

Estimating Individual Preferences in the Music CD Market

Sotaro KATSUMATA

Graduate School of Economics, the University of Tokyo

[E-mail: ee077021@mail.ecc.u-tokyo.ac.jp](mailto:ee077021@mail.ecc.u-tokyo.ac.jp)

Abstract: This research proposes a model to analyze individual customer preferences using purchase records such as Point-of-Sales (POS) data. To some extent, we can identify the interests of customers from their demographics. Consumers are, however, essentially heterogeneous. It is difficult to determine individual customer behavior in detail through aggregate-level estimation. In this paper, we use a Markov chain Monte Carlo (MCMC) method to construct a hierarchical model for tackling this problem. The model encompasses both “commonality” and “heterogeneity.” We apply this MCMC method to the music CD market, where customers have some commonalities although they are heterogeneous. This empirical analysis shows that a hierarchical Bayes (HB) model has a high predictive performance as compared to the naïve forecasting and aggregate-level models.

Keywords: customer relationship management, customer heterogeneity, hierarchical Bayes model

1. Introduction

In classical economics, the individual heterogeneity of preferences is treated as an “error term” or a “disturbance.” In marketing, on the other hand, it is assumed that consumers are heterogeneous. The recent developments in information technology have facilitated the collection of Point-of-Sales (POS)

data along with customer profiles, and firms can now record the individual time-series purchase data of their customers. At the same time, this has led to a need for new methods to analyze the vast amounts of POS data. Although firms collect massive amounts of data, ironically, a problem they often face is that the individual figures of purchases are not sufficient for analysis. For example, an ordinary household

does not purchase detergents numerous times in a year. When we construct a model using classical regression, we need to aggregate the data because of the scarcity of the number of individual samples. However, when using this method, the customer identification information of the data gets ignored.

In cases where it is not possible to estimate the individual parameters of the model, a Bayesian estimation through the use of a Markov chain Monte Carlo (MCMC) method can be utilized to obtain parameters. In an MCMC method, we suppose that all parameters are distributed probability distributions and derive estimates from the random numbers generated from these distributions. Using a MCMC method, we can easily expand models and construct complex models that are difficult to estimate through the most likelihood method. There are many existing application researches in the field of marketing (e.g., Abe, in press; Rossi and Allenby, 2003).

In this paper, we analyze a vast amount of data using an MCMC method and investigate the process for the application of hierarchical Bayes (HB) models for database analysis.

2. Data and modeling procedure

2.1. Data overview

We use POS data with customer identification information in a music CD retail store provided by the 2005 Data Analysis Competition hosted by Joint Association Study Group of Management Science.

These records are collected over a period of two years. Additionally, we obtain customer demographics such as age and gender.

2.2. Modeling procedure

We need to consider the following three factors when analyzing the music CD retail selling database. First, we must consider the complexity of customer preferences as customers never purchase music CDs that do not interest them. Although we can estimate their preferences on the basis of age and gender to some extent, there is significant variability even within the same demographic group. Hence, it is difficult to construct a practicable model without considering customer heterogeneity. Second, we need to consider the number of purchases by a customer; the number of customers who purchase music CDs over twenty times in a year is remarkably low. Therefore, it is difficult to accurately estimate choice behavior using custom-made (individual) models. Third, the number of products needs to be considered. There are several hundreds to thousands of artists in the music CD market. Hence, it is difficult to apply ordinary discrete choice models (e.g., logit and probit models).

In this paper, we consider the three issues mentioned above. At first, we adopt a hierarchical model using an MCMC method to tackle the first and second problems. The hierarchical structure successfully solves the abovementioned problems as it can add the aggregate-level tendency of demographic variables for an individual model. It is

Estimating individual preferences in the music CD market

especially effective for analyzing markets like the music CD market, which has few samples and individual variations. However, in previous researches involving hierarchical modeling, models using discrete choice methods have been used. These models cannot be applied to this case without modification. Therefore, we adopt a dimension reduction method that is used for perceptual mapping (e.g., Churchill, 1999) and a joint space map (e.g., DeSarbo and Wu, 2001; DeSarbo, Ramaswamy, Wedel, and Bijmolt, 1996). Using this method, we assign each artist multidimensional continuous values of attribution in order to deal with the third problem.

3. Model construction

Purchase records are collected over a period of two years. We use the first half of the data collected (the first year) as a learning period to estimate the parameters of models, and the latter half of the data (the second year) as a calibration period to validate the predictive performance of the model. In other words, we predict the purchases of the coming year using the history of the preceding one.

3.1. Objective customers and artists

The total number of purchases recorded in the first year is 605,593 and the number of artists whose music is purchased in this period is 8,545. However, the top 500 artists account for 81% of the total sales. If the sales figures are too low, we cannot assure

good accuracy of artist score estimation. Therefore, we use the sales of the top 500 artists for estimation.

In the first year, 161,805 customers purchased the music CDs of these top 500 artists. However, most of the customers make purchases only once or twice (74,276 customers buy once and 32,410 customers buy twice). In this paper, we only consider the 55,119 customers who purchased CDs over three times in our analysis. In addition, later in this paper, we will discuss the follow-up estimation for the excluded samples.

3.2. Data division

We divide the above-mentioned 55,119 customers into two datasets, namely, A and B. We derive artist attributes using dataset A and construct models to estimate individual preferences using dataset B. This procedure aims to avoid a loop of analysis caused by the derivation of artist attributes for dependent and explanatory variables from the same data source. We separate datasets A and B completely and treat artist attributes as an exogenous variable.

3.3. Derivation of artist attributes

In this section, we describe the process of the derivation of artist attributes. From dataset A, attributes characterizing artists are extracted and each artist is rated according to these attributes.

At first, we prepare a matrix of the size $27,559 \times 500$. The number of artist k 's CDs purchased by customer j is contained in the (j, k) element of the matrix. The rows of this matrix signify customers'

co-purchases. By assuming that the artists whose music CDs are purchased by a particular customer have similar attributes, we can obtain artist attributes reflecting customer preferences from this co-purchase matrix.

Artist attributes are obtained from this co-purchase matrix by reducing dimensions. Although reducing data gives rise to some errors, it is difficult to use a co-purchase matrix to obtain attributes without modification. We use factor analysis (using the most likelihood and varimax rotation methods (refer Harman, 1976)) and obtain 11 factors whose eigenvalues exceed 2.

In general, dimension reducing methods such as factor analysis are used for mapping in order to compare the relative position of each brand or product. Although we do not use mapping in this paper, we adopt a similar technique—namely, factor analysis—to treat reduced variables as brand and product value attributes.

We denote the factor loadings of artist k as f_k .

Figure 1 shows attributes of representative artists.

3.4. Construction of the hierarchical Bayes (HB) Model

When an individual n chooses an artist k 's CD as his/her t -th purchase, let $x_{nt} = f_k$, where x_{nt} is a j -dimensional continuous variable. Assuming that x_{nt} is a revealed variable of his/her latent preference y_n with error term ε_{nt} , we get,

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

We assume that there is individual heterogeneity of variance. It is appropriate to make this assumption since some people may depict variety-seeking behavior, while others may show loyalty to a particular artist(s).

In addition, since individual purchase samples are scarce, we supplement our information using the demographic variables r_n and their parameters Q . We can describe this relation as follows:

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

r_n contains (1) number of purchases in learning period (T_n), (2) age, and (3) gender (male = 0, female = 1). We exclude individuals whose gender information is missing. The percentage of exclusion is less than 0.1%.

$$r_n = \begin{pmatrix} 1 \\ \log(T_n) \\ \text{Age}_n \\ \text{Gender}_n \end{pmatrix} \quad (3)$$

We can obtain a model by assembling the above two equations as follows:

$$\begin{cases} x_{nt} = y_n + \varepsilon_{nt}, & \varepsilon_{nt} \sim \mathcal{N}(0, \Sigma_n) \\ y_n = Qr_n + u_n, & u_n \sim \mathcal{N}(0, \Gamma) \end{cases} \quad (4)$$

Customer preference y is a vector variable that supplements individual purchase behavior with aggregate attitude from demographic variables.

We estimate the parameters of this hierarchical model using an MCMC method. For more information on prior distributions, posterior distributions,

Estimating individual preferences in the music CD market

Figure 1. Attributes of representative artists

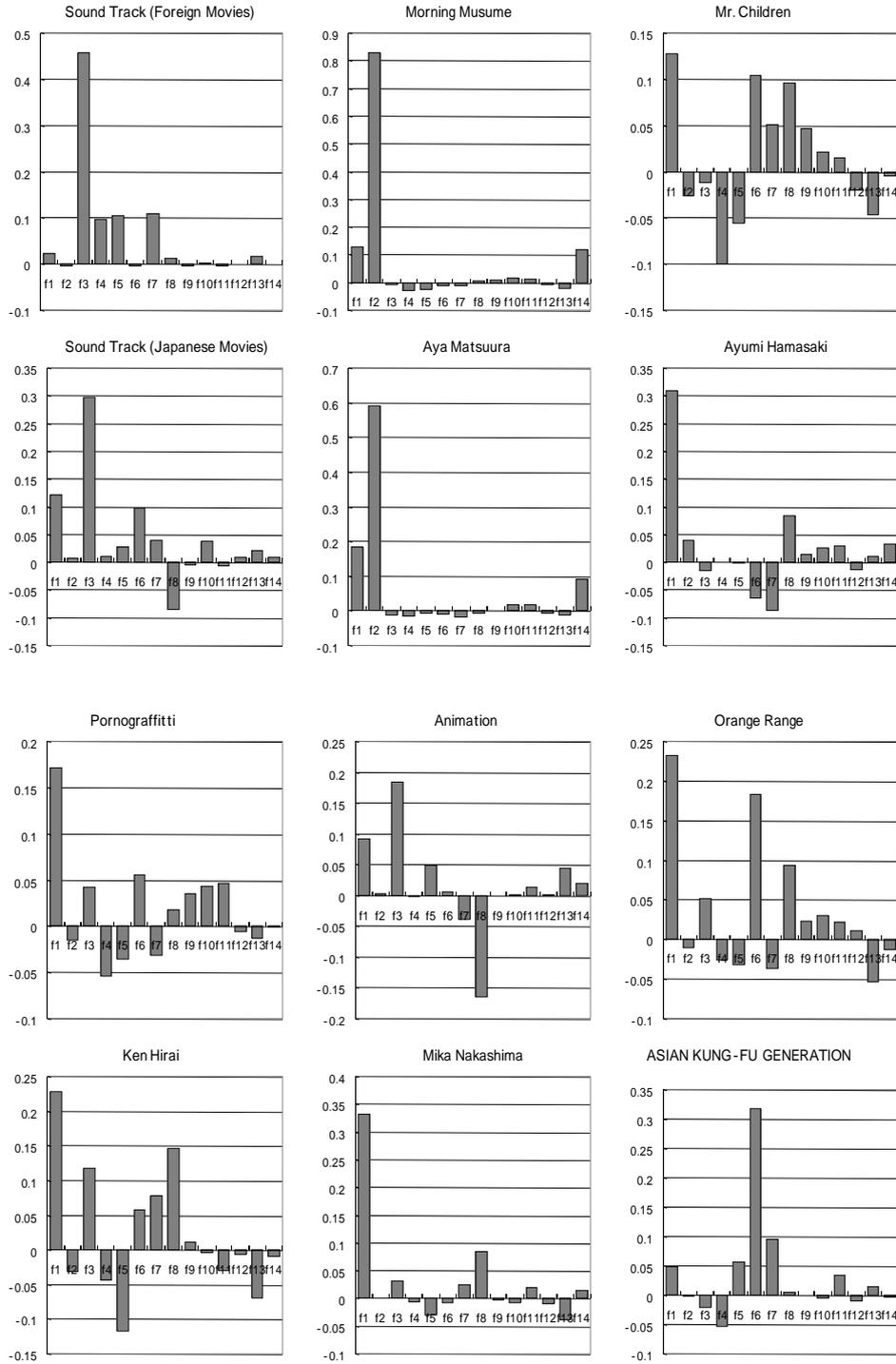
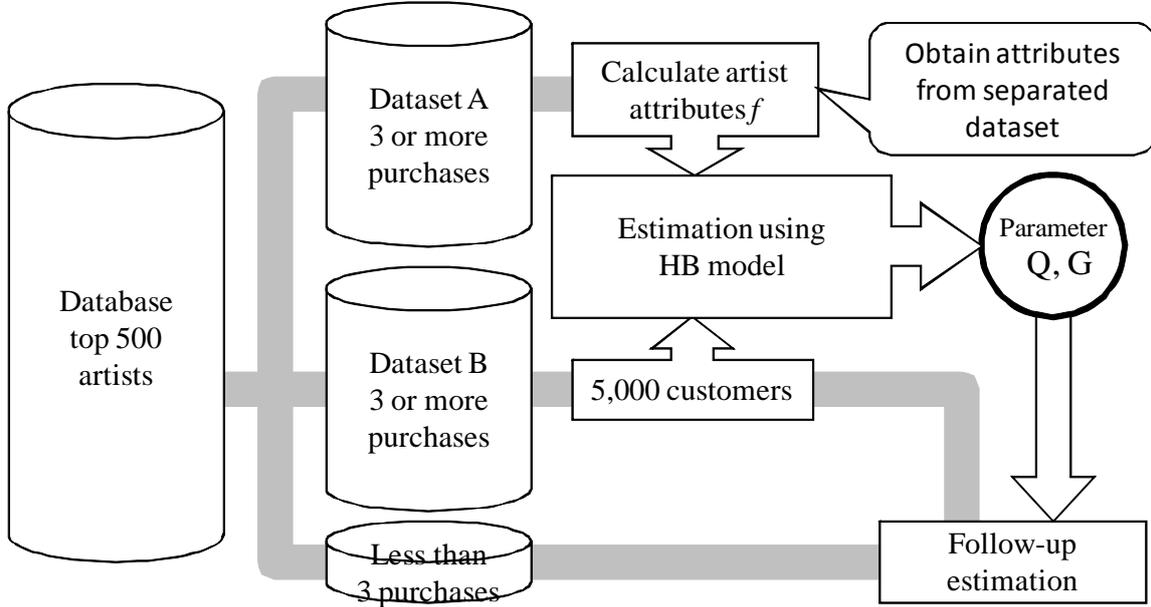


Figure 2. Modeling procedure



and hyper parameters, please refer to the appendix.

3.5. Sampling

We choose 5,000 customers from dataset B, consisting of 27,559 customers ($N = 5,000$), in order to estimate parameters. However, we can subsequently estimate excluded customers by using these parameters. The sampling procedure is as follows:

1. $Q | \Gamma, y_n$
2. $\Gamma | Q, y_n$
- 3.1. $y_n | Q, \Gamma, \Sigma_n$
- 3.2. $\Sigma_n | y_n$

1–2 are samplings of upper layers. 3.1–3.2 are samplings for each individual. The detailed explanations of posterior distributions can be found in the appendix.

In the simulation, we burn-in these procedures for 1,000 iterations, and then save 10,000 samples.

Figure 2 shows the modeling procedures at this point.

3.6. Comparison models

We also estimate the following two models collaterally to compare their forecasting performances with the HB model.

Naïve forecasting model: The artist whom the customer purchased most frequently during the first year will be purchased again by him/her in the second year.

However, this forecasting model cannot rank people who purchase artist k 's CDs the same number of times or those who do not purchase them at all.

Aggregate model: An aggregate model

Estimating individual preferences in the music CD market

estimates y directly from the second layer of the HB model using ordinary least squares (OLS). This model cannot estimate parameters for individuals or classify customers on the basis of their demographic variables (number of purchases, age, and gender). Since the same parameters Q and Γ are common across customers, customers with the same demographic profile have the same preferences.

$$y_j = Qr_j + u_j, \quad j=1, \dots, \sum_{n=1}^N T_n \quad (5)$$

4. Results

4.1. Predictive performance

4.1.1. Predicting individual preferences

In this section, we compare the predictive performances of the HB and comparison models. Initially, we determine the customers who have a high likelihood of purchasing the music CDs of the rock band “ASIAN KUNG-FU GENERATION” in the validation period (the latter year of data collection).

We use the mean of the samples of y_n as the predictive values of the HB model.

$$\bar{y}_n = \frac{1}{M} \sum_{m=1}^M y_n^{(m)}, \quad (6)$$

where M is number of samples (in this paper, $M = 10,000$). Value m in the bracket on the top-right of $y_n^{(m)}$ denotes the value obtained at m -th sampling.

\bar{y}_n is a vector variable that is of the same size as the artist attributes. The greater the approximate

closeness to an artist’s attributes, the higher is the likelihood of purchasing the artist’s CD. Next, we obtain the squared differences between the attributes of ASIAN KUNG-FU GENERATION and individual preferences. We can derive \hat{e}_{nk} —squared differences between customer n and artist k —from the following equation.

$$\hat{e}_{nk} = (f_k - \bar{y}_n)'(f_k - \bar{y}_n) \quad (7)$$

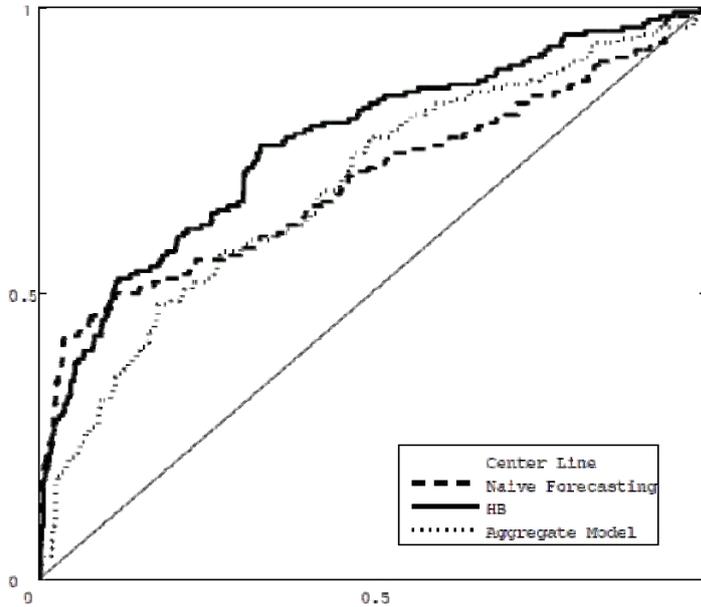
In the aggregate model, substitute the OLS estimator of the equation and calculate the squared differences.

Figure 3 is the cumulative gain chart (Berry and Linoff, 1997, 2000) that arranges customers in the ascending order of their squared differences on the horizontal axis and the cumulative sales rate on the vertical axis. In naïve forecasting, customers who purchase CDs the same number of times are arranged in a random order.

Since all models exceed a center line, these three models possess predictive power to a certain degree. The naïve forecasting model shows good performance on the left side. This means that the customers who purchase the music CDs of “ASIAN KUNG-FU GENERATION” in the first year also purchase its CDs in the next year. However, the naïve forecasting model cannot predict customers who do not purchase the band’s CDs in the first year. Unfortunately, this category of people forms the majority.

Figure 3. Cumulative gain chart of

ASIAN KUNG-FU GENERATION



4.1.2. Predictive performances compared by gain area

We obtain the predictive performance of another artist and calculate the area between the cumulative gain chart and center line. Let this area be S . This indicator may take a negative value if the prediction is worse than random guessing.

We obtain S for all artists. Since this indicator shows the relative predictive performance of a model, we compare these models by rank. We compare the S of the three models and rank these models for each artist. When S takes a negative value, we classify the artist as “N/A (not applicable)” Furthermore, we exclude artists whose CDs are not purchased in the validation period.

Table 1 shows the summations that indicate the ranks of the predictive performances for all the artists. The HB model shows the best performance of 230 artists, while the naïve forecasting model shows the best performance for 159 artists. The HB model shows “N/A” for 83 artists and this figure is the lowest among the three models. On the whole, it can be concluded that the HB model shows the best performance, followed by the naïve forecasting model. Since the aggregate model discriminates individual heterogeneity only on the basis of demographic variables, it cannot describe the diversity of the music CD market.

Figure 4. Area S

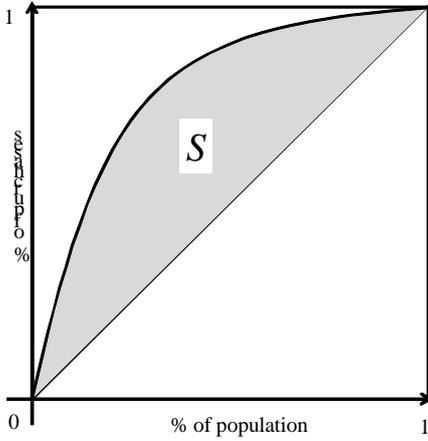


Table 2 shows the comparison of the HB model and the naïve forecasting model. This table also indicates that the performance of the HB model is better than that of the naïve forecasting model.

Table 3 shows the comparison of the predictive performances for artists that were not purchased during the first year (learning period). With the naïve forecasting model, the rule becomes equivalent to random guess. Note that the performance of the naïve forecasting model deteriorates severely, whereas that of the HB model stays the same.

As described above, the HB model can maintain a better quality of forecasting among prospects.

4.2. Application

4.2.1. Application of second layer

Parameter Q of the second layer of the HB model is a matrix whose size is the dimension of factors (dependent variables) \times the number of demographic variables. The (j, d) element of the parameter shows

Table 1. Rank of predictive performance

| | HB | Naïve | Aggregate |
|-----|-----|-------|-----------|
| 1st | 230 | 159 | 29 |
| 2nd | 125 | 101 | 129 |
| 3rd | 17 | 82 | 127 |
| N/A | 83 | 113 | 170 |

Table 2. Comparison of HB Model and naïve forecasting

| | HB | Naïve |
|-----|-----|-------|
| 1st | 252 | 163 |
| 2nd | 120 | 179 |
| N/A | 83 | 113 |

Table 3. Comparison of HB Model and naïve forecasting (prospects)

| | HB | Naïve |
|-----|-----|-------|
| 1st | 174 | 41 |
| 2nd | 21 | 45 |
| N/A | 260 | 269 |

the impact of the d -th explanatory variable on the j -th factor.

Table 4 shows the sample mean of parameter Q . It allows us to observe which variables affect which particular factors. In this table, “***” denotes significance at 99% level. We calculate these indicators from the MCMC samples. Significance at the 99% level implies that 0.99% of the highest posterior density interval (HPD/HPDI) does not cross 0. “*” indicates 95% and “·,” 90%. We can view the aggregate level tendency of customer preference. Further, using this information, we can perform any actions for prospects.

Q and Γ can also be applied to customers

excluded from the estimation and to new customers who purchase a CD for the first time. For example, if we obtain the sample mean $\bar{Q}, \bar{\Gamma}$ of Q and Γ estimated by N customers and use this sample mean as a hyper parameter for the first layer, then we can estimate a single-layer Bayes model for customer i and obtain i 's preferences. The sampling procedure is the same as that of the HB model. The difference in the case of the aggregate model is that while the parameters of the aggregate model are the same among individuals, a single-layer Bayes model estimates parameters individually.

1. $y_n | \bar{Q}, \bar{\Gamma}, \Sigma_n$
2. $\Sigma_n | y_n$

4.2.2. Promotion of new artists and targeting customers

In cases of target segments that are already defined, we can use the information on artist attributes and

customer preferences to identify target customers. For example, we can define target customers who have like ‘‘Ayumi Hamasaki’’ and ‘‘Mika Nakashima’’ in the following manner.

Let the new artist’s attributes, $f_{\{\text{New Artist}\}}$, be

$$f_{\{\text{New Artist}\}} = \frac{1}{2} (f_{\{\text{Mika Nakashima}\}} + f_{\{\text{Ayumi Hamasaki}\}}) \quad (8).$$

Thus, we can obtain a customer set whose preferences approximate the new artist’s attributes and promote the new artist to these customers. However, it should be noted that this procedure may be futile when averaging too many artists.

4.2.3. Applications for recommendation systems

In the music CD market, the top 500 artists account for 80% of the total sales. However, it is unlikely that ordinary customers know all these 500 artists. It is possible that there are some unknown artists that approximate their preference. Recommendations are especially effective in stimulating these latent needs. When firms estimate the preferences of their customers, they can rank artists for each customer and develop preference information into recommendation systems. This system is not of the type that recommends a product whose characteristics correspond to those of another—for example, recommending B to customers who purchased product A. Instead, this model also considers customers’ purchase records in the past year and compliments this information using

Table 4. Parameters \bar{Q}

| | Intercept | Number of Purchases | Age | Gender |
|-----|-------------------|---------------------|-------------------|--------|
| f1 | 10.3 | 8.6 [·] | - 3.3 | - 11.6 |
| f2 | - 291.6** | - 3.9 | 118.5** | 9.6 |
| f3 | 61.3 | 3.1 | 5.7 | 9.6 |
| f4 | - 12.4 | 2.6 | 12.8 | 7.4 |
| f5 | 75.4 [·] | 6.0 | - 17.9 | - 12.4 |
| f6 | - 47.1 | - 3.4 | 26.4 [·] | - 6.1 |
| f7 | 55.0 | 1.0 | - 7.8 | - 1.6 |
| f8 | 164.0** | - 1.2 | - 43.4 | - 7.0 |
| f9 | 42.5 | - 1.7 | - 9.0 | - 4.4 |
| f10 | 53.7 | 2.3 | - 18.9 | 7.1 |
| f11 | 2.7 | 1.6 | - 1.1 | - 1.2 |

Note: For simplify to show, multiple 10^3

Estimating individual preferences in the music CD market

demographic variables such as age and gender. Ansari, Essegai, and Kohli (2000) applied a model using an MCMC method for Internet recommendation systems.

5. Conclusion

It is difficult to develop a model that considers both the aggregate tendency of demographics and the heterogeneity of customers using the classical framework of econometric methods. Furthermore, while POS data records a vast amount of data, many researchers face the problem of scarcity of data when they try to analyze the data on an individual basis. The HB model is a breakthrough model as it deals with both the “heterogeneity” and “commonality” of customers.

The useful feature of the model is that it has predictive power for purchases by prospective customers (customers who do not purchase in the first one year) that cannot be predicted by the naïve forecasting model. There is no effective technique for taking into account a large number of prospects only through the naïve forecasting model. However, the HB model can estimate the latent needs of both prospective and new customers.

When we use condensed attributes for dependent variables, as the result shows, the model has a high predictive performance. Given the fact that it is difficult to use non-condensed data for modeling, this method is more practicable. However, we need to compare this method with other rotation

and estimation methods of factor analysis, and other dimension reducing methods like principal component analysis, fuzzy clustering, and MDS for elaboration. Moreover, we have to determine a method of modeling using non-condensed data.

In addition, we need to expand the model to contain a time series variation of artist attributes and customer preferences. Although the proposal model assumes that these variables are stable over a span of two years, it is possible for them to change over time. Hence, it is desirable to exclude the assumption of stability when analyzing more long term data.

Acknowledgement

The author would like to thank Professor Makoto Abe for advice and feedback, Joint Association Study Group of Management Science for providing the data.

This paper was first published as Katsumata, S. (2007) "Ongaku CD koubai ni okeru kojiri shiko no suitei," Communications of the Operations Research Society of Japan, Vol. 52, pp. 725–731. (In Japanese).

References

Abe, M. (in press). “COUNTING YOUR CUSTOMERS” ONE BY ONE: A Hierarchical Bayes Extension to the Pareto/NBD Model. *Marketing Science*.

- Ansari, A., Essegair, S., & Kohli, R. (2000). Internet recommendation systems. *Journal of Marketing Research*, 37, 363–375.
- Berry, J. A., & Linoff, G. (1997). *Data mining techniques: For marketing, sales, and customer relationship management*. New York: Wiley.
- Berry, J. A., & Linoff, G. (2000). *Mastering data mining: The art and science of customer relationship management*. New York: Wiley.
- Churchill, G. A., Jr. (1999). *Marketing research: Methodological foundations* (7th ed.). Fort Worth, TX, Tokyo: Dryden Press.
- DeSarbo, W. S., & Wu, J. (2001). The joint spatial representation of multiple variable batteries collected in marketing research. *Journal of Marketing Research*, 38, 244–253.
- DeSarbo, W. S., Ramaswamy, V., Wedel, M., & Bijmolt, T. (1996). A spatial interaction model for deriving joint space maps of bundle compositions and market segments from pick-any/J data: An application to new product options. *Marketing Letters*, 7, 131–145.
- Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian inference*. London: Chapman & Hall.
- Harman, H. H. (1976). *Modern factor analysis* (3rd ed., rev.). Chicago: University of Chicago Press.
- Koop, G. (2003). *Bayesian econometrics*. Chichester, UK: Wiley.
- McCulloch, R. E., & Rossi, P. E. (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64, 207–240.
- McCulloch, R. E., Polson, N. G., & Rossi, P. E. (2000). A Bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics*, 99, 173–193.
- Rossi, P. E., & Allenby, G. M. (2003). Bayesian statistics and marketing. *Marketing Science*, 22, 304–328.
- Rossi, P. E., Allenby, G. M., & McCulloch, R. (2005). *Bayesian statistics and marketing*. New York: Wiley.
- Rowe, D. B. (2002). *Multivariate Bayesian statistics*. Boca Raton, FL: Chapman & Hall.

Appendix. Posterior distributions and generation of random numbers

A.1. Prior distributions, likelihood functions, and posterior distributions

Detailed derivations of posterior distributions are found in Rossi and Allenby (2003) and Koop (2003).

Hereafter, N is the number of customers; T_n , the number of purchases of the n -th customer; D , the dimension of the explanatory variable r ; J , the dimension of y_n . Further, I_K is an identity matrix whose size is $K \times K$, and $O_{M \times N}$ is a zero matrix of size $M \times N$.

We stack y_n and r_n and arrange the following matrices Y and R .

$$Y = \begin{pmatrix} y'_1 \\ \vdots \\ y'_N \end{pmatrix}, \quad R = \begin{pmatrix} r'_1 \\ \vdots \\ r'_N \end{pmatrix}$$

Estimating individual preferences in the music CD market

Likelihood function and prior distributions

$$x_{nt} \sim \mathcal{N}_J(y_n, \Sigma_n)$$

$$Q \sim \mathcal{N}_{J \times D}(Q_0, \Gamma, \Lambda_0)$$

$$\Gamma^{-1} \sim \mathcal{W}(g_0, G_0)$$

$$y_n \sim \mathcal{N}_J(Q r_n, \Gamma)$$

$$\Sigma_n^{-1} \sim \mathcal{W}(J, 100I_J)$$

$$Q_0 = O_{J \times D}$$

$$\Lambda_0 = I_D$$

$$g_0 = J$$

$$G_0 = 100 I_J$$

$$s_0 = J$$

$$S_0 = 100 I_J$$

Starting values

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

$$\Gamma^{(0)} = I_J$$

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

$$\Sigma_n^{(0)} = I_J, n = 1, \dots, N$$

Posterior distributions

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

$$\Lambda_1 = (R' \Sigma_n^{-1} R + \Lambda_0)^{-1}$$

$$Q_1 = (Y' \Sigma_n^{-1} R + \Lambda_0^{-1} Q_0) \Lambda_1$$

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

$$g_1 = g_0 + N$$

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

$$y_n | Q, \Gamma, \Sigma_n \sim \mathcal{N}_J(y_{n1}, \Sigma_{n1})$$

$$\Sigma_{n1} = (\Gamma^{-1} + T_n \Sigma_n^{-1})^{-1}$$

$$y_{n1} = \Gamma_{n1} \left(\Gamma^{-1} Q r_n + \sum_{t=1}^{T_n} \Sigma_n^{-1} x_{nt} \right)$$

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

$$s_{n1} = s_0 + T_n$$

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

A.2. Generation of random numbers

We can generate random numbers from matrix normal distribution (matrix distribution) and Wishart distribution by following these procedures, for which Gamerman (1997), Rossi and Allenby (2003) and Rowe (2002) provide detailed explanations.

Matrix normal distribution: $\mathcal{N}_{M \times N}(M, A, B)$

(1) Generate the $M \times N$ matrix Z , where $z_{ij} \sim \mathcal{N}(0, 1)$.

(2) Calculate lower triangle matrices L_A and L_B that satisfy $L_A' L_A = A$, $L_B' L_B = B$, then $X = L_A' Z L_B + M \sim \mathcal{N}_{M \times N}(M, A, B)$.

Wishart distribution: $\mathcal{W}(v, V)$

(1) Generate a lower triangle matrix T , where its diagonal elements $T_{ii} \sim \chi^2(v)$ and non-diagonal elements, $T_{ij} \sim \mathcal{N}(0, 1)$ (where, $i > j$).

(2) $TT' \sim \mathcal{W}(v, I)$.

(3) Decompose V and obtain U , where $V = U'U$.

(4) We get $X = U' T T' U \sim \mathcal{W}(v, V)$.

[Received September 30, 2008; accepted October 7, 2008]

Katsumata

ABAS: ANNALS OF BUSINESS ADMINISTRATIVE SCIENCE

Vol. 7 December 2008

CHIEF EDITOR

Junjiro Shintaku, *the University of Tokyo*

EDITORIAL BOARD

Makoto Abe, *the University of Tokyo*

Takahiro Fujimoto, *the University of Tokyo*

Makoto Kasuya, *the University of Tokyo*

Nobuo Takahashi, *the University of Tokyo*

Technical Editor: Maki Nishida, *Global Business Research Center*