バイオメディカル・ファジィ・システム学会誌　Vol.20, No.1, pp.15－22（2018）

[Original article]

（2018 年 5 月 1 日　Accepted）

顧客クラスが変化する推薦システムにおける半教師付き学習

前田　康成[1]，山内　翔[1]，鈴木　正清[1]，松嶋　敏泰[2]

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：　従来研究では，顧客クラスが変化する推薦システムを表現する確率モデルとしてマルコフ決定過程が採用され，マルコフ決定過程の真のパラメータが既知の仮定のもとで検討されている．本研究では，真のパラメータが未知の仮定のもとで顧客クラスが変化する推薦システムにおける半教師付き学習方法を提案する．学習データは完全データと不完全データによって構成される．提案方法では EM アルゴリズムを用いる．シミュレーションによって提案方法の有効性を示す．

キーワード：推薦システム，マルコフ決定過程，顧客クラス，半教師付き学習

Semi-supervised Learning on Recommender System with Transitions of User Classes

Yasunari MAEDA[1], Sho YAMAUCHI[1], Masakiyo SUZUKI[1], Toshiyasu MATSUSHIMA[2]

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

Abstract: *Markov decision processes(MDP) are applied to recommender system with transitions of user classes. In previous research the true parameters of MDP are known. In this research we propose a semi-supervised learning method for recommender system with transitions of user classes under the condition that the true parameters are unknown. Learning data consist of complete data and incomplete data. In the proposed method EM algorithm is used. The effectiveness of the proposed method is shown by some simulations.*

Keywords: *recommender system, Markov decision processes, user class, semi-supervised learning*

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の商品推薦などに利用される推薦システムを扱う．従来研究[1][2][3][4]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[5]に基づきベイズ基準のもとで最大化している．食品や薬品などの購買は顧客の健康状況に依存し，趣味に関する購買は顧客の当該時点での趣味に依存する．このような商品に関する顧客のクラスは時間の経過とともに変化することが想定されるが，従来研究[3]では顧客のクラスの変化は考慮していない．

他方，顧客クラスの変化を考慮した推薦システムにおける売上高のベイズ基準のもとでの最大化も検討されているが，従来研究[4]では各種確率分布を支配する真のパラメータが既知と仮定されている．

そこで，本研究ではより現実に近い真のパラメータが未知の場合を対象にする．顧客のクラスも既知の完全データと顧客のクラスは未知の不完全データからなる学習データを受け取ったもとで未知パラメータを学習する半教師付き学習方法を検討する．EMアルゴリズム[6]を用いて半教師付き学習を実施し，商品の推薦については真のパラメータが既知の場合の従来研究[4]で提案されている推薦方法にEMアルゴリズムによる学習結果を適用する．提案方法の有効性をシミュレーションによって確認する．

2. 準備

ここでは，本研究で用いる各種記号などのうち，従来研究[4]と共通部分の定義を行う．

本研究では従来研究[1][2][3][4]と同様に推薦システムをマルコフ決定過程（MDP）[7]を用いてモデル化する．従来研究のMDPによる推薦システムのモデル化では，顧客が商品を購入する確率が当該顧客に対する推薦履歴や当該顧客の購入履歴に依存すると仮定している．例えば購入履歴のみに依存する単純な例として，書籍販売サイトにおいて全10巻のシリーズ物の5巻までを購入済の顧客であれば，次回アクセス時に同シリーズ物の6巻を購入する確率が大きい傾向にあることが挙げられる．

c_i，$c_i \in C$ は顧客のクラスを示し，$C = \{c_1, c_2, \ldots, c_{|C|}\}$ は顧客クラスの集合である．顧客クラスはさまざまなものが表現できる．例えば，書籍販売サイトであれば歴史小説，イギリス文学，アメリカ文学などの顧客の趣味によるクラス表現が考えられる．また，食品／健康食品販売サイトであれば顧客の嗜好や健康状況によるクラス表現も考えられる．m_i，$m_i \in M$ は推薦対象の商品を示し，$M = \{m_1, m_2, \ldots, m_{|M|}\}$ は商品集合である．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，$N = \{n_1, n_2, \ldots, n_{|N|}\}$ は反応集合である．$1 \le i \le |M|$ 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．$r(n_i)$，$1 \le i \le |M|$ 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．本研究では商品の推薦を T 回行う T 期間のMDP問題を扱う．換言すると，通販サイトへの初回のアクセスも含めて当該顧客が同通販サイトに今後 T 回アクセスするという設定のもとで，T 回の売上高の総和の期待値を最大化する．なお，本研究では有限の T が既知と仮定している．本来であれば，実際の推薦システムでは扱う商品の性質，対象とする顧客の性質，利用する計算機環境に合わせて適切な T を決定する必要がある．具体的な検討は今後の課題とするが，利便性を考慮すると，顧客に対して1度だけ大きな T に関する大規模のMDP問題を解くのではなく，顧客のアクセスの度にあまり大きくない適当な大きさの T に関する小規模のMDP問題を解いて当該MDP問題の初期の選択を採用することを繰り返すのが妥当な運用方法だと思われる．

y_t，$y_t \in M$ は t 回目の推薦商品を示し，MDP の t 期の行動に相当する．z_t，$z_t \in N$ は t 回目の推薦に対する顧客の t 期の反応を示す．w_t，$w_t \in C$ は t 期の顧客のクラスを示し，未知である．(y_{t-1}, z_{t-1}) は t 回目の推薦を受ける際の顧客の状態で，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．本研究では過去1期間の履歴のみに依存するモデルを採用しているが，依存する履歴の長さを拡張することは容易である．

従来研究[4]と同様に顧客クラスはマルコフ連鎖に従って変化し，$p(w_{t+1}|w_t,\theta^*)$は顧客クラスの遷移確率を示す．遷移確率は顧客クラスの変化する確率であり，顧客の趣味／嗜好／健康状況によるクラス表現の場合であれば，当該顧客の趣味／嗜好／健康状況が変化する確率に相当する．本研究のモデルでは顧客が通販サイトなどにアクセスするタイミングを1つの期として扱い，顧客クラスの変化は次の期に進む際に起きると仮定している．換言すると，通販サイトへのアクセスとアクセスの間で顧客クラスが変化すると仮定している．$p(w_1|\theta^*)$は顧客クラスの初期確率である．従来研究[4]では初期確率の代わりにベイズ統計学における事前確率を導入している．$p(z_t|y_{t-1},z_{t-1},y_t,w_t,\theta^*)$は，1期前の推薦商品と反応が$y_{t-1}$，$z_{t-1}$でクラス$w_t$の顧客が推薦商品$y_t$に対して反応が$z_t$となる確率で，顧客の状態遷移確率を示し，MDPの状態遷移確率に相当する．θ^*は確率分布を支配する真のパラメータで従来研究では既知，本研究では未知である．$w^ty^tz^t$はt期までの系列$w^ty^tz^t=w_1y_1z_1\cdots w_ty_tz_t$を示す．

3. 真のパラメータθ^*既知の従来研究における推薦方法

各種確率分布を支配する真のパラメータθ^*既知の場合を検討した従来研究[4]におけるベイズ最適な推薦方法を以下に掲載する．本研究でも，半教師付き学習後のパラメータの推定値を代入することによって，当該推薦方法を利用する．

当該推薦方法では動的計画法（DP）を用いて，T期目から遡りながら計算して，ベイズ最適な推薦方法を算出する．

T期目まで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．

$$V(y^{T-1}z^{T-1},T)=\max_{y_T\in M}\sum_{z_T\in N}p(z_T|y^{T-1}z^{T-1},y_T,\theta^*)r(z_T), \quad (1)$$

ただし，

$$p(z_T|y^{T-1}z^{T-1},y_T,\theta^*)=\sum_{w_T\in C}p(w_T|y^{T-1}z^{T-1})p(z_T|y_{T-1},z_{T-1},y_T,w_T,\theta^*), \quad (2)$$

$$p(w_T|y^{T-1}z^{T-1})=\sum_{w_{T-1}\in C}p'(w_{T-1}|y^{T-1}z^{T-1})p(w_T|w_{T-1},\theta^*), \quad (3)$$

$$p'(w_{T-1}|y^{T-1}z^{T-1})=\frac{p\left(w_{T-1}|y^{T-2}z^{T-2}\right)g}{\sum_{w\in C}p\left(w|y^{T-2}z^{T-2}\right)g'}, \quad (4)$$

$$g=p(z_{T-1}|y_{T-2},z_{T-2},y_{T-1},w_{T-1},\theta^*), \quad (5)$$

$$g'=p(z_{T-1}|y_{T-2},z_{T-2},y_{T-1},w,\theta^*), \quad (6)$$

$p(z_T|y^{T-1}z^{T-1},y_T,\theta^*)$は顧客の反応確率の事後確率$p(w_T|y^{T-1}z^{T-1})$による期待値で，$V(y^{T-1}z^{T-1},T)$は最後の期の期待利得の最大値である．式(1)の右辺を最大化する行動y_Tがベイズ最適な推薦商品である．

t期目（$1\leq t<T$）まで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．

$$V(y^{t-1}z^{t-1},t)=\max_{y_t\in M}\sum_{z_t\in N}p(z_t|y^{t-1}z^{t-1},y_t,\theta^*)\left(r(z_t)+V(y^tz^t,t+1)\right). \quad (7)$$

式(7)の右辺を最大化する行動y_tがベイズ最適な推薦商品である．

4. EMアルゴリズムによる半教師付き学習方法の提案

各種確率分布を支配する真のパラメータが未知の場合の半教師付き学習方法を，EMアルゴリズムを適用することによって提案する．なお，本研究で扱っている未知情報を伴うMDP（部分観測MDP）は，MDPと隠れマルコフモデル（HMM）[8]を融合させたモデルとも解釈できる．HMMの視点から解釈すると，以下4.2節のαおよびβの計算方法はHMMにおける前向きアルゴリズムおよび後向きアルゴリズムの拡張に相当する．また，本章のEMアルゴリズムの適用結果はHMMにおけるバウム・ウェルチアルゴリズムの拡張に相当する．半教師付き学習は自然言語処理における文書分類問題などに対して多くの従来研究[9]があり，本研究は文書分類問題などにおける半教師付き学習の成果を推薦システムに適用したとも解釈できる．

4.1　学習データ（学習系列）

顧客クラスが既知のL_c本の完全データの学習系列$W'Y'Z'$と，顧客クラスが未知のL_u本の不完全データの学習系列YZを受け取ったもとで半教師付き学習を実施する．ただし，

$$W'Y'Z'=W_1'Y_1'Z_1'W_2'Y_2'Z_2'\cdots W_{L_c}'Y_{L_c}'Z_{L_c}', \quad (8)$$

W_i'，Y_i'，Z_i'はそれぞれi本目の完全データの顧客クラス系列，推薦商品系列，顧客の反応系列である．

$$W_i'Y_i'Z_i'=y_{i,0}'z_{i,0}'w_{i,1}'y_{i,1}'z_{i,1}'\cdots w_{i,T_i'}'y_{i,T_i'}'z_{i,T_i'}', \quad (9)$$

$w_{i,j}'$，$y_{i,j}'$，$z_{i,j}'$はそれぞれi本目の完全データのj期目の顧客クラス，推薦商品，反応を示し，T_i'はi本目の完全データの長さ（当該データにおける推薦回数）である．なお，$y_{i,0}'$，$z_{i,0}'$は1期目のMDPの状態を構成するための情報（1期前の推薦と反応）であり，推薦回数には含めない．

$$YZ=Y_1Z_1Y_2Z_2\cdots Y_{L_u}Z_{L_u}, \quad (10)$$

Y_i，Z_iはそれぞれi本目の不完全データの推薦商品系列，

顧客の反応系列である．

$$Y_i Z_i = y_{i,0} z_{i,0} y_{i,1} z_{i,1} \cdots y_{i,T_i} z_{i,T_i}, \tag{11}$$

$y_{i,j}$，$z_{i,j}$はそれぞれi本目の不完全データのj期目の推薦商品，反応を示し，T_iはi本目の不完全データの長さ（当該データにおける推薦回数）である．なお，$y_{i,0}$，$z_{i,0}$は1期目のMDPの状態を構成するための情報（1期前の推薦と反応）であり，推薦回数には含めない．

4.2 EMアルゴリズムの不完全データへの適用

ここでは，不完全データに対するEMアルゴリズムの適用について説明する．完全データについては，次節で不完全データへのEMアルゴリズムの適用結果を拡張することによって対応する．

不完全データの生起確率は以下のとおりである．

$$p(YZ|\theta) = \sum_W p(WYZ|\theta), \tag{12}$$

ただし，Wは不完全データに対応する未観測（未知）の顧客クラス系列である．

$$\begin{aligned} p(WYZ|\theta) = & \\ & \prod_{i=1}^{L_u} p(w_{i,1}|\theta) p(z_{i,1}|y_{i,0}, z_{i,0}, y_{i,1}, w_{i,1}, \theta) \\ & \prod_{t=2}^{T_i} p(w_{i,t}|w_{i,t-1}, \theta) \\ & p(z_{i,t}|y_{i,t-1}, z_{i,t-1}, y_{i,t}, w_{i,t}, \theta). \end{aligned} \tag{13}$$

いくつか簡略化した書き方を用いる．

$$\rho_i = p(c_i|\theta). \tag{14}$$

$$a_{ij} = a(c_i, c_j) = p(c_j|c_i, \theta). \tag{15}$$

$$\begin{aligned} b_{ijklq} &= b(m_j, n_k, m_l, c_i, n_q) \\ &= p(n_q|m_j, n_k, m_l, c_i, \theta). \end{aligned} \tag{16}$$

$$\alpha_{v,t}(i)\beta_{v,t}(i) = p(Y_v, Z_v, w_{v,t} = c_i|\theta), \tag{17}$$

ただし，

$$\alpha_{v,t}(i) = p(y_{v,1} \cdots y_{v,t}, z_{v,1} \cdots z_{v,t}, w_{v,t} = c_i|\theta), \tag{18}$$

$$\begin{aligned} &\beta_{v,t}(i) = \\ &p(y_{v,t+1} \cdots y_{v,T_v}, z_{v,t+1} \cdots z_{v,T_v}|w_{v,t} = c_i, \theta), \end{aligned} \tag{19}$$

$2 \le t \le T_v$のtに対して，$\alpha_{v,t}(j)$は再帰的に計算できる．

$$\begin{aligned} \alpha_{v,t}(j) = & \left(\sum_{i=1}^{|C|} \alpha_{v,t-1}(i) a_{ij}\right) \\ & b(y_{v,t-1}, z_{v,t-1}, y_{v,t}, c_j, z_{v,t}). \end{aligned} \tag{20}$$

$$\begin{aligned} \alpha_{v,1}(i) &= p(y_{v,1}, z_{v,1}, w_{v,1} = c_i|\theta) \\ &= \rho_i b(y_{v,0}, z_{v,0}, y_{v,1}, c_i, z_{v,1}), \end{aligned} \tag{21}$$

$$\begin{aligned} p(Y_v, Z_v|\theta) &= p(y_{v,1} \cdots y_{v,T_v}, z_{v,1} \cdots z_{v,T_v}|\theta) \\ &= \sum_{i=1}^{|C|} \alpha_{v,T_v}(i). \end{aligned} \tag{22}$$

$1 \le t \le T_v - 1$のtに対して，

$$\begin{aligned} &\beta_{v,t}(i) = \\ &\sum_{j=1}^{|C|} a_{ij} b(y_{v,t}, z_{v,t}, y_{v,t+1}, c_j, z_{v,t+1}) \beta_{v,t+1}(j). \end{aligned} \tag{23}$$

$$\beta_{v,T_v}(i) = 1. \tag{24}$$

不完全データのv本目の推薦商品系列Y_vと顧客の反応系列Z_vが得られたもとで，t期目の顧客クラス$w_{v,t}$がc_iである確率は次式のとおりである．

$$\begin{aligned} \gamma_{v,t}(i) &= p(w_{v,t} = c_i|Y_v, Z_v, \theta) \\ &= \frac{p(Y_v, Z_v, w_{v,t} = c_i|\theta)}{p(Y_v, Z_v|\theta)} = \frac{\alpha_{v,t}(i)\beta_{v,t}(i)}{\sum_{i=1}^{|C|} \alpha_{v,T_v}(i)} \end{aligned} \tag{25}$$

不完全データのv本目の推薦商品系列Y_vと顧客の反応系列Z_vが得られたもとで，t期目の顧客クラス$w_{v,t}$がc_iかつ$t+1$期目の顧客クラス$w_{v,t+1}$がc_jである確率は次式のとおりである．

$$\begin{aligned} \xi_{v,t}(i,j) &= p(w_{v,t} = c_i, w_{v,t+1} = c_j|Y_v, Z_v, \theta) \\ &= \frac{p(Y_v, Z_v, w_{v,t} = c_i, w_{v,t+1} = c_j|\theta)}{p(Y_v, Z_v|\theta)} \\ &= \frac{\alpha_{v,t}(i) a_{ij} b(y_{v,t}, z_{v,t}, y_{v,t+1}, c_j, z_{v,t+1}) \beta_{v,t+1}(j)}{\sum_{i=1}^{|C|} \alpha_{v,T_v}(i)}. \end{aligned} \tag{26}$$

EMアルゴリズムにおいて最大化する$Q(\theta^0, \theta)$関数を次式に示す．θはθ^0の更新後のパラメータである．

$$\begin{aligned} Q(\theta^0, \theta) &= \sum_W p(W|Y, Z, \theta^0) \log p(W, Y, Z|\theta) \\ &= \sum_W p(W|Y, Z, \theta^0) \sum_{v=1}^{L_u} \log p(w_{v,1}) \\ &\quad + \sum_W p(W|Y, Z, \theta^0) \sum_{v=1}^{L_u} \sum_{t=1}^{T_v-1} \log a(w_{v,t}, w_{v,t+1}) \\ &\quad + \sum_W p(W|Y, Z, \theta^0) \\ &\qquad \sum_{v=1}^{L_u} \sum_{t=1}^{T_v} \log b(y_{v,t-1}, z_{v,t-1}, y_{v,t}, w_{v,t}, z_{v,t}) \\ &= Q(\theta^0, \rho) + Q(\theta^0, A) + Q(\theta^0, B), \end{aligned} \tag{27}$$

ただし，

$$Q(\theta^0, \rho) = \sum_W p(W|Y, Z, \theta^0) \sum_{v=1}^{L_u} \log p(w_{v,1}), \tag{28}$$

$$\begin{aligned} &Q(\theta^0, A) = \\ &\sum_W p(W|Y, Z, \theta^0) \sum_{v=1}^{L_u} \sum_{t=1}^{T_v-1} \log a(w_{v,t}, w_{v,t+1}), \end{aligned} \tag{29}$$

$$\begin{aligned} &Q(\theta^0, B) = \sum_W p(W|Y, Z, \theta^0) \\ &\quad \sum_{v=1}^{L_u} \sum_{t=1}^{T_v} \log b(y_{v,t-1}, z_{v,t-1}, y_{v,t}, w_{v,t}, z_{v,t}), \end{aligned} \tag{30}$$

$Q(\theta^0, \rho)$の最大化を以下で説明する．

$$\begin{aligned} Q(\theta^0, \rho) &= \sum_W p(W|Y, Z, \theta^0) \sum_{v=1}^{L_u} \log p(w_{v,1}) \\ &= \sum_{i=1}^{|C|} \Big(\sum_{v=1}^{L_u} \sum_{W, w_{v,1} = c_i} \\ &\qquad \prod_{l=1}^{L_u} p(w_{l,1} \cdots w_{l,T_l}|Y_l, Z_l, \theta^0)\Big) \log \rho_i \\ &= \sum_{i=1}^{|C|} \sum_{v=1}^{L_u} \gamma_{v,1}(i) \log \rho_i = \sum_{i=1}^{|C|} d_i \log \rho_i, \end{aligned} \tag{31}$$

ただし，

$$d_i = \sum_{v=1}^{L_u} \gamma_{v,1}(i). \tag{32}$$

$\sum_{i=1}^{|C|} \rho_i = 1$の制約のもとで，$f = \sum_{i=1}^{|C|} d_i \log \rho_i$を最大化する$\rho_i$をラグランジュの未定乗数法によって求める．ラグランジュの未定乗数法より，

$$L = f - \lambda\left(\sum_{i=1}^{|C|} \rho_i - 1\right), \tag{33}$$

として，ρ_iで偏微分して0とおくと以下のようになる．

$$\frac{\partial L}{\partial \rho_i} = \frac{\partial f}{\partial \rho_i} - \lambda = \frac{d_i}{\rho_i} - \lambda = 0. \tag{34}$$

$$\lambda \rho_i = d_i. \tag{35}$$

式(35)の両辺をiについて（顧客クラス集合の全要素について）足し合わせて制約を考慮すると以下のようになる．

$$\lambda = \sum_{i=1}^{|C|} d_i. \tag{36}$$

よって，最大化するρ_iの推定値$\hat{\rho}_i$は次式のとおりである．

$$\hat{\rho}_i = \frac{d_i}{\lambda} = \frac{d_i}{\sum_{i=1}^{|C|} d_i} = \frac{\sum_{v=1}^{L_u} \gamma_{v,1}(i)}{\sum_{i=1}^{|C|}\sum_{v=1}^{L_u} \gamma_{v,1}(i)} = \frac{\sum_{v=1}^{L_u} \gamma_{v,1}(i)}{L_u}. \tag{37}$$

式(37)の分子はL_u本の不完全データ系列において先頭の顧客クラスがc_iの頻度の期待値である．

$Q(\theta^0, A)$の最大化を以下で説明する．

$$\begin{aligned} Q(\theta^0, A) &= \sum_W p(W|Y,Z,\theta^0) \sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1} \log a(w_{v,t}, w_{v,t+1}) \\ &= \sum_{i=1}^{|C|}\sum_{j=1}^{|C|}\Big(\sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\sum_{W,w_{v,t}=c_i,w_{v,t+1}=c_j} \\ &\qquad \prod_{l=1}^{L_u} p\big(w_{l,1}\cdots w_{l,T_l}\big|Y_l, Z_l, \theta^0\big)\Big)\log a_{ij} \\ &= \sum_{i=1}^{|C|}\sum_{j=1}^{|C|}\Big(\sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\xi_{v,t}(i,j)\Big)\log a_{ij} \\ &= \sum_{i=1}^{|C|}\Big(\sum_{j=1}^{|C|} d_{ij}\log a_{ij}\Big), \end{aligned} \tag{38}$$

ただし，

$$d_{ij} = \sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\xi_{v,t}(i,j). \tag{39}$$

$\sum_{j=1}^{|C|} a_{ij} = 1$の制約のもとで，各$i$に対して$f = \sum_{j=1}^{|C|} d_{ij}\log a_{ij}$を最大化する$a_{ij}$をラグランジュの未定乗数法によって求める．ラグランジュの未定乗数法より，

$$L = f - \lambda\Big(\sum_{j=1}^{|C|} a_{ij} - 1\Big), \tag{40}$$

として，a_{ij}で偏微分して0とおくと以下のようになる．

$$\frac{\partial L}{\partial a_{ij}} = \frac{\partial f}{\partial a_{ij}} - \lambda = \frac{d_{ij}}{a_{ij}} - \lambda = 0. \tag{41}$$

$$\lambda a_{ij} = d_{ij}. \tag{42}$$

式(42)の両辺をjについて（顧客クラス集合の全要素について）足し合わせて制約を考慮すると以下のようになる．

$$\lambda = \sum_{j=1}^{|C|} d_{ij}. \tag{43}$$

よって，最大化するa_{ij}の推定値$\hat{a}_{ij}$は次式のとおりである．

$$\begin{aligned} \hat{a}_{ij} &= \frac{d_{ij}}{\lambda} = \frac{d_{ij}}{\sum_{j=1}^{|C|} d_{ij}} = \frac{\sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\xi_{v,t}(i,j)}{\sum_{j=1}^{|C|}\sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\xi_{v,t}(i,j)} \\ &= \frac{\sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\xi_{v,t}(i,j)}{\sum_{v=1}^{L_u}\sum_{t=1}^{T_v-1}\gamma_{v,t}(i)}. \end{aligned} \tag{44}$$

式(44)の分子はL_u本の不完全データ系列において顧客クラスc_iからc_jへ遷移した頻度の期待値，分母は遷移元が顧客クラスc_iの頻度の期待値である．

$Q(\theta^0, B)$の最大化を以下で説明する．

$$\begin{aligned} Q(\theta^0, B) &= \sum_W p(W|Y,Z,\theta^0) \\ &\quad \sum_{v=1}^{L_u}\sum_{t=1}^{T_v} \log b\big(y_{v,t-1}, z_{v,t-1}, y_{v,t}, w_{v,t}, z_{v,t}\big) \\ &= \sum_{i=1}^{|C|}\sum_{j=1}^{|M|}\sum_{k=1}^{|N|}\sum_{l=1}^{|M|}\sum_{q=1}^{|N|}\Big(\sum_{v=1}^{L_u}\sum_{t=1}^{T_v} \\ &\quad \delta\big(y_{v,t-1}, z_{v,t-1}, y_{v,t}, z_{v,t}, m_j, n_k, m_l, n_q\big) \\ &\quad \sum_{W,w_{v,t}=c_i}\prod_{u=1}^{L_u} p\big(w_{u,1}\cdots w_{u,T_u}\big|Y_u, Z_u, \theta^0\big)\Big) \\ &\quad \log b_{ijklq} \\ &= \sum_{i=1}^{|C|}\sum_{j=1}^{|M|}\sum_{k=1}^{|N|}\sum_{l=1}^{|M|}\sum_{q=1}^{|N|}\Big(\sum_{v=1}^{L_u}\sum_{t=1}^{T_v} \\ &\quad \delta\big(y_{v,t-1}, z_{v,t-1}, y_{v,t}, z_{v,t}, m_j, n_k, m_l, n_q\big)\gamma_{v,t}(i)\Big) \\ &\quad \log b_{ijklq} \\ &= \sum_{i=1}^{|C|}\sum_{j=1}^{|M|}\sum_{k=1}^{|N|}\sum_{l=1}^{|M|}\Big(\sum_{q=1}^{|N|} d_{ijklq}\log b_{ijklq}\Big), \end{aligned} \tag{45}$$

ただし，

$$\delta\big(y_{v,t-1}, z_{v,t-1}, y_{v,t}, z_{v,t}, m_j, n_k, m_l, n_q\big) = \begin{cases} 1, & y_{v,t-1} = m_j \text{かつ} z_{v,t-1} = n_k \text{かつ} \\ & y_{v,t} = m_l \text{かつ} z_{v,t} = n_q\,; \\ 0, & \text{それ以外}, \end{cases} \tag{46}$$

$$\begin{aligned} d_{ijklq} &= \sum_{v=1}^{L_u}\sum_{t=1}^{T_v} \\ &\quad \delta\big(y_{v,t-1}, z_{v,t-1}, y_{v,t}, z_{v,t}, m_j, n_k, m_l, n_q\big)\gamma_{v,t}(i). \end{aligned} \tag{47}$$

$\sum_{q=1}^{|N|} b_{ijklq} = 1$の制約のもとで，各$(i,j,k,l)$の組に対して$f = \sum_{q=1}^{|N|} d_{ijklq}\log b_{ijklq}$を最大化する$b_{ijklq}$をラグランジュの未定乗数法によって求める．ラグランジュの未定乗数法より，

$$L = f - \lambda\Big(\sum_{q=1}^{|N|} b_{ijklq} - 1\Big), \tag{48}$$

として，b_{ijklq}で偏微分して0とおくと以下のようになる．

$$\frac{\partial L}{\partial b_{ijklq}} = \frac{\partial f}{\partial b_{ijklq}} - \lambda = \frac{d_{ijklq}}{b_{ijklq}} - \lambda = 0. \tag{49}$$

$$\lambda b_{ijklq} = d_{ijklq}. \tag{50}$$

式(50)の両辺をqについて（顧客の反応集合の全要素について）足し合わせて制約を考慮すると以下のようになる．

$$\lambda = \sum_{q=1}^{|N|} d_{ijklq}. \tag{51}$$

よって，最大化するb_{ijklq}の推定値$\hat{b}_{ijklq}$は次式のとおりである．

$$\hat{b}_{ijklq} = \frac{d_{ijklq}}{\lambda} = \frac{d_{ijklq}}{\sum_{q=1}^{|N|} d_{ijklq}} = \frac{\sum_{v=1}^{L_u} \sum_{t=1}^{T_v} G}{\sum_{q=1}^{|N|} \sum_{v=1}^{L_u} \sum_{t=1}^{T_v} G}, \quad (52)$$

ただし，

$$G = \delta(y_{v,t-1}, z_{v,t-1}, y_{v,t}, z_{v,t}, m_j, n_k, m_l, n_q)\gamma_{v,t}(i). \quad (53)$$

式(52)の分子はL_u本の不完全データ系列において顧客クラスc_iの顧客が1期前に商品m_jを推薦されて反応n_kを返し，当該期に商品m_lを推薦されて反応n_qを返す頻度の期待値である．分母は，顧客クラスc_iの顧客が1期前に商品m_jを推薦されて反応n_kを返し，当該期に商品m_lを推薦される頻度の期待値である．

4.3　完全データへの対応

不完全データにEMアルゴリズムを適用する際のパラメータの更新式，式(37)，式(44)，式(52)を完全データにも対応するように拡張する．

式(37)の分子はL_u本の不完全データ系列において先頭の顧客クラスがc_iの頻度の期待値である．他方，完全データでは顧客クラスは既知なので，確定的な値になる．よって，完全データも追加したρ_iの推定は以下のようになる．

$$\hat{\rho}_i = \frac{g_i + \sum_{v=1}^{L_u} \gamma_{v,1}(i)}{L_c + L_u}, \quad (54)$$

ただし，g_iは完全データ中で先頭の顧客クラスがc_iである頻度を示す．

式(44)の分子はL_u本の不完全データ系列において顧客クラスc_iからc_jへ遷移した頻度の期待値，分母は遷移元が顧客クラスc_iの頻度の期待値である．他方，完全データでは顧客クラスは既知なので，確定的な値になる．よって，完全データも追加したa_{ij}の推定は以下のようになる．

$$\hat{a}_{ij} = \frac{g_{ij} + \sum_{v=1}^{L_u} \sum_{t=1}^{T_v - 1} \xi_{v,t}(i,j)}{\sum_{j=1}^{|C|} g_{ij} + \sum_{v=1}^{L_u} \sum_{t=1}^{T_v - 1} \gamma_{v,t}(i)}. \quad (55)$$

ただし，g_{ij}は完全データ中で顧客クラスc_iからc_jへ遷移した頻度を示す．

式(52)の分子はL_u本の不完全データ系列において顧客クラスc_iの顧客が1期前に商品m_jを推薦されて反応n_kを返し，当該期に商品m_lを推薦されて反応n_qを返す頻度の期待値である．分母は，顧客クラスc_iの顧客が1期前に商品m_jを推薦されて反応n_kを返し，当該期に商品m_lを推薦される頻度の期待値である．他方，完全データでは顧客クラスは既知なので，確定的な値になる．よって，完全データも追加したb_{ijklq}の推定は以下のようになる．

$$\hat{b}_{ijklq} = \frac{g_{ijklq} + \sum_{v=1}^{L_u} \sum_{t=1}^{T_v} G}{\sum_{q=1}^{|N|} g_{ijklq} + \sum_{q=1}^{|N|} \sum_{v=1}^{L_u} \sum_{t=1}^{T_v} G}, \quad (56)$$

ただし，

$$G = \delta(y_{v,t-1}, z_{v,t-1}, y_{v,t}, z_{v,t}, m_j, n_k, m_l, n_q)\gamma_{v,t}(i). \quad (57)$$

g_{ijklq}は完全データ中で顧客クラスc_iの顧客が1期前に商品m_jを推薦されて反応n_kを返し，当該期に商品m_lを推薦されて反応n_qを返した頻度を示す．

提案方法におけるEMアルゴリズムの具体的な手順は，最初に各パラメータρ_i, a_{ij}, b_{ijklq}の初期値を設定する．次に，式(54)，式(55)，式(56)を用いて各パラメータを推定し，ρ_i，a_{ij}，b_{ijklq}を更新する．推定結果が収束するまで推定と更新を繰り返す．なお，式(54)，式(55)，式(56)の計算には式(20)，式(21)，式(22)，式(23)，式(24)，式(25)，式(26)を用いる．

実際の商品の推薦については3章記載の真のパラメータ既知でクラス未知の従来研究[4]の推薦方法にEMアルゴリズムによる推定結果を代入して実施する．

5. シミュレーション例

提案方法の有効性を確認するため，シミュレーション例を報告する．

顧客クラス数$|C| = 3$，商品数$|M| = 4$，推薦回数$T = 4$として，真のパラメータρ_i，a_{ij}，b_{ijklq}と商品の利得$r(m_i)$を一様乱数によって200パターン設定した．なお，商品の利得$r(m_i)$は100以上1000以下の100きざみで設定した．

EMアルゴリズムの初期値も一様乱数で設定し，EMアルゴリズムの繰り返し回数は事前実験より収束するのに十分と想定される1000とした．

学習データの完全データ数L_cと不完全データ数L_uについては，それぞれ10，100，1000としてデータ数の9組の組合せを設定した．学習データの系列長はすべて上記の推薦回数$T = 4$と同じにした．学習データは各シミュレーションパターンの真のパラメータに従って生成した．

比較対象として，真のパラメータ既知の場合の従来方法に真のパラメータを代入した場合のシミュレーションも行った．なお，従来方法に代入するクラスの事前確率としてρ_iを用いた．完全データ，不完全データおよびシミュレーション中の0期の推薦商品と顧客の反応も一様乱数で生成した．

上記200パターンのシミュレーション結果を表1に報

告する．表1中の数値は真のパラメータ既知の場合の総利得に対する提案方法による総利得の割合の百分率である．当該百分率は各パターンに対して推薦回数4回のMDPによる試行を1000回実施した200パターンでの平均である．

表1.　シミュレーション結果

	$L_u = 10$	$L_u = 100$	$L_u = 1000$
$L_c = 10$	87.98	88.19	93.85
$L_c = 100$	88.02	88.81	94.96
$L_c = 1000$	96.64	96.94	98.33

シミュレーション結果より，完全データが少量の場合でも不完全データが増加すれば総利得が大きくなることが確認できる．また，学習データの増加にともない総利得が大きくなっており，十分なデータ数があれば真のパラメータ既知の場合の総利得に近い利得が獲得できていることも確認した．

よって，小規模なシミュレーションであるが，EMアルゴリズムを用いた提案方法による半教師付き学習の効果が確認できた．

6.　まとめと今後の課題

従来から顧客が属する顧客クラスが変化する推薦問題において，顧客クラスが未知という条件のもとで売上高の最大化が検討されている．しかし，従来研究では各種確率分布を支配する真のパラメータは既知と仮定されている．

そこで，本研究ではより現実に近い真のパラメータ未知の場合を対象に，顧客クラスが既知である完全データと顧客クラスが未知である不完全データで構成される学習データ（推薦と購買の履歴データ）に対する半教師付き学習方法を検討した．提案方法はEMアルゴリズムに基づく方法であり，学習結果である未知パラメータの推定結果を真のパラメータが既知の場合の従来の推薦方法に代入することによって商品の推薦を実施する．

小規模なシミュレーションであるが，提案方法によるパラメータの推定結果を従来の推薦方法に代入した場合と真のパラメータ既知の場合の獲得利得を比較することにより，提案方法の有効性を確認した．

提案方法のより詳細な性質に関する検討や，より現実の推薦システムに近い問題設定に関する検討は今後の課題としたい．

本研究で扱っている確率モデルはMDPと隠れマルコフモデル（HMM）[8]を融合させたモデルであり，HMMの視点から本研究を解釈すると，4.2節のαおよびβの計算方法はHMMの前向きアルゴリズムおよび後向きアルゴリズムの拡張に相当し，4章のEMアルゴリズムの適用結果はHMMのバウム・ウェルチアルゴリズムの拡張に相当する．また，本研究は文書分類問題などにおける半教師付き学習の成果[9]を推薦システムに適用したとも解釈できる．

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の検討内容は適用可能である．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 に関する学習問題を対象としている．従来研究[10]では，未知情報を伴う MDP を医療検査・治療方針の決定支援に適用している．当該従来研究では，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 の行動として，ヘルスケアに関する患者のコスト最小化を検討しているが，各種確率分布を支配する真のパラメータは既知のもとで検討している．当該従来研究の問題設定を各種確率分布の真のパラメータが未知の場合における半教師付き学習問題に拡張する際には，本研究の検討内容が適用可能である．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．

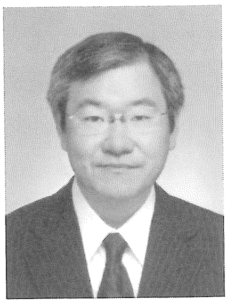
謝辞

本研究の一部は JSPS 科研費 JP16K00417 の助成による．

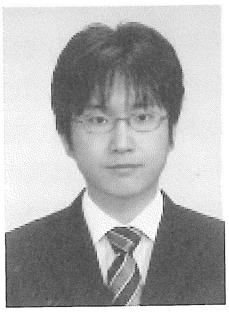
参考文献

[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] 前田康成，鈴木正清，松嶋敏泰：顧客クラスが変化する推薦システムに関する一考察，電気学会論文誌C, Vol.137, No.6, pp.815-816, 2017.

[5] J.O. Berger : Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, New York, 1980.
[6] A.P. Dempster, N.M. Laird, and D.B. Rubin, : Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society, Series B, Vol.39, No.1, pp.1-38, 1977.
[7] 金子哲夫：マルコフ決定理論入門，槙書店，東京，1973.
[8] X.D. Huang, Y. Ariki, and M.A. Jack, : Hidden Markov Models for Speech Recognition, Edinburgh University Press, Edinburgh, 1990.
[9] K. Nigam, A.K. Mccallum, S. Thrun, and T. Mitchell, : Text Classification from Labeled and Unlabeled Documents using EM, Machine Learning, Vol.39, pp.103-134, 2000.
[10] 前田康成，山内翔，鈴木正清，高野賢裕，松嶋敏泰：マルコフ決定過程を用いたヘルスケア支援に関する一考察，バイオメディカル・ファジィ・システム学会誌, Vol.19, No.2, pp.21-27, 2017.



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.

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・応用数理学科教授．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.