The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Structural Change Point Detection by Sequential Probability Ratio Test and Chow Test

*Tetsuo Hattori, *Katsunori Takeda, **Izumi Tetsuya, ***Hiromichi Kawano

**Graduate School of Engineering, Kagawa University / 2217-20 Hayashi, Takamatsu City, Kagawa 761-0396, Japan*
***Micro Technica Co., Ltd., / Yamagami BLD. 3-12-2 Higashi-ikebukuro, Toshima-ku, Tokyo 170-0013, Japan*
**** NTT Advanced Technology s, / Musashino-shi Nakamachi 19-18, Tokyo 180-0006, Japan*
(Tel : 81-87-864-2221; Fax : 81-87-864-2262)
(hattori@eng.kagawa-u.ac.jp)

Abstract: Time series analysis is used in various fields. The problem of predicting time series can be classified into three. The first problem is how to make a prediction model, that adequately represents the characteristics of the past time series data. The second problem is how to correctly and promptly detect the structural change, when the estimated prediction model does not meet the data. The third problem is how to quickly find the new prediction model after the change. This paper focuses on the second problem and proposes a method based on a probability ratio test. This paper also shows some experimental results comparing with a conventional method, Chow test, and presents the effectiveness.

Keywords: Time series analysis, sequential probability ratio test (SPRT), Chow Test

I. INTRODUCTION

Time series analysis is used in various fields such as not only in economics but also in pattern recognition, where the analysis is used for contour (or shape) analysis and sound signal processing, etc.

The problem of time series can be classified into three types, in a practical sense. The first problem is how to generate a prediction model that adequately represents the characteristics of the early time series data. For this purpose, various kinds of model such as Box-Jenkins method [1], and Kalman Filter [2], neuro network [3], fuzzy model [4], Chaos model [5] have been proposed. Also, model selection criteria have been proposed such as AIC [6], Cp [7], CV [8] and CMV [9]. The second problem is how to detect the structural change of the time series, as soon and correctly as possible, when the estimated prediction model does not meet the real data. The third problem is, after the change detection, how to quickly estimate the time series model again.

This paper focuses on the second problem and proposes a novel method, that is based on a sequential probability ratio test (SPRT)[10], [11], for quick detection of the structural change point in time series. And, this paper also describes the features of the method from numerical experimentation results and also shows its effectiveness in comparison with the Chow Test [12].

II. Structural change point detection

For the early structural change detection problem, we propose an application of Sequential Probability Ratio Test (SPRT) that has been mainly used in the field of quality control.

1. SPRT

The SPRT is used for testing a null hypothesis H_0 (e.g. the quality is under pre-specified limit 1%) against hypothesis H_1 (e.g. the quality is over pre-specified limit 1%). And it is defined as follows:

At each stage of successive events $Z_1, Z_2, \cdots Z_i$ that are respectively corresponding to observed time series data, the probability ratio λ_i is computed.

$$\lambda_i = \frac{P(Z_1 \mid \mathbf{H}_1) \cdot P(Z_2 \mid \mathbf{H}_1) \cdots P(Z_i \mid \mathbf{H}_1)}{P(Z_1 \mid \mathbf{H}_0) \cdot P(Z_2 \mid \mathbf{H}_0) \cdots P(Z_i \mid \mathbf{H}_0)} \quad (1)$$

where $P(Z \mid H_0)$ denotes the distribution of Z when H_0 is true and $P(Z \mid H_1)$ denotes the distribution of Z when H_1 is true.

Two positive constants C_1 and C_2 ($C_1 < C_2$) are chosen. If $C_1 < \lambda_i < C_2$, the experiment is continued by taking an additional observation. If $C_2 < \lambda_i$, the process is terminated with the rejection of H_0 (acceptance of H_1). If $\lambda_i < C_1$, the process is terminated with the acceptance of H_0.

$$C_1 = \frac{\beta}{1-\alpha} \quad C_2 = \frac{1-\beta}{\alpha} \qquad (2)$$

where α means type I error (reject a true null hypothesis), and β means type II error (accept a null hypothesis as true one when it is actually false).

2. Application to change point detection

For simplicity, we describe our proposed method for structural change detection of time series, by taking an example. We assume that the time series is generated in the following equation as a function of time t.

$$y_t = \beta_1 \cdot t + \beta_0 + \varepsilon \qquad (3)$$

where $\varepsilon \approx N(0,\sigma^2)$, i.e., the error ε is a random variable subject to the Normal Distribution with the average 0 and the variation σ^2.

Then we also assume that the structural change occurred at the time point tc* (what we call change point), involving the change of equation coefficients β_1, β_0. Concretely speaking, such data is generated by the following equations.

$$y_t = \beta_{11} \cdot t + \beta_{10} + \varepsilon \quad (t \le tc^*) \qquad (4)$$

$$y_t = \beta_{21} \cdot t + \beta_{20} + \varepsilon \quad (tc \le t^*) \qquad (5)$$

where tc* is called a change point in the structural change.

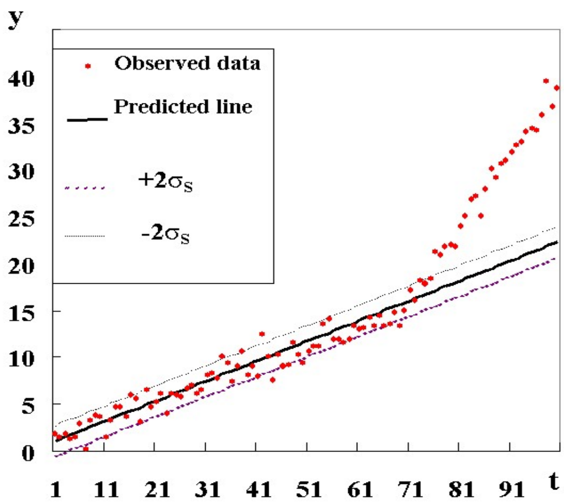

Fig.1. Example of time series data where the change point tc* =70.

In Fig.1, a time series data based on the above Equations (4) and (5) is plotted, where $\beta_{11} = 0.2$, $\beta_{10} = 1.0$, $\beta_{21} = 0.8$, $\beta_{20} = -41.0$, and tc* =70, $\varepsilon \approx N(0,1)$. Moreover, σ_S means the standard deviation of error between early-observed data $\{y_t \mid 1 \le t \le 40\}$ and the predicted line obtained from the early data.

The concrete procedure of structural change detection is as follows:

Step1: Make a prediction expression and set the tolerance band (***a***) (e.g. ***a***=**2**σ_S) that means permissible error margin between the predicted data and the observed one.

Step2 : Set up the null hypothesis H_0 and alternative hypothesis H_1.

H_0 : Change has not occurred yet.

H_1 : Change has occurred.

Set the values α, β and compute C_1 and C_2, according to Equation (2). Initialize i = 0, $\lambda_0 = 1$.

Remark:

The statement of the null hypothesis H_0, "Change has not occurred yet.", means in statistical sense. It means that the generation probability for the data to go out from the tolerance band is less than (or equal to) θ_0 (for instance, 1%). Similarly, the statement of the alternative hypothesis H_1, "Change has occurred." means that the generation probability for the data to go out from the tolerance band is greater than (or equal to) θ_1 (for instance, 99%). Additionally, we suppose that θ_1 is considerably greater than θ_0.

Step3: Incrementing i (i = i+1), observe the following data y_i. Evaluate the error $|\varepsilon_i|$ between the data y_i and the predicted value from the aforementioned prediction expression.

Step4: Judge as to whether the data y_i goes in the tolerance band or not, i.e., the ε_i is less than (or equal to) the permissible error margin or not. If it is Yes, then set $\lambda_i = 1$ and return to Step3. Otherwise, advance to Step5.

Step5: Calculate the probability ratio λ_i, using the following Equation (6) that is equivalent to Equation (1)

$$\lambda_i = \lambda_{i-1} \frac{P(\varepsilon_i \mid \mathbf{H}_1)}{P(\varepsilon_i \mid \mathbf{H}_0)} \qquad (6)$$

where, if the data y_i goes in the tolerance band, $P(\varepsilon_i \mid H_0) = \theta_0$ and $P(\varepsilon_i \mid H_1) = \theta_1$, otherwise,
$P(\varepsilon_i \mid H_0) = (1-\theta_0)$ and $P(\varepsilon_i \mid H_1) = (1-\theta_1)$.

Step6: Execution of testing.

(i) If the ratio λ_i is greater than C_2 (= (1-β)/α), dismiss the null hypothesis H_0, and adopt the alternative hypothesis H_1, and then End.

(ii) Otherwise, if the ratio λ_i is less than C_1 (= β/(1-α)), adopt the null hypothesis H_0, and dismiss the alternative hypothesis H_1, and then set $\lambda_i = 1$ and return to Step3.

(iii) Otherwise (in the case where $C_1 \le \lambda_i \le C_2$), advance to Step7.

Step7: Observe the following data y_i incrementing i. Evaluate the error $|\varepsilon_i|$ and judge whether the data y_i goes in the tolerance band, or not. Then, return to Step5 (calculation of the ratio λ_i).

III. Experimentation in comparison with Chow Test

We have experimented with the proposed method for both artificial and real time series data, in comparison with the well-known Chow Test [12]. Those artificial time series data are generated based on aforementioned Equations (4) and (5). In this section, we show the experimental results.

1. Chow Test

The well known Chow Test checks the significant differences among residuals for three Regression Lines, where regression Line 1 obtained from the data before a change point tc, Line 2 from the data after tc, and Line 3 from the whole data so far, by setting up hypothesis of change point at each point in the whole data. Fig.2 shows the conceptual image of Chow Test.

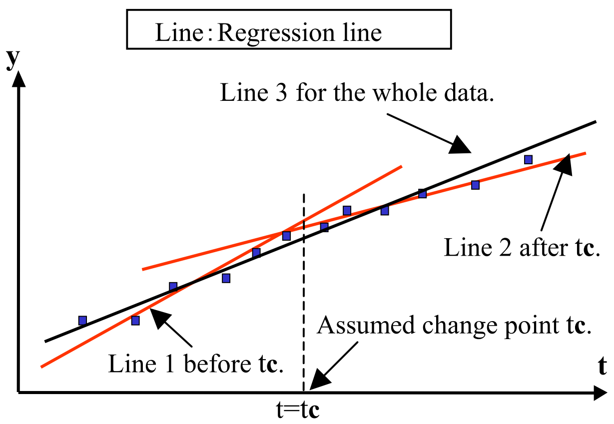

Fig.2. Chow Test in the situation where a hypothesis is set up that the structural change has occurred at t=tc.

2. Experimentation

The experimental results from artificially generated time series data are illustrated in Fig.3-6, where the standard deviation σ of error ε and so on in the generated data are varied and horizontal axis shows observing time t (detection operation has started from t=41). The vertical axis shows the detected change point tc, whose value is the average of experimentation results for 100 sets of generated time series data where the true change point is tc*=70. Fig.3-6 show that Chow Test outputs the change point at the time when every data is observed after t=40. This means that Chow Test cannot detect the change point properly. And, the time point when Chow Test can work well is long late enough after the true change occurs.

On the other hand, the proposed method detects the change points a, b, c, where each of them correspond to different parameter values of θ_0, θ_1 that are used in the SPRT as **a**:(θ_0=0.1, θ_1=0.9), **b**:(θ_0=0.2, θ_1=0.8), and **c**:(θ_0=0.3, θ_1=0.7).

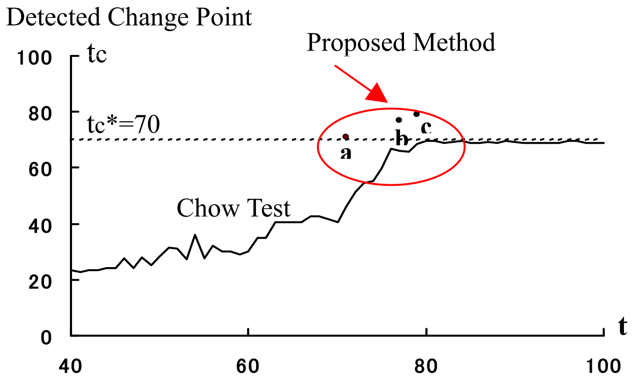

Fig.3. Relation between the observing time and detected change point. (β_{11}=0.2, β_{21}=0.4, σ=0.5. Detection operation starts at t=41.)

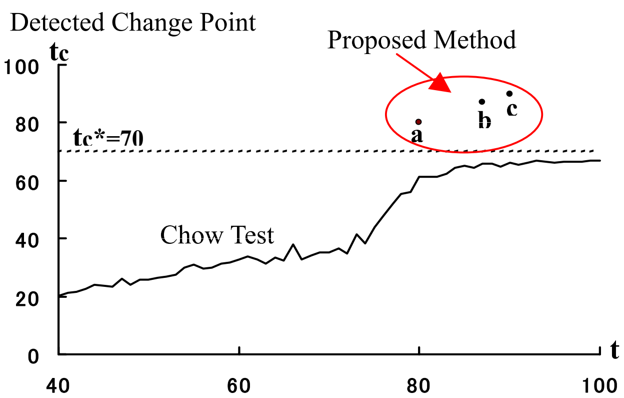

Fig.4. Relation between the observing time t and detected change point tc. (β_{11}=0.2, β_{21}=0.4, σ=1.0. Detection operation starts at t=41.)

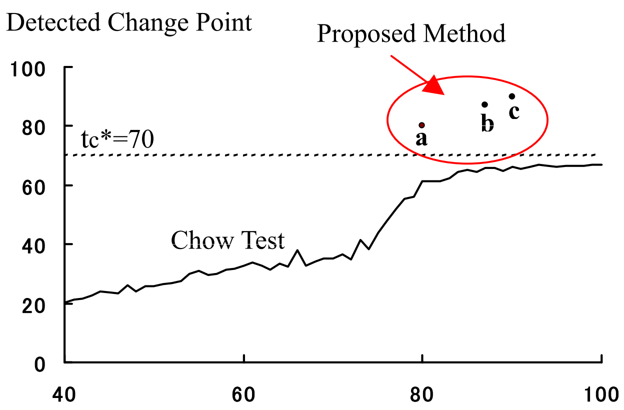

Fig.5. Relation between the observing time and detected change point. (β_{11}=0.2, β_{21}=0.4, σ=1.5. Detection operation starts at t=41.)

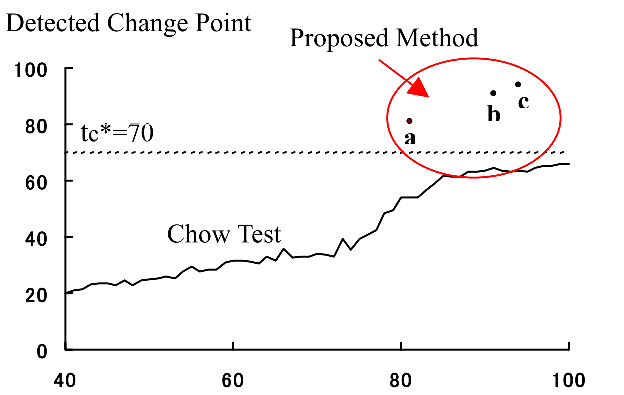

Fig.6. Relation between the observing time and detected change point. (β_{11}=0.2, β_{21}=0.3, σ=1.5. Detection operation starts at t=41.)

IV. CONCLUSION

We have proposed a sequential probability ratio test for structural change point detection of time series data. We have presented the algorithm and procedure how to apply a probability ratio test to change detection problem. Based on numerical experimentation results, we have found the effectiveness of SPRT in comparison with Chow Test, as follows:

It can quickly and accurately detect the structural change point when a forecasting model cannot forecast the time series.

It can detect correctly the structural change point even when the trend in time series changes little (Chow test can not detect the change point when the trend changes little).

Moreover the SPRT is effective in the sense as follows. (1)Differently from the Chow Test, it does not need to set the change point in a priori. (2)Unlike the Bay's method, it does not need to give the distribution of time series data. (3) It can early detect the structural change point by sequential processing. (4) It is a meta-level method so that we can apply it to any prediction model.

REFERENCES

[1]C.G.E.P.Box and G.M.Jenkins (1976), Time Series Analysis: Forecasting and Control, Prentice Hall.
[2] Peter J. Brockwell, Richard A. Davis (2003), Introduction to Time Series and Forecasting, Springer; 2nd edition.
[3] Robert J. Jannarone (1997), Concurrent Learning and Information Processing: A Neuro-Computing System That Learns During Monitoring, Forecasting, and Control, Kluwer Academic Pub.
[4] George Bojadziev, Maria Bojadziev (2007), Fuzzy Logic for Business, Finance, and Management (Advances in Fuzzy Systems -Applications and Theory), World Scientific Pub Co. Inc..
[5] H. Tong (1995), Chaos and Forecasting (Nonlinear Time Series and Chaos, Vol 2), World Scientific Pub Co. Inc..
[6]Akaike,H.(1973), Information theory and an extension of the maximum likehood principle, Proc. 2nd International Symposium on Information Theory, Akademia Kiado, budapest, pp.269-281.
[7]Mallows, C.L. (1973), Some remarks of Cp, Technometrics, Vol.15, No.4, pp.661-675.
[8] Allen,D.M. (1971), Mean squqre error of prediction as a criterion of selecting variables, Technometrics, Vol.13, pp.469-475.
[9] Hjorth,U. (1982), Model selection and forward validation, J. Statist., Vol.9, pp.95-105.
[10] A. Wald (1947), Sequential Analysis, John Wiley & Sons.
[11]Hiromichi Kawano, Tetsuo Hattori, Ken Nishimatsu (2008), Structural Change Point Detection Method of Time Series Using Sequential Probability Ratio Test－Comparison with Chow Test in the ability of early detection－(in Japanese), IEEJ Trans. EIS, Vol.128, No.4, pp.583-592.
[12] Chow,G.C. (1960), Tests of Equality Between Sets of Coefficients in Two Linear Regressions, Econometrica, Vol.28, No.3.