

SEQUENTIAL DECISION PROBLEM WITH PARTIAL MAINTENANCE ON A PARTIALLY OBSERVABLE MARKOV PROCESS

TORU NAKAI

Received February 22, 2010

ABSTRACT. In the present paper, a sequential decision problem on a partially observable Markov process is set up which takes into account a partial maintenance. We develop an optimal maintenance policy for the products. During their life cycle, a condition of this item changes, which causes some troubles. For a small trouble, it is possible to handle individually, but it might be necessary to replace a faulty component. The decision-maker does not observe a condition directly, but information is obtained through a magnitude of a trouble. A state of an item changes according to a Markovian transition rule based on TP_2 . The decision-maker decides a level of repair with cost which varies with the level. This problem is how much to expend to maintain this item to minimize the total expected cost. A dynamic programming formulation implies a recursive equation about expected cost obtainable under the optimal policy, and the purpose of this paper is to observe monotonic properties for this value.

1 Introduction A sequential decision problem on a Markov process in which states are closely related to outcome is treated in Nakai [11]. In [11], a state can be changed by expending an additional amount within a range of the budget, and it also changes according to a Markovian transition rule based on the total positivity of order two (TP_2). In the present paper, a sequential decision problem on a partially observable Markov process is set up which takes into account a partial maintenance to minimize the total expected cost.

We develop an optimal maintenance policy for the products such as electrical devices, cars and so on. During their life cycle, a condition of this item changes, which causes some troubles. For a small trouble, it is possible to handle individually, but it might be necessary to replace a faulty component. The decision-maker does not observe this condition, but information is obtained through a magnitude of a trouble. This condition is considered as an element of a state space $(0, \infty)$ of a Markov process. These states change according to a Markovian transition rule based on TP_2 , which plays an important role in the Bayesian learning procedure for a partially observable Markov process. For a state $s \in (0, \infty)$, as s approaches to 0, this item complied with user, and it is not sufficiently complied with their demands as s becomes larger. Associated to each state s , there exists a random variable X_s which represents a magnitude of a trouble, and information about unobservable state is obtained through a realized value of this random variable. These X 's are assumed to be *i.i.d.* random variables with finite mean. After observing this value, the decision-maker improves information about unobservable state of the process by employing a Bayesian learning procedure. All information is summarized by a probability distribution on the state space such as a log-normal distribution for example. On the other hand, the decision-maker decides a level of repair or maintenance with cost which varies with this level. This problem is how much to expend to maintain this item to minimize the total expected cost, which is formulated as a sequential decision problem with partial maintenance

2000 *Mathematics Subject Classification.* Primary 90C39, Secondary 90C40.

Key words and phrases. Sequential Decision Problem, Dynamic Programming, Total Positivity, Bayesian Learning Procedure, Partially Observable Markov Process.

on a partially observable Markov process. A dynamic programming formulation implies a recursive equation about expected cost obtainable under the optimal policy, and the purpose of this paper is to observe monotonic properties for this value. A problem treated in Nakai [11] is one of a problem called partially observable Markov decision processes, and a problem treated in section 3 is also one of them. For these problem, there are many researches as Monahan[7, 8, 9], Grosfeld-Nir[4], Albright[1], White[13], Itoh and Nakamura[5], Cao and Guo[2], Ohnishi, Kawai and Mine[12], Fernandez-Gaucherand, Arapostathis and Marcus[3], for example.

2 Sequential Decision Problem on a Partially Observable Markov Process A sequential decision problem with partial maintenance is set up to develop an optimal maintenance policy for some products. During a life cycle of a product, a condition of this item changes, which causes some troubles. This condition is considered as an element of a state space $(0, \infty)$ of a Markov process. For a state $s \in (0, \infty)$, as s approaches to 0, this item complied with user, and it is not sufficiently complied with their demands as s becomes larger. Associated to each state s , there exists a non-negative random variable X_s representing a magnitude of a trouble for this item in state s . These X 's are assumed to be *i.i.d.* random variables with finite mean, and X_s is stochastically increasing random variable with respect to s , *i.e.* a magnitude of a trouble increases stochastically as s becomes larger. It is assumed that the random variable X_s is absolutely continuous with density function $f_s(x)$. A state of this item is not directly observable to the decision-maker, and information about this state will be obtained through a realized value of these random variables.

A state of the process changes according to a Markov process with transition rule $\mathbf{P} = (p_s(t))_{s,t \in (0, \infty)}$, which is independent to the random variables X_s . This transition rule satisfies a property called TP_2 (Assumption 1).

Assumption 1 If $s < t$, then $\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \geq 0$ for any u, v where $u < v$.

Assumption 1 implies that a probability of moving from a current state to 'worse' states increases with deterioration in a current state and decreases with improvement in a current state. From this fact, as s becomes larger, a probability to make a transition into a class of larger values increases.

2.1 Partially Observable Markov Process and Information Information about unobservable state is assumed to be a probability distribution $\boldsymbol{\mu}$ on the state space $(0, \infty)$ with density function $\mu(s)$. Let \mathcal{S} be a set of all information about unobservable state, then

$$\mathcal{S} = \{ \boldsymbol{\mu} = (\mu(s)) \mid \int_0^\infty \mu(s) ds = 1, \mu(s) \geq 0 (s \in (0, \infty)) \}.$$

Among informations in \mathcal{S} , a partial order is defined by using a TP_2 property, *i.e.* for two probability distributions $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ on $(0, \infty)$, if $\mu(s')\nu(s) \geq \mu(s)\nu(s')$ for any $s, s' (s \leq s', s, s' \in (0, \infty))$ and $\mu(s')\nu(s) > \mu(s)\nu(s')$ at least one pair of s and s' , then $\boldsymbol{\mu}$ is said to be larger than $\boldsymbol{\nu}$, or simply $\boldsymbol{\mu} \succ \boldsymbol{\nu}$. By this definition, if $\boldsymbol{\mu} \succ \boldsymbol{\nu}$, then information under $\boldsymbol{\nu}$ is better than one under $\boldsymbol{\mu}$ since a state becomes worse as s increases. This stochastic order is also said to be TP_2 . Concerning this order relation, Lemma 1 is obtained from Kijima and Ohnishi[6].

Lemma 1 If $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ in \mathcal{S} , then $\int_0^\infty h(x) dF_{\boldsymbol{\mu}}(x) \geq \int_0^\infty h(x) dF_{\boldsymbol{\nu}}(x)$ for a non-decreasing non-negative function $h(x)$ of x , where $F_s(x)$ is a probability distribution function of X_s and $F_{\boldsymbol{\mu}}(x) = \int_0^\infty \mu(s) F_s(x)$ is a weighted distribution function.

2.2 Learning Procedure Associated to each state s , there exists a random variable X_s representing a magnitude of a trouble when a state of an item is s . It is assumed that these random variables X_s satisfy Assumption 2.

Assumption 2 *If $s \leq s'$, then $X_{s'} \geq X_s$, i.e. $f_s(x')f_{s'}(x) \leq f_s(x)f_{s'}(x')$ for any x and x' where $x \leq x'$.*

By Assumption 2, if $s \leq s'$, a random variable X_s is larger than $X_{s'}$ by means of the likelihood ratio, and a random variable X_s takes on smaller values as s becomes smaller.

Let prior information about unobservable state of the process be μ , then we improve this information in sequence as follows.

1. Observe a realized value of random variables $\{X_s\}_{s \in (0, \infty)}$.
2. If a realized value is x , improve information by employing the Bayes' theorem as $\mu^x = (\mu^x(s)) \in \mathcal{S}$, i.e.

$$(1) \quad \mu^x(s) = \frac{\mu(s)f_s(x)}{\int_0^\infty \mu(s)f_s(x)ds}$$

3. This process will make a transition to a new state according to the transition rule $P = (p_s(t))_{s,t \in (0, \infty)}$.
4. After making a transition into a new state, information at the next instant becomes $\overline{\mu^x} = (\overline{\mu^x}(s))$ as

$$(2) \quad \overline{\mu^x}(s) = \int_0^\infty \mu^x(t)p_t(s)dt.$$

Regarding a relationship between prior information μ and posterior information $\overline{\mu^x}$, Lemma 2 is obtained under Assumptions 1 and 2 as Nakai[10].

Lemma 2 *If $\mu \succ \nu$, then $\mu^x \succ \nu^x$ and $\overline{\mu^x} \succ \overline{\nu^x}$ for any realized value x . For any μ , μ^x and $\overline{\mu^x}$ increases as x increases, i.e. $\mu^{x'} \succ \mu^x$ and $\overline{\mu^{x'}} \succ \overline{\mu^x}$ where $x < x'$.*

Lemma 2 implies that an order relation among prior information μ is preserved in posterior information μ^x and $\overline{\mu^x}$. Furthermore, for the same prior information μ , posterior information $\overline{\mu^x}$ becomes worse by means of the likelihood ratio as x increases.

3 Sequential Decision Problem with Partial Maintenance A sequential decision problem on a partially observable Markov process is formulated as a partial maintenance is allowed for an item. When a magnitude of a trouble is x , it costs $c(x)$ immediately and the decision-maker decides a level of repair, i.e. she/he chooses a proportion α to maintain this item ($0 < \alpha \leq 1$). If she/he chooses α , a state s changes to a new state αs with cost $C(\alpha)$. This α corresponds to a level of repair for this item, and a cost $C(\alpha)$ varies with the level. When $\alpha = 1$, the decision-maker decides to do nothing at this time and $C(\alpha) = 0$. It is assumed that $C(\alpha)$ is non-increasing and non-negative bounded function of α . Let $u(s)$ be a terminal reward when a state is s , and $u(s)$ is a non-decreasing and non-negative convex function of s . $c(x)$ is a non-negative and non-decreasing function of x . This is one of a problem called partially observable Markov decision problem, and similar problem is treated in Nakai [11] for an additive case.

In order to analyze this sequential decision problem, three models are formulated in sequence. Initially, a sequential decision problem with partial maintenance is formulated

which does not take into account a stochastic transition among states and the decision-maker directly observes a current state. When there are n stages to go and a current state is s , let $v_n^*(s)$ be a total expected cost obtainable under the optimal policy, and let $v_n^*(s|x)$ be a total expected cost obtainable under the optimal policy conditional on x . The principle of the optimality implies following recursive equations.

$$v_n^*(s) = E_{X_s}[v_n^*(s|X_s)]$$

$$(3) \quad v_n^*(s|x) = c(x) + \min_{0 < \alpha \leq 1} \{C(\alpha) + v_{n-1}^*(\alpha s)\},$$

where $v_1^*(s|x) = c(x) + \min_{0 < \alpha \leq 1} \{C(\alpha) + u(\alpha s)\}$. When a state of the process is s and the decision-maker chooses a proportion α to maintain this item ($0 < \alpha \leq 1$), a state is improved as αs with immediate cost $C(\alpha)$ ($0 < \alpha \leq 1$).

Since $u(s)$ is an increasing function of s , $v_1^*(s|x)$ is also an increasing function of s . By using an induction principle on n , $v_n^*(s|x)$ is also an increasing function of s since $v_{n-1}^*(\alpha s)$ is an increasing function of s . This property implies that $v_n^*(s)$ is also an increasing function of s . Since $C(\alpha) + u(\alpha s)$ is a convex function of α , an optimal decision for $n = 1$ is obtained as a minimizing point of an equation $C(\alpha) + u(\alpha s) = 0$ ($0 < \alpha \leq 1$).

Secondly, a state changes according to a Markov process and the decision-maker directly observes a current state, *i.e.* it is a sequential decision problem with partial maintenance on a Markov process. When there are n stages to go and a current state is s , let $\hat{v}_n(s)$ be a total expected cost obtainable under the optimal policy, and let $\hat{v}_n(s|x)$ be a total expected cost obtainable under the optimal policy conditional on a realized value x . The principle of the optimality implies the recursive equations as

$$\hat{v}_n(s) = E_{X_s}[\hat{v}_n(s|X_s)]$$

$$(4) \quad \hat{v}_n(s|x) = c(x) + \min_{0 < \alpha \leq 1} \{C(\alpha) + \int_0^\infty p_{\alpha s}(t)\hat{v}_{n-1}(t)dt\},$$

where $\hat{v}_1(s|x) = c(x) + \min_{0 < \alpha \leq 1} \{C(\alpha) + \hat{v}(\alpha s)\}$. From these equations, it is easy to show Lemma 3 by using an induction principle on n and Lemma 1.

Lemma 3 $\hat{v}_n(s)$ is a non-decreasing function of s , and $\hat{v}_n(s|x)$ is a non-decreasing function of s and x .

Finally, a sequential decision problem with partial maintenance is formulated as a partially observable Markov decision problem, *i.e.* it is possible to change a current state by making a decision α and the decision-maker cannot observe a current state directly. Let μ be prior information about unobservable state of the process. The decision-maker improve information about unobservable state by using a realized value of a random variable X_s . When prior information is μ , let μ^x be posterior information improved by a Bayesian learning procedure defined by Equation (1). For this improved information, the decision-maker chooses a decision α and a state becomes αs when a current state is s . Therefore, information about state of the process becomes μ_α^x by this decision. After that, time moving forward by one unit, and this process will make a transition to a new state according to the transition rule $(p_s(t))_{s,t \in (0,\infty)}$, and information becomes $\overline{\mu}_\alpha^x$ at the next instant. It is also possible to formulate and analyze this model by other order similarly.

When information about unobservable state of the process is μ , let $v_n(\mu)$ be a total expected cost obtainable under the optimal policy when there are n stages to go. Conditional

on x , $v_n(\boldsymbol{\mu}|x)$ is a total expected cost obtainable under the optimal policy when there are n stages to go and information is $\boldsymbol{\mu}$. The principle of optimality implies a recursive equation as

$$(5) \quad \begin{aligned} v_n(\boldsymbol{\mu}) &= \int_0^\infty v_n(\boldsymbol{\mu}|x) dF_{\boldsymbol{\mu}}(x) \\ v_n(\boldsymbol{\mu}|x) &= c(x) + \min_{0 < \alpha \leq 1} \{C(\alpha) + v_{n-1}(\overline{\boldsymbol{\mu}}_\alpha^x)\} \end{aligned}$$

where $v_0(\boldsymbol{\mu}) = \int_0^1 u(t) d\boldsymbol{\mu}(t)$. In order to treat monotonic properties of $v_n(\boldsymbol{\mu})$ and $v_n(\boldsymbol{\mu}|x)$, some preliminary properties will be considered in subsequent section.

3.1 Preliminary Results Concerning Prior and Posterior Information In Nakai [11], a sequential expenditure problem on a partially observable Markov process with state space $(-\infty, \infty)$ is formulated, in which some conditions are treated to obtain monotonic properties about the total expected reward obtainable under the optimal policy. In a problem treated in [11], when a current state is s , a state of this process moves to a new state $s(y) = s + d(y)$ by taking a decision $y (> 0)$, where $d(y)$ is non-decreasing function of y with $d(0) = 0$. In order to observe some properties concerning an optimal value for a problem treated here, we summarize several properties concerning prior and posterior information obtained in Nakai [11]. We introduce following notations for prior and posterior informations as [11].

$\boldsymbol{\mu}$: probability distribution $\boldsymbol{\mu} = (\mu(s))$ on the state space $(-\infty, \infty)$ as prior information

$\overline{\boldsymbol{\mu}}$: probability distribution $\overline{\boldsymbol{\mu}} = (\overline{\mu}(s))$ after making a transition to a new state where $\overline{\mu}(s) = \int_0^\infty \mu(t) p_t(s) dt$

$\boldsymbol{\mu}^x$: probability distribution $\boldsymbol{\mu}^x = (\mu^x(s))$ improved by using a realized value x according to the Bayes' theorem as Eq. (1)

$\boldsymbol{\mu}_y$: probability distribution $\boldsymbol{\mu}_y = (\mu_y(s))$ after taking a decision with y as $\mu_y(s) = \mu(s - d(y))$

For a probability distribution on the state space $(-\infty, \infty)$ like information $\boldsymbol{\mu} = (\mu(s))$, consider a condition called (G') as follows.

(Condition G')

A probability density function $\mu(s)$ on $(-\infty, \infty)$ satisfies $\frac{\mu(s)}{\mu(s')} \geq \frac{\mu(t)}{\mu(t')}$ for any $s < t$ and $s' < t'$ where $s - s' = t - t' < 0$.

This condition is called as a gradually condition in Nakai [11]. Let $\boldsymbol{\mu}$ be a normal distribution, then this $\boldsymbol{\mu}$ satisfies Condition G' by simple calculations. It is easy to show that if $\boldsymbol{\mu}$ satisfies Condition G' then $\boldsymbol{\mu}_y$ satisfies Condition G' for any y . In order to show Lemma 4, Assumption 3 is induced for the transition probability. Whenever $p_s(t)$ is a density function of normal distribution $N(s, \sigma^2)$, this $p_s(t)$ satisfies this assumption.

Assumption 3 *If $u < v$, then $p_u(s)p_v(t') - p_u(t)p_v(s') \geq p_v(s)p_u(t') - p_v(t)p_u(s')$ for any $s < s'$ and $t < t'$.*

Lemma 4 *Under Assumption 3, if $\boldsymbol{\mu}$ satisfies Condition G' then $\overline{\boldsymbol{\mu}}$ also satisfies Condition G' .*

As for posterior information μ^x improved by a Bayesian learning procedure for any x , Assumption 4 is induced for density functions of X_s to show Lemma 5.

Assumption 4 A probability density function $f_s(x)$ of a random variable X_s satisfies $\frac{f_s(x)}{f_{s'}(x)} \geq \frac{f_t(x)}{f_{t'}(x)}$ for any $s < t$ and $s' < t'$ where $s - s' = t - t' < 0$.

If $f_s(x)$ is a density function of a normal distribution $N(s, \sigma^2)$, then these $f_s(x)$ satisfy Assumption 4.

Lemma 5 Under Assumption 4, if μ satisfies Condition G' , then μ^x satisfies Condition G' for any x .

In the discussions in [11], order of learning about unobservable state, making a decision and a transition among state according to the transition rule is considered as follows.

1. Observe a realized value x of the random variable.
2. Improve information about it as μ^x by using the Bayes' theorem.
3. Expend an additional amount y within a range of a budget, *i.e.* take a decision y .
4. Time moving forward by one unit.
5. Process will make a transition to a new state according to \mathbf{P} , and information about new state becomes $\overline{\mu}_y^x = (\overline{\mu}_y^x(s))$.

It is also possible to formulate and analyze this problem by other order. According to Nakai [11], some monotonic properties concerning the relationships between prior and posterior information are obtained in Lemma 6 by Lemmas 2, 4 and 5.

Lemma 6 Let μ and ν be prior information in \mathcal{S} which satisfy Condition G' . If $y > y'$, then $\mu_y^x \succeq \mu_{y'}^x$ and $\overline{\mu}_y^x \succeq \overline{\mu}_{y'}^x$ for any x , and if $\mu \succeq \nu$, then $\mu_y \succeq \nu_y$, $\mu_y^x \succeq \nu_y^x$ and $\overline{\mu}_y^x \succeq \overline{\nu}_y^x$ for any y . If $x > x'$, then $\mu_y^x \succeq \mu_{y'}^{x'}$ and $\overline{\mu}_y^x \succeq \overline{\mu}_{y'}^{x'}$ for any y .

3.2 Log-normal Distribution and Information Since the state space is $(0, \infty)$ in this sequential decision problem, a log-normal distribution is a typical distribution for information μ , random variables X_s and the transition rule \mathbf{P} . By this reason, essential properties concerning log-normal distribution are summarized in this subsection.

Let X be a log-normal distribution with distribution function $\Psi(x|\mu, \sigma^2)$ and density function $\psi(x|\mu, \sigma^2)$ where

$$\Psi(x|\mu, \sigma^2) = \int_{-\infty}^{\log x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \quad \psi(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}},$$

and, therefore, $E[X] = \mu + \frac{\sigma^2}{2}$ and $\psi(x|\mu, \sigma^2) = \frac{\phi(\log x|\mu, \sigma^2)}{x}$ for a density function $\phi(x|\mu, \sigma^2)$ of a normal distribution $N(\mu, \sigma^2)$.

For density functions $\psi(x|\mu, \sigma_1^2)$ and $\psi(x|\mu, \sigma_2^2)$ of log-normal distribution with $\sigma_2^2 \leq \sigma_1^2$,

$$\begin{vmatrix} \psi(s|u, \sigma_1^2) & \psi(s|v, \sigma_2^2) \\ \psi(t|u, \sigma_1^2) & \psi(t|v, \sigma_2^2) \end{vmatrix} = \frac{1}{st} \begin{vmatrix} \phi(\log s|u, \sigma_1^2) & \phi(\log s|v, \sigma_2^2) \\ \phi(\log t|u, \sigma_1^2) & \phi(\log t|v, \sigma_2^2) \end{vmatrix} \geq 0$$

for any $s, t > 0$, since $\begin{vmatrix} \phi(x|\mu, \sigma_1^2) & \phi(x|\mu, \sigma_2^2) \\ \phi(x'| \nu, \sigma_1^2) & \phi(x'| \nu, \sigma_2^2) \end{vmatrix} \geq 0$ for any $x, x', \mu, \nu, \sigma_1$ and σ_2 where $x \leq x', \mu \leq \nu$ and $\sigma_2^2 \leq \sigma_1^2$, *i.e.* normal distributions $N(\mu, \sigma^2)$ has a TP_2 property.

Therefore, the log-normal distributions with density function $\psi(x|\mu, \sigma^2)$ have a property called TP₂ when σ^2 is a non-increasing function of μ .

In a sequential decision problem treated in Nakai [11], a state becomes $s(y) = s + d(y)$ by taking a decision $y (> 0)$ when a state of the process is s , and, therefore, some monotonic properties are obtained under Condition G'. For this problem, a state of the process becomes αs by taking a proportion α when a state of the process is s , *i.e.* $s(\alpha) = \alpha s$. From this reason, Condition G is induced for probability distributions on the state space $(0, \infty)$ like μ in \mathcal{S} to analyze monotonic properties for this problem.

(Condition G)

A probability density function $\mu(s)$ satisfies $\frac{\mu(s)}{\mu(s')} \geq \frac{\mu(t)}{\mu(t')}$ for any $0 < s < t$ and $0 < s' < t'$ where $\frac{s}{s'} = \frac{t}{t'} < 1$.

Since information is defined as a probability distribution μ on the state space $(0, \infty)$, if μ is a log-normal distribution with density function $\psi(x|\mu, \sigma^2)$, then this μ satisfies Condition G since $\frac{\mu(s)}{\mu(s')} \geq \frac{\mu(t)}{\mu(t')}$ is equivalent to $\frac{\phi(\log \alpha + \log s'|\mu, \sigma^2)}{\phi(\log s'|\mu, \sigma^2)} \geq \frac{\phi(\log \alpha + \log t'|\mu, \sigma^2)}{\phi(\log t'|\mu, \sigma^2)}$ for $s = \alpha s'$ and $t = \alpha t'$ where $0 < \alpha \leq 1$. This comes from a fact that a density function $\phi(x|\mu, \sigma^2)$ of $N(\mu, \sigma^2)$ satisfies the condition G'.

3.3 Relationship between Prior and Posterior Information All information about unobservable state of the process is summarized by a probability distribution on the state space $(0, \infty)$, and a state moves to a new state αs by taking a decision α . In the present paper, we use the following notations instead of Nakai [11].

μ : probability distribution on $(0, \infty)$ as prior information

$\bar{\mu}$: probability distribution after making a transition to a new state according to the transition rule $\mathbf{P} = (p_s(t))_{s,t \in (0, \infty)}$

μ^x : probability distribution improved by employing a Bayesian learning procedure after observing a magnitude x of a trouble

μ_α : probability distribution on $(0, \infty)$ after taking a decision α where $0 < \alpha \leq 1$

In this problem, whenever prior information in μ , first observe a magnitude x of a trouble as information and improve information about unobservable state as μ^x by employing the Bayes' theorem. After that, the decision-maker chooses a decision α , and information about unobservable state becomes μ_α^x . Finally, this process will make a transition to a new state according to $\mathbf{P} = (p_s(t))_{s,t \in (0, \infty)}$, and, at the next instant, information about unobservable state becomes $\bar{\mu}_\alpha^x$.

Under Assumptions 5 and 6 instead of Assumptions 3 and 4, if μ satisfies Condition G then $\bar{\mu}$, μ^x and μ_α also satisfy Condition G by a method similar to one used in Lemmas 4 and 5, *i.e.* Lemmas 18 and 20 of Nakai [11]. A transition rule $(p_s(t))_{s,t \in (0, \infty)}$ which satisfies Assumption 3 also satisfies Assumption 5. It is easy to show that μ_α satisfies Condition G for any α whenever μ satisfies Condition G.

Assumption 5 *If $u < v$, then $p_u(s)p_v(t') - p_u(t)p_v(s') \geq p_v(s)p_u(t') - p_v(t)p_u(s')$ for any $0 < s < t$ and $0 < s' < t'$ where $\frac{s}{s'} = \frac{t}{t'} < 1$.*

Assumption 6 *A probability density function $f_s(x)$ of X_s ($s \in (0, \infty)$) satisfies a condition that $\frac{f_s(x)}{f_{s'}(x)} \geq \frac{f_t(x)}{f_{t'}(x)}$ for any $0 < s < t$ and $0 < s' < t'$ where $\frac{s}{s'} = \frac{t}{t'} < 1$.*

Lemma 7 *Under Assumption 5, if μ satisfies Condition G then $\bar{\mu}$ also satisfies Condition G. Under Assumption 6, if μ satisfies Condition G, then μ^x satisfies Condition G for any x .*

When a probability distribution ($p_s(t)$) for any $0 < s < \infty$ is a log-normal distribution on the state space with density function $p_s(t) = \psi(t|\log s, \sigma^2)$, these random variables satisfy Assumption 3 since $\psi(\alpha x|\mu, \sigma^2) = \frac{\phi(\log \alpha x|\mu, \sigma^2)}{\alpha x}$ and a density function of normal distribution $N(s, \sigma^2)$ satisfies Assumption 3, and, therefore, these random variables satisfy Assumption 5. On the other hand, let T_s be a random variable with density function $p_s(t) = \psi(t|\log s - \sigma^2, 2\sigma^2)$ for any $0 < s < \infty$, then these ($p_s(t)$) $_{s,t \in (0, \infty)}$ also satisfy Assumption 5 and $E[T_s] = s$. When the random variables X_s are assumed to be a log-normal distribution with density function $\psi(\alpha x|\mu, \sigma^2)$, then these random variables satisfy Assumption 4 since a density function of normal distribution $N(s, \sigma^2)$ satisfies Assumption 4.

On the other hand, let a probability distribution ($p_s(t)$) for any $0 < s < \infty$ be a log-normal distribution where $p_s(t) = \psi(t|\log s - \sigma^2, 2\sigma^2)$ and let X_s be a log-normal distribution with density function $\psi(x|\log s, \bar{\sigma}^2)$, then $\bar{\mu}$, μ^x and μ_α are also log-normal distributions by simple calculations for log-normal prior information μ with $\psi(s|\mu, \hat{\sigma}^2)$, and, therefore, Lemma 7 is valid for this case.

Concerning the relationship between prior and posterior information, the following monotonic properties are derived by a method similar to one used in Nakai [11].

Lemma 8 *Let μ be information which satisfies Condition G. If $1 \geq \alpha > \beta > 0$, then $\mu_\alpha \succeq \mu_\beta$, $\mu_\alpha^x \succeq \mu_\beta^x$ and $\bar{\mu}_\alpha^x \succeq \bar{\mu}_\beta^x$ for any observation x .*

Lemma 9 *Let μ and ν be information in \mathcal{S} . If $\mu \succeq \nu$, then $\mu_\alpha \succeq \nu_\alpha$, $\mu_\alpha^x \succeq \nu_\alpha^x$ and $\bar{\mu}_\alpha^x \succeq \bar{\nu}_\alpha^x$ for any decision α and observation x .*

Lemma 10 *Let μ be information in \mathcal{S} . If $x > x'$, then $\mu_\alpha^x \succeq \mu_\alpha^{x'}$ and $\bar{\mu}_\alpha^x \succeq \bar{\mu}_\alpha^{x'}$ for any decision α ($0 < \alpha \leq 1$).*

3.4 Monotonic Property of a Sequential Decision Problem with Partial Maintenance Finally, we will treat monotonic properties of a sequential decision problem with partial maintenance where a state changes according to a partially observable Markov process.

Associated to each state s , there exists a random variable X_s which represents a magnitude of a trouble when a state is s , and to observe a realized value of these random variables corresponds to an observation process for this partially observable Markov process. Let μ be prior information as a probability distribution on the state space $(0, \infty)$ which satisfies Condition G.

This sequential decision problem with partial maintenance is formulated as follows.

1. First observe a magnitude x of a trouble as information, which is a realized value of random variable X_s .
2. According to x , improve information by employing the Bayes' theorem as a learning procedure, and let information about unobservable state be μ^x .
3. For this information μ^x , the decision-maker chooses a decision α , and information about unobservable state becomes μ_α^x .
4. Time moving forward by one unit, and this process will make a transition to a new state according to ($p_s(t)$) $_{s,t \in (0, \infty)}$, and information about new state becomes $\bar{\mu}_\alpha^x$.

Let $v_n(\boldsymbol{\mu})$ be a total expected cost obtainable under the optimal policy when there are n stages to go and information about unobservable state is $\boldsymbol{\mu}$. Conditional on x , $v_n(\boldsymbol{\mu}|x)$ is a total expected cost obtainable under the optimal policy when there are n stages to go and information is $\boldsymbol{\mu}$. By the principle of optimality, this $v_n(\boldsymbol{\mu})$ satisfies Equation (5) as

$$\begin{aligned} v_n(\boldsymbol{\mu}) &= \int_0^\infty v_n(\boldsymbol{\mu}|x) dF_{\boldsymbol{\mu}}(x) \\ v_n(\boldsymbol{\mu}|x) &= c(x) + \min_{0 < \alpha \leq 1} \{C(\alpha) + v_{n-1}(\overline{\boldsymbol{\mu}}_\alpha^x)\} \end{aligned}$$

with $v_0(\boldsymbol{\mu}) = \int_0^1 u(t) d\boldsymbol{\mu}(t)$. Whenever prior information is $\boldsymbol{\mu}$; first observe a magnitude x of a trouble; according to x , improve information by employing the Bayesian learning procedure and let information about unobservable state be $\boldsymbol{\mu}^x$; for this information, the decision-maker chooses a decision α , and information about unobservable state becomes $\boldsymbol{\mu}_\alpha^x$; time moving forward by one unit, and this process will make a transition to a new state according to $(p_s(t))_{s,t \in (0, \infty)}$, and information about new state becomes $\overline{\boldsymbol{\mu}}_\alpha^x$; after that, the expected total cost obtainable under the optimal policy is $v_{n-1}(\overline{\boldsymbol{\mu}}_\alpha^x)$ by the principle of the optimality. It is also possible to formulate and analyze this model by other order similarly.

By the monotonic property between prior and posterior information in Lemmas 8 to 10, when $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ satisfy Condition G, if $\boldsymbol{\mu} \succ \boldsymbol{\nu}$, then $\overline{\boldsymbol{\mu}}_\alpha^x \succeq \overline{\boldsymbol{\nu}}_\alpha^x$ for any decision α and realized value x . This fact and Lemma 1 imply Property 1 by the induction principle on n .

Property 1 Suppose that $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ satisfy Condition G. If $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$, then $v_n(\boldsymbol{\mu}) \geq v_n(\boldsymbol{\nu})$ and $v_n(\boldsymbol{\mu}|x) \geq v_n(\boldsymbol{\nu}|x)$, i.e. $v_n(\boldsymbol{\mu})$ and $v_n(\boldsymbol{\mu}|x)$ is a non-decreasing function of $\boldsymbol{\mu}$.

In the present paper, a sequential decision problem with partial maintenance is formulated as a partially observable Markov decision process. Especially, monotonic properties between prior and posterior information are obtained based on the TP₂ property and Condition G. The problem is how much proportion to spend to maintain and to improve this item. Under some assumptions about the transition rule and information process, there exist some monotonic properties concerning the expected value obtainable under the optimal policy as Proposition 1.

Acknowledgement This research was partially supported by the Grant-in-Aid for Scientific Research of the Japan Society for the Promotion of Science and Technology.

REFERENCES

- [1] S. C. Albright: Structural results for partially observable Markov decision processes. *Oper. Res.* 27 (1979), 1041–1053.
- [2] X. Cao and X. Guo: Partially observable Markov decision processes with reward information: basic ideas and models. *IEEE Trans. Automat. Control* 52 (2007), 677–681.
- [3] E. Fernandez-Gaucherand, A. Arapostathis and S. I. Marcus: On the average cost optimality equation and the structure of optimal policies for partially observable Markov decision processes. *Ann. Oper. Res.* 29 (1991), 439–469.
- [4] A. Grosfeld-Nir: A two-state partially observable Markov decision process with uniformly distributed observations. *Oper. Res.* 44 (1996), 458–463.
- [5] H. Itoh and K. Nakamura: Partially observable Markov decision processes with imprecise parameters. *Artificial Intelligence* 171 (2007), 453–490.
- [6] M. Kijima and M. Ohnishi: Stochastic Orders and Their Applications in Financial Optimization, *Math. Methods of Oper. Res.*, 50, 351–372, (1999).

- [7] G. E. Monahan: Optimal stopping in a partially observable Markov process with costly information. *Oper. Res.* 28 (1980), 1319–1334.
- [8] G. E. Monahan: Optimal stopping in a partially observable binary-valued Markov chain with costly perfect information. *J. Appl. Probab.* 19 (1982), 72–81.
- [9] G. E. Monahan: Optimal selection with alternative information. *Naval Res. Logist. Quart.* 33 (1986), 293–307.
- [10] T. Nakai: A Generalization of Multivariate Total Positivity of Order Two with an Application to Bayesian Learning Procedure, *J. of Inf. & Opt. Sci.*, **23**, 163–176, (2002).
- [11] T. Nakai: A Sequential Decision Problem based on the Rate Depending on a Markov Process, *Recent Advances in Stochastic Operations Research 2* (Eds. T. Dohi, S. Osaki and K. Sawaki), World Scientific Publishing, 11–30, 2009.
- [12] M. Ohnishi, H. Kawai and H. Mine: An optimal inspection and replacement policy under incomplete state information. *European J. Oper. Res.* 27 (1986), 117–128.
- [13] D. J. White: Structural properties for contracting state partially observable Markov decision processes. *J. Math. Anal. Appl.* 186 (1994), 486–503

Department of Mathematics, Faculty of Education, Chiba University, Inage, Chiba 263-8522, Japan.

e-mail: t-nakai@faculty.chiba-u.jp