Relation between Optimal Stopping Solution and NSPR for Structural Change Point Detection Problem

Tetsuo Hattori
Kagawa University, 2217-20 Hayashi Cho
Takamatsu City, Kagawa 761-0396, Japan

Yoshihide Koyama, Yusuke Kawakami, Yoshiro Imai, Yo Horikawa
Kagawa University, 2217-20 Hayashi Cho
Takamatsu City, Kagawa 761-0396, Japan

Hiromichi Kawano
NTT advanced technology Company Ltd, Mitaka, Tokyo, Japan
E-mail: {hattori, imai, horikawa}@eng.kagawa-u.ac.jp, hiromichi.kawano@ntt-at.co.jp

Abstract

Previously, we have proposed a novel method using New Sequential Probability Ratio (NSPR) for the structural change point detection (SCPD) of ongoing time series data instead of using SPRT (Sequential Probability Ratio Test). Moreover, we have formulated the SCPD problem in time series data as an Optimal Stopping one using the concept of DP (Dynamic Programming) and also have shown the solution theorem in the form of Inequality. In this paper, we discuss the relation between the solution of Optimal Stopping and NSPR for the SCPD Problem.

Keywords: Image processing, Curvature, Variance estimation, Histogram matching, HMGD

1. Introduction

For ongoing time series analysis, three stages are considered: prediction model construction, structural change detection (and/or disparity detection between the model and observing data), and renewal of prediction model. Above all, it is important to detect the change point as quickly and also correctly as possible Especially in the second stage, in order to renew the accurate prediction model as soon as possible after the detection.

As the structural change detection, or change point detection (CPD), some methods have been proposed [1-4]. The standard well known method is Chow Test that is used in econometrics[2]. It does a statistical test by setting the hypothesis that the change has occurred at time *t* for all of data obtained so far.

Meanwhile, we have previously formulated the structural change detection method in time series as an Optimal Stopping Problem with an action cost, using the concept of DP (Dynamic Programming) [5-6]. Moreover for the change point detection problem, we have proposed a model introduced SPRT (Sequential Probability Ratio Test) as a New Sequential Probability Ratio (NSPR) test method[7-8].

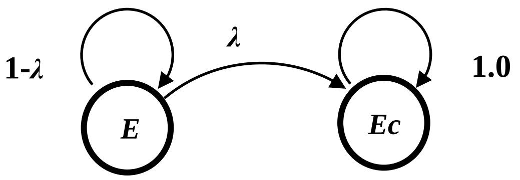
In this paper, we present that there is a relation between the NSPR and the optimal solution theorem for CPD that dealt as an Optimal Stopping Problem and describe it concretely.

2. Definitions and Equations

2.1. *Structural Change Model* [5-6]

We assume that the structural change is Poisson occurrence of average λ, and that, once the change has occurred during the observing period, the structure does not go back to the previous one. The reason why we set such a model is that we focus on the detection of the first structural change in the sequential processing (or sequential test). The concept of the structural change model is shown in Fig. 1.

Moreover, we introduce a more detailed model. Let R be the probability of the failing when the structure is unchanged. Let Rc be the probability of the failing when the structure change occurred. We consider that Rc is greater than R, i.e., Rc>R. The detailed model for the State Ec and E are illustrated as similar probabilistic finite state automatons in Fig.2 and Fig.3, respectively.



Ec **:** State that the structural change occurred.
E **:** State that the structure is unchanged.
λ : Probability of the structural change occurrence. (Poisson Process.)

Fig.1. Structural change model.

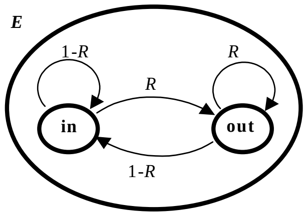


Fig.2. Internal model of the State *E*.

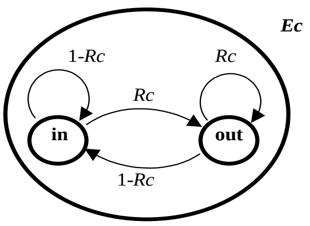


Fig.3. Internal model of the State *E*c.

2.2. *Optimal Stopping Formulation and its Slution Theorem* [5-6]

Let the cost(n) be $a \cdot n$ as a linear function for n, where a is the loss caused by the failing in one time. And for simplicity, let T and A denote the Total_cost and cost (A), respectively. Then, the evaluation function is denoted as the following equation (1).

$$T = A + a \cdot n \tag{1}$$

We recursively define a function $ET(n,N)$ to obtain the optimum number of times n that minimizes the expectation value of the evaluation function of Equation (3), using the concept of DP (Dynamic Programming). Let N be the optimum number. Let the function $EC(n,N)$ be the expectation value of the evaluation function at the time when the failing has occurred in continuing n times, where n is less than or equal to N, i.e., $0 \le n \le N$.

Thus the function is recursively defined as follows.

$$\text{(if } n = N\text{)} \quad ET(n,N) = A + a \cdot N \tag{2}$$

$$\text{(if } n < N\text{)} \quad ET(n,N) = P(\bar{S}_{n+1} \mid S^n) \cdot a \cdot n + (1 - P(\bar{S}_{n+1} \mid S^n))\, ET(n+1, N) \tag{3}$$

where S^n means the state of failing in continuing n times, $\bar{S}_{n+1}$ the state of hitting at the (n+1) th observed data, and $P(\bar{S}_{n+1} \mid S^n)$ means the conditional probability that the state $\bar{S}_{n+1}$ occurs after the state S^n.

The first term in the right-hand side (RHS) of the equation (3) indicates the expectation value of the evaluation function at the time when hitting happens at the (n+1)th data after the continuing n times failing. The second term in the RHS of the equation (3) indicates the expectation value of the evaluation function for the time when failing happens at the (n+1)th data after continuing n times failing.

Then, from the definition of the function $ET(n,N)$, the goal is to find the N that minimizes $ET(0,N)$, because the N is the same as n that minimizes the expectation value of the evaluation function of (1).

[Optimal Solution Theorem (OST)]

The N that minimizes $ET(0,N)$ is given as the largest number n that satisfies the following Inequality (4).

$$a < (A + a) \cdot P(\bar{S}_n \mid S^{n-1}) \tag{4}$$

where the number N+1 can also be the optimum one that minimizes $ET(0,N)$, i.e., $ET(0,N) = ET(0,N+1)$, only if

$$a = (A + a) \cdot P(\overline{S}_{N+1} \mid S^N)$$

2.3. *New Sequential Probability Ratio (NSPR) Based on Structural Change Model* [7-8]

Let $a_1 a_2, \ldots, a_i, \ldots a_n \quad a_i \in \{\mathrm{IN}, \mathrm{OUT}\}$ be a string (or symbol sequence) obtained from the observed data.

Let θ_i and $\tilde{\theta}_i$ be the conditional probability that outputs the observed data (or above symbol sequence, $C_n = a_1 a_2 \ldots a_n$ in the state $\boldsymbol{S_0}$ and $\boldsymbol{S_1}$, respectively. That is, it means that $\theta_i \in \{R, 1-R\}$ and $\tilde{\theta}_i \in \{R_c, 1-R_c\}$, respectively.

And let $P(a_1 \ldots a_n, \mathrm{H}_0)$ and $P(a_1 \ldots a_n, \mathrm{H}_1)$ be the joint probability of the symbol sequence C_n happens with the event H_0 (the structural change is not occurred) and H_1 (the change is occurred), respectively.

Then, the following equations hold.

$$\begin{aligned} P(a_1 \ldots a_n, \mathrm{H}_0) &= P(C_n, \mathrm{H}_0) \\ &= (1-\lambda)^n \theta_1 \ldots \theta_n = (1-\lambda)^n \prod_{i=1}^{n} \theta_i \end{aligned} \tag{5}$$

$$\begin{aligned} P(a_1 \ldots a_n, \mathrm{H}_1) &= P(C_n, \mathrm{H}_1) \\ &= \lambda \prod_{i=1}^{n} \tilde{\theta}_i + ((1-\lambda)\theta_1)(\lambda \prod_{i=2}^{n} \tilde{\theta}_i) \\ &\quad + ((1-\lambda)^2 \theta_1 \theta_2)(\lambda \prod_{i=3}^{n} \tilde{\theta}_i) + \ldots \\ &= \sum_{k=1}^{n} ((1-\lambda)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\lambda \prod_{i=k}^{n} \tilde{\theta}_i) \end{aligned} \tag{6}$$

New Sequential Probability Ratio (NSPR) Λ_n that we propose is represented using the aforementioned equations as following Eq.(7).

$$\begin{aligned} NSPR\ \Lambda_n &= \frac{P(\mathrm{H}_1 \mid a_1 \ldots a_n)}{P(\mathrm{H}_0 \mid a_1 \ldots a_n)} = \frac{P(\mathrm{H}_1 \mid C_n)}{P(\mathrm{H}_0 \mid C_n)} = \frac{P(C_n, \mathrm{H}_1)}{P(C_n, \mathrm{H}_0)} \\ &= \frac{\sum_{k=1}^{n} ((1-\lambda)^{k-1} \cdot \prod_{j=0}^{k-1} \theta_j)(\lambda \prod_{i=k}^{n} \tilde{\theta}_i)}{(1-\lambda)^n \prod_{i=1}^{n} \theta_i} \end{aligned} \tag{7}$$

If the NSPR is greater than 1.0, we can regard that the structural change has been occurred before the present time.

3. Relation between NSPR and OST

We show the relations using the probability in the Optical Solution Theorem, considering $R_c >> R$.

$$\begin{aligned} P(\overline{S_{n+1}} \mid S^n) &= (1-R)(1 - P(E_{cn} \mid S^n)) + (1-R_c) P(E_{cn} \mid S^n) \\ &= (1-R) - P(E_{cn} \mid S^n)(R_c - R) \end{aligned} \tag{8}$$

Therefore, we have

$$P(E_{cn} \mid S^n) = \frac{(1-R) - P(\overline{S_{n+1}} \mid S^n)}{(R_c - R)} \tag{9}$$

Similarly,

$$P(E \mid S^n) = \frac{P(\overline{S_{n+1}} \mid S^n) - (1-R_c)}{(R_c - R)} \tag{10}$$

Since

$$P(E_{cn} \mid S^n) = P(H_1 \mid S^n) \text{ and } P(E \mid S^n) = P(H_0 \mid S^n)$$

we have

$$NSPR\,\Lambda_n = \frac{P(H_1 \mid S^n)}{P(H_0 \mid S^n)} = \frac{(1-R) - P(\overline{S_{n+1}} \mid S^n)}{P(\overline{S_{n+1}} \mid S^n) - (1-R_c)} \tag{11}$$

Since $P(\overline{S_{n+1}} | S^n)$ is a monotonous decreasing function with respect to n, NSPR becomes an increasing one. From the OST, the optimal N is the maximum n that satisfies (4), so the N is the the maximum n that satisfies the following Inequality (12).

$$NSPR\ \Lambda_n = \frac{P(H_1 | S^n)}{P(H_0 | S^n)} < \Theta \equiv \frac{\left(\frac{A}{A+a}\right) - R}{R_c - \left(\frac{A}{A+a}\right)} \tag{12}$$

If $R = 0.05$, $R_c = 0.9$, $A = 10, a = 1$, $\Theta = 21.5$.

Meanwhile, since NSPR is a probability ratio, NSPR makes unnecessary the restriction condition of n times continuous failures (or “OUT”) in the Optimal Stopping Formulation.

4. **Conclusion**

We have described the relation between New Sequential Probability Ratio (NSPR) method and the Optimal Solution Theorem for CPD that dealt as an Optimal Stopping Problem. From this relation, we can use NSPR as well in the case where we have to consider some constraints with respect to loss cost.

References

1. R.Jana and S.Dey, Change detection in Teletraffic Models, *IEEE Trans. Signal Processing*, vol.48, No.3, pp.846-853, 2000.
2. Chow,G.C., Tests of Equality Between Sets of Coefficients in Two Linear Regressions, *Econometrica*, Vol.28, No.3, pp.591-605, 1960.
3. S.MacDougall, A.K. Nandi and R.Chapman, Multisolution and hybrid Bayesian algorithms for automatic detection of change points, *Proc. of IEEE Visual Image Signal Processing*, vol.145, No.4, pp.280-286, 1998.
4. E.S.Page, Continuous inspection schemes, *Biometrika*, Vol.41, pp.100-115, 1954.
5. Tetsuo Hattori, Katsunori Takeda, Izumi Tetsuya, Hiromichi Kawano, "Early Structural Change Detection as an Optimal Stopping Problem (I) --- Formulation Using Dynamic Programming with Action Cost ----", *Proc. of the 15th International Symposium on Artificial Life and Robotics (AROB15th'10)*, ISBN 978-4-9902880-4-4, pp.763-766, 2010.
6. Hiromichi Kawano, Tetsuo Hattori, Katsunori Takeda, Izumi Tetsuya, "Early Structural Change Detection as an Optimal Stopping Problem (II) --- Solution Theorem and its Proof Using Reduction to Absurdity ----", *Proc. of the 15th International Symposium on Artificial Life and Robotics (AROB15th'10)*, ISBN 978-4-9902880-4-4, pp.767-772, 2010.
7. Yoshihide Koyama, Tetsuo Hattori, Hiromichi Kawano, “Model Introduced SPRT for Structural Change Detection of Time Series (I)”, *Journal of Robotics, Networks and Artificial Life* (ISSN: 2352-6386), Vol.1, No.1, pp.55-60 (Atlantis Press, Jun. 2014)
8. Yoshihide Koyama, Tetsuo Hattori, Katsunori Takeda, Hiromichi Kawano, “Model Introduced SPRT for Structural Change Detection of Time Series (II)”, *Journal of Robotics, Networks and Artificial Life* (ISSN: 2352-6386), Vol.1, No.3, pp.237-243 (Atlantis Press, Dec. 2014)