電気学会論文誌C（電子・情報・システム部門誌）
IEEJ Transactions on Electronics, Information and Systems
Vol.132 No.10 pp.1719–1720 DOI: 10.1541/ieejeiss.132.1719

研究開発レター

パラメータ未知のマルコフ決定過程の保全への適用

非会員 前田 康成* 非会員 後藤文太朗* 非会員 升井 洋志*
非会員 桝井 文人* 非会員 鈴木 正清*

Applying Markov Decision Processes with Unknown Parameters to Maintenance

Yasunari Maeda*, Non-member, Fumitaro Goto*, Non-member, Hiroshi Masui*, Non-member, Fumito Masui*, Non-member, Masakiyo Suzuki*, Non-member

（2012年1月5日受付，2012年4月26日再受付）

In this research we apply Markov decision processes with unknown parameters to maintenance problem. We propose a new maintenance method which maximizes total reward with reference to a Bayes criterion.

キーワード：保全，マルコフ決定過程，未知パラメータ，統計的決定理論，ベイズ基準
Keywords: maintenance, Markov decision processes, unknown parameter, statistical decision theory, Bayes criterion

1. まえがき

本研究では，製造工場の設備の保全問題を扱う。従来研究(1)~(3)ではマルコフ決定過程（MDP）(4)を用いてモデル化し，設備の状態が既知の場合(1)(2)と，未知の場合(3)を検討している。本研究では，既知の場合を検討する。従来研究(1)~(3)では，MDPの真のパラメータが既知の場合のコストの最小化を行っている。しかし，実際には真のパラメータは未知であり，何らかの推定値をパラメータ既知の場合の最適な方法に代入しても最適性は保証されるとは限らない。また，実際の経営現場では，コスト以外に設備状態に依存した収益（出荷製品数の変動などによる）も考慮する必要がある。

そこで，本研究では収益も考慮した真のパラメータ未知のMDPを用いて，統計的決定理論(5)に基づいて定式化し，有限期間の総利得をベイズ基準のもとで最大化するという意味で最適な保全計画を求める方法を提案する。

2. 準 備

ここでは，本研究で用いる各種記号などの定義を行う。

$s_i \in S$ は設備の状態を示し，MDPの状態に相当する。$S = \{s_1, s_2, \ldots, s_{|S|}\}$ は状態集合である。$a_i \in A$ は修理を示し，MDPの行動に相当する。$A = \{a_1, a_2, \ldots, a_{|A|}\}$ は修理（行動）集合である。$m(a_i)$ は修理 a_i のコスト，$o(s_i)$ は状態 s_i の設備の1期間の稼働コスト，$b(s_i)$ は状態 s_i の設備の1期間の製造（稼働）によって得られる収益を示す。

$p(s_k \mid s_i, a_j, \theta^*)$ は状態 s_i の設備に修理 a_j を実施して状態 s_k に遷移する保全状態遷移確率を示す。修理は各期の最初に状態観測後に所要時間無しで実施可能で，修理前後の状態はともに既知（観測可能）とする。$\theta^* \in \Theta$ は真のパラメータで未知である。Θ は連続パラメータの集合で既知である。$p(s_j \mid s_i, \psi^*)$ は状態 s_i の設備が1期間稼働後に状態 s_j に遷移する稼働状態遷移確率を示す。$\psi^* \in \Psi$ は真のパラメータで未知である。Ψ は連続パラメータの集合で既知である。保全状態遷移確率と稼働状態遷移確率がMDPの状態遷移確率に相当する。

$x_t \in S$ は t 期の最初に観測される状態を示す確率変数，$x'_t \in S$ は t 期の修理後の状態を示す確率変数である。ただし，$x_1 \in S$ は初期状態を示す定数で既知とする。$y_t \in A$ は t 期に実施（選択）する修理（行動）を示す。$r(x_t, y_t, x'_t, x_{t+1}) = -m(y_t) - o(x'_t) + b(x'_t)$ は状態 x_t で修理 y_t を実施し状態 x'_t へ遷移し，さらに1期間の稼働後に状態 x_{t+1} に遷移した際に得られるMDPの t 期の利得である。T は有限の既知の定数で，本研究では T 期間の総利得の最大化を目的とする。

本研究で扱うMDPは S，A，θ^*，ψ^*，$r(\cdot,\cdot,\cdot,\cdot)$ で構成される。各期の最初に状態 x_t 観測後に修理 y_t を実施し，修理後の状態 x'_t のもとで1期間稼働する。これを1期から T 期まで繰り返す。次章では，総利得の最大化を統計的決定理論に基づいて定式化し，ベイズ基準のもとで最大化する。

3. 収益を考慮したベイズ最適な保全

〈3·1〉 定 式 化 統計的決定理論に基づいて定式化する。効用関数 $U(d(\cdot,\cdot,\cdot), x^{T+1}y^T x'^T, \theta^*, \psi^*)$ を次式で定義する。

$$U(d(\cdot,\cdot,\cdot), x^{T+1}y^T x'^T, \theta^*, \psi^*) = \sum_{t=1}^{T} r(x_t, y_t, x'_t, x_{t+1})$$
$$= \sum_{t=1}^{T} (-m(y_t) - o(x'_t) + b(x'_t)), \cdots\cdots (1)$$

* 北見工業大学情報システム工学科
〒090-8507 北海道北見市公園町165
Dept. of Computer Science, Kitami Institute of Technology
165, Koen-cho, Kitami-shi, Hokkaido 090-8507, Japan

ただし，$d(\cdot,\cdot,\cdot)$ は状態とそれまでの遷移系列と期を受け取って，選択する行動を返す決定関数で $y_t = d(x_t, x^t y^{t-1} x'^{t-1}, t) \in A$，$x^{T+1} y^T x'^T$ は系列 $x_1 y_1 x'_1 x_2 y_2 x'_2 \cdots x_{T+1}$ を示す。この効用関数は真のパラメータ θ^*，ψ^* のもとで，決定関数 $d(\cdot,\cdot,\cdot)$ を用いて，$x^{T+1} y^T x'^T$ と遷移した場合の総利得である。

次に期待効用 $EU(d(\cdot,\cdot,\cdot),\theta^*,\psi^*)$ を次式で定義する。

$$EU(d(\cdot,\cdot,\cdot),\theta^*,\psi^*) = \sum_{t=1}^{T} \sum_{x^{t+1} y^t x'^t} p(x^{t+1} y^t x'^t \mid d(\cdot,\cdot,\cdot),\theta^*,\psi^*)$$
$$(-m(y_t) - o(x'_t) + b(x'_t)), \cdots\cdots\cdots\cdots\cdots\cdots (2)$$

ただし，期待効用は真のパラメータ θ^*，ψ^* のもとで，ある決定関数 $d(\cdot,\cdot,\cdot)$ を用いた場合の総利得の期待値である。

パラメータの事前分布 $p(\theta)$，$p(\psi)$ に対して期待値をとるベイズ期待効用 $BEU(d(\cdot,\cdot,\cdot), p(\theta), p(\psi))$ を次式で定義する。

$$BEU(d(\cdot,\cdot,\cdot), p(\theta), p(\psi))$$
$$= \int_\Theta \int_\Psi p(\theta) p(\psi) EU(d(\cdot,\cdot,\cdot)),\theta^*,\psi^*) d\psi d\theta. \cdots (3)$$

(3) 式のベイズ期待効用を最大にする決定関数がベイズ基準のもとで総利得を最大にするという意味で最適な保全計画に相当する。(3) 式のベイズ期待効用を書き下すと，真のパラメータ未知の一般的な MDP の場合 [6] と同様に T 期間の入れ子構造になる。この入れ子構造に動的計画法（DP）を適用し，最適な保全計画を求める方法を次節で提案する。

〈3·2〉 提案方法　DP を用いて，T 期目から遡りながら計算して，ベイズ最適な保全計画を求める。

T 期目のすべての状態 x_T と遷移系列 $x^T y^{T-1} x'^{T-1}$ の組に対する処理は以下のとおりである。

$$V(x_T, x^T y^{T-1} x'^{T-1}, T) = \max_{y_T \in A} \Big(-m(y_T) + \sum_{x'_T \in S} \int_\Theta p(\theta \mid x^T y^{T-1} x'^{T-1})$$
$$p(x'_T \mid x_T, y_T, \theta) d\theta \big(-o(x'_T) + b(x'_T) \big) \Big), \cdots\cdots\cdots (4)$$

ただし，$p(\theta \mid x^T y^{T-1} x'^{T-1})$ は 1 期から T 期の遷移系列が $x^T y^{T-1} x'^{T-1}$ の場合の事後分布で，$V(x_T, x^T y^{T-1} x'^{T-1}, T)$ は最後の期間の期待利得の最大値である。ベイズ最適な決定は次式となる：$d^*(x_T, x^T y^{T-1} x'^{T-1}, T) = argV(x_T, x^T y^{T-1} x'^{T-1}, T)$.

t 期目（$1 \le t \le T-1$）のすべての状態 x_t と遷移系列 $x^t y^{t-1} x'^{t-1}$ の組に対する処理は以下のとおりである。

$$V(x_t, x^t y^{t-1} x'^{t-1}, t) = \max_{y_t \in A} \Big(-m(y_t) + \sum_{x'_t \in S} \int_\Theta p(\theta \mid x^t y^{t-1} x'^{t-1})$$
$$p(x'_t \mid x_t, y_t, \theta) d\theta \big(-o(x'_t) + b(x'_t) + \sum_{x_{t+1} \in S} \int_\Psi p(\psi \mid x^t y^{t-1} x'^{t-1})$$
$$p(x_{t+1} \mid x'_t, \psi) d\psi V(x_{t+1}, x^{t+1} y^t x'^t, t+1) \big) \Big), \cdots\cdots\cdots (5)$$

ただし，$V(x_t, x^t y^{t-1} x'^{t-1}, t)$ は t 期以降の総利得の期待値の最大値で，ベイズ最適な決定は次式となる：$d^*(x_t, x^t y^{t-1} x'^{t-1}, t) = argV(x_t, x^t y^{t-1} x'^{t-1}, t)$.

上記のように T 期から遡りながら DP を適用し，T 期間のベイズ最適な保全計画を求める。(4) 式と (5) 式中には積分計算があるが，事前分布としてディリクレ分布を仮定すると，積分計算部分は簡易な四則演算で実施できる [7]。

〈3·3〉 シミュレーション結果　提案方法の有効性を確認するため，シミュレーション結果を報告する。

$S = \{s_1, s_2\}$，$A = \{a_1, a_2\}$，$x_1 = s_1$，$T = 4$，$m(a_1) = 0$，$m(a_2) = 100$，$o(s_1) = 10$，$o(s_2) = 50$，$b(s_1) = 400$，$b(s_2) = 100$，$p(s_1 \mid s_1, a_1, \theta^*) = 1.0$，$p(s_1 \mid s_1, a_2, \theta^*) = 1.0$，$p(s_1|s_2, a_1, \theta^*) = 0.0$，$p(s_1|s_2, a_2, \theta^*) = 0.99$，$p(s_1|s_1, \psi^*) = 0.8$，$p(s_1|s_2, \psi^*) = 0.0$ という条件のもとで，提案方法による保全計画とコスト最小化の保全計画（提案方法の (4) 式，(5) 式から $b(x'_T)$，$b(x'_t)$ を削除した場合に相当）を求めた。なお，事前分布はジェフリーズの事前分布の考え方 [5][7] に従って設定した。10000 回のシミュレーション結果から T 期間の平均総利得（総収益と総コストの差分の平均に相当）と平均総コストを算出した。提案方法による保全計画の平均総利得は 1254.178，平均総コストは 93.732 だった。コスト最小化の保全計画の平均総利得は 1201.538，平均総コストは 82.172 だった。このように，総利得最大化の保全計画と総コスト最小化の保全計画は同一ではない。

一例ではあるが，上記のシミュレーション結果より，保全問題において収益を考慮することの有効性と，提案方法の有効性が確認できた。

4. まとめと今後の課題

本研究では，製造工場における設備の保全問題を扱った。MDP を用いた従来研究は多くあるが，真のパラメータ未知の場合が未検討である。また，最適化の対象も従来研究ではコストのみで，収益は考慮されていない。そこで，本研究では最適化の対象を収益も含めるように拡張したもとで，統計的決定理論に基づいて，真のパラメータ未知の MDP の有限期間の総利得をベイズ基準のもとで最大化する問題として定式化した。また，DP を用いてベイズ最適な保全計画を求める方法を提案した。さらに，一例ではあるがシミュレーションによって，保全問題において収益も考慮することの有効性と，提案方法の有効性を確認した。

提案方法の計算量の軽減や近似方法に関する検討は今後の課題としたい。また，需要予測の考慮など収益に関するより深い検討についても今後の課題としたい。

文　献

（1） N. Douer and U. Yechiali: “Optimal Repair and Replacement in Markovian Systems”, Commun. Statist.-Stochastic Models, Vol.10, No.1, pp.253–270 (1994)

（2） N. Tamura and T. Miyamura: “A Markovian Deteriorating Model Considering Maintenance and Imperfect Repair”, *IEICE Trans. A*, Vol.J84-A, No.2, pp.197–207 (2001) (in Japanese)
田村信幸・宮村鐵夫：「改良保全と不完全修理を考慮したマルコフ的劣化モデル」，信学論 A，Vol.J84-A, No.2, pp.197–207 (2001)

（3） M. Ohnishi, H. Kawai, and H. Mine: “An Optimal Inspection and Replacement policy under incomplete state information”, European Journal of Operational Research, Vol.27, pp.117–128 (1986)

（4） 金子哲夫：マルコフ決定理論入門，槙書店 (1973)

（5） J.O. Berger: Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, New York (1980)

（6） J.J. Martin: Bayesian Decision Problems and Markov Chains, John Wiley & Sons, (1967)

（7） T. Matsushima, and S. Hirasawa: “A Bayes coding algorithm for Markov models”, Technical Report of IEICE, IT95-1, pp.1–6 (1995)