



**The Journal of  
Mathematics and Computer Science**

Available online at

<http://www.TJMCS.com>

The Journal of Mathematics and Computer Science Vol .2 No.1 (2011) 100-110

## Numerical taxonomy analysis with trapezoidal fuzzy data

Ali mohammadi<sup>1,\*</sup>, javad shohani<sup>2</sup>, rajabali borzooei<sup>3</sup>

*1-amohammadi\_ie@yahoo.com*

*2- shohani@math.usbv.ac.ir.*

*3-borzooei@sbu.ac.ir*

Received: September 2010, Revised: December 2010

Online Publication: January 2011

### Abstract

Numerical taxonomy analysis is one of the best method of grading, classifying and comparing countries or different regions according to their development levels and modernity, that it can be used for different grading too. In this paper, the numerical taxonomy method with triangular fuzzy data that has been introduced by Mr. mohammadi and his colleagues in 2010, is expended to the method of numerical taxonomy with trapezoidal fuzzy data. So, if alternatives values, are place in diverse indicators of triangular fuzzy values, the output of expanded method of this paper will be the same as the numerical taxonomy method with triangular fuzzy data that has been introduced by Mr. mohammadi and his colleagues.

**Keywords:** numerical taxonomy, trapezoidal fuzzy number, development level.

---

<sup>1,\*</sup> Isalamic Azad University-Bojnourd Branch.

<sup>2</sup>, Mathematics PHD student, Faculty of Mathematics, University of Sistan & Baluchestan.

<sup>3</sup>, Pure Mathematics and Faculty Member of Shahid Behshti University.

### 1.Introduction

One of the methods of areas grading as development is taxonomy analyze. Taxonomy analyze is implemented in different sciences for different grading. Its special kind is numerical taxonomy that is used for different grading and grouping of townships (provinces and counties) as development and also is used as a locating technique for an industry or an economic production.

Generally, this method is an excellent one for grading, classifying, and comparing countries or different areas according to modernity and development level of theirs, and also is a method that divides a group in to similar subgroups and makes an acceptable indicator to survey rate of areas development in programmer's hand.

This method was suggested by Anderson in 1765 for the first time and was presented as a tool to classify and grade development between different nations by professor Helowing from economy superior school in Unesco in 1969. ([1],[ 2])

In this method probably some indicators used for grading various alternatives are linguistic value that are equal to trapezoidal fuzzy numbers as experts option, so, in this paper ,the taxonomy method with crisp data is expanded to the numerical taxonomy with trapezoidal fuzzy.

### 2.Definitions

**Definition 1-** Normalizing for interval numbers:

Suppose that decision making matrix is given as following one:

Indicators townships	1	2	...	m
1	$[X_{11}^l, X_{11}^r]$	$[X_{12}^l, X_{12}^r]$	...	$[X_{1m}^l, X_{1m}^r]$
2	$[X_{21}^l, X_{21}^r]$	$[X_{22}^l, X_{22}^r]$	...	$[X_{2m}^l, X_{2m}^r]$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
n	$[X_{n1}^l, X_{n1}^r]$	$[X_{n2}^l, X_{n2}^r]$	...	$[X_{nm}^l, X_{nm}^r]$

To normalize this matrix data we use following relation:

$$N_{ij} = \begin{cases} \frac{[x_{ij}^l, x_{ij}^r]}{\sum_{i=1}^n (x_{ij}^r - x_{ij}^l)} & \text{if } \sum_{i=1}^n x_{ij}^r \neq \sum_{i=1}^n x_{ij}^l \\ \frac{[x_{ij}^l, x_{ij}^r]}{\sum_{i=1}^n x_{ij}^l} & \text{if } \sum_{i=1}^n x_{ij}^r = \sum_{i=1}^n x_{ij}^l \end{cases} \tag{1}$$

Where  $[x_{ij}^l, x_{ij}^r]$ , is value of the ij th element of data matrix (value of the j for the township i ).

In this case, the matrix N witch it's elements are made up of the interval numbers  $N_{ij}$  , is normalized matrix of basic decision making matrix.[3]

### Definition 2-

Suppose that R is a relation in the set A, so, we call R associate relation in A whenever R is transitive And the function is trinity principal it means, for each x and y belonged to the set A, always one and just one of the relations  $x=y$ ,  $x>y$ , or  $x<y$  is true .so, we call (A, R) as the order set of A with the associate Relation R .[4]

### Definition 3-

We define the greater relation R that is expressed by the symbol  $\succ$  below:

$$\succ : ([a,b],[c,d]) \rightarrow \begin{cases} [a,b] & \text{if : } \frac{a+b}{2} > \frac{c+d}{2} \\ [c,d] & \text{if : } \frac{a+b}{2} < \frac{c+d}{2} \end{cases} \quad (2)$$

Attention: the relation  $\succ$ , according to the upper definition of associate relation, is a associate relation.[3]

**Definition 4-**

We define  $\dot{Y}$  for defining the distance [a,b] from [c,d] as following :

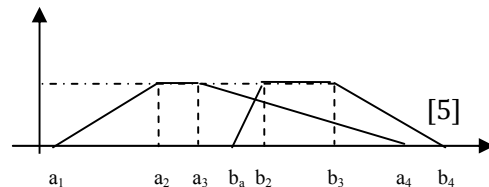
$$\dot{Y}([a,b],[c,d]) = \begin{cases} \frac{a+b-c-d}{2} & \text{if : } [a,b] \succ [c,d] \\ \frac{c+d-a-b}{2} & \text{if : } [c,d] \succ [a,b] \end{cases} \quad (3)$$

Where  $\succ$  is the greater relation that is defined above.[3]

**Definition 5 -**The distance of two fuzzy numbers:

If  $\tilde{A} = (a_1, a_2, a_3, a_4)$ ,  $\tilde{B} = (b_1, b_2, b_3, b_4)$  are two trapezoidal fuzzy numbers, then the distance of  $\tilde{B}$  from  $\tilde{A}$  is achieved by following relation:

$$S(\tilde{B}, \tilde{A}) = \frac{(b_1 + b_2 + b_3 + b_4) - (a_1 + a_2 + a_3 + a_4)}{4} \quad (4)$$



**Definition 6 -**The value of anti symmetric of a trapezoidal fuzzy number:

Suppose that  $\tilde{A} = (a_1, a_2, a_3, a_4)$ , is a trapezoidal fuzzy number ,the value of anti symmetric in this fuzzy trapezoidal number is computed below:

$$d_A = ||a_1 - a_2| - |a_3 - a_4|| \quad (5)$$

**3.The method of numerical taxonomy with interval data:**

first, suppose that we want to grade and classify n cities according to m indicators that are defined by interval values and clarify each city development.

**First stage- making data matrix:**

At this step we put township in rows and indicators in columns of matrix that we receive a matrix with n rows and m columns finally which is shown below:

Indicators \ townships	1	2	...	m
1	$[X_{11}^l, X_{11}^r]$	$[X_{12}^l, X_{12}^r]$	...	$[X_{1m}^l, X_{1m}^r]$
2	$[X_{21}^l, X_{21}^r]$	$[X_{22}^l, X_{22}^r]$	...	$[X_{2m}^l, X_{2m}^r]$
.	.	.	.	.

.	.	.	.	.
<b>n</b>	[X <sub>n1</sub> <sup>l</sup> , X <sub>n1</sub> <sup>r</sup> ]	[X <sub>n2</sub> <sup>l</sup> , X <sub>n2</sub> <sup>r</sup> ]	...	[X <sub>nm</sub> <sup>l</sup> , X <sub>nm</sub> <sup>r</sup> ]

**Second stage-making normal matrix:**

As different indicators are measured by different units, so, to omit different units and substitute united indicator, we will do what we have below:

$$N_{ij} = \frac{[x_{ij}^l, x_{ij}^r]}{\sum_{i=1}^n (x_{ij}^r - x_{ij}^l)} \tag{6}$$

Where [x<sub>ij</sub><sup>l</sup>, x<sub>ij</sub><sup>r</sup>] is the value of ij th element of data matrix (the value of the indicator j for the township i).

In this case, the matrix N which is made up of N<sub>ij</sub> s as its elements in normalized matrix.

**Third stage - computing the distance between townships:**

At this step, to compute the distance of the township a from b we will use following relation:

$$( N_{aj}, N_{bj} ) \tag{7}$$

$$d_{ab} = \sum_{j=1}^m \dot{Y}$$

Where N<sub>aj</sub> is the value of j th indicator for the township a in normalized matrix. And d<sub>ab</sub> is the value of the distance between a, b township that following relations are obvious:

$$d_{ab} = d_{ba} \cdot d_{bb} = d_{aa} = 0$$

**Forth stage- assigning the shortest distances:**

At this step first we try to make the matrix which clarifies distances between townships ( D ), then , in each raw , the lowest value , we write it in separated column( column d ) . if a and b have the shortest distance , so there is the most similarity between a and b( a , the township located in raw , b , the township located in column).

In expression, the township b is called model or "pattern" for the township a and the township a, is called "shadow" township.

**Fifth stage- drawing chart:**

We connect townships which have the most similarity between their characteristics (townships which have the shortest distances with a vector). direction of the vector is the same as the pattern or model township and it's length equals the length of the shortest distance between the two townships .first ,it is probable that all townships are not connected according to the shortest distances, so here , we will use the shortest distance of second level( the second shortest distance) and we connect all townships in a chart.

**Sixth stage- finding Compatible Township:**

In order to find the compatible township, we compute upper limit distances (+d) and lower limit distances (-d) according to the following relations:

$$d+ = \bar{d} + 2sd \tag{8} \quad d- = \bar{d} - 2sd \tag{9}$$

What the average of the shortest distance equals  $\bar{d} = \sum_{i=1}^n \frac{di}{n}$  (10),  $sd = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n}}$  (11) is standard

variation. Compatible townships are ones which their distances are between d+ and d- limits and take place in the same group. Two townships that their combined distance is more than d+ , it is noted that , there is not any similarity as development between then, for instance there is not similarity between ZABOL and TEHRAN as development . So we can separate all connections which have more length than

upper limit (or upper critical distance) because length of these distance is more than the one can be computed as a united chart. Also we can separate all connections that have less length than lower limit (or lower critical distance) because the length of these distances are less that the one that shows any difference between two townships, on the other hand, in such states like this, two townships are as similar as each other that can't be considered as two different townships, it means their models and shadows are the same. The chart that is made by omitting connections between d+ and d- is optimum.

**Seventh stage- grading township as development:**

If at the sixth stage all townships don't take place in one compatible group. We will make data belonging to the matrix X (data matrix) for compatible townships and then we compute normal matrix again and we write it (if all townships are compatible, normal matrix is the old one).

We suppose to continue this loop till all townships which get the last stage are in one compatible group , so , at the last computed normal matrix , in each column we find the least value if it is increased function of that indicator ( the interval that it's center ( summation of first and end of the interval divide 2 ) is the most value) and if it is decreased function of that indicator , we will find the least value ( the interval that it's central value is the least ones and we write it in a raw which is named I and we call it ideal township value. after finding ideal value for all indicators , we get development pattern for each township from following relation.

$$C_{io} = \sum_{j=1}^m \dot{Y}(x_{ij}, I_{oj}) \tag{12}$$

$C_{io}$  is development pattern that shows the distance of the township I from the ideal township ,  $I_{oj}$  is value of the j th indicator of the ideal township and  $x_{ij}$  is the value of the indicator j of the township i.

the less  $C_{io}$  will be, the more developed .

**Eighth stage- computing townships development level:**

If we rate the township I's development level with  $f_i$  , then , the value of  $f_i$  is computed with the following relations:

$$f_i = \frac{C_{io}}{C_o} = \frac{\text{Pattern development}}{\text{Maximum of pattern development}} \tag{13}$$

$$C_o = C_{io} + 2S_{io} \tag{14}$$

$$S_{io} = \sqrt{\sum_{i=1}^n \frac{(C_{io} - \bar{C}_{io})^2}{n}} \tag{15}$$

$$\bar{C}_{io} = \frac{\sum_{i=1}^n C_{io}}{n} \tag{16}, \text{ where } C_o \text{ is the value of upper limit of development pattern for the township i.}$$

Development level is between zero and one ( $0 \leq f_i \leq 1$ ) .the closer  $f_i$  is to zero, the more development the township is and the closer to one, the more undeveloped the township is considered. After computing  $C_{io}$  and  $f_i$  , we arrange townships according to development level and we put them in the chart like following one:

Row	Township	$C_{io}$	$C_o$	$f_i$
1				

2				
.				
.				
n				

So, township are graded according to considered indicators (m indicators) and the development level of is assigned. [3]

**4. Suggested method taxonomy for trapezoidal fuzzy data**

To solve the problems which have indicators with trapezoidal fuzzy value , first we choose a cutting surface ( $\alpha$  ) separately for each fuzzy indicator , then we cut all trapezoidal fuzzy numbers related to different choices in the indicator , in the surface  $\alpha$ . then ,easily it can be possible to solve the problem by getting help from numerical taxonomy analyze method for interval data in such problems like these choosing the cutting surface ( $\alpha$ ) is very critical due to different intervals choices caused by different  $\alpha$ 's surfaces that may cause different conclusions.

With considering such importance in choosing of the cutting surface ( $\alpha$ ), in this article, for fuzzy indicators according to two following indicators, the suitable cutting surface ( $\alpha$ ) is chosen.

**Indicator1- attention value of membership function symmetric of a trapezoidal fuzzy number:**

As you know if a trapezoidal fuzzy number doesn't have symmetric membership function, it can't introduce an equal interval to this number easily. So, to solve this problem we introduce an equal symmetric trapezoidal fuzzy number to this number according to the distance of this number from zero like below:

Suppose that  $\tilde{A} = (a_1, a_2, a_3, a_4)$  is an in anti symmetric trapezoidal fuzzy number ( it mean  $d_A \neq 0$  ), the symmetric trapezoidal fuzzy number  $\tilde{A}' = (a'_1, a'_2, a'_3, a'_4)$  is defined as below:

$$s = \frac{a_1 + a_2 + a_3 + a_4}{4} \tag{17}$$

$$l = \min\{|a_2 - a_1|, |a_4 - a_3|\} \tag{18}$$

$$t = \frac{a_3 - a_2}{2} \tag{19}$$

$$a'_1 = s - l - t, a'_2 = s - t, a'_3 = s + t, a'_4 = s + l + t$$

**Indicator2- attention to membership grade rate of support of a trapezoidal fuzzy number:**

The more important, upper membership grade is, the closer value a should be chosen to one, otherwise, the value of a should be near zero.

Now according to indicators 1 and 2, with following relation we can obtain equivalent interval to the trapezoidal fuzzy number  $\tilde{A} = (a_1, a_2, a_3, a_4)$ :

$$A_\alpha = [x^l, x^r] = [a_1 + \alpha(a_2 - a_1), a_4 - \alpha(a_4 - a_3)] \tag{20}$$

**5. Example:** we intend to grade three township a,b,c according to three indicators . Suppose that data chart of these township according to indicators is below. In this state we operate like below:

**Stage 1 - making data matrix:**

Indicator \ Townships	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>
A	(-1, 1, 2, 3)	(1, 2, 2.5, 3)	(-5, 1, 1.5, 2)
B	(-1, 1, 2, 6)	(0, 2, 3, 7)	(80, 90, 91, 100)
C	(0, 2, 3, 4)	(1, 2, 3, 4)	(0, 0, 1, 2)

**Stage 2-** making normalized matrix:

**First step:** figuring out anti symmetric trapezoidal fuzzy number equal to each in symmetric trapezoidal fuzzy number in data matrix.

At this step first, we compute the value of anti symmetric for all of data matrix values (the first stage matrix), then we introduce fuzzy number that is anti symmetric for each value which is equal to it. For instant, the value of the indicator I<sub>1</sub> is anti symmetric for township b:

So the equal symmetric fuzzy number to this number is obtained like below:

$$l = \min \{ |a_2 - a_1|, |a_4 - a_3| \} = \min \{ |1 - (-1)|, |6 - 2| \} = 2$$

$$s = \frac{a_1 + a_2 + a_3 + a_4}{4} = \frac{(-1) + 1 + 2 + 6}{4} = \frac{8}{4} = 2$$

$$t = \frac{a_3 - a_2}{2} = \frac{1}{2} = 0.5$$

$$a'_1 = s - l - t = -0.5, a'_2 = s - t = 1.5, a'_3 = s + t = 2.5, a'_4 = s + l + t = 4.5$$

Other cases are obtained in similar way.

Indicator \ Townships	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>
A	(0.25, 0.75, 1.75, 2.75)	(1.38, 1.88, 2.38, 2.88)	(-0.88, -0.38, 0.12, 0.62)
B	(-0.5, 1.5, 2.5, 4.5)	(0.5, 2.5, 3.5, 5.5)	(80.75, 89.75, 90.75, 99.75)
C	(0.75, 1.75, 2.75, 3.75)	(1, 2, 3, 4)	(0.25, 0.25, 1.25, 1.25)

**Second step-**

importance rate of membership grade (assigning cutting surface of  $\alpha_i$  s):

Suppose that  $\alpha_{I_1} = 0.2$  ,  $\alpha_{I_2} = 0.5$  ,  $\alpha_{I_3} = 0.8$

**Third step-** assigning equivalent interval numbers to fuzzy numbers of data matrix:

For example for the township C in the indicator I<sub>1</sub>:

(1, 2, 3, 4) is equivalent to:

$$A_\alpha = [x^l, x^r] = [a_1 + \alpha(a_2 - a_1), a_4 - \alpha(a_4 - a_3)] = [1 + 0.5(2 - 1), 4 - 0.5(4 - 3)] = [1.5, 3.5]$$

And other case is obtained in similar way .Results of this stage computations are brought in the matrix below:

Indicator \ Townships	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>
A	[0.5, 2.25]	[1.63, 2.63]	[-0.63, 0.37]
B	[0.5, 3.5]	[1.5, 4.5]	[85.25, 95.25]

assigning

c	[1.25,3.25]	[1.5,3.5]	[0.25,1.25]
---	-------------	-----------	-------------

**Fourth step-** normalizing third step matrix:

As a sample, for the indicator  $I_2$  in the township A :

$$N_{ij} = \frac{[x_{ij}^1, x_{ij}^2]}{\sum_{i=1}^n (x_{ij}^2 - x_{ij}^1)} = \frac{[1.63, 2.63]}{(2.63 - 1.63) + (4.5 - 1.5) + (3.5 - 1.5)} = [0.27, 0.44]$$

And other cases are obtained in similar way computations. Results of this stage are brought below:

Indicator \ Townships	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>
A	[0.07,0.33]	[0.27,0.44]	[-0.05,0.03]
B	[0.07,0.52]	[0.25,0.75]	[7.1,15.88]
C	[0.19,0.48]	[0.25,0.58]	[0.02,0.1]

**Third stage:** computation of townships distances

As an example, the distance of a, b townships are computed below:

$$d_{ab} = \sum_{j=1}^m \dot{Y}(z_{aj}, z_{bj}) = \dot{Y}([0.07, 0.33], [0.07, 0.52]) + \dot{Y}([0.27, 0.44], [0.25, 0.75]) +$$

$$\dot{Y}([-0.05, 0.03], [7.1, 15.88]) = 11.74$$

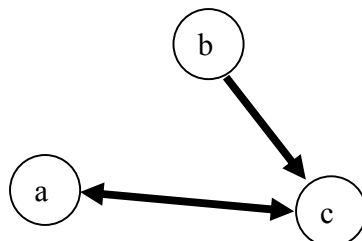
And other values are achieved in the same way. Computations results of this stage are brought below:

$$d_{bc} = 11.55 \quad , \quad d_{ac} = 0.27$$

**Fourth stage-** computing of the shortest:

Indicator \ Townships	A	B	C	D
A	0	11.74	0.27	0.27
B	11.74	0	11.55	11.55
C	0.27	11.55	0	0.27

**Fifth stage-** drawing chart:





**Sixth stage-** finding compatible townships:

$$\bar{d} = \sum_{i=1}^n \frac{d_i}{n} = \frac{0.27 + 11.55 + 0.27}{3} = 4.03 \quad , \quad sd = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n}} = 5.32$$

$$d- = \bar{d} - 2sd = -6.61 \quad , \quad d+ = \bar{d} + 2sd = 14.67$$

Since all of  $d_i$ 's are between the values d- and d+, so, all of townships are compatible.

Seventh stage- grading of townships as development:

At this step, first we figure out the ideal townships values brought in the following chart in the row which is related to the township I:

Indicator Townships	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>
A	[0.07,0.33]	[0.27,0.44]	[-0.05,0.03]
B	[0.07,0.52]	[0.25,0.75]	[7.1,15.88]
C	[0.19,0.48]	[0.25,0.58]	[0.02,0.1]
I	[0.19,0.48]	[0.25,0.75]	[7.1,15.88]

Then for each township we try to compute development pattern:

$$C_{io_a} = \check{Y} ([0.07,0.33], [0.19,0.48]) + \check{Y} ([0.27,0.44], [0.25,0.75]) + \check{Y} ([-0.05,0.03], [7.1,15.88]) = 11.78$$

And other are achieved the same.

$$C_{io_b} = 0.04 \quad , \quad C_{io_c} = 11.51$$

So, the township b has the superior preference and the townships C and A are in next preference respectively.

**Eighth sage** - computing of townships development grade:

$$\bar{C}_{io} = \frac{C_{io_a} + C_{io_b} + C_{io_c}}{3} = 7.78$$

$$S_{io} = \sqrt{\frac{\sum_{i=1}^n (C_{io} - \bar{C}_{io})^2}{n}} = 5.47$$

Now, we compute  $f_i$  from the following relation:

For instance, for the township a:

$$f_i = \frac{C_{io_a}}{C_o} = \frac{11.78}{22.72} = 0.52 \text{ , } C_o = C_{io_a} + 2S_{io} = 11.78 + 10.94 = 22.72$$

It is computed for other cases the same and its conclusion are brought in following chart:

Township	$C_{io}$	$C_o$	$f_i$	grade
A	11.78	22.72	0.518	third
B	0.04	10.98	0.0036	first
C	11.51	22.45	0.512	second

Finally the township B is in superior preference and the townships C and A are respectively in the next preferences.

## 6. Conclusion:

In this paper, the numerical taxonomy method is expanded to numerical taxonomy with trapezoidal fuzzy data. Using the proposed method's results in real problems considers its efficiency and rationality. One of considerable results captured from this paper is that if trapezoidal fuzzy numbers are changed to triangular fuzzy numbers, the proposed model will be altered to the model which has been introduced by Mr Mohammadi and his colleagues. [3]

Of course the results of using the proposed model in this paper can be affected by diverse factors that the most important ones are such as:

- 1- Correctness scope of fuzzy numbers estimations of the indicators for different choices.
- 2- Attention to membership grade importance scope in choosing cutting surface.
- 3- In symmetric scope of trapezoidal fuzzy numbers that is due to the first estimation.
- 4- Correct choice of the cutting surface a.
- 5- Made estimations in interval taxonomy method.

## References:

- 1- ziary,K.A., " principles and methods of regional planning".
- 2- moaiery ,F., " assigning strategy of kordestan province industrial development as respect performance view".
- 3- mohammadi,a., shohani,j.,jahanshahi,h., "numerical taxonomy analysis with triangular fuzzy data",the 10th conference of fuzzy systems,iran,shahid beheshti university,2010.
- 4- jamali,A.R. , " public topology ", payam- e – nour university.
- 5- jahanshahi,H., "the course pamphlet of fuzzy set theory and it' usages in decision making ", emam hossain university.
- 6- bidabady,B., " taxonomy analyze", azerbaiejan budjest and planning organization .1984.

- 7- *azerbaiejan budjest and planning organization, "introduction of numerical taxonomy simply , a method for grouping / and assigning grade and aims of development".1984.*
- 8- *basiry,M.H., and nabi jurdy,F.S. , " economical- strategic grading of Iranian developing mine material, with taxonomy grading method to develop investment in iran".*
- 9- *nourani,M., " fuzzy numbers and their comparison methods ",am thesis, kerman university ,1999.*
- 10- *keler,J.j, keler,u.,s., ann,b.u.," fuzzy set theory,principles and applications ", translation of Dr mohammad hossain fazel zarandy*
- 11- *Zimmermann,H.J., "Fuzzy Set Theory and its Applications" ,2nd edn. Kluwer Academic Publishers ,Boston/Dordrecht/London, 1991.*