電子情報通信学会論文誌 2012/9 Vol. J95-D No. 9

収益を考慮した保全への状態が未知のマルコフ決定過程の適用

前田 康成[†a)] (正員)　後藤文太朗[†]
升井 洋志[†]　桝井 文人[†] (正員)
鈴木 正清[†] (正員)

Applying Markov Decision Processes with Unknown States to Maintenance Considering Profits

Yasunari MAEDA[†a)], *Member*, Fumitaro GOTO[†], Hiroshi MASUI[†], *Nonmembers*, Fumito MASUI[†], and Masakiyo SUZUKI[†], *Members*

[†] 北見工業大学情報システム工学科, 北見市
Dept. of Computer Science, Kitami Institute of Technology, 165 Koen-cho, Kitami-shi, 090-8507 Japan
a) E-mail: maeda@cs.kitami-it.ac.jp

あらまし 本研究では，設備の状態が未知の保全問題において，コストと製造による収益を合わせて考慮する．マルコフ決定過程を用いて，統計的決定理論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保全計画を求める方法を提案する．

キーワード 保全，マルコフ決定過程，未知状態，統計的決定理論，ベイズ基準

1. まえがき

本研究では，製造工場の設備の保全問題を研究対象とする．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，各期での設備の状態が既知の場合[1]と，未知の場合[2],[3]を検討している．本研究では，各期での設備の状態が未知の場合を検討する．従来研究[1]〜[3]では，コストの最小化を行っている．しかし，実際にはコストと製造による収益を併せて考慮する必要がある．

そこで，本研究では収益も考慮した各期の状態が未知の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（製造による収益とコストの差分）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．提案方法では，操業データである修理の履歴情報と不良品数の履歴情報を用いて事後分布の計算を行う．

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．

本研究における記号の表記方法として，h_i のように記号に右下付の添字が付与されている場合には添え字付の一つの記号 h を示し，h^i のように記号に右上付の添字が付与されている場合には $h^i = h_1 h_2 \cdots h_i$ という記号 h の系列を示す．なお，記号 h は表記方法の説明のために仮に記述した例であり，本研究では利用しない記号である．

s_i，$s_i \in S$ は設備の状態を示し，MDPの状態に相当する．S，$S = \{s_1, s_2, \ldots, s_{|S|}\}$ は設備の状態集合で，$|S|$ 種類の設備状態を想定する．a_i，$a_i \in A$ は修理を示し，MDPの行動に相当する．A，$A = \{a_1, a_2, \ldots, a_{|A|}\}$ は修理の集合で，$|A|$ 種類の修理を想定する．$m(a_i)$ は修理 a_i のコスト，o は設備の1期間の稼動コスト，N は1期間に製造される製品数，b は製品1個当りの収益，b' は不良品1個当りのコストを示す．

$p(s_k \mid s_i, a_j, \theta^*)$ 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所要時間無しで実施可能で，修理前後の状態は共に未知（観測不可能）とする．θ^*，$\theta^* \in \Theta$ は真のパラメータで既知である．$p(s_j \mid s_i, \psi^*)$ は状態 s_i の設備が1期間稼動後に状態 s_j に遷移する稼動状態遷移確率を示す．ψ^*，$\psi^* \in \Psi$ は真のパラメータで既知である．保全状態遷移確率と稼動状態遷移確率がMDPの状態遷移確率に相当する．

$p(e \mid s_i, \xi^*)$ は状態 s_i の設備が不良品を製造する確率（不良率）を示す．ξ^*，$\xi^* \in \Xi$ は真のパラメータで既知である．

x_t，$x_t \in S$ は t 期の修理前の設備の状態，x'_t，$x'_t \in S$ は t 期の修理後の状態，x_1 は初期状態を示す．状態は全て未知であるが，初期状態の事前分布 $p(s_i)$ は既知である．

y_t，$y_t \in A$ は t 期に実施（選択）する修理（行動）を示す．w_t は t 期の設備稼働によって製造された N 個の製品中の不良品数で既知（観測可能）である．

z_t，$z_t = -m(y_t) - o + (N - w_t)b - w_t b'$ は t 期に修理 y_t を実施して N 個の製品中に不良品が w_t 個あった場合に，t 期に得られるMDPの利得である．利得は，製造による収益から各種コストを減算することで算出される．T は有限で，本研究では T 期間の総利得の最大化を目的とする．

上記のモデルでは，各期の最初に修理を実施し，修理後の状態のもとで1期間稼動する．設備状態が未知で，他は全て既知である．次章では，このモデルを用いて，次式の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．

$$\sum_{t=1}^{T} z_t = \sum_{t=1}^{T}(-m(y_t) - o + (N - w_t)b - w_t b'). \quad (1)$$

3. 収益を考慮したベイズ最適な保全

3.1 定 式 化

定式化するにあたって，最初に効用関数 $U(d(\cdot,\cdot), x^{T+1}y^T x'^T w^T, \theta^*, \psi^*, \xi^*)$ を次式で定義する．

$$U(d(\cdot,\cdot), x^{T+1}y^T x'^T w^T, \theta^*, \psi^*, \xi^*) = \sum_{t=1}^{T}(-m(y_t) - o + (N - w_t)b - w_t b'), \quad (2)$$

ただし，$d(\cdot,\cdot)$ は t 期に系列 $y^{t-1}w^{t-1} = y_1 w_1 \cdots y_{t-1} w_{t-1}$ と期を示す整数 t を受け取って，t 期に選択する行動（修理）y_t を返す決定関数である．決定関数 $d(\cdot,\cdot)$ が保全計画に相当するので，よい決定関数 $d(\cdot,\cdot)$ を求めることが，よい保全計画を求めることである．式 (2) の効用関数は，ある決定関数 $d(\cdot,\cdot)$ を用いて，系列 $x^{T+1}y^T x'^T w^T$ に対応する事象が起きた場合の総利得である．

次に，設備の初期状態が x_1 という条件のもとで決定関数 $d(\cdot,\cdot)$ を用いた場合の総利得の期待値である期待効用 $EU(d(\cdot,\cdot), x_1, \theta^*, \psi^*, \xi^*)$ を次式で定義する．

$$\begin{aligned} &EU(d(\cdot,\cdot), x_1, \theta^*, \psi^*, \xi^*) \\ &= \sum_{t=1}^{T} \sum_{x^t x'^t w^t} p(x^t x'^t w^t \mid d(\cdot,\cdot), x_1, \theta^*, \psi^*, \xi^*) \\ &\quad (-m(y_t) - o + (N - w_t)b - w_t b') \\ &= \sum_{t=1}^{T} \sum_{x^t x'^t w^t} \prod_{i=1}^{t} \Big(p(x_i | x'_{i-1}, \psi^*) p(x'_i \mid x_i, y_i, \theta^*) \\ &\quad \frac{N!}{(N-w_i)!w_i!} p(e|x'_i, \xi^*)^{w_i} (1 - p(e|x'_i, \xi^*))^{N-w_i} \Big) \\ &\quad (-m(y_t) - o + (N - w_t)b - w_t b'), \end{aligned} \quad (3)$$

ただし，初期状態が x_1 という条件のもとなので $p(x_1 \mid x'_0, \psi^*) = 1$ とする．また，$N!/(N-w_i)!w_i!$ は N 個の製品中に w_i 個の不良品が含まれる組合せ数である．本研究では設備状態が未知（観測不可能）の場合を扱うが，既知の場合であれば，式 (3) の期待効用を最大にする決定関数が最適な保全計画に相当する．

本研究では設備状態は未知で，初期状態も未知である．そこで，設備の初期状態 x_1 の事前分布 $p(x_1)$ に対して期待値をとるベイズ期待効用 $BEU(d(\cdot,\cdot), p(x_1))$ を次式で定義する．

$$BEU(d(\cdot,\cdot), p(x_1)) = \sum_{x_1 \in S} p(x_1) EU(d(\cdot,\cdot), x_1, \theta^*, \psi^*, \xi^*). \quad (4)$$

式 (4) 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．次節でベイズ最適な保全計画を動的計画法（DP）を用いて求める方法を提案する．

3.2 提 案 方 法

式 (4) のベイズ期待効用を書き下すと，T 期間の入れ子構造の形になる．この T 期間の入れ子構造に DP を適用して，T 期目からさかのぼりながら計算すると，ベイズ最適な保全計画を求めることができる．下記の提案方法中にある設備状態の事後分布は次式で計算される．

$$\begin{aligned} p(x_{t+1} \mid y^t w^t) &= \sum_{x_t} p(x_t \mid y^t w^t) p(x_{t+1} \mid x_t) \\ &= \sum_{x_t} \frac{p(x_t | y^{t-1} w^{t-1}) \sum_{x'_t} G(x_t)}{\sum_{s_i} p(s_i | y^{t-1} w^{t-1}) \sum_{x'_t} G(s_i)} p(x_{t+1} | x_t), \end{aligned} \quad (5)$$

ただし，

$$\begin{aligned} &G(s_i) \\ &= p(x'_t | s_i, y_t, \theta^*) p(e|x'_t, \xi^*)^{w_t} (1 - p(e|x'_t, \xi^*))^{N-w_t}, \end{aligned} \quad (6)$$

$$\begin{aligned} &p(x_{t+1} \mid x_t) \\ &= \sum_{x'} p(x'_t | x_t, y_t, \theta^*) p(x_{t+1} | x'_t, \psi^*). \end{aligned} \quad (7)$$

以下に，ベイズ最適な保全計画を求める提案方法を示す．

ステップ 1：T 期目の全ての系列（修理と不良品数の履歴）の候補 $y^{T-1}w^{T-1}$ に対する処理は以下のとおりである．

$$\begin{aligned} V(y^{T-1}w^{T-1}, T) &= \max_{y_T \in A} \sum_{x_T} \sum_{x'_T} \sum_{w_T} \frac{N!}{(N - w_T)! w_T!} \\ &\quad p(x_T | y^{T-1} w^{T-1}) p(x'_T | x_T, y_T, \theta^*) \\ &\quad p(e|x'_T, \xi^*)^{w_T} (1 - p(e|x'_T, \xi^*))^{N-w_T} \\ &\quad (-m(y_T) - o + (N - w_T)b - w_T b'), \end{aligned} \quad (8)$$

$$\begin{aligned} &d^*(y^{T-1}w^{T-1}, T) \\ &= \arg\max_{y_T \in A} \sum_{x_T} \sum_{x'_T} \sum_{w_T} \frac{N!}{(N - w_T)! w_T!} \\ &\quad p(x_T | y^{T-1} w^{T-1}) p(x'_T | x_T, y_T, \theta^*) \\ &\quad p(e|x'_T, \xi^*)^{w_T} (1 - p(e|x'_T, \xi^*))^{N-w_T} \end{aligned}$$

$$(-m(y_T) - o + (N - w_T)b - w_T b'), \quad (9)$$

ただし，$p(x_T \mid y^{T-1}w^{T-1})$ は 1 期から $T-1$ 期の修理と不良品数の履歴が系列 $y^{T-1}w^{T-1}$ の場合の設備状態の事後分布，$V(y^{T-1}w^{T-1}, T)$ は最後の期間の期待利得の最大値，このときの行動（修理）選択 $d^*(y^{T-1}w^{T-1}, T)$ がベイズ最適な決定であり，ベイズ最適な保全計画における修理に相当する．

ステップ 2：t 期目（$1 \leq t \leq T-1$）の全ての系列の候補 $y^{t-1}w^{t-1}$ に対する処理は以下のとおりである．

$$\begin{aligned} V(y^{t-1}w^{t-1}, t) = \max_{y_t \in A} \sum_{x_t} \sum_{x'_t} \sum_{w_t} \frac{N!}{(N-w_t)!w_t!} \\ p(x_t \mid y^{t-1}w^{t-1}) p(x'_t \mid x_t, y_t, \theta^*) \\ p(e \mid x'_t, \xi^*)^{w_t} (1 - p(e \mid x'_t, \xi^*))^{N-w_t} \big(-m(y_t) - o \\ + (N - w_t)b - w_t b' + V(y^t w^t, t+1)\big), \end{aligned} \quad (10)$$

$$\begin{aligned} d^*(y^{t-1}w^{t-1}, t) \\ = \arg\max_{y_t \in A} \sum_{x_t} \sum_{x'_t} \sum_{w_t} \frac{N!}{(N-w_t)!w_t!} \\ p(x_t \mid y^{t-1}w^{t-1}) p(x'_t \mid x_t, y_t, \theta^*) \\ p(e \mid x'_t, \xi^*)^{w_t} (1 - p(e \mid x'_t, \xi^*))^{N-w_t} \big(-m(y_t) - o \\ + (N - w_t)b - w_t b' + V(y^t w^t, t+1)\big), \end{aligned} \quad (11)$$

ただし，$p(x_t \mid y^{t-1}w^{t-1})$ は 1 期から $t-1$ 期の修理と不良品数の履歴が系列 $y^{t-1}w^{t-1}$ の場合の設備状態の事後分布，$V(y^{t-1}w^{t-1}, t)$ は t 期以降の総利得の期待値の最大値，このときの行動選択 $d^*(y^{t-1}w^{t-1}, t)$ がベイズ最適な決定であり，ベイズ最適な保全計画における修理に相当する．

上記のように T 期からさかのぼりながら DP を適用し，T 期間のベイズ最適な保全計画 $d^*(\cdot, \cdot)$ を求める．最終的に求まる $t = 1$ のとき（1 期）の $V(y^0w^0, 1)$ が T 期間のベイズ最適な保全計画 $d^*(\cdot, \cdot)$ を用いた場合の T 期間の総利得の期待値である．なお，1 期ではまだ修理や不良品数の履歴は存在しないので y^0w^0 は空列である．

3.3 数値実験例

提案方法の有効性を確認するために，一例ではあるが数値実験例を報告する．

$S = \{s_1, s_2\}$，$A = \{a_1, a_2\}$，$N = 4$，$T = 5$，$m(a_1) = 0$，$m(a_2) = 50$，$o = 10$，$b = 400$，$b' = 100$，$p(s_1 \mid s_1, a_1, \theta^*) = 1.0$，$p(s_1 \mid s_1, a_2, \theta^*) = 1.0$，$p(s_1 \mid s_2, a_1, \theta^*) = 0.0$，$p(s_1 \mid s_2, a_2, \theta^*) = 0.99$，$p(s_1 \mid s_1, \psi^*) = 0.8$，$p(s_1 \mid s_2, \psi^*) = 0.0$，$p(e \mid s_1, \xi^*) = 0.01$，$p(e \mid s_2, \xi^*) = 0.1$，$p(s_1) = 0.5$，$p(s_2) = 0.5$ の条件のもとで，提案方法による保全計画 $d^*(\cdot, \cdot)$ とコスト最小化の保全計画 $d'^*(\cdot, \cdot)$（$b = 0$ と仮定）を求めた．$b = 0$ として提案方法を適用すると，コストに -1 を乗じたものの最大化になるので，最終的に求まる保全計画はベイズ基準のもとでコスト最小化を実現する保全計画に相当する．コスト最小化の場合も提案方法を適用しているが，$b = 0$ と仮定した場合であることを明示的に示すために $d'^*(\cdot, \cdot)$ という表記にした．提案方法の保全計画 $d^*(\cdot, \cdot)$ による期待利得 $V(y^0w^0, 1)$ は 7625.27，コストの期待値（保全計画 $d^*(\cdot, \cdot)$ による $\sum_{t=1}^{T}(m(y_t) + o + w_t b')$ の期待値）は 234.95 だった．コスト最小化の保全計画 $d'^*(\cdot, \cdot)$ による期待利得（保全計画 $d'^*(\cdot, \cdot)$ による $\sum_{t=1}^{T}(-m(y_t) - o + (N - w_t)b - w_t b')$ の期待値）は 7425.68，コストの期待値（保全計画 $d'^*(\cdot, \cdot)$ による $\sum_{t=1}^{T}(m(y_t) + o + w_t b')$ の期待値）は 174.36 だった．このように，製造による収益とコストの差分である利得の期待値を最大にする提案方法による保全計画と，コストの期待値を最小にする保全計画は異なる保全計画である．

一例ではあるが，上記の数値実験例によって，設備状態が未知の保全問題において，製造による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. むすび

本研究では，製造工場における設備の保全問題を扱った．従来研究 [2], [3] では MDP を用いてコストの最小化が検討されているが，製造による収益は考慮されていない．そこで，本研究では収益も考慮するように拡張したもとで，統計的決定理論に基づいて，状態が未知の 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定式化を行った．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．一例ではあるが数値実験によって，製造による収益を考慮することの有効性と，提案方法の有効性を確認した．

従来研究 [2], [3] では製造工場に限定せずに設備状態が未知の保全問題を扱っており，何らかの監視情報が得られたもとで未知の設備状態を推定している．本研究では，製造工場における監視情報の一例として不良品数を採用した．不良品数という情報は製造工場では一般的によく扱われるので，不良品数を監視情報として利用する提案方法は製造工場で広く適用可能な方法だと考えられる．他方，紙や布生地などを製造する製

造工場や，石油精製などの化学処理工場，製鉄所などを対象にする場合には不良品数以外の監視情報を検討する必要がある．そのような不良品数以外の監視情報に関する検討は今後の課題としたい．

提案方法では，式 (10) などの演算を修理と不良品数の履歴の候補数と同じ回数実施する必要があり，計算量は指数オーダである．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，本研究では製造による収益を単純にモデル化したが，実際には需要と供給の関係や完成品の在庫管理など，より複雑なモデル化が必要である．更に企業における意思決定組織を考えると，需要予測などに基づいて製造数を決定する組織と，工場における保全計画を決定する組織は別組織であることが多いと考えられる．より現実に近いモデル化や，組織横断的な意思決定方法に関する検討も今後の課題としたい．

本研究では，修理の際の状態遷移確率，設備稼動の際の状態遷移確率，製造に伴う不良率を支配する真のパラメータ θ^*，ψ^*，ξ^* を既知と仮定したが，実際にはこれらのパラメータは未知である．経験的に得られた推定値を真のパラメータの代わりに用いることが一般的であるが，理論的に最適性を保証するためには，真のパラメータ未知のもとでの検討が必要である．真のパラメータ未知の一般的な MDP については，学習理論の分野で能動学習として従来から研究されている．収益を考慮した保全問題における能動学習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] 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.

[3] 高橋將人，鈴木和幸，"複数修理保全モデルにおける最適保全方策に関する一考察，" 信学論 (A)，vol.J80-A, no.4, pp.677–683, April 1997.

[4] 金子哲夫，マルコフ決定理論入門，槙書店，東京，1973.

[5] J.O. Berger, Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, New York, 1980.

(平成 24 年 2 月 22 日受付，4 月 17 日再受付)