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

Early Structural Change Detection as an Optimal Stopping Problem (II) --- Solution Theorem and its Proof Using Reduction to Absurdity ----

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

**NTT Advanced Technology / Musashino-shi Nakamachi 19-18, Tokyo 180-0006, Japan*
***Graduate School of Engineering, Kagawa University / 2217-20 Hayashi, Takamatsu City, Kagawa 761-0396, Japan*
****Micro TechnicaCo., Ltd./ Yamagami BLD. 3-12-2 Higashi-ikebukuro, Toshima-ku, Tokyo 170-0013, Japan*
(Tel : 81-087-864-2221; Fax : 81-087-864-2262)
(hattori@eng.kagawa-u.ac.jp)

***Abstract*:** Change point detection (CPD) problem in time series is to find that a structure of generating data has changed at some time point by some cause. We formulated the structural change detection in time series as an optimal stopping problem using the concept of DP (Dynamic Programming) and presented the optimal solution and the correctness by numerical calculation. In this paper, we present the solution theorem and its proof using reduction to absurdity.

***Keywords*:** time series data, structural change point detection, optimal stopping problem

I. INTRODUCTION

Change point detection (CPD) problem in time series is to find that a structure of generating data has changed at some time point by some cause. In the literature [1],[2], we have proposed a new and practical method based on an evaluation function of loss cost. In addition, we have formulated the CPD problem as an optimal stopping problem and have given the algorithm for the optimum solution and also have presented the effectiveness using numerical experimental results.

In this paper, we show the solution theorem and its proof using reduction to absurdity.

II. FORMULATION

1. Definition of Function

Let the cost(n) be an as a linear function for n, where a is the loss caused by the failing (i.e., missing the forecast, or fail in prediction) in one time. And for simplicity, let C and A denote the Total cost and cost (A), respectively. Then, the evaluation function is denoted as the following Equation (1).

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

We introduce a function $EC(n,N)$ to obtain the optimum number of times n that minimizes the expectation value of C, using the concept of DP. Let N be the optimum number. Let the function $EC(n,N)$ be the evaluation function at the time when the failing has occurred in continuing n times ($n \le N$). Then the function can be defined in the following form.

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

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

where S_n is the state of failing in continuing n times, $\bar{S}_{n+1}$ is the state of unfailing (or hitting) for the (n+1)-th time 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.

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

2. Minimization of the Evaluation Function

For the aforementioned $EC(0,N)$, the following theorem holds and gives the n that minimizes the expectation value of the evaluation function of (1).

A. Theorem

The N that minimizes $EC(0,N)$ 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 $EC(0,N)$, i.e., $EC(0,N) = EC(0,N+1)$, only if $a = (A + a) \cdot P(\bar{S}_{N+1} \mid S_N)$.

Proof.

We derive a contradiction with two assumptions under a premise as follows.

Premise: a number N' is the largest number n that satisfies the Inequality (4).

Assumption 1:
There exists a number N'' such that
$N'' < N'$ and $EC(0, N'') < EC(0, N')$.

Assumption 2:
There exists a number N'' such that
$N' < N''$ and $EC(0, N') > EC(0, N'')$.

We give the proof of this theorem by three steps. At Step 1, we prove fundamental lemmas: Lemma 1-1 and Lemma 1-2. At Step 2, we prove two lemmas: Lemma 2-1, and Lemma 2-2. Using those lemmas, we show that the above Assumption 1 contradicts the Premise. Similarly, at Step 3, we show that the Assumption 2 contradicts the Premise, using two lemmas: Lemma 3-1 and Lemma 3-2, which are proved in Appendix.

B. Step 1

Lemma 1-1.

Let E_{cn} be the event that the structural change occurs once during the period of observation in continuing *n* times. Let $P(E_{cn} | S_n)$ be the conditional probability that the E_{cn} happens under the condition that failing occurs in continuing *n* times.

$P(E_{cn} | S_n)$ is an increase function for *n*.

Proof.

We derive some useful equations for this proof.

The event E_{cn} is given in (5).

$$E_{cn} = \bigcup_{i=0}^{n-1} \left(E^i \cap E_c^{n-i} \right) \qquad (5)$$

where *E* is the event that there is no structural change, E_c is the event that the structural change occurred, and E^n is defined as $E^n = \bigcap_{i=1}^{n} E^i$.

The probability of the event E_{cn} defined in (5) is given as follows.

$$P(E_{cn}) = P\left(\bigcup_{i=0}^{n-1} \left(E^i \cap E_c^{n-i} \right) \right) = \sum_{i-0}^{n-1} P\left(E^i \cap E_c^{n-i} \right)$$

$$= \sum_{i=0}^{n-1} (1-\lambda)^i \lambda \qquad (6)$$

λ **:** Probability of the structural change occurrence. The joint event of E_{cn} and S_n, and its probability are given by (7) and (8), respectively. Let *R* be the probability of the failing when the structure is unchanged. Let *Rc* be the probability of the failing when the structural change occurred.

$$S_n \cap E_{cn} = S_n \cap \left(\bigcup_{i=0}^{n-1} \left(E^i \cap E_c^{n-i} \right) \right) \qquad (7)$$

$$P\left(S_n \cap E_{cn} \right) = \sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i} \qquad (8)$$

Therefore, using (6) and (8), we have

$$P(S_n | E_{cn}) = \frac{P(S_n \cap E_{cn})}{P(E_{cn})}$$

$$= \frac{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}{\sum_{i=0}^{n-1} (1-\lambda)^i \lambda} \qquad (9)$$

According to the Bayes' theorem, the posterior probability $P(E_{cn} | S_n)$ is given by the following (10).

$$P(E_{cn} | S_n) = \frac{P(S_n \cap E_{cn})}{P(S_n \cap E_{cn}) \cup P(S_n \cap E^n)}$$

$$= \frac{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i} + (1-\lambda)^n R^n}$$

$$= \frac{1}{1 + \dfrac{(1-\lambda)^n R^n}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}} = \frac{1}{1 + D(n)} \qquad (10)$$

where $D(n) = \dfrac{(1-\lambda)^n R^n}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}$.

The *D*(*n*) is also expressed as the following (11).

$$D(n) = \frac{(1-\lambda)^n \left(\dfrac{R}{R_c} \right)^n}{\lambda \sum_{i=0}^{n-1} (1-\lambda)^i \left(\dfrac{R}{R_c} \right)^i} = \frac{\mathbf{X}^n}{\lambda \sum_{i=0}^{n-1} \mathbf{X}^i}$$

$$= \frac{1}{\lambda \left(\dfrac{1}{\mathbf{X}^n} + \dfrac{1}{\mathbf{X}^{n-1}} + \cdots + \dfrac{1}{\mathbf{X}} \right)} \qquad (11)$$

where $\mathbf{X} = (1-\lambda) \dfrac{R}{R_c}$.

Since $0 \le \lambda < 1$, $0 < 1-\lambda \le 1$, and $R_c > R$, then $0 < X < 1$. So, the *D*(*n*) becomes a monotonous decrease for *n*. Therefore, the probability $P(E_{cn} | S_n)$ of (10) is a monotonous increase function for *n*.

Lemma 1-1 is proved.

Remark: Lemma 1-1 indicates that, if the number of times of the failing *n* increases, the probability that the structural change has occurred increases. This meets our intuition clearly.

Lemma 1-2.

The conditional probability $P(\overline{S}_{n+1} | S_n)$ is a decrease function for *n*.

Proof.

We have

$$P(\bar{S}_{n+1} \mid S_n) = (1-R)(1-P(E_{cn} \mid S_n)) + (1-R_c)P(E_{cn} \mid S_n) \quad (12)$$

The first term in the RHS of (12) shows the probability that the hitting occurs for the (n+1)-th time observed data when the structure is unchanged. The second term shows the probability that the hitting occurs for the (n+1)-th time observed data when the structure changed. From (12), we have

$$P(\bar{S}_{n+1} \mid S_n) = 1 - R + P(E_{cn} \mid S_n)(R - R_c) \quad (13)$$

By Lemma 1-1, $P(E_{cn} \mid S_n)$ is an increase function, and $R < R_c$, therefore, $P(\bar{S}_{n+1} \mid S_n)$ is a decrease function for n. (The decreasing situation by the numerical computing is shown in Fig.1.)

Lemma 1-2 is proved.

Remark: Lemma 1-2 indicates that, if the number of times of continuous failing increases, the probability of the hitting for the next observed data after those continuous failing decreases. This is intuitively clear, because, by Lemma 1-1, the probability of the structural change increases if the number of times of the continuous failing increases.

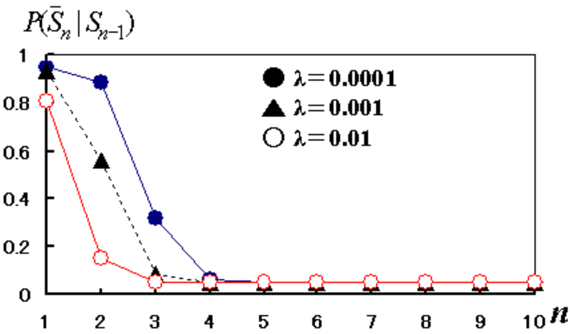

Fig.1. The probability $P(\bar{S}_n \mid S_{n-1})$ for three kinds of λ (Occurrence probability of structural change) in the case of Rc =0.95.

C. Step 2

We derive a contradiction for proving the Theorem using the following Lemma 2-1 and Lemma 2-2, where the notation is the same as aforementioned.

Lemma 2-1.

If $N'' < N'$, then $EC(N'', N') < EC(N'', N'')$

Proof.

Since $N'' < N'$, we can represent N'' as $N'' = N' - m$, where $m = 1,,, N'$. Then we have the equivalent inequality to this lemma as follows.

$$EC(N'-m, N') < EC(N'-m, N'-m) \quad (14)$$

We prove the Inequality (14) by mathematical induction on m.

(i) If $m = 1$, applying the Equation (3) to the left hand side (LHS) of Inequality (14), we have

$$\begin{aligned} EC(N'-1, N') &= P(\bar{S}_{N'} \mid S_{N'-1}) \cdot a(N'-1) \\ &+ (1 - P(\bar{S}_{N'} \mid S_{N'-1})) \cdot EC(N', N') \\ &= P(\bar{S}_{N'} \mid S_{N'-1}) \cdot a(N'-1) \\ &+ (1 - P(\bar{S}_{N'} \mid S_{N'-1})) \cdot (A + a \cdot N') \\ &= A + aN' - (A+a) \cdot P(\bar{S}_{N'} \mid S_{N'-1}) \end{aligned} \quad (15)$$

By the Premise,

$$(A+a)P(\bar{S}_{N'} \mid S_{N'-1}) > a \quad (16)$$

Therefore, next inequality holds on the RHS of (15).

$$\begin{aligned} A + aN' - (A+a) \cdot P(\bar{S}_{N'} \mid S_{N'-1}) &< A + aN' - a \\ &= A + a(N'-1) \end{aligned}$$

Then we have,

$$A + aN' - (A+a) \cdot P(\bar{S}_{N'} \mid S_{N'-1}) < EC(N'-1, N'-1) \quad (17)$$

This proves the Inequality (14) for m=1.

(ii) Assume that the Inequality (14) holds for $m = k$. In case of $m = k+1$, applying Equation (3) to the LHS of the Inequality (14), we have,

$$\begin{aligned} EC(N'-k-1, N') &= P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \cdot a \cdot (N'-k-1) \\ &+ (1 - P(\bar{S}_{N'-k} \mid S_{N'-k-1})) \cdot EC(N'-k, N') \end{aligned} \quad (18)$$

By the assumption for $m = k$, the next inequality holds.

$$EC(N'-k, N') < EC(N'-k, N'-k).$$

Therefore, for the RHS of (18), we have

$$\begin{aligned} & P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \cdot a \cdot (N'-k-1) \\ &+ (1 - P(\bar{S}_{N'-k} \mid S_{N'-k-1})) \cdot EC(N'-k, N') \\ &< P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \cdot a \cdot (N'-k-1) \\ &+ (1 - P(\bar{S}_{N'-k} \mid S_{N'-k-1})) \cdot EC(N'-k, N'-k) \end{aligned} \quad (19)$$

Applying Equation (2) to the RHS of (19), we have

$$\begin{aligned} & P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \cdot a \cdot (N'-k-1) \\ &+ (1 - P(\bar{S}_{N'-k} \mid S_{N'-k-1})) \cdot (A + a \cdot (N'-k)) \\ &= A + a \cdot (N'-k) - (a+A) \cdot P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \end{aligned} \quad (20)$$

By Lemma 1-2, $P(\bar{S}_n \mid S_{n-1})$ is a decrease function for n and by the Premise for N', we have

$$\begin{aligned} (A+a)P(\bar{S}_{N'-k} \mid S_{N'-k-1}) &> (A+a)P(\bar{S}_{N'} \mid S_{N'-1}) \\ &> a \end{aligned}$$

Therefore, next inequality is obtained for the RHS of (20).

$$\begin{aligned} & A + a \cdot (N'-k) - (a+A) \cdot P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \\ &< A + a \cdot (N'-k) - a \end{aligned} \quad (21)$$

By Equation (2), the RHS of (21) is equal to

$EC(N' - k - 1, N' - k - 1)$.

Thus, we have the following (22), and this implies that Inequality (14) holds for the case $m = k + 1$.

$$EC(N' - k - 1, N') < EC(N' - k - 1, N' - k - 1) \quad (22)$$

This proves the Lemma 2-1.

Lemma 2-2

If $N'' < N'$, then, for m ($0 < m \le N''$),

$$EC(N'' - m, N') < EC(N'' - m, N'') \quad (23)$$

Proof.

We prove this by mathematical induction for *m*.

(i) First, for *m*=1, we prove the following inequality.

$$EC(N'' - 1, N') < EC(N'' - 1, N'') \quad (24)$$

Applying Equation (3) to the LHS of the Inequality (24), we have

$$EC(N'' - 1, N') = P(\bar{S}_{N''} \mid S_{N''-1}) \cdot a \cdot (N'' - 1)$$
$$+ (1 - P(\bar{S}_{N''} \mid S_{N''-1})) \cdot EC(N'', N') \quad (25)$$

By applying Lemma 2-1 to the RHS of (25), the following inequality is obtained.

$$EC(N'' - 1, N') < P(\bar{S}_{N''} \mid S_{N''-1}) \cdot a \cdot (N'' - 1)$$
$$+ (1 - P(\bar{S}_{N''} \mid S_{N''-1})) \cdot EC(N'', N'') \quad (26)$$

Applying the Equation (3) to the RHS of Inequality (24), $EC(N'' - 1, N'')$ is equal to the RHS of (26). Then, the Inequality (24) holds, and this establishes the Lemma 2-2 for $m = 1$.

(ii) Assuming that the Lemma 2-2 holds for $m = k$, we prove it for the case of $m = k + 1$. The LHS of Inequality (23) is expressed as (27) using Equation (3).

$$EC(N'' - k - 1, N')$$
$$= P(\bar{S}_{N''-k} \mid S_{N''-k-1}) \cdot a \cdot (N'' - k - 1)$$
$$+ (1 - P(\bar{S}_{N''-k} \mid S_{N''-k-1})) \cdot EC(N'' - k, N') \quad (27)$$

Here, recalling the assumption that, for $m = k$,

$$EC(N'' - k, N') < EC(N'' - k, N'')$$

We can obtain the following inequality from (27).

$$EC(N'' - k - 1, N')$$
$$< P(\bar{S}_{N''-k} \mid S_{N''-k-1}) \cdot a \cdot (N'' - k - 1)$$
$$+ (1 - P(\bar{S}_{N''-k} \mid S_{N''-k-1})) \cdot EC(N'' - k, N'') \quad (28)$$

The RHS of Inequality (23) for $m = k + 1$,

$EC(N'' - k - 1, N'')$ is equal to the RHS of (28), by Equation (3). Therefore, we have

$$EC(N'' - k - 1, N') < EC(N'' - k - 1, N'') \quad (29)$$

This completes the proof of the Lemma 2-2.

By putting $m = N''$ in the Lemma 2-2, we have

$EC(0, N') < EC(0, N'')$ in case of $N'' < N'$.

This inequality contradicts the Assumption 1: There exists a number N'' such that $N'' < N'$ and $EC(0, N'') < EC(0, N')$.

D. Step 3

Similarly to the Step 2, the following Lemma 3-1 and Lemma 3-2 hold, as described in subsequent Appendix.

Lemma 3-1.

If $N' < N''$, then $EC(N', N'') \ge EC(N', N')$,

Where the equality holds only if $N'' = N' + 1$ and $a = (A + a) \cdot P(\bar{S}_{N'+1} \mid S_{N'})$.

Lemma 3-2.

If $N' < N''$, then, for m ($0 < m \le N'$)

$$EC(N' - m, N'') \ge EC(N' - m, N'),$$

where the equality holds only if $N'' = N' + 1$ and $a = (A + a) \cdot P(\bar{S}_{N'+1} \mid S_{N'})$.

By putting $m = N'$ in the Lemma 3-2, we have $EC(0, N'') \ge EC(0, N')$ in case of $N' < N''$.

This contradicts Assumption 2: There exists a number N'' such that $N' < N''$ and $EC(0, N') > EC(0, N'')$.

After all, $EC(0, N') \le EC(0, N'')$ ($N'' < N'$ or $N' < N''$), where the equality holds only if $N'' = N' + 1$ and $a = (A + a) \cdot P(\bar{S}_{N'+1} \mid S_{N'})$.

It means that N' minimizes $EC(0, N)$. And, when $a = (A + a) \cdot P(\bar{S}_{N'+1} \mid S_{N'})$, $N' + 1$ also minimizes $EC(0, N)$, i.e., $EC(0, N') = EC(0, N' + 1)$.

This completes the proof of Theorem. Fig.2 shows that the relation between the expectation *EC*(0,*N*) and *N*.

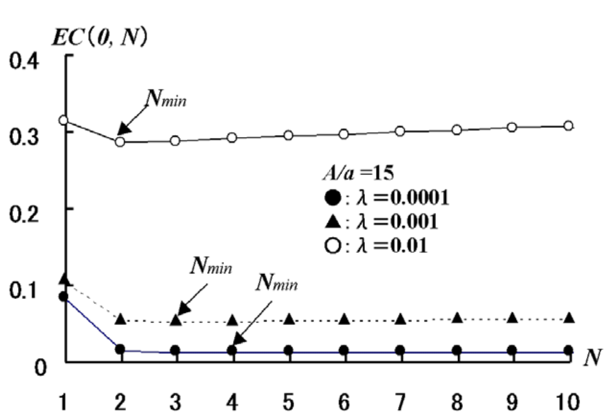

Fig.2. The expectation *EC*(0,*N*) and *N* with the ratio of ***A***/***a*** fixed (***A***=1.5, ***a*** =0.1) and varying λ. ***Nmin*** means the optimum number that minimizes *EC*(0,*N*).

III. CONCLUSION

We have proposed a sequential processing method for structural change detection of time series data as an optimal stopping problem with a cost evaluation function. We have presented the solution theorem and its proof using reduction to absurdity.

Appendix

Proof of Lemma 3-1 and Lemma 3-2

A1. Lemma 3-1

If $N' < N''$, then $EC(N', N'') \geq EC(N', N')$,

Where the equality holds only if $N'' = N'+1$ and if $a = (A+a) \cdot P(\overline{S}_{N'+1} \mid S_{N'})$.

Proof.

Since $N' < N''$, we can let $N'' = N'+m$ (m: natural number), and we have $EC(N', N'+m) \geq EC(N', N')$.

We prove this by mathematical induction for *m*.

(i) For *m*=1, we prove that

$$EC(N', N'+1) \geq EC(N', N') \quad \text{(a1)}$$

By applying the Equation (3) to the LHS of (a1), we have

$$EC(N', N'+1) = P(\overline{S}_{N'+1} \mid S_{N'}) \cdot a \cdot N' + (1 - P(\overline{S}_{N'+1} \mid S_{N'})) \cdot EC(N'+1, N'+1)$$
$$= P(\overline{S}_{N'+1} \mid S_{N'}) \cdot a \cdot N' + (1 - P(\overline{S}_{N'+1} \mid S_{N'})) \cdot (A + a \cdot (N'+1))$$
$$= A + a \cdot (N'+1) - (A+a) \cdot P(\overline{S}_{N'+1} \mid S_{N'}) \quad \text{(a2)}$$

By the Premise that N' is the largest number *n* that satisfies $a < (A+a) \cdot P(\overline{S}_n \mid S_{n-1})$, and by the fact that $P(\overline{S}_n \mid S_{n-1})$ decreases for *n*, we have

$a \geq (A+a) \cdot P(\overline{S}_{N'+1} \mid S_{N'})$ for $N'+1$.

Then,

$$EC(N', N'+1) = A + a \cdot (N'+1) - (A+a) \cdot P(\overline{S}_{N'+1} \mid S_{N'})$$
$$\geq A + a \cdot (N'+1) - a = A + a \cdot N' = EC(N', N').$$

This implies that (a1) holds.

We can see that the equality in (a1) holds only if $(A+a) \cdot P(\overline{S}_{N'+1} \mid S_{N'}) = a$, otherwise, $EC(N', N'+1) > EC(N', N')$.

However,
even if $EC(N', N'+1) = EC(N', N')$, i.e.,
if $(A+a) \cdot P(\overline{S}_{N'+1} \mid S_{N'}) = a$,

it holds that $(A+a) \cdot P(\overline{S}_{N'+2} \mid S_{N'+1}) < a$ (a3)

because the probability $P(\overline{S}_{n+1} \mid S_n)$ decreases for *n*.

Here, we prove the following Proposition.

Proposition.

$$EC(N', N'+2) > EC(N', N') \quad \text{(a4)}$$

Even if $EC(N', N'+1) = EC(N', N')$.

Proof.

Using the Equation (3),

$$EC(N', N'+2) = P(\overline{S}_{N'+1} \mid S_{N'}) \cdot a \cdot N' + (1 - P(\overline{S}_{N'+1} \mid S_{N'})) \cdot EC(N'+1, N'+2) \quad \text{(a5)}$$

By applying the Equation (3) to $EC(N'+1, N'+2)$ in (a5), we have

$$EC(N'+1, N'+2) = P(\overline{S}_{N'+2} \mid S_{N'+1}) \cdot a \cdot (N'+1) + (1 - P(\overline{S}_{N'+2} \mid S_{N'+1})) \cdot EC(N'+2, N'+2)$$
$$= P(\overline{S}_{N'+2} \mid S_{N'+1}) \cdot a \cdot (N'+1) + (1 - P(\overline{S}_{N'+2} \mid S_{N'+1})) \cdot (A + a \cdot (N'+2))$$
$$= A + a \cdot (N'+2) - (A+a) \cdot P(\overline{S}_{N'+2} \mid S_{N'+1}) \quad \text{(a6)}$$

Using (a3), (a6) and the Equation (2), we have

$EC(N'+1, N'+2) > A + a \cdot (N'+1) = EC(N'+1, N'+1)$.

Thus, $EC(N'+1, N'+2) > EC(N'+1, N'+1)$ (a7)

Moreover, using (a5), (a7) and the Equation (3),

$$EC(N', N'+2) > P(\overline{S}_{N'+1} \mid S_{N'}) \cdot a \cdot N' + (1 - P(\overline{S}_{N'+1} \mid S_{N'})) \cdot EC(N'+1, N'+1)$$
$$= EC(N', N'+1)$$

Then, using (a1) that we have already proved, we obtain $EC(N', N'+2) > EC(N', N')$.

Thus we have proved the above Proposition.

(ii) Accordingly, we can assume that
$EC(N', N'+k) > EC(N', N')$ for $m = k$ (>1).

And, we go on to prove that, for $m = k+1$,
$EC(N', N'+k+1) > EC(N', N')$ holds.

By applying Equation (3) to the above LHS, we have

$$EC(N', N'+k+1) = P(\overline{S}_{N'+1} \mid S_{N'}) \cdot a \cdot N'$$
$$(1 - P(\overline{S}_{N'+1} \mid S_{N'})) \cdot EC(N'+1, N'+k+1) \quad \text{(a8)}$$

Let $\alpha = P(\overline{S}_{N'+1} \mid S_{N'}) \cdot a \cdot N'$, $\beta = (1 - P(\overline{S}_{N'+1} \mid S_{N'}))$,
then we have $0 \leq \alpha < 1$, $0 < \beta \leq 1$, and

$$EC(N', N'+k) = \alpha + \beta \cdot EC(N'+1, N'+k)$$
$$EC(N', N'+k+1) = \alpha + \beta \cdot EC(N'+1, N'+k+1) \quad \text{(a9)}$$

Thus, $EC(N', N'+k+1) > EC(N', N'+k)$ holds if and only if $EC(N'+1, N'+k+1) > EC(N'+1, N'+k)$.

Similarly,
$EC(N'+1, N'+k+1) > EC(N'+1, N'+k)$ if and only if
$EC(N'+2, N'+k+1) > EC(N'+2, N'+k)$.

Consequently, we have the following equivalent relation.
$EC(N', N'+k+1) > EC(N', N'+k)$ if and only if
$EC(N'+k, N'+k+1) > EC(N'+k, N'+k)$ (a10)

By applying the Equation (2) and (3) to the LHS of (a10), we have

$$EC(N'+k, N'+k+1) = P(\bar{S}_{N'+k+1} \mid S_{N'+k}) \cdot a \cdot (N'+k)$$
$$+ (1 - P(\bar{S}_{N'+k+1} \mid S_{N'+k})) \cdot EC(N'+k+1, N'+k+1)$$
$$= P(\bar{S}_{N'+k+1} \mid S_{N'+k}) \cdot \{a \cdot (N'+k) - (A + a \cdot (N'+k+1)\}$$
$$+ A + a \cdot (N'+k+1)$$
$$= A + a \cdot (N'+k+1) - (A+a) \cdot P(\bar{S}_{N'+k+1} \mid S_{N'+k}) \quad \text{(a11)}$$

Since $(A+a) \cdot P(\bar{S}_{N'+k+1} \mid S_{N'+k}) < a$, and by the Equation (2), we have

The last RHS of (a11) $> A + a \cdot (N'+k+1) - a$

$$= A + a \cdot (N'+k) = EC(N'+k, N'+k) \quad \text{(a12)}$$

Thus, it establishes the Inequality (a10), i.e.,
$EC(N'+k, N'+k+1) > EC(N'+k, N'+k)$.

Since this inequality is equivalent to the following
$EC(N', N'+k+1) > EC(N', N'+k)$,
and by the assumption of induction for $m = k$,
$EC(N', N'+k) > EC(N', N')$, we have
$EC(N', N'+k+1) > EC(N', N')$.

This completes the proof of Lemma 3-1.

A2. Lemma 3-2

If $N' < N''$, then for m ($0 < m \le N'$),

$$EC(N'-m, N'') \ge EC(N'-m, N') \quad \text{(b1)}$$

where the equality holds only if $N'' = N'+1$ and
$a = (A+a) \cdot P(\bar{S}_{N'+1} \mid S_{N'})$.

Proof.

We prove this by mathematical induction for m .

(i) If $m = 1$, applying the Equation (3) to the LHS of (b1), we have

$$EC(N'-1, N'') = P(\bar{S}_{N'} \mid S_{N'-1}) \cdot a \cdot (N'-1)$$
$$+ (1 - P(\bar{S}_{N'} \mid S_{N'-1})) \cdot EC(N', N'') \quad \text{(b2)}$$

Using the Lemma 3-1, we have

$$EC(N'-1, N'') \ge P(\bar{S}_{N'} \mid S_{N'-1}) \cdot a \cdot (N'-1)$$
$$+ (1 - P(\bar{S}_{N'} \mid S_{N'-1})) \cdot EC(N', N') \quad \text{(b3)}$$

where the equality holds only if $N'' = N+1$ and
$(A+a) \cdot P(\bar{S}_{N'+1} \mid S_{N'}) = a$.

By the Equation (3), the RHS of (b3) equals to
$EC(N'-1, N')$. Thus we have

$$EC(N'-1, N'') \ge EC(N'-1, N') \quad \text{(b4)}$$

This establishes the Lemma 3-2 for $m = 1$, where the equality holds only if $N'' = N+1$ and
$(A+a) \cdot P(\bar{S}_{N'+1} \mid S_{N'}) = a$.

(ii) Assuming that Lemma 3-2 holds for $m = k$, we prove it for $m = k+1$. By applying the Equation (3) to the LHS of (b1), we have

$$EC(N'-k-1, N'')$$
$$= P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \cdot a \cdot (N'-k-1)$$
$$+ (1 - P(\bar{S}_{N'-k} \mid S_{N'-k-1})) \cdot EC(N'-k, N'') \quad \text{(b5)}$$

By the above assumption,
$EC(N'-k, N'') \ge EC(N'-k, N')$.

Then, from (b5), we have

$$EC(N'-k-1, N'')$$
$$\ge P(\bar{S}_{N'-k} \mid S_{N'-k-1}) \cdot a \cdot (N'-k-1)$$
$$+ (1 - P(\bar{S}_{N'-k} \mid S_{N'-k-1})) \cdot EC(N'-k, N') \quad \text{(b6)}$$

By the Equation (3), the RHS of (b6) equals to
$EC(N'-k-1, N')$. Thus we have

$$EC(N'-k-1, N'') \ge EC(N'-k-1, N') .$$

This establishes the Lemma 3-2 for $m = k+1$, where the equality holds only if $N'' = N+1$ and
$(A+a) \cdot P(\bar{S}_{N'+1} \mid S_{N'}) = a$.

The proof of Lemma 3-2 is completed.

REFERENCES

[1] Hiromichi Kawano, et al (2004), Structural Change Detection in Time Series Based on DP with Action Cost, Proc. of IEEE ISCIT2004, pp.114-119.

[2] Tetsuo Hattori, et al (2010), Early Structural Change Detection as an Optimal Stopping Problem (I), AROB 15th '10.