論文・事例研究

音楽CD購買における個人嗜好の推定

勝又壮太郎

1. はじめに

一般的な経済学において，個人の異質性は「誤差」や「撹乱項」として扱われている．しかしながら，マーケティングにおいては，個人は異質であると考えられている．近年の情報技術の発達によって，ID付きPOSデータなどの収集が容易になり，個人ごとの時系列購買履歴を記録することができるようになった．それに伴って，大量のデータを解析する手法が求められるようになっている．しかし，そのような膨大なデータが記録される一方で，個人の購買数は，1年間のデータを蓄積させても，分析ができるほど多くないケースも少なくない．例えば家庭用洗剤などは，普通は，1人（一家計）で1年間に何十回も購買されることはほとんどない．このようなデータを旧来の回帰モデル等でモデル化する場合，個人（家計）ごとの標本数が少ないため，個人差を考えない集計レベルでのモデル化が行われることが多い．しかし，それでは個人ごとの異質性をとらえることはできず，また，IDをつけて記録をしている個人の情報を捨ててしまうことになる．

こうした，従来の手法では推定できない小標本のデータを推定する方法として注目されているのが，MCMC（Markov Chain Monte Carlo：マルコフ連鎖モンテカルロ）法による推定手法を用いるベイズモデルである．MCMC法は，パラメータに分布を仮定し，その分布に従う乱数を発生させる，かなり従来の手法とは異なった推定法である．柔軟なモデルの拡張が可能であり，最尤法では推定が困難な，複雑なモデル化も可能である．マーケティング分野への応用も盛んに論じられている[1][11]．

本研究では，階層ベイズ（HB：Hierarchical Bayes）モデルを用いて，MCMC法の特性を生かした，大容量データベース解析への応用方法を探っていく．

2. データ概要と分析動機

2.1 データ概要

データは，平成17年度データ解析コンペティションにおいて提供された音楽CD販売店のID付きPOSデータである．集計期間は2年間．記録されている顧客のデモグラフィック変数は，年齢（生年月日）と性別である．

2.2 分析目的とプロセス

音楽CD小売店の分析に関して，以下の3点の特徴に留意する必要がある．まず1点は，顧客の嗜好の複雑性である．音楽CDという嗜好製品は，個人の嗜好が複雑で，自分に興味が無いアーティストは，いくらプロモーションをしても買われない．年齢や性別などからある程度推定可能な場合もあるが，同年齢・同性でも嗜好のばらつきは大きく，個人差を考慮しないモデル化は難しい．2点めは，購買枚数である．1年間に10枚以上購買する顧客はほとんどおらず，個人単位の分析では，精緻な選択行動の推定は難しい．3点めは，選択肢の多さである．数百～数千の選択肢（アーティスト）があり，ロジット，プロビットなどの通常の離散選択モデルをあてることは非常に困難である．

本研究では，この3点を考慮して分析を行う．まず，1点めと2点めの解決策として，MCMC法による階層モデルを用いる[3]．階層モデルは，個人単位の分析に，デモグラフィック変数などの集計的な傾向情報を付加するもので，個人差が大きく，かつ標本数の少ないデータなどに特に有効である．しかしながら，これらの階層モデルの研究では，離散選択モデルが用いられることが多く，3点めにあるように，膨大な選択肢をもつ音楽CDのデータには応用することができない．そこで，3点めの解決策として，これらの選択肢（アーティスト）に，知覚マップ[6]，ジョイント・スペース[2]などで用いられている縮約手法を用いて多

かつまた　そうたろう
東京大学 大学院経済学研究科
〒 113-0033 文京区本郷 7-3-1
受付 06.7.20　採択 07.7.26

次元連続型の属性を与え，分析を行う．

3. モデル構築

2年間の購買データのうち，前半1年間のデータを学習期間としてパラメータを推定し，後半1年間のデータを検証期間として，モデルの精度を検証する．すなわち，直前1年間の購買履歴から，今後1年間に顧客が購買するアーティストの傾向を推定する．

3.1 分析対象の抽出

前半1年間で購買記録は全部で605,593件，購買されたアーティストの数は8,545人（組）．しかしながら，全売り上げの約81%を，売り上げ上位500のアーティストが占めている．売り上げ枚数が余りにも少ないと，アーティスト属性の数値化において，良い精度が期待しにくいので，対象を，上位500アーティストに限定した．

同じく前半1年間で，上で限定した500アーティストを購買した顧客は全部で161,805人．そのほとんどは，購買回数が1回～2回といった，ごく少数の購買しかしていない顧客である（購買回数1回：74,276人，購買回数2回：32,410人）．今回，モデル構築のために，3回以上の購買記録がある顧客55,119人を分析対象とした．ただし，分析から外れた顧客への追加推定についても後述する．

3.2 データの分割

ここで，前半1年間の，分析対象55,119人を，ランダムに2分割し，それぞれ集合A，集合Bとおく．そして，集合Aのデータからアーティスト属性を算出し，集合Bのデータを用いて個人嗜好の推定のためのモデル構築を行う．この処理は，同じデータから被説明変数となるアーティストの属性を算出し，同じデータの個人嗜好を推定してしまうという，分析のループを避けるために行う．アーティスト属性を算出する集合Aと個人嗜好の推定を行う集合Bを完全に分離させ，アーティスト属性を，外部から与えられた数値として扱うことができるようにしている．

3.3 アーティスト属性の数値化

3.3節では，アーティスト属性の数値化について説明する．顧客の併買行動に注目して，アーティストを分類する．3.2節で分割した集合Aを用いて属性を求める．

まず，購買記録を，行列にまとめる．縦軸に分析対象の個人を取り，横軸に500人（組）のアーティストをとった27,559×500の行列を作成する．行列の第ik成分は，顧客iが1年間に購買したアーティストkのCDの枚数となる．この行列の横軸は，個人の併買行動を示している．同じ個人に購買されているアーティストは属性が近いと仮定し，個人の併買行列からアーティストの属性を算出することで，顧客の嗜好を反映した数値化をすることができる．

こうして得られた併買行列を集約して属性を算出する．データの集約には誤差を伴うが，併買行列をそのまま属性として利用することは困難なので，データの集約を行う．ここでは，集約手法として最も一般的に用いられている因子分析（最尤法，バリマックス回転）を採用した[8]．今回は，相関係数行列の固有値が2を超えた11次元の因子を抜き出した．

因子分析などの縮約手法は，前述のように，マッピングのために使われることが多く[6][2]，選択肢の相対的な座標位置からブランド・製品属性の考察がなされている．本研究では，11次元の因子を抽出するため，マッピングを目的とはしないが，因子分析などで

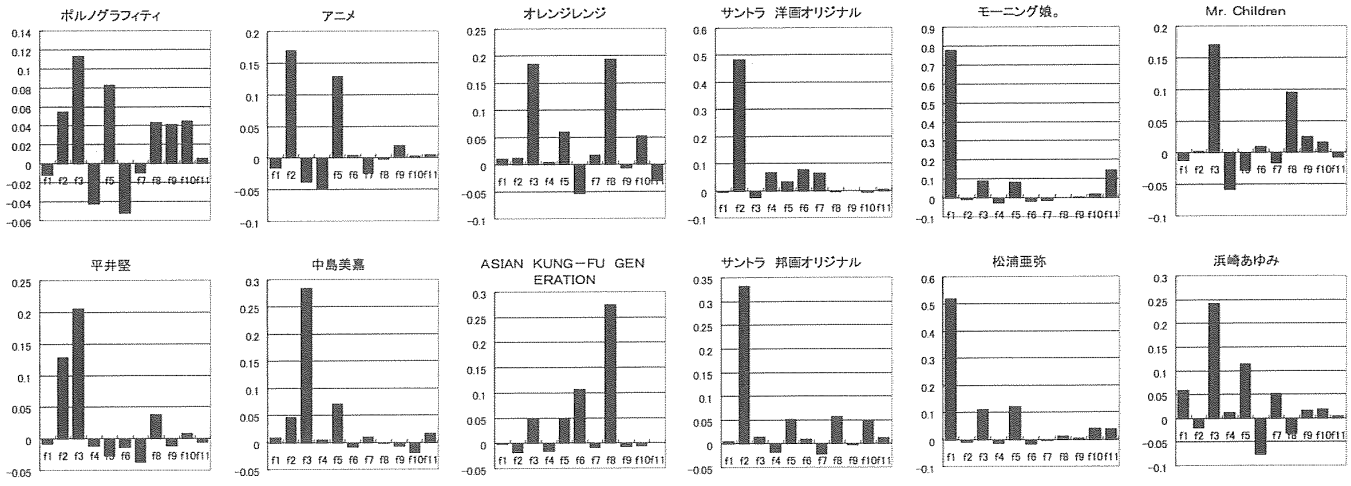

図1 代表的なアーティストの属性グラフ

縮約した数値を，ブランド・製品属性とする基本的な考え方は同様である．

アーティスト k の因子負荷を，f_k と表記する．代表的なアーティストの属性は，図1のようになっている．

3.4 階層ベイズ（HB）モデルの構築

個人 n の第 t 回めの購買において，アーティスト k のCDが選ばれたとき，$x_{nt}=f_k$ とおく．ここで，x_{nt} は，k 次元の連続数で表される個人の潜在的嗜好 y_n が誤差項 ε_{nt} を伴って顕在したものであると考えると，

$$x_{nt}=y_n+\varepsilon_{nt},\quad \varepsilon_{nt}\sim\mathcal{N}(0,\Gamma_n) \tag{1}$$

とおくことができる．また，分散・共分散項 Γ_n は人によって異なると仮定する．幅広い興味をもつ人もいれば，狭い嗜好をもつ人もいるという仮定は，現実のケースを考えても違和感はない．

さらに，個人の購買回数が少ないので，y_n に対して，デモグラフィック変数による情報を補完する．デモグラフィック変数 r_n とそのパラメータ Q で y_n を説明する式は，以下のようになる．

$$y_n=Qr_n+u_n,\quad u_n\sim\mathcal{N}(0,V) \tag{2}$$

r_n は，顧客 n の，学習期間中（1年間）の購買回数（T_n），年齢，性別（男=0，女=1）[1] を用いた．

$$r_n=\begin{pmatrix}1\\ \log(T_n)\\ \log\ (\text{年齢}_n)\\ \text{性別}_n\end{pmatrix} \tag{3}$$

これらの2式をまとめて，モデルを以下のように記述することができる．

$$\begin{cases}\text{第1層：}x_{nt}=y_n+\varepsilon_{nt},\quad \varepsilon_{nt}\sim\mathcal{N}(0,\Gamma_n)\\ \text{第2層：}y_n=Qr_n+u_n,\quad u_n\sim\mathcal{N}(0,V)\end{cases} \tag{4}$$

顧客個人の嗜好 y_n は，顧客の購買行動に，デモグラフィックからの集計的な傾向を補完した値となる．

この2階層のモデルのパラメータ推定は，MCMC法を使って行う．事前分布，事後分布，ハイパーパラメータについては，付録に記載する．

3.5 サンプリング方法

今回，メモリサイズの問題で，分析対象全員をまとめて分析することができなかった．したがって，分割した残りの集合B，27,559人から5,000人をランダムに抽出し（$N=5,000$），モデルのパラメータを推定した．ただし，抽出から外れた顧客に対しても，追加分析をかけることができる．詳細は3.6節で述べる．

[1] 性別が欠損の個人は，分析対象から除外している．割合としては，全体の0.1%未満だった．

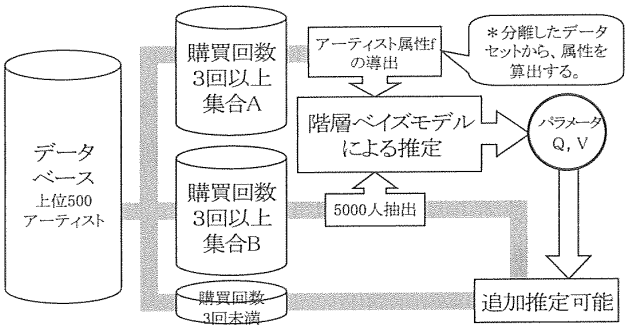

図2　分析のフロー

サンプリングの手順は次のとおりである．

【1】　$Q|V,\{y_n\}$

【2】　$V|Q,\{y_n\}$

【3.1】　$y_n|Q,V,\Gamma_n$

【3.2】　$\Gamma_n|y_n$

【1】-【2】は第2層のサンプリングである．【3.1】-【3.2】は，各個人ごとにサンプリングを行う．事後分布の詳細は付録にて記述する．

MCMC法によるサンプリングは，11,000回行った．そのうち，はじめの1,000回を初期値の影響がなくなるまでの稼動検査期間（Burn-in Period）として捨て，後の10,000回をサンプルとしてとった．

ここで，これまでの分析のフローを図2に示す．

3.6 比較モデル

階層ベイズ（HB）モデルの精度を検証するために，次の2モデルを並行して計算し，予測精度を比較する．

単純予測：単純予測は，「前半1年めにアーティスト k のCDをたくさん買った人ほど，次の年にアーティスト k のCDを買う可能性が高い」という予測である．ただし，同一購買枚数の顧客や，購買がなかった顧客を順位付けすることはできない．

集計モデル：顧客のデモグラフィック属性のみから，嗜好の傾向を推定するモデルである．購買記録を個人ごとに分けず，まとめたデータでモデルを構築する．ここで，T は全員の購買回数の総和なので，$T=\sum_{n=1}^{N}T_n$ である．また，5,000人でモデルを構築しているので，$N=5,000$ である．

$$x_j=Br_j+u_j,\quad j=1,\cdots,T \tag{5}$$

第 j 回めの購買においてアーティスト k が購買されたとき，$x_j=f_k$ となる．それを購買した顧客は r_j というデモグラフィック変数をもつ，という状態をあらわしている．デモグラフィック変数の回帰係数 B は，最小二乗法で求め，$\hat{B}$ とおく．集計モデルから推定した顧客 n の嗜好 y_n^{Agg} は，この回帰係数と，顧

客 n のデモグラフィック変数 r_n から，次の式で求める．

$$Y_n^{Agg}=\hat{B}r_n \tag{6}$$

したがって，集計モデルでは，デモグラフィック変数（年齢，性別，購買枚数）がすべて同じ顧客は，同じ嗜好をもっていると判別される．

4. 推定結果

4.1 予測精度の比較

4.1.1 個人嗜好の予測

HB モデルと，比較モデルとの予測精度を検証する．ここでは，検証期間として推定から外した後半1年間でアーティスト「ASIAN KUNG-FU GENERATION」の CD を購買する傾向の高い顧客を予測する．

HB モデルの推定値は，サンプリングした y_n の平均を用いる．

$$\bar{y}_n=\frac{1}{M}\sum_{m=1}^{M} y_n^{(m)} \tag{7}$$

ここで，M はサンプリングの回数，今回は $M=10{,}000$ となる．また，$y_n^{(m)}$ 右上部にある括弧内の数値 m は，m 回目のサンプリング時に得た値という意味である．

こうして算出した $\bar{y}_n$ は，アーティスト属性と同じ次元のベクトルである．この値がアーティストの属性に近いほど，そのアーティストの CD を購買する傾向が高いといえる．ここで，「ASIAN KUNG-FU GENERATION」の属性との二乗差を算出する．HB モデルによる顧客 n のアーティスト k との二乗差 $\hat{e}_{nk}$ は，次の式から求めることができる．

$$\hat{e}_{nk}=(f_k-\bar{y}_n)'(f_k-\bar{y}_n) \tag{8}$$

比較モデルとして計算した「集計モデル」は，最小二乗推定量をパラメータにとり，そのパラメータを代入した嗜好の推定値を使って，同様に二乗差を算出する．

算出した二乗差を用いて，各モデルごとに，横軸にその値が小さい順に顧客を並べ，縦軸に2年目の購買者の累積割合をとって，累積ゲイン図[5]を描いたところ，図3のようになった[2]．センターラインを越えれば予測能力が一定程度ある，と考えると，どのモデルも予測能力はあるといえる．「単純予測」は，1年目に「ASIAN KUNG-FU GENERATION」を購買した顧客の並ぶ左方で予測精度が高くなっている．これは，「1年目に ASIAN KUNG-FU GENERATION を購買した顧客が2年目にも ASIAN KUNG-FU GENERATION を購買する」という単純予測が，かなり高い精度で当たる，ということを示している．

しかしながら，1年目に購買していない大多数の顧客については，ランダムに並んでいるだけなので，ほとんど予測能力がない．

4.1.2 ゲイン面積による精度比較

同様にして，他のアーティストに対しても，予測精度を計算する．モデルごとに，累積ゲイン図を描いたときにセンターラインを上回った面積 S を算出する．これは，予測精度が悪いときには，負の値も取り得る．

全アーティストについてゲイン面積 S を算出した．このゲイン面積 S は，相対的なモデルの当てはまりのよさを示す指標であるため，順位による比較を行う．アーティストごとに，4つのモデルのゲイン面積 S を比較し，順位付けを行う．ただし，値がマイナスになったものは，「予測能力なし」とみなし，順位をつけていない．また，抽出した標本の中で，2年目に購買が記録されなかったアーティストは，検証をすることができないので，除外している．

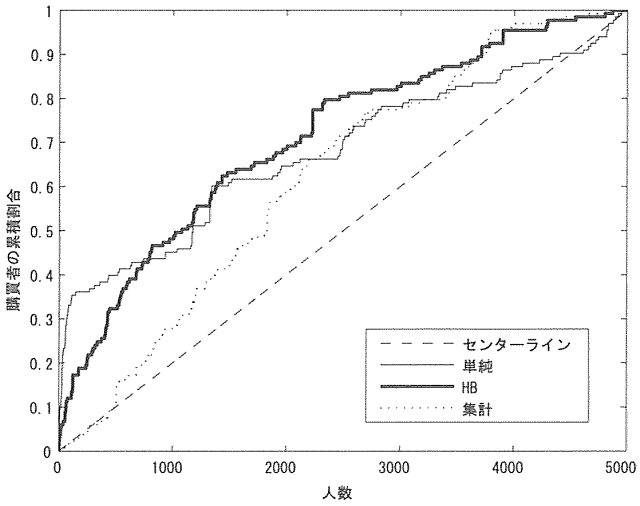

図3 ASIAN KUNG-FU GENERATION の累積ゲイン図

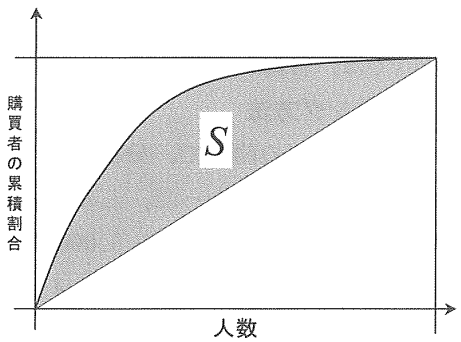

図4 ゲイン面積 S

[2] 単純予測では，同一購買枚数の顧客の順番はランダムに並んでいる．

表1は，アーティストごとに予測精度の順位を算出し，足し合わせたものである．HBモデルによる予測精度が最も高かったアーティストが230人（組），単純予測による予測精度が最も高かったものが159人（組）であった．ゲイン面積Sが負の値をとり，予測能力がない，と判定された数は，HBモデルが83人（組）と，最も少なかった．全体的に見ると，HBモデルが最も予測精度が高く，単純予測がそれに続く，という結果になった．デモグラフィック変数でしか個人差を判定できない集計モデルは，音楽CDの多様な嗜好を考慮していないため，精度が上がらなかった．

表2は，HBモデルと単純予測の二者間を比較したものである．ここでもHBモデルの方が予測能力が高いという結果となった．

表3は，1年めに対象のアーティストを購買しなかった顧客だけを抽出して，比較をしたものである．単純予測が大きく精度を落とす反面，HBモデルは高い予測精度を保持していることがわかる．このように，「HBモデル」は，特に，「単純予測」では予測不可能な，1年目に対象アーティストのCDを購買しなかった見込み顧客に対しても，一定水準以上の予測精度を維持することができるという結果となった．

4.2 推定結果の活用

4.2.1 第2層の活用

HBモデルの第2層，パラメータQは，因子（被説明変数）の次元×デモグラフィック説明変数の数，という大きさをもつ行列である．このパラメータの第jd成分は，第j因子に第d説明変数が与える影響の強さを表している．

表4は，パラメータQのサンプル平均値$\bar{Q}$の値である．これを見ると，どの因子にどの変数が影響しているのかが分かる．有意性[3]は，「＊＊」が99%，「＊」が95%，「・」が90%を示す．Qを見ることで，集計レベルの大まかな傾向を把握することができ，見込み顧客に対しても何らかのアクションを起こすことができる．

第2層で求めたパラメータQ, Vは，見込み顧客だけでなく，抽出対象から外れた顧客や，1枚目のCDをはじめて買った新規顧客の嗜好を把握するために活用することができる．5,000人を抽出して計算したQ, Vの，例えば平均値$\bar{Q}, \bar{V}$などを，パラメータとしてではなく，所与のハイパーパラメータとして事前分布に組み込み，特定の顧客nに対して，第1層のみの単層ベイズモデルを計算することで，嗜好を計算できる．このような柔軟性もベイズモデルの有用性の一つといえる．サンプリングの手順および事後分布は，階層モデルの一部と同様である[4]．

【1】 $y_n|\bar{Q}, \bar{V}, \Gamma_n$

【2】 $\Gamma_n|y_n$

まず，【1】で，顧客nの潜在的嗜好y_nを，$\bar{Q}, \bar{V}$をハイパーパラメータに加えてサンプリングする．そして，【2】で，顧客nの分散・共分散項Γ_nをサンプリングする．この2ステップだけを繰り返せば良い．

表1　予測精度の順位

	HB	単純	集計
1位	230	159	29
2位	125	101	129
3位	17	82	127
予測能力なし	83	113	170

表2　HBと単純予測の比較

	HB	単純
1位	252	163
2位	120	179
予測能力なし	83	113

表3　HBと単純予測の比較（見込み顧客）

	HB	単純
1位	174	41
2位	21	45
予測能力なし	260	369

表4　$\bar{Q}$の値（数値は見やすいように10^3をかけている）

	切片	購買枚数	年齢	性別
f1	10.32	7.91・	-2.81	-12.36・
f2	-291.59**	-3.97	118.56**	9.68
f3	63.00	3.25	5.00	10.14
f4	-10.96	2.36	12.68	7.07
f5	74.95・	6.32	-17.93	-12.71・
f6	-45.01	-3.38	25.90・	-6.92
f7	55.10	0.97	-7.72	-1.85
f8	157.68**	-1.86	-41.11*	-7.16
f9	41.30	-1.84	-8.58	-4.23
f10	53.95	2.80	-19.32	7.36
f11	0.83	1.74	-0.57	-1.52

[3] ここでの有意性は，標準誤差を算出して求めたものではなく，サンプルの何%が片方の符号に偏ったかを示している．すなわち，10,000個のサンプルのうち，9,900個以上が正の値のときは，「99%有意で正」としている．

[4] 「集計モデル」との差異について補足する．「集計モデル」は，個人嗜好をデモグラフィック変数とその係数のみから推定するが，このモデルでは，それらの情報だけではなく，個人の購買履歴も考慮したうえで個人嗜好の推定を行っている．

4.2.2 新人アーティストプロモーションのターゲティング

新人アーティストの売り込みに際して，あらかじめターゲット顧客が想定されている場合，ターゲット顧客の特定にも，アーティスト属性と顧客属性を使うことができる．例えば，「浜崎あゆみ」と「中島美嘉」の中間をターゲットとするときは，

$$f_{\text{新人}}=\frac{1}{2}(f_{\text{浜崎あゆみ}}+f_{\text{中島美嘉}}) \tag{9}$$

とおいて，$f_{\text{新人}}$ と属性の近い顧客からプロモーションをかけていくことができる．また，上記の式以外でも，中間属性として柔軟に値を与えることもできる．ただし，あまり多くのアーティストの中間ばかりを取ると，属性が平均化され，結果として特徴のない層へプロモーションを掛けてしまう可能性もあるため，注意が必要である．

4.2.3 レコメンデーションシステムとしての活用

上位500のアーティストが売り上げの約80%を占めているという音楽CD業界だが，顧客がすべてのアーティストを熟知しているとは考えにくく，その顧客の嗜好に合っているアーティストだとしても，本人が知らないというケースは少なからずあるだろう．そのような埋もれた需要を掘り起こすために，レコメンデーションは非常に効果的である．顧客の嗜好を把握することができれば，そこから個人ごとにアーティストを順位付けすることができ，レコメンデーションシステムに組み込むこともできる．本モデルでは，「Aを買った人にはBを推薦する」という1商品に対応するレコメンデーションではなく，顧客の過去1年間の購買記録と，性別や年齢などのデモグラフィック変数を組み込んだ，より精緻な推定に基づいたレコメンデーションが可能になる[5]．

5. おわりに

デモグラフィックによる集計的な傾向を考慮しつつ，顧客の異質性を仮定するモデルは，従来の枠組みではモデル化することは困難であった．また，大量のレコード数を保持するPOSデータも，個人単位の分析となると，途端に標本数と情報の少なさに悩まされていた．階層ベイズモデルは，「個人ごとの嗜好の把握」，「集計傾向による情報の補完」の双方を取り入れた画期的なモデルといえる．

また，単純予測では予測できない1年目に未購買であった見込み顧客に対しても，高い精度を上げることができるという特性は，非常に有用である．単純予測によるプロモーションだけでは，大多数の未購買顧客に対してアクションをおこす手段がない．階層ベイズモデルなら，未購買の見込み顧客，新規顧客に対しても嗜好や潜在的需要の推定ができる．

本論文では集約したデータを被説明変数に用いて推定を行ったが，結果が示すように，高い予測精度を上げることができた．計算上，集約しないデータを用いてモデルを構築することがかなり困難であることを考えると，集約データを用いた予測は実用的であるといえる．ただし，モデルの精緻化のため，推定法や回転法を変えた因子分析による精度比較や，その他の集約手法の比較や検討が，今後の課題である．さらなる課題として，より精度の高いモデルを考えるのであれば，集約しないデータを用いたモデルの構築が望ましいといえる．

他にも，モデルの拡張として，属性や嗜好の時系列変化を考えることができる．今回の分析では，2年間に，アーティストの属性や個人の嗜好は変化しないという仮定を置いているが，当然個人の嗜好やアーティストの音楽性などは変化しうるもので，より長期間の分析を考えるなら，これらの数値が動的に変化していくという仮定を置くことが望ましい．

謝辞　本論文の執筆にあたり，阿部誠先生（東京大学）より，多くの助言をいただきました．深く感謝を申し上げます．また，データを提供していただいたデータ解析コンペティション事務局および企業の方に，併せて感謝を申し上げます．

A. 付録　事後分布および確率乱数の発生

A.1 事前分布，初期値，事後分布

事後分布の導出に関する詳細な記述は，文献[3][12][14]を参照．以下，N は分析人数，T_n は個人 n の購買回数．D は説明変数 r の数，J は y_n の次元．また，I_K は $K\times K$ の単位行列，$O_{M\times N}$ は $M\times N$ の零行列．

尤度関数

$$f(x_n|y_n,\Gamma_n)\propto\prod_{t=1}^{T_n}|\Gamma|^{-1/2}\times\exp-\frac{1}{2}(x_{nt}-y_n)'\Gamma^{-1}(x_{nt}-y_n),\quad n=1,\cdots,N$$

$$f(y|Q,V)\propto\prod_{n=1}^{N}|V|^{-1/2}$$

[5] MCMC法をレコメンデーションシステムに応用した例として文献[4]がある．

$$\times \exp -\frac{1}{2}(y_n - Qr_n)' V^{-1}(y_n - Qr_n)$$

事前分布

$\Gamma_n^{-1} \sim \mathcal{W}(g_0, G_0)$, $n=1, \cdots, N$

$Q \sim \mathcal{N}_{J\times D}(Q_0, V, \Delta_0)$

$V^{-1} \sim \mathcal{W}(s_0, S_0)$

事前分布のハイパーパラメータは，以下の値をおく．

$Q_0 = O_{J\times D}$, $\Delta_0 = I_D$

$g_0 = J$, $G_0 = 100 I_J$

$s_0 = J$, $S_0 = 100 I_J$

初期値

$Q^{(0)} = O_{J\times D}$

$V^{(0)} = I_J$

$y_n^{(0)} = \mathbf{0}$, $n=1, \cdots, N$

$\Gamma_n^{(0)} = I_J$, $n=1, \cdots, N$

事後分布

$Q|V, Y \sim \mathcal{N}_{J\times D}(Q_1, V, \Lambda_1)$

$\Lambda_1 = (R'R + \Delta_0^{-1})^{-1}$

$Q_1 = (Y'R + Q_0\Delta_0^{-1})\Lambda_1$

ここで，$Y = \begin{pmatrix} y_1' \\ \vdots \\ Y_N' \end{pmatrix}$, $R = \begin{pmatrix} r_1' \\ \vdots \\ r_N' \end{pmatrix}$.

$V^{-1}|Q, \{y_n\} \sim \mathcal{W}(g_1, G_1)$

$g_1 = g_0 + N$

$G_1^{-1} = G_0^{-1} + \sum_{i=1}^{N}(y_i - Qr_i)(y_i - Qr_i)'$
$+ (Q - Q_0)(Q - Q_0)'$

$y_n|Q, V, \Gamma_i \sim \mathcal{N}_J(y_{n1}, V_{n1})$

$V_{n1} = (V^{-1} + T_n\Gamma_n^{-1})^{-1}$

$y_{n1} = V_{n1}(V^{-1}Qr_n + \sum_{t=1}^{T_n}\Gamma_n^{-1}x_{nt})$

$\Gamma_n^{-1}|y_n \sim \mathcal{W}(s_{n1}, S_{n1})$

$s_{n1} = s_0 + T_n$

$S_{n1}^{-1} = S_0^{-1} + \sum_{t=1}^{T_n}(x_{nt} - y_n)(x_{nt} - y_n)'$

A.2 乱数の発生

行列正規分布（Matrix Distribution, Matrix Normal Distribution）およびウィシャート分布からの乱数の発生方法は以下のとおりである．詳細については文献[7][12][13]．

行列正規分布：$\mathcal{N}_{M\times N}(M, A, B)$

(1) $z_{ij} \sim \mathcal{N}(0,1)$ を要素とするサイズ $M\times N$ 行列 Z を発生させる．

(2) $L_A'L_A = A$, $L_B'L_B = B$ となる下三角行列 L_A, L_B を得たとき，$X = L_AZL_B' + M \sim \mathcal{N}_{M\times N}(M, A, B)$ となる．

ウィシャート分布：$\mathcal{W}(\nu, V)$

(1) 対角項に $T_{ii}^2 \sim \chi^2(\nu)$ となる T_{ii} をとり，非対角項に $T_{ij} \sim \mathcal{N}(0,1)$（ただし $i>j$）をとる下三角行列 T を発生させる．

(2) $TT' \sim \mathcal{W}(\nu, I)$ となる．

(3) $V = U'U$ となる U を導出する．

(4) $X = U'TT'U \sim \mathcal{W}(\nu, V)$ となる．

参考文献

[1] 阿部誠：“消費者行動のモデル化：消費者の異質性，”オペレーションズ・リサーチ, 48 (2003), 121-129.

[2] 阿部誠：“プロダクト空間とブランド空間を考慮したジョイント・スペース・マップ—北米ピックアップ・トラック市場への応用—，”オペレーションズ・リサーチ, 48 (2003), 719-728.

[3] 阿部誠, 近藤文代：「マーケティングの科学—POSデータの解析—」, 朝倉書店 (2005).

[4] A. Ansari, S. Essegair and R. Kohli: “Internet Recommendation Systems,” *Journal of Marketing Research*, 37 (2000), 363-375.

[5] M. J. A. Berry, G. S. Linoff (著), 江原淳, 金子武久, 斎藤史朗, 佐藤栄作, 清水聰, 寺田英治, 守口剛 (共訳)：「マスタリング・データマイング—CRMのアートとサイエンス—理論編」, 海文堂 (2002).

[6] 古川一郎, 守口剛, 阿部誠：「マーケティング・サイエンス入門—市場対応の科学的マネジメント—」, 有斐閣 (2003).

[7] D. Gamerman: *Markov Chain Monte Carlo : Stochastic Simulation for Bayesian Inference*, Chapman & Hall (1997).

[8] H. H. Harman: *Modern Factor Analysis—Third Edition Revised—*, The University of Chicago Press: Chicago (1976).

[9] R. E. McCulloch and P. E. Rossi: “An exact likelihood analysis of the multinomial probit model,” *Journal of Econometrics*, 64 (1994), 207-240.

[10] R. E. McCulloch, N. G. Polson and P. E. Rossi: “A Bayesian analysis of the multinomial probit model with fully identified parameters,” *Journal of Econometrics*, 99 (2000), 173-193.

[11] P. E. Rossi and G. M. Allenby: “Bayesian Statistics and Marketing,” *Marketing Science*, 22 (2003), 304-328.

[12] P. E. Rossi, G. M. Allenby and R. McCulloch: *Bayesian Statistics and Marketing*, Wiley (2005).

[13] D. B. Rowe: *Multivariate Bayesian Statistics*, Chapman & Hall (2002).

[14] 和合肇（編著）：「ベイズ計量経済分析」, 東洋経済新報社 (2005).