FUZZY TIME SERIES ANALYSIS AND PREDICTION PROCESS

Eva Michaliková
Dep. of Economy and Management
Fak. of Civil Engineering
Slovac Technical University
Radlinského 11
813 68 Bratislava
Czechoslovakia

Abstract :

The purpose of this paper is to present some ideas on the appliеation of fuzzy sets to prediction process.

First, the deterministic analysis of fuzzy time series is given. Then, the basic principles of fuzzy time series prediction are introduced. Finally, an example illustrates the usefulness of the basic ideas and techniques.

Keywords : Fuzzy time series analysis, Periodogram, Prediction theory

I. Introduction

The time series analysis, modeling and prediction are one of the most progresive developing fields of mathematics where the outputs have wide and quick applicability using computers /see Box - Jenkins (2)/ .

Generally, all time series analysis methods presented till now deal with the time series which values are the real numbers. In practice, there are often situations when the value X_t are described linquistically (foe example - the potatoes crop in 1983 can be characterized as "good"). Then one nature quastion arises - How to analyse such time series ?
Two tasks should be underlined to get the answer :
- the time series analysis and
- the formalization of linquistic expressions

Zadeh (1) introduces concept of linquistic variables using fuzzy sets theory. Therefor in this paper we analyse the linquistic expressions as the fuzzy sets. The goal of our contribution is to explain fuzzy time series analysis with linquistic variables and it's application to prediction process.

Since, the topic represents very large and complicate problem, we restrict ourselves to deal only with deterministic part of fuzzy time series, mostly from the fuzzy data processing methodics point of view. Since, there are two the most often used methods for deterministic part verification of the crisps time series - LEAST SQUARE METHOD and PERIODOGRAM, we present the second one - Periodogram in section II.

Section III of the paper is concerned with fuzzy time series analysis and prediction. Finally, the example illustrates the deterministic analysis of concrete fuzzy time series with it's prediction.

II. PERIODOGRAM

Definition : Assume that we have the finite series of random quantities $Y_1, \dots Y_n$. We define the function $I(\lambda)$ as

$$I(\lambda) = \frac{1}{2\pi N} \left| \sum_{t=1}^{N} Y_t \cdot e^{-it\lambda} \right|^2 \qquad (1)$$

The function $I(\lambda)$ is called PERIODOGRAM of series $Y_1, \dots Y_n$.

Assume that $$Y_t = \sum_{k=1}^{p} a_k \cdot e^{it\lambda k} + X_t \qquad t = 1, 2 \dots \quad (2)$$

where $a_1, \dots a_k$ are non zero constants

$\lambda_1, \dots \lambda_p$ are independent numbers from interval $(\pi, -\pi)$ and $\{X_t\}$ is sequence of noncorrelated random quantities with sero mean value and the same dispersion $\sigma^2 > 0$.

The constants λ_k , k = 1, ... p express the frequencies occured in analysed data. In N - data situation the possible (identificationed) frequencies are

$$\lambda_i = \frac{2\pi i}{N} \qquad i = 1, \dots \left[\frac{N}{2}\right]$$

where $[x]$ denotes integer nearest to x

then the relevant periods are $\left[\frac{N}{i}\right]$.

According Box - Jenkins (2), and Anděl (3) the periodogram values in our "interesting" points are

$$I(\lambda_i) = \frac{N}{2\pi} \left(a_i^2 + b_i^2 \right) \qquad (3)$$

where

$$a_i = 2/N \sum_{t=1}^{N} Y_t \cos \frac{2\pi i}{N} t \qquad (4)$$

$$b_i = 2/N \sum_{t=1}^{N} Y_t \sin \frac{2\pi i}{N} t \qquad (5)$$

for i = 1, 2 ...

III. FUZZY TIME SERIES ANALYSIS AND PREDICTION

As mentioned about, using periodogram we analyse the time series periodicity.

The time series periodical member is estimated as

$$Y(t) = a_0 + \sum_{i=1}^{q} \left(a_i \cos \frac{2\pi i t}{N} + b_i \sin \frac{2\pi i t}{N} \right) \qquad (6)$$

where $$a_0 = \frac{1}{N} \sum_{t=1}^{N} Y_t$$

then the prediction at time N+1 is

$$Y(N+1) = a_0 + \sum_{i=1}^{q} \left(a_i \cos \frac{2\pi i (N+1)}{N} + b_i \sin \frac{2\pi i (N+1)}{N} \right) =$$

= Y(1)

Before the fuzzy time series is analysed the conjunction of fuzzy sets has to be made by their multiplying according Zadeh (1)

$$A.B = \sum_{u_i \, U} \mu_A(u_i).\mu_B(u_i) \; / \; u_i$$

Owing to the basic set gets immense expansion and some membership functions attain the minimal values. Setting the α- level which states the minimal permissible membership function value this undesirable effect can be prevented.

Applying the periodogram for N values which satisfied the α- level we can obtain the relevant prediction with its membership function. The total prediction is then fuzzy set.

Since, some prediction values can occur many times, their association in the sence of fuzzy values disjunction is needed. For determining membership function we use fuzzy disjunction according Zadeh (1)

$$A \oplus B = (A'. B') = \sum_{u_i \, U} (1 - (1-\mu_A(u_i))(1-\mu_B(u_i)))/ \; u_i =$$

$$= \sum_{u_i \, U} (\mu_A(u_i) + \mu_B(u_i) - \mu_A(u_i)\mu_B(u_i))/ \; u_i$$

IV. ILLUSTRATIVE EXAMPLE

A simple example is given to illustrate the fuzzy time series analysis and prediction.

Example : The potatoes crop in Czechoslovakia in 1983- 87 was linquistically described as following

1983	1984	1985	1986	1987
good	very good	bad	very good	good

Determine the potatoes crop prediction for 1988 year if

α - level is 0.1 . Assume that the trend is equal zero.

Substituing the fuzzy sets representing linquistic terms we get :

good (0.3, 25), (1., 30), (0.3, 35)

very good (0.3, 35), (1., 40), (1., 45)

bad (1., 15), (0.9, 20), (0.3, 25)

Let the following values hold for setting α- level

(0.3, 25), (1., 40), (1., 15), (1., 40), (1., 30)

and the total membership function μ_i = 0.3 (after multiplying)

according (4) and (5) we get

$$a_0 = \frac{1}{5} \sum_{t=1}^{5} Y_t = \frac{1}{5} (25+40+15+40+30) = 30$$

$$a_1 = \frac{2}{5} \sum_{t=1}^{5} Y_t . \cos \frac{2\pi}{5} t = 2.236$$

$$b_1 = \frac{2}{5} \sum_{t=1}^{5} Y_t . \sin \frac{2\pi}{5} t = 0.172$$

$$a_2 = \frac{2}{5} \sum_{t=1}^{5} Y_t . \cos \frac{4\pi}{5} t = -2.236$$

$$b_2 = \frac{2}{5} \sum_{t=1}^{5} Y_t . \sin \frac{4\pi}{5} t = -13.037$$

then for prediction we get notation

$$f_i = 30 + 2.236\cos \frac{2\pi t}{5} + 0.172\sin \frac{2\pi t}{5} - 2.236\cos \frac{4\pi t}{5} - 13.037\sin \frac{4\pi t}{5}$$

where t=6 because it's prediction for 6th year.

Solwing the equation just about the prediction for 1988 is 25 with membership function 0.3 .

The outputs from computer after solwing all conjunctions shows the Table 1 .

Associated membership function

$$\mu(25) = 1 - (1-0.3)^4(1-0.27)^4 = 0.93$$

$$\mu(30) = 1 - (1-0.3)^{16}(1-0.27)^{12}(1-1)^3(1-0.9)^4 = 1.0$$

$\mu(35) = 1 - (1-0.3)^4(1-0.27)^3 = 0.91$

u_i	partial (u_i)	(u_i)
25	0.3,0.2,0.27,0.27,0.3,0.3,0.27,0.27	0.93
30	0.3,0.3,0.27,0.3,0.3,1.0,0.3,0.3,1.0,0.3, 0.27,0.27,0.9,0.27,0.27,0.9,0.27,0.3,0.3, 0.3,0.3,1.0,0.3,0.3,1.0,0.3,0.27,0.27,0.4 0.9,0.27,0.27,0.9,0.27,0.3,0.3	1.0
35	0.3,0.3,0.27,0.27,0.3,0.3,0.27	0.91

TABLE 1 : Primary fuzzy prediction

The total prediction for 1988 is :(0.93, 25),(1.0, 30), (0.91, 35), so between good and bad. We can say that the expected potatoes crop in 1988 will be average.

V. CONCLUSION

In this paper the basic principles have been proposed to deal with prediction theory using fuzzy sets.
The illustrative example presents the possibility of the linquistic approach to this process.

In the case of suprem disjunction analogy the prediction for 1988 year would be (0.3, 25),(1.0, 30),(0.3, 35) which means more less good potatoes crop.

NOTE : In a crisps case we use the first member for the prediction of N+1 member.

In a fuzzy case the situation is a little bit different - the possible values of the first member don't change, but the membership function values will be in prediction different.

The biger contribution of the ideas presented about seems to be the possibility to unreveal hidden periodicities and their importance would be then represented by fuzzy numbers.

This paper is not a final version, it only shows the outline of the usefulness of the new ideas. However, there are a lots of problems to be solved in both lines - in real modeling process and fuzzy sets theory as well. Notice, the paper puts both of them togrther.

REFERENCES

(1) L.A. Zadeh : The concept of a linquistic variable and its application to approximate reasoning, 1965

(2) G.E.P. Box - G.M. Jenkins : Time series analysis forecasting and control, Holden Day, San Francisco, 1970

(3) J. Anděl : Statistic time series analysis, SNTL Praque, 1976

(4) R. Mesiar, B. Michaliková, I. Rakovská : Fuzzy time series analysis, Student scientific project, Bratislava, 1983