

## 推薦システムにおける新規顧客問題に関する一考察

前田 康成<sup>1</sup>, 山内 翔<sup>1</sup>, 鈴木 正清<sup>1</sup>, 松嶋 敏泰<sup>2</sup>

1) 北見工業大学・地域未来デザイン工学科      2) 早稲田大学・応用数理学科

**要約:** 推薦システムにおける従来研究では確率モデルとしてマルコフ決定過程が採用されている。当該従来研究ではベイズ基準のもとで総利得が最大化されているが、新規顧客に対する質問方法は検討されていない。また、新規顧客に対する質問方法を検討している従来研究もあるが、当該従来研究では総利得の最大化は検討されていない。

そこで、本研究ではマルコフ決定過程を推薦システムにおける新規顧客に対する質問に適用し、総利得をベイズ基準のもとで最大化する質問方法を提案する。提案アルゴリズム中では、最適化手法として動的計画法を用いる。

**キーワード:** 推薦システム, 新規顧客問題, マルコフ決定過程, 統計的決定理論, ベイズ基準

## A Note on a New Customer Problem in Recommender System

Yasunari MAEDA<sup>1</sup>, Sho YAMAUCHI<sup>1</sup>, Masakiyo SUZUKI<sup>1</sup>, Toshiyasu MATSUSHIMA<sup>2</sup>

1) School of Regional Innovation and Social Design Engineering, Kitami Institute of Technology  
2) Department of Applied Mathematics, Waseda University

**Abstract:** In previous research of recommender system Markov decision processes are used in order to represent recommender system. Total reward is maximized with reference to a Bayes criterion. In the previous research a questionnaire method for a new customer is not studied. In other previous research a questionnaire method for a new customer is studied, but the total reward is not maximized.

In this research we apply Markov decision processes to questionnaire for a new customer in recommender system. We propose a new questionnaire method which maximizes the total reward with reference to the Bayes criterion. In the proposed method dynamic programming is used.

**Keywords:** recommender system, new customer problem, Markov decision processes, statistical decision theory, Bayes criterion

---

Yasunari MAEDA

165 Koen-cho, Kitami-shi, Hokkaido, 090-8507, Japan

Phone: +81-157-26-9328, Fax: +81-157-26-9344, E-mail: maedaya@mail.kitami-it.ac.jp

## 1. はじめに

本研究では、インターネット上で通信販売を行っている通販サイトなどにおいて顧客に商品を推薦する際に利用されている推薦システムを研究対象とする。

従来研究[1]~[3]では、商品の売上高の最大化を目的とした推薦方法が研究されている。従来研究[3]では、商品の好みなどが似た顧客が同一クラスに属すると仮定し、クラス未知の顧客に対して当該顧客の売上高を統計的決定理論[4]に基づきベイズ基準のもとで最大化している。推薦システムでは、システム側で何も情報を持っていないクラスが未知の新規の顧客への対応に関する問題（新規顧客問題）も重要である。従来研究[3]でも、事前確率の工夫次第で新規顧客対応も可能である。しかし、新規顧客から当該顧客に関する情報を入手する質問などの対応は検討されていない。

他方、従来研究[5]では新規顧客からなるべく多くの情報を入手するための質問方法が提案されているが、当該顧客の売上高を最大化する理論的保証はない。

そこで、本研究では新規顧客に対して当該顧客の売上高を統計的決定理論に基づきベイズ基準のもとで最大化するための質問方法を提案する。売上高の最大化を目的とする従来研究同様にマルコフ決定過程（MDP）[6]を確率モデルとして採用し、初回の商品推薦の前に新規顧客の売上高を最大化するための質問を行うように従来研究の問題設定を拡張する。統計的決定理論に基づく定式化後に、動的計画法（DP）を用いるベイズ最適な質問および推薦商品の計算方法を提案し、その有効性を数値計算例によって確認する。

## 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。本研究では新規顧客に対して $T'$ 回の質問後に $T$ 回の商品の推薦を行う

$c_i, c_i \in C$ は顧客のクラスを示し、 $C = \{c_1, c_2, \dots, c_{|C|}\}$ は顧客クラスの集合である。 $q_i, q_i \in Q$ は新規顧客に関する情報を入手するための質問を示し、 $Q = \{q_1, q_2, \dots, q_{|Q|}\}$ は質問集合である。 $a_i, a_i \in A$ は質問に対する回答を示し、 $A = \{a_1, a_2, \dots, a_{|A|}\}$ は回答集合である。 $m_i, m_i \in M$ は推薦対象の商品を示し、 $M = \{m_1, m_2, \dots, m_{|M|}\}$ は商品集合である。本研究では、

従来研究と同様にMDPを用いてモデル化し、質問 $q_i$ の実施および商品 $m_i$ の推薦がMDPにおける行動選択に相当する。 $n_i, n_i \in N$ は推薦に対する顧客の反応を示し、 $N = \{n_1, n_2, \dots, n_{|N|}\}$ は反応集合である。 $1 \leq i \leq |M|$ では $n_i = m_i$ であり、顧客の反応が商品 $m_i$ の購入に相当する。 $|N| = |M| + 1$ であり、 $n_{|M|+1}$ は顧客が何も購入しなかったことを示す。 $r(n_i), 1 \leq i \leq |M|$ は商品 $m_i$ の売上高を示し、顧客が何も購入しなかった場合には $r(n_{|M|+1}) = 0$ である。 $r(n_i)$ はMDPの利得に相当する。

本研究では新規顧客に対して $T'$ 回の質問後に $T$ 回の商品の推薦を行うための、 $T' + T$ 期間のMDP問題を扱う。なお、質問に対しては利得は発生しない。

$y_t$ はMDPにおける $t$ 期の行動を示し、質問期間（ $1 \leq t \leq T'$ ）であれば $t$ 回目の質問 $y_t \in Q$ 、推薦期間（ $T' + 1 \leq t \leq T' + T$ ）であれば $t - T'$ 回目の商品の推薦 $y_t \in M$ に相当する。なお、MDP問題としては $T' + T$ 期間の問題形式になるが、実際の推薦システムとしては新規顧客のアクセス時に $T'$ 回の質問と1回目の推薦（ $y_{T'+1}$ ）を実施し、当該顧客の2回目のアクセス時に2回目の推薦（ $y_{T'+2}$ ）を実施する。

$z_t$ はMDPにおける推薦システムの行動 $y_t$ に対する顧客の回答または反応を示す。質問期間（ $1 \leq t \leq T'$ ）であれば $t$ 回目の質問 $y_t \in Q$ に対する回答 $z_t \in A$ 、推薦期間（ $T' + 1 \leq t \leq T' + T$ ）であれば $t - T'$ 回目の商品の推薦 $y_t \in M$ に対する反応 $z_t \in N$ に相当する。

$w, w \in C$ は新規顧客のクラスを示し、未知である。クラスの事前確率 $p(w)$ は既知とする。 $x_t = (y_{t-1}, z_{t-1})$ は $t$ 回目（ $t \geq T' + 2$ ）の推薦を受ける際の顧客の状態であり、MDPの $t$ 期の状態に相当する。状態は1期前の推薦商品と顧客の反応で構成される。上記の状態定義を採用しているため、厳密には質問期間および1回目の推薦の際に状態が未定義になるが、DPを用いる提案アルゴリズムでは系列 $y^{t-1} z^{t-1}$ を計算に用いるので状態が未定義の期が存在しても問題はない。

$p(a_k | q_i, c_j, \psi^*)$ はクラス $c_j$ に属する顧客が質問 $q_i$ に対して回答 $a_k$ を返す確率を示す。 $p(\text{pass} | \xi^*)$ は質問期間において新規顧客が質問に回答せずに推薦期間に

移行してしまう確率である。一般的に新規顧客対応の質問に対して顧客が対応するかどうかは不確実であるため、本研究ではこのような確率を導入する。

$p(n_o | m_i, n_j, m_k, c_l, \theta^*)$  はクラス  $c_l$  に属する顧客が1期前の推薦と反応が  $m_i$  と  $n_j$  のもとで商品  $m_k$  を推薦されたときに反応  $n_o$  を返す確率を示す。確率  $p(n_o | m_i, n_j, m_k, c_l, \theta^*)$  は以下のように MDP における状態遷移確率に相当する。

$$\begin{aligned} p(x_{t+1} | x_t, y_t, w, \theta^*) &= p(y_t, z_t | y_{t-1}, z_{t-1}, y_t, w, \theta^*) \\ &= p(z_t | y_{t-1}, z_{t-1}, y_t, w, \theta^*), \end{aligned} \quad (1)$$

ただし、 $t \geq T' + 2$ 。

$p(n_k | m_i, c_j, \theta^*)$  は1回目の推薦 ( $t = T' + 1$ ) において、クラス  $c_j$  に属する顧客が商品  $m_i$  を推薦されたときに反応  $n_k$  を返す確率である。確率  $p(n_k | m_i, c_j, \theta^*)$  は以下のように MDP における状態  $x_{T'+2}$  の生起確率に相当する。

$$\begin{aligned} p(x_{t+1} | y_t, w, \theta^*) &= p(y_t, z_t | y_t, w, \theta^*) \\ &= p(z_t | y_t, w, \theta^*). \end{aligned} \quad (2)$$

$\theta^*, \psi^*, \xi^*$  は各種確率分布を支配する真のパラメータで本研究では既知と仮定する。 $x_{T'+2}^{T'+T} y^{T'+T} z^{T'+T}$  は推薦期間終了後の系列を示し、 $x_{T'+2}^{T'+T} = x_{T'+2} x_{T'+3} \cdots x_{T'+T}$ 、 $y^{T'+T} = y_1 \cdots y_{T'+T}$ 、 $z^{T'+T} = z_1 \cdots z_{T'+T}$ 。新規顧客が質問対応を実施せずに

推薦期間に移行した場合には、系列が短くなるが、表記を簡便にするため表記上は全質問に回答した場合と同じ表記を行っている。ただし、以下の章で提案するアルゴリズム中では適切に質問期間から推薦期間に移行した計算を行っている。質問期間では状態  $x_t$  が未定義のため、質問と回答のみの系列  $y^t z^t$  になる。 $d(\cdot, \cdot)$

は期を示す  $t$  と当該期までの系列を受け取って当該期の行動  $y_t$  を返す決定関数である。

### 3. 売上高をベイズ基準のもとで最大化する新規顧客対応

#### 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  $U(d(\cdot, \cdot), x_{T'+2}^{T'+T} y^{T'+T} z^{T'+T}, w, \theta^*, \psi^*, \xi^*)$  を次式で定義する。

$$\begin{aligned} U(d(\cdot, \cdot), x_{T'+2}^{T'+T} y^{T'+T} z^{T'+T}, w, \theta^*, \psi^*, \xi^*) \\ = \sum_{i=T'+1}^{T'+T} r(z_i). \end{aligned} \quad (3)$$

効用  $U(d(\cdot, \cdot), x_{T'+2}^{T'+T} y^{T'+T} z^{T'+T}, w, \theta^*, \psi^*, \xi^*)$  はパラメータ  $\theta^*, \psi^*, \xi^*$  のもとでクラス  $w$  に属する新規顧客に対して決定関数  $d(\cdot, \cdot)$  を使って、 $T'$  回の質問と  $T$  回の商品の推薦  $y^{T'+T}$  を行つて、新規顧客の質問への回答と推薦への反応が  $z^{T'+T}$  だった場合の総売上（総利得）を示す。

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

$$\begin{aligned} EU(d(\cdot, \cdot), w, \theta^*, \psi^*, \xi^*) \\ = \sum_{x_{T'+2}^{T'+T} z^{T'+T}} p(x_{T'+2}^{T'+T} z^{T'+T} | d(\cdot, \cdot), w, \theta^*, \psi^*, \xi^*) \\ U(d(\cdot, \cdot), x_{T'+2}^{T'+T} y^{T'+T} z^{T'+T}, w, \theta^*, \psi^*, \xi^*). \end{aligned} \quad (4)$$

期待効用  $EU(d(\cdot, \cdot), w, \theta^*, \psi^*, \xi^*)$  はパラメータ  $\theta^*, \psi^*, \xi^*$  のもとでクラス  $w$  に属する新規顧客に対して

決定関数  $d(\cdot, \cdot)$  を使って質問と商品の推薦を行う場合の総売上（総利得）の期待値を示す。

新規顧客のクラス  $w$  は未知なので事前確率  $p(w)$  を導入し、期待効用の期待値であるベイズ期待効用

$BEU(d(\cdot, \cdot), p(w), \theta^*, \psi^*, \xi^*)$  を次式で定義する。

$$\begin{aligned} BEU(d(\cdot, \cdot), p(w), \theta^*, \psi^*, \xi^*) \\ = \sum_{w \in C} p(w) EU(d(\cdot, \cdot), w, \theta^*, \psi^*, \xi^*). \end{aligned} \quad (5)$$

式(5)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売上を最大にするという意味で最適な質問および推薦方法に相当し、次式で定義される。

$$d_B(\cdot, \cdot) = \arg \max_{d(\cdot, \cdot)} BEU(d(\cdot, \cdot), p(w), \theta^*, \psi^*, \xi^*). \quad (6)$$

式(5)のベイズ期待効用を書き下すと、 $T' + T$  期間の入れ子構造になる。この入れ子構造に DP を適用し、最適な質問および推薦商品を算出する方法を次々節で提案する。

### 3.2 事後確率の計算

次節で DP を用いた提案方法を説明するが、ここでは次節の提案方法の中で実施する事後確率の計算について先に説明する。

$t$  期,  $1 \leq t \leq T' + 1$  に使用する事後確率  $p(w|y^{t-1}z^{t-1})$

は次式で計算できる。

$$p(w|y^{t-1}z^{t-1}) = \frac{p(w|y^{t-2}z^{t-2})p(z_{t-1}|y_{t-1}, w, \psi^*)}{\sum_{w'} p(w'|y^{t-2}z^{t-2})p(z_{t-1}|y_{t-1}, w', \psi^*)}, \quad (7)$$

ただし,  $p(w|y^0z^0) = p(w)$ . 質問期間の  $t$  期,  $1 \leq t \leq T'$

に顧客が質問を回避して推薦期間に進んだ場合には推薦期間の最初の期である  $T' + 1$  期で使用する事後確率

は  $p(w|y^{T'}z^{T'}) = p(w|y^{t-1}z^{t-1})$  とする。

$t$  期,  $t = T' + 2$  に使用する事後確率

$p(w|x_{T'+2}^t y^{t-1} z^{t-1})$  は次式で計算できる。

$$\begin{aligned} p(w|x_{T'+2}^t y^{t-1} z^{t-1}) \\ = \frac{p(w|y^{t-2}z^{t-2})p(z_{t-1}|y_{t-1}, w, \theta^*)}{\sum_{w'} p(w'|y^{t-2}z^{t-2})p(z_{t-1}|y_{t-1}, w', \theta^*)}. \end{aligned} \quad (8)$$

$t$  期,  $T' + 3 \leq t \leq T' + T$  に使用する事後確率

$p(w|x_{T'+2}^t y^{t-1} z^{t-1})$  は次式で計算できる。

$$\begin{aligned} p(w|x_{T'+2}^t y^{t-1} z^{t-1}) \\ = \frac{p(w|x_{T'+2}^{t-1} y^{t-2} z^{t-2})p(z_{t-1}|x_{t-1}, y_{t-1}, w, \theta^*)}{\sum_{w'} p(w'|x_{T'+2}^{t-1} y^{t-2} z^{t-2})p(z_{t-1}|x_{t-1}, y_{t-1}, w', \theta^*)}, \end{aligned} \quad (9)$$

ただし,

$$\begin{aligned} p(z_{t-1}|x_{t-1}, y_{t-1}, w, \theta^*) \\ = p(z_{t-1}|y_{t-2}, z_{t-2}, y_{t-1}, w, \theta^*). \end{aligned} \quad (10)$$

### 3.3 提案方法

DP を用いて,  $T' + T$  期目から遡りながら計算して、ベイズ最適な質問および推薦商品を算出する。

$T' + T$  期目のすべての遷移系列  $x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}$

に対する処理は以下のとおりである。

$$\begin{aligned} V(x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}, T' + T) = \max_{y_{T'+T} \in M} \\ \sum_{z_{T'+T} \in N} p(z_{T'+T} | x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}, y_{T'+T}, \theta^*) \\ r(z_{T'+T}), \end{aligned} \quad (11)$$

ただし,

$$\begin{aligned} p(z_{T'+T} | x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}, y_{T'+T}, \theta^*) \\ = \sum_{w \in C} p(w | x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}) \\ p(z_{T'+T} | y_{T'+T-1}, z_{T'+T-1}, y_{T'+T}, w, \theta^*), \end{aligned} \quad (12)$$

$p(z_{T'+T} | x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}, y_{T'+T}, \theta^*)$  は顧客の反応確率の事後確率  $p(w | x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1})$  による期待値で、

$V(x_{T'+2}^{T'+T} y^{T'+T-1} z^{T'+T-1}, T' + T)$  は最後の期の期待利得の最大値である。式(11)の右辺を最大化する行動  $y_{T'+T}$  がベイズ最適な最後の推薦商品である。

$t$  期目 ( $T' + 2 \leq t \leq T' + T - 1$ ) のすべての遷移系列

$x_{T'+2}^t y^{t-1} z^{t-1}$  に対する処理は以下のとおりである。

$$\begin{aligned} V(x_{T'+2}^t y^{t-1} z^{t-1}, t) \\ = \max_{y_t \in M} \sum_{z_t \in N} p(z_t | x_{T'+2}^t y^{t-1} z^{t-1}, y_t, \theta^*) \\ (r(z_t) + V(x_{T'+2}^{t+1} y^t z^t, t+1)), \end{aligned} \quad (13)$$

ただし,

$$\begin{aligned} p(z_t | x_{T'+2}^t y^{t-1} z^{t-1}, y_t, \theta^*) \\ = \sum_{w \in C} p(w | x_{T'+2}^t y^{t-1} z^{t-1}) p(z_t | y_{t-1}, z_{t-1}, y_t, w, \theta^*), \end{aligned} \quad (14)$$

$V(x_{T'+2}^t y^{t-1} z^{t-1}, t)$  は  $t$  期以降の期待総利得の最大値である。式(13)の右辺を最大化する行動  $y_t$  がベイズ最適な  $t$  期の推薦商品である。

$t$  期目 ( $t = T' + 1$ ) のすべての遷移系列  $y^{t-1} z^{t-1}$  に

対する処理は以下のとおりである。

$$V(y^{t-1}z^{t-1}, t) = \max_{y_t \in M} \sum_{z_t \in N} p(z_t | y^{t-1}z^{t-1}, y_t, \theta^*) \quad (15)$$

$$(r(z_t) + V(x_{T'+2}^{t+1}y^t z^t, t+1)),$$

ただし,

$$p(z_t | y^{t-1}z^{t-1}, y_t, \theta^*) \quad (16)$$

$$= \sum_{w \in C} p(w | y^{t-1}z^{t-1}) p(z_t | y_t, w, \theta^*),$$

$V(y^{t-1}z^{t-1}, t)$  は  $t$  期以降の期待総利得の最大値である。  
式(15)の右边を最大化する行動  $y_t$  がベイズ最適な  $t$  期の推薦商品である。

$t$  期目 ( $1 \leq t \leq T'$ ) のすべての遷移系列  $y^{t-1}z^{t-1}$  に対する処理は以下のとおりである。

$$V(y^{t-1}z^{t-1}, t) = \max_{y_t \in Q - y^{t-1}} \left( (1 - p(\text{pass} | \xi^*)) \right. \quad (17)$$

$$\sum_{z_t \in A} p(z_t | y^{t-1}z^{t-1}, y_t, \psi^*) V(y^t z^t, t+1)$$

$$\left. + p(\text{pass} | \xi^*) V(y^{T'} z^{T'}, T'+1) \right),$$

ただし,

$$p(z_t | y^{t-1}z^{t-1}, y_t, \psi^*) \quad (18)$$

$$= \sum_{w \in C} p(w | y^{t-1}z^{t-1}) p(z_t | y_t, w, \psi^*),$$

$Q - y^{t-1}$  は未実施の質問集合である。 $V(y^{t-1}z^{t-1}, t)$  は  $T'+1$  期以降の期待総利得の最大値である。式(17)では新規顧客が質問に回答せずに推薦期間に移行してしまう場合も考慮している。式(17)の右边を最大化する行動  $y_t$  がベイズ最適な  $t$  期の質問である。

### 3.4 数値計算例

提案方法の有効性を確認するため、数値計算例を報告する。

顧客クラス数  $|C|=2$  , 商品数  $|M|=2$  , 質問回数

$T'=1$  , 推薦回数  $T=3$  , 質問数  $|Q|=2$  , 回答数  $|A|=2$  ,

質問の回避確率  $p(\text{pass} | \xi^*) = 0.2$  , クラスの事前確率

$p(w)$  は等確率とした。商品推薦に対する顧客の反応確

率  $p(n_o | m_i, n_j, m_k, c_l, \theta^*)$  ,  $p(n_k | m_i, c_j, \theta^*)$  , 質問に対

する回答確率  $p(a_k | q_i, c_j, \psi^*)$  をそれぞれ表 1～表 3 とした。

表 1.  $p(n_o | m_i, n_j, m_k, c_l, \theta^*)$  の例

$i$	$j$	$k$	$l$	$o=1$	$o=2$	$o=3$
1	1	1	1	0.9	0.05	0.05
1	1	2	1	0.2	0.2	0.6
1	2	1	1	0.2	0.2	0.6
1	2	2	1	0.05	0.9	0.05
1	3	1	1	0.8	0.1	0.1
1	3	2	1	0.1	0.1	0.8
2	1	1	1	0.9	0.05	0.05
2	1	2	1	0.2	0.2	0.6
2	2	1	1	0.2	0.2	0.6
2	2	2	1	0.05	0.9	0.05
2	3	1	1	0.8	0.1	0.1
2	3	2	1	0.1	0.1	0.8
1	1	1	2	0.9	0.05	0.05
1	1	2	2	0.2	0.2	0.6
1	2	1	2	0.8	0.1	0.1
1	2	2	2	0.2	0.2	0.6
1	3	1	2	0.6	0.1	0.3
1	3	2	2	0.1	0.6	0.3
2	1	1	2	0.2	0.2	0.6
2	1	2	2	0.1	0.8	0.1
2	2	1	2	0.2	0.2	0.6
2	2	2	2	0.05	0.9	0.05
2	3	1	2	0.6	0.1	0.3
2	3	2	2	0.1	0.6	0.3

表 2.  $p(n_k | m_i, c_j, \theta^*)$  の例

$i$	$j$	$k=1$	$k=2$	$k=3$
1	1	0.8	0.1	0.1
2	1	0.1	0.1	0.8
1	2	0.6	0.1	0.3
2	2	0.1	0.6	0.3

表3.  $p(a_k | q_i, c_j, \psi^*)$  の例

$i$	$j$	$k=1$	$k=2$
1	1	0.9	0.1
2	1	0.6	0.4
1	2	0.1	0.9
2	2	0.4	0.6

上記の設定のもとで、提案方法、従来研究[3]と同様に質問をせずに期待総利得を最大化する方法、従来研究[5]と同様にエントロピー関数を利用した情報量最大化による質問方法の3方法を比較した。なお、質問選択にエントロピー関数を利用する場合も推薦期間には提案方法などと同様にベイズ最適な商品（ベイズ基準のもとで総売上を最大化する商品）を選択するようにした。その結果、期待総利得は提案方法 185.395、質問なし 171.125、エントロピー関数を利用した質問選択 171.225 となった。

本計算例では、新規顧客対応の質問を実施しない従来研究[3]による期待総利得よりも、質問を実施する本研究の提案方法および従来研究[5]による期待総利得の方が大きい。よって、本計算例から新規顧客対応の質問を実施することの有効性が確認できる。また、本計算例では期待総利得のベイズ基準のもとでの最大化を実施する本研究の提案方法による期待総利得は、エントロピー関数を利用した情報量最大化を実施する従来研究[5]による期待総利得よりも大きい。この期待総利得の差の存在は、期待総利得最大化のための質問と情報量最大化のための質問は本計算例のように一致しない場合もあることを意味する。よって、推薦システムの本来の目的である売上高（期待総利得）の最大化に対しては、従来研究[5]による情報量最大化の質問よりも本研究の提案方法による期待総利得最大化の質問がより有効であることが確認できる。以上より、小規模な計算例であるが、売上高の最大化を目的とする提案方法による質問選択の有効性が確認できた。

なお、本計算例は従来研究[5]による情報量最大化の質問と本研究の提案方法による期待総利得最大化の質問が大きく異なる例である。言い換えると、情報量の最大化が期待総利得の最大化にほとんど寄与しない例である。そのため、情報量最大化の期待総利得と提案方法による期待総利得の差が大きく、情報量最大化の期待総利得と質問なしの期待総利得が近い値になって

いる。本計算例と異なる設定の中には、情報量最大化の期待総利得と期待総利得最大化の期待総利得が近い値になる（情報量の最大化が期待総利得の最大化に寄与する）設定も存在することが予想される。

#### 4. まとめと今後の課題

本研究では、新規顧客に対して当該顧客の売上高を統計的決定理論に基づきベイズ基準のもとで最大化するための質問と推薦商品を DP を用いて算出する方法を提案した。また、小規模な数値計算例であるが、売上高の最大化を目的とする提案方法による質問選択の有効性を確認した。

本研究では、顧客が属するクラスは固定（時間の経過とともに変化しない）と仮定した。しかし、実際の通販サイトにおいては顧客のクラスが時間の経過とともに変化する場合も想定される。顧客のクラスが変化する場合の検討については、今後の課題としたい。

次に本研究内容の健康・医療分野への適用可能性について述べる。本研究ではインターネット上の通販サイトなどにおける推薦システムを対象としたが、通販サイトで扱う商品が医療関連商品や医療保険、医療サービスであっても本研究と同様に売上高の最大化が可能である。また、本研究を数理工学の視点から整理すると、本研究では未知情報を伴う MDP における期待総利得の最大化を行っている。患者の真の健康状況を未知情報、患者に対する検査項目や治療方針（医療サービス）を MDP の行動、良好な健康状態における滞在時間（確率）などを利得としたモデル化を想定すると、本研究の成果を医療検査・治療方針の決定支援に適用することも可能だと考えられる。これらの健康・医療分野への適用に関しては、別の機会に改めて報告したい。

#### 謝辞

本研究の一部は JSPS 科研費 JP16K00417 の助成による。

#### 参考文献

- [1] G. Shani, D. Heckerman, and R. I. Brafman: An MDP-Based Recommender System, *Journal of Machine Learning Research*, Vol.6, pp.1265-1295, 2005.
- [2] 桑田修平, 前田康成, 松嶋敏泰, 平澤茂一: 推薦システムのための状態遷移確率の構造を未知としたマルコフ決定過程, *情報処理学会論文誌 数理モデル化と応用*, Vol.6, No.1, pp.20-30, 2013.

[3] 岩井秀輔, 宮希望, 前田康成, 松嶋敏泰: 推薦対象ユーザのクラスが未知の推薦問題におけるマルコフ決定過程を用いた推薦システムに関する一考察, 信学技報IT, Vol.114, No.138, pp.49-54, 2014.

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

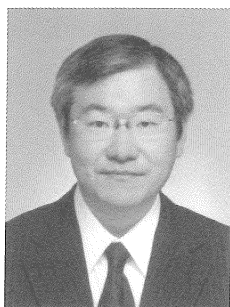
[5] K. Yu, A. Schwaighofer, V. Tresp, X. Xu, and H. P. Kriegel: Probabilistic memory-based collaborative filtering, IEEE Transactions on Knowledge and Data Engineering, Vol.16, No.1, pp.56-69, 2004.

[6] 金子哲夫: マルコフ決定理論入門, 槇書店, 東京, 1973.



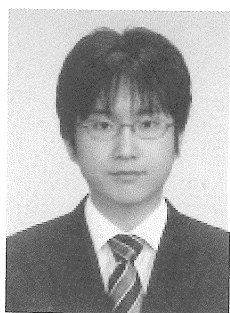
松嶋敏泰 (まつしまとしやす)

昭53 早大・理工・工業経営卒. 昭55 同大学院修士課程了. 同年, 日本電気(株) 入社. 昭61 早大・理工学研究科・博士後期課程入学. 現在, 早大・応用数理学教授. 知識情報処理及び情報理論とその応用に関する研究に従事. 工学博士. IEEE 等各会員.



前田康成 (まえだやすなり)

平成7年早大・理工卒. 平成9年同大学院理工学研究科修士課程修了. 日本電信電話(株), 東日本電信電話(株), 北見工大助手, 助教, 准教授を経て平成28年同大学教授, 現在に至る. 博士(工学). 統計的決定理論の学習問題への応用に関する研究に従事. 電子情報通信学会等各会員.



山内翔 (やまうちしょう)

1988年10月27日生. 2011年3月北海道大学工学部情報エレクトロニクス学科卒業. 2014年同大学院情報科学研究科複合情報学専攻複雑系工学講座博士後期課程修了. 2014年より日本学術振興会特別研究員. 2016年より北見工業大学助教. 自律ロボットシステムの研究に従事.



鈴木正清 (すずきまさきよ)

昭57 北大・工・電子卒. 昭62 同大学院博士課程了. 同年同大応用電気研究所助手. 平5 北見工大・助教授, 平8 北大・電子研助教授. センサアレー信号処理, 鮭追跡システムの開発, 国際会議運営支援システムの開発, 電子波包絡の回路モデルの研究に従事. 平13より北見工大教授. 工博.