An Extended Optimal Stopping Method for Structural Change Point Detection Problem

Yoshihide Koyama
Kagawa University, 2217-20 Hayashi Cho
Takamatsu City, Kagawa 761-0396, Japan

Tetsuo Hattori, Yoshiro Imai, Yo Horikawa, Yusuke Kawakami
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 and formulated the SCPD (Structural Change Point Detection) problem in time series data as an Optimal Stopping one using the concept of DP (Dynamic Programming). And also we have shown the solution theorem in the form of inequality. In this paper, based on the relation between the solution of Optimal Stopping and NSPRT (New Sequential Probability Ratio Test), we present the extended Optimal Stopping Method in order to obtain more practical one, considering a loss cost and an action cost involved by failure prediction.

Keywords: Time series data, Structural change point detection, Optimal stopping problem, NSPRT (New Sequential Probability Ratio Test)

1. Introduction

There are three stages to be considered for ongoing time series analysis as follows;
(i) the stage of prediction model construction,
(ii) the stage of structural change detection (and/or disparity detection between the model and observing data),
(iii) the stage of renewal of prediction model.

Above all, it is very important to detect the change point as quickly and correctly as possible in the second stage, in order to renew the accurate prediction model as soon as possible after the change detection.

For the problem of the structural change point detection (SCPĐ) problem, or change point detection (CPĐ), some methods have been proposed. The standard well known method is Chow Test used in econometrics. Chow Test carries out a statistical test by setting the hypothesis that the change has occurred at time t for all of data obtained so far.

© The 2017 International Conference on Artificial Life and Robotics (ICAROB 2017), Jan. 19-22, Seagaia Convention Center, Miyazaki, Japan
As for the SCPD problem, we have previously formulated the change detection method as a solving method for Optimal Stopping Problem with an action cost, using the concept of DP (Dynamic Programming) \(^5^6\). Moreover, 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 extend the Optimal Stopping Problem as a more general problem, and we show the extended solution. Moreover, we present the relation between NSPR and the extended optimal solution theorem for SCPD.

## 2. Definitions and Equations

### 2.1. Structural Change Model \(^5^6\)

We assume that the structural change is Poisson occurrence of average \(\lambda\), 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 \(Re\) be the probability of the failing when the structure change occurred. We consider that \(Re\) is greater than \(R\), i.e., \(Re>R\). The detailed model for the State \(Ec\) and \(E\) are illustrated as similar probabilistic finite state automaton in Fig.2 and Fig.3, respectively.

![Fig.1. Structural change model.](image)

### 2.2. Optimal Stopping Formulation and its Solution 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 one-time prediction failure. And for simplicity, let \(T\) and \(A\) denote the Total loss cost and an involved action cost, respectively. Then, the evaluation function is denoted as the following equation (1).

\[
T = A + a \cdot n
\]  
(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., \(0 \leq n \leq N\).

Thus the function is recursively defined as follows.
if \( n = N \) \[ ET(n,N) = A + a \cdot N \] (2)

if \( n < N \) \[ ET(n,N) = P(S_{n+1} | S^n) \cdot a \cdot n + (1 - P(S_{n+1} | S^n))ET(n+1,N) \] (3)

where \( S^n \) means the state of failing in continuing \( n \) times, \( S_{n+1} \) the state of hitting at the \( (n+1) \)th observed data, and \( P(S_{n+1} | S^n) \) means the conditional probability that the state \( S_{n+1} \) 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).

\[
a < (A + a) \cdot P(S_n | S^{n-1})
\] (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

\[
a = (A + a) \cdot P(S_{N+1} | S^N)
\]

2.3. **Extended Optimal Stopping Formulation**

We extend the meaning of the loss \( a \) as the average of loss cost \( a^* \) caused by the “so far \( n \)-time prediction failure”. The “so far \( n \)-time prediction failure” does not mean the continuous \( n \)-time prediction failure\(^9\). And \( A \) means the loss cost involved by some action, e.g., renewal and/or disposal of some equipments and/or some systems after the structural change point detection.

Then, \( S^n \) does not mean “the state of failing in continuing \( n \) times”, but means “the state of so far \( n \)-times prediction failure”. Thus we obtain an extended OST as follows.

\[
\text{[Extended Optimal Solution Theorem (EOST)]}
\]

The \( N \) that minimizes \( ET(0,N) \) is given as the largest number \( n \) that satisfies the following Inequality (5).

\[
a^* < (A + a^*) \cdot P(S_n | S^{n-1})
\] (5)

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

\[
a^* = (A + a^*) \cdot P(S_{N+1} | S^N)
\]

2.4. **New Sequential Probability Ratio (NSPR) Based on Structural Change Model**

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

Let \( \theta_i \) and \( \tilde{\theta}_i \) be the conditional probability that outputs the observed data (or above symbol sequence, \( C_n = a_1a_2\ldots a_n \) in the state \( S_0 \) and \( S_1 \), respectively. That is, it means that \( \theta_i \in \{ R, 1-R \} \) and \( \tilde{\theta}_i \in \{ R, 1-R_c \} \), respectively.

And let \( P(a_1, \ldots, a_n, H_0) \) and \( P(a_1, \ldots, a_n, H_1) \) 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 New Sequential Probability Ratio (NSPR) \( \Lambda_n \) is represented as the following Eq.(6).

\[
\text{NSPR} \; \Lambda_n = \frac{P(H_1 | a_1, \ldots, a_n)}{P(H_0 | a_1, \ldots, a_n)} = \frac{P(H_1 | a_n)}{P(H_0 | a_n)} = \frac{P(C_n, H_1)}{P(C_n, H_0)}
\]

\[
= \sum_{i=1}^{n} \left( (1-\lambda) \prod_{j=0}^{i-1} \theta_j \right) \left( \lambda \prod_{j=0}^{n-1} \tilde{\theta}_j \right)
\]

\[
(1-\lambda)^n \prod_{i=1}^{n} \tilde{\theta}_i
\]

(6)

From Eq.(6), we have the following recursive equation (7).

\[
\Lambda_n = \frac{1}{1 - \gamma} (\Lambda_{n-1} + \gamma (\tilde{\theta}_n / \theta_n))
\]

where \( \theta_0 = 1, \tilde{\theta}_0 = 1, \Lambda_0 = 0 \)

(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 EOST

We show the relations using the probability in the Extended Optical Solution Theorem, considering $R_e \gg R$.

$$P\left(\overrightarrow{S_n} \mid S^n\right) = \left(1 - R\right)P\left(E_{\text{sn}} \mid S^n\right) = \left(1 - R\right)\left(1 - P\left(E_{\text{on}} \mid S^n\right)\right) \equiv \left(1 - R\right) - P\left(E_{\text{on}} \mid S^n\right)$$

Therefore, we have

$$P\left(E_{\text{on}} \mid S^n\right) = \frac{\left(1 - R\right) - P\left(\overrightarrow{S_n} \mid S^n\right)}{R_e - R}$$

Since

$$P\left(E_{\text{on}} \mid S^n\right) = P\left(H_1 \mid S^n\right), \quad P\left(E \mid S^n\right) = P\left(H_0 \mid S^n\right)$$

we have

$$\Lambda_n = \frac{P\left(H_1 \mid S^n\right)}{P\left(H_0 \mid S^n\right)} = \frac{\left(1 - R\right) - P\left(\overrightarrow{S_n} \mid S^n\right)}{P\left(\overrightarrow{S_n} \mid S^n\right) - (1 - R_e)}$$

If $P\left(\overrightarrow{S_n} \mid S^n\right)$ is a monotonous decreasing function with respect to $n$, NSPR becomes an increasing one. From the EOST, the optimal $N$ is the maximum $n$ that satisfies (5), so the $N$ is the maximum $n$ that satisfies the following Inequality (11).

$$\Lambda_n = \frac{P\left(H_1 \mid S^n\right)}{P\left(H_0 \mid S^n\right)} < \Theta = \frac{A}{A + a} - \frac{R}{R_e - \frac{A}{A + a}}$$

If $R = 0.05$, $R_e = 0.9$, $A = 10$, $a^* = 1$, then we have the threshold $\Theta = 21.5$.

4. Conclusion

We have proposed the extended Optimal Stopping Formulation for the SCPD problem and its extended solution theorem. Moreover, we have presented the relation between New Sequential Probability Ratio (NSPR) method and the Extended Optimal Solution Theorem. From this relation, we can use the NSPR as an extended optimal stopping method for structural change point detection in ongoing time series data.

References