$-2 + 9r$	$6 + r$

i	$\int_0^1 (p_i^2 + P_i^2)dr$	$\int_0^1 (q_i^2 + Q_i^2)dr$	$\int_0^1 (p_i q_i + P_i Q_i)dr$	$\int_0^1 (p_i Q_i + P_i q_i)dr$
1	9.66	32.66	16.50	15.5
2	0.58	108.66	21.00	20.00
3	43.33	55.33	46.66	36.00

and

$$f_{12}(b) = |-3b^2 - 9b + 76|,\ g_{12}(b) = |-3b^2 - 9b + 76|$$

$f_{13}(b) = |33.66b - 60.33b + 22.66|$, $g_{13}(b) = |33.66b^2 - 41b + 22.66|$

$f_{23}(b) = |36.66b^2 - 51.33b - 53.33|$, $g_{23}(b) = |36.66b^2 - 32b - 53.33|$.

Case 1: We search the minimizing points for the function $\sum_{\substack{i,j=\overline{1,3}\\ i<j}} f_{ij}(b)$.

Accordingly to the facts proved in the preceding chapters, namely Section 1, case 1.2, the set of feasible points is $S = S_2 = \{-6.75;-0.69;0.53;1.25;2.09;3.75\}$ (notice: for this example we obtain $S_1 = S_3 = \phi$) and the minimum is attained for $b = b^* = 2.09 > 0$. The complete solution is $(a,b) = (a^*, b^*) = (-0.9;2.09)$.

Case 2: We search the minimizing point for $\sum_{\substack{i,j=\overline{1,3}\\ i<j}} g_{ij}(b)$. Using the theoretical results obtained in Section 1, case 2.2, the set of feasible points is $S' = \{-6.75;-0.84;1.71;3.75\}$ and the minimum is attained in $b^{**} = 1.71 > 0$. This point don't fulfill the restriction concerning the sign of parameter b. From the appearance of the decreasing function $\sum_{\substack{i,j=\overline{1,3}\\ i<j}} g_{ij}(b)$ on $(-\infty,1.71)$ we conclude that the estimators for $b < 0$ are the real numbers
$\beta \in V_\varepsilon(0) \cap (-\infty,0)$ where $V_\varepsilon(0) = (-\varepsilon,\varepsilon)$ and ε depends by the desired threshold of error.
If $\beta_1, \beta_2 \in V_\varepsilon(0) \cap (-\infty,0)$, $|\beta_1| < |\beta_2|$ then β_1 is a better estimator.
At last, we have $\sum_{\substack{i,j=\overline{1,3}\\ i<j}} f_{ij}(2.09) = 87.44$ and $151.99 = \sum_{\substack{i,j=\overline{1,3}\\ i<j}} g_{ij}(0) < \sum_{\substack{i,j=\overline{1,3}\\ i<j}} g_{ij}(\beta)$ for all $\beta \in V_\varepsilon(0) \cap (-\infty,0)$. Thus $\sum_{\substack{i,j=\overline{1,3}\\ i<j}} f_{ij}(2.09) < \sum_{\substack{i,j=\overline{1,3}\\ i<j}} g_{ij}(\beta)$ and the final solution for this problem is $(a,b) = (-0.9;2.09)$.

3. Conclusions

From the preceding theoretical facts and numerical example we obtain the following conclusions:

1) For $b^* > 0$, $b^{**} < 0$, we evaluate $\sum_{i<j} |D_i^2 - D_j^2|$ and $\sum_{i<j} |d_i^2 - d_j^2|$.

If $\sum_{i<j} |D_i^2 - D_j^2| < \sum_{i<j} |d_i^2 - d_j^2|$ thus the solution is b^*.

If $\sum_{i<j} |D_i^2 - D_j^2| > \sum_{i<j} |d_i^2 - d_j^2|$ thus the solution is b^{**}.

2) For $b^* > 0$, $b^{**} > 0$ or $b^* < 0$, $b^{**} < 0$ it is necessary to make small supplementary calculations which implies the special properties of the functions $\sum_{\substack{i,j=\overline{1,3} \\ i<j}} f_{ij}(b)$, $\sum_{\substack{i,j=\overline{1,3} \\ i<j}} g_{ij}(b)$, as we shown in Section 2.

References

1. Arthanary T. S., Yadolah Dodge, **Mathematical Programming in Statistics,** John Wiley and Sons, New York, 1980
2. Wu Cong-Xin and Ma Ming: **Embedding problem of fuzzy number space: Part I**, Fuzzy Sets and Systems 44, 1991, p. 33-38
3. Wu Cong-Xin and Ma Ming: **Embedding problem of fuzzy number space: Part III**, Fuzzy Sets and Systems 46, 1992, p. 281-286
4. P. Diamond and P. Kloeden, **Metric spaces of fuzzy sets,** Fuzzy Sets and Systems 35, 1990, p. 241-249
5. P. Diamond, **Fuzzy least squares,** Inform. Sci. 46, 1988, p. 141-157
6. P. Diamond and P. Kloeden, **Metric spaces of fuzzy sets, Corrigendum,** Fuzzy Sets and Systems 45, 1992, p. 123
7. R. Goetschel, W. Voxman, **Elementary Calculus,** Fuzzy Sets and Systems 18, 1986, p. 31-43
8. I.M. Hammerbacher and R. R. Yager, **Predicting television revenues using fuzzy subsets,** TIMS Stud. Management Sci. 20, 1984, p. 469-477
9. A. Katsaras and D. B. Liu, **Fuzzy vector spaces and fuzzy topological vector spaces,** J. Math. Anal. Appl. 58, 1977, p. 135-146
10. Ma Ming, M. Friedman, A. Kandel, **General fuzzy least squares,** Fuzzy Sets and Systems 88, 1997, p. 107-118
11. C. V. Negoiţă and D. A. Ralescu, **Applications of Fuzzy Sets to Systems Analysis,** Wiley, New York, 1975
12. H. Prade, **Operations research with fuzzy data,** in: P. P. Wang and S. K. Chang, Eds., **"Fuzzy sets: Theory and Application to Policy Analysis and Information Systems",** plenum, New York, 1980, p. 115-169
13. M. L. Puri and D. A. Ralescu, **Differentials for fuzzy functions,** J. Math. Anal. Appl. 91, 1983, p. 552-558
14. H. Tanaka, H. Isibuchi and S. Yoshikawa, **Exponential possibility regression analysis,** Fuzzy Sets and Systems 69, 1995, p. 305-318
15. H. Tanaka, S. Uejima and K. Asai, **Linear regression analysis with fuzzy model,** IEEE Trans. Systems Man Cybernet SMC-12, 1982, p. 903-907
16. H. J. Zimmermann, **Fuzzy programming and linear programming with several objective functions,** Fuzzy Sets and Systems 1, 1978, p. 45-55
17. R. R. Yager, **Fuzzy prediction based upon regression models,** Inform. Sci. 26, 1982, p. 45-63

---

[1] Sudradjat Supian was born in Tasikmalaya, Indonesia on May 19 1958. After completing secondary school at SMA St. Maria Bandung in 1977 he continued his study at the Department of Mathematics, Universitas Padjadjaran Indonesia which he graduated in March 1983.
In 1986, he was accepted within the Department of Mathematics, Universitas Padjadjaran Indonesia as a junior lecturer. In July 1986 he went on with his study at Department of Industrial Engineering and Management Institute of Technology of Bandung (ITB) with the purpose of obtaining a master degree in Industrial Engineering and Management, which he succeeded in October 1989. In September 2003, he got the opportunity to pursue PhD degree at the University of Bucharest.
He started PhD in September 2004 at the Operations Research and Statistics Faculty of Mathematics and Computer Science, University of Bucharest under supervision of Prof. dr. Vasile Preda.

[2] Arthanary T. S., Yadolah Dodge, **Mathematical Programming in Statistics,** John Wiley and Sons, New York, 1980

[3] Ma Ming, M. Friedman, A. Kandel, **General fuzzy least squares,** Fuzzy Sets and Systems 88, 1997, p. 107-118

[4] R. Goetschel, W. Voxman, **Elementary Calculus,** Fuzzy Sets and Systems 18, 1986, p. 31-43

[5] Ma Ming, M. Friedman, A. Kandel, **General fuzzy least squares,** Fuzzy Sets and Systems 88, 1997, p. 107-118

[6] ibidem