

Calculation and Visualization of Dynamic Price Elasticities

Hirokazu Tajima
Tokyo Keizai University
E-Mail: htajima@tku.ac.jp

ABSTRACT

Applying the Hierarchical Bayesian Regression model to weekly aggregated sales history data from 92 retail stores located around Tokyo, I calculated price elasticities by item, week, and store. These elasticities are more stable than figures calculated using the Hierarchical Regression or Bayesian Regression models. Furthermore, using Google Earth, I visualized the calculated price elasticities of these 92 stores over 67 weeks, providing a better understanding about heterogeneity of price elasticities across time and space.

Keywords: Price Elasticity, Hierarchical Bayesian Regression Model, Markov Chain Monte Carlo (MCMC) method, Google Earth

INTRODUCTION

Currently consumers in Japan are more price-conscious, especially when purchasing commodity goods in supermarkets. In response to this, many retailers cut shelf price to gain business. It is however necessary to optimize the price and/or promotions in addition to cutting prices in order to maintain profits and sustain growth. Because consumer behavior depends on location, competitors, and other factors, store customization is a very important consideration (Levy and Weitz, 2011).

Chintagunta, Dube, and Singh (2008) measured the effects of price discrimination by a supermarket chain in Chicago, and proposed optimal stores pricing. In contrast, Dobson and Waterson (2008) proposed uniform pricing rather than optimal pricing specific to each store.

Price elasticities by store and time provide important, fundamental information for price customization. A popular way to calculate price elasticity by store and week is the estimation of parameters using a regression model, that is, a model with parameters depending on store and week, using daily sales data aggregated by store. Because of limited storage capacity, however, some retailers retain only weekly

historical sales data. However, estimating parameters using a regression model and input by store and week using stores' weekly sales data, could be unstable (Blattberg & George, 1991; Montgomery, 1997).

Therefore, in this study, I applied the MCMC method using Gibbs' sampling to estimate the parameters of the Hierarchical Bayesian Regression model and calculated weekly price elasticities by store. I also dynamically visualized weekly price elasticities by store, using Google Earth.

BAYESIAN UPDATING USING MCMC METHOD

Calculation of Posterior Expectation

Given parameter θ and observed data $y = (y_1, \dots, y_n)$, and using the definition of conditional probability, Bayes' theorem holds: $\Pr(\theta|y) = \Pr(\theta)\Pr(y|\theta)/\Pr(y)$, where $\Pr(\theta)$ is a prior distribution of parameter θ including all a priori information, $\Pr(y|\theta)$ is a likelihood function of the observed data given the parameter θ , and $\Pr(\theta|y)$ is a posterior distribution of y . The expectation of posterior distribution $\Pr(\theta|y)$

$$E[\theta] = \int \theta \Pr(\theta|y) d\theta = \frac{1}{\Pr(\theta)} \int \theta \Pr(\theta) \Pr(y|\theta) d\theta$$

is called a Bayesian estimator. The advantages of Bayesian estimation are the ability to use prior information, parameter updating, and so on.

To estimate (calculate the expectation of) posterior distribution we must solve the integral stated above, which is usually very difficult. If solving the integral is difficult or impossible, then conjugating prior, asymptotic expansion and Monte-Carlo integration are used to estimate the posterior distribution.

Monte-Carlo Method and MCMC Method

If the random sample following the posterior distribution $\{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(N)}\}$ is available, we can approximate the expectation of posterior distribution by the Monte-Carlo integral $\sum_{n=1}^N \theta^{(n)} / N$. By the law of large numbers, the Monte-Carlo integral is a valid approximation of expectation (Tienery & Kadane, 1986). However, when random sampling of the posterior distribution is not available and if the parameter is high-dimensional, random sampling is very inefficient. In this case, using a sample following Markov chains derived from posterior distribution, we calculate the posterior expectation using the Monte-Carlo integral. This method of calculating

posterior expectation is called the MCMC (Markov chain Monte Carlo) method. Note that sampling based on the MCMC method is not random and we cannot apply the law of large numbers. However, because Markov chains used in the MCMC method are ergodic, converging to posterior approximation of posterior expectation by the Monte-Carlo integral is valid (Rossi, Allenby, & McCulloch, 2005). The MCMC method is also applicable if solving the integral of posterior expectation or random sampling from posterior distribution is difficult.

Gibbs sampling and Metropolis-Hasting sampling are popular algorithms of the MCMC method, and we applied Gibbs sampling (Figure 1) in this study (Tajima, 2012).

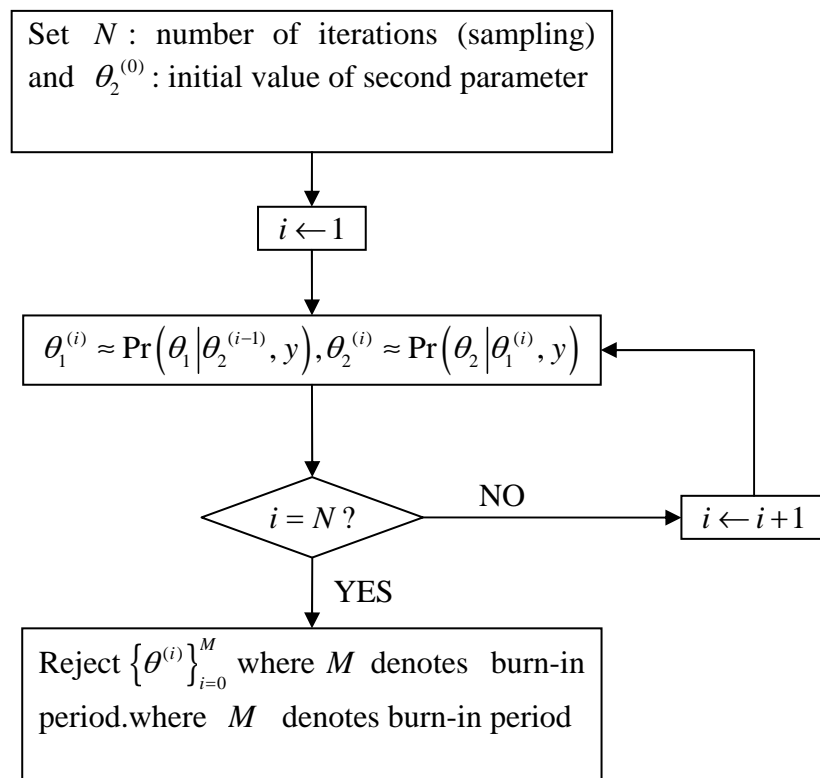


Figure 1 Gibbs Sampling

Gibbs sampling requires two conditions: (1) the full conditional distribution is known, and (2) sampling from the full conditional distribution is available. Although sample $\{\theta^{(i)}\}_{i=0}^N$ is not random, the transition probability from $\theta^{(i)}$ to $\theta^{(i+1)}$ is $K(\theta^{(i)}, \theta^{(i+1)} | y) = \Pr(\theta_1^{(i+1)} | \theta_1^{(i)}, \theta_2^{(i)}, y) \Pr(\theta_2^{(i+1)} | \theta_1^{(i+1)}, \theta_2^{(i)}, y)$; hence, it states that $\{\theta^{(i)}\}_{i=0}^N$ is a sample from a Markov chain with kernel K . As $N \rightarrow \infty$, the distribution of $\theta^{(i)}$

will converge to the posterior distribution of θ . Upon completion, the first M samples are rejected, where M is the burn-in period. The resulting sample $\{\theta^{(i)}\}_{i=M+1}^N$ is used to calculate the posterior expectation.

Because we can explicitly calculate the full conditional probability of the Hierarchical Regression model (Table 1), this study uses Gibbs sampling.

Full Conditional Probability of Hierarchical Bayesian Regression Model The Hierarchical Regression model assumes that parameters vary across households or stores but independently and identically follow the same distribution; that is, the Hierarchical Regression model assumes the parameters' hierarchical structure. Thus, the parameter for the said store is estimated using not only its own data but that of all other stores for the given time period. Furthermore, by estimating the parameters using the Bayesian approach, we can use historical data from both the said store and from all of the other stores. Thus, the Hierarchical Bayesian Regression model uses all the historical data from all of the stores for parameter estimation.

Table 1 Hierarchical Bayesian Regression Model vs. Non-Hierarchical and Non-Bayesian model

Hierarchy of Parameters(also using data of other stores)	Parameter Estimation using Bayesian Approach (also using past data)	
	NO	YES
NO	Regression Model	Bayesian Regression Model
YES	Hierarchical Regression Model	Hierarchical Bayesian Regression Model

The Hierarchical Regression model consists of two models: the within-subject model which directly describes the response structure within stores, and the other is the between-subjects model describing the relationship between stores. That is, the within-subject model is a regression with parameters specific to a certain store. This model is expressed as follows:

$$y_{ht} = x'_{ht}\theta_h + \varepsilon_{ht} \approx N(x'_{ht}\theta_h, \sigma_h^2), \text{ where } h = 1, \dots, H \text{ and } t = 1, \dots, T_h \text{ denote store and week, respectively. } x_{ht}, y_{ht} \text{ is the observed data, and } \theta_h, \sigma_h^2 \text{ is the parameter.}$$

As stated above, the between-subject model determines the relationship between the within-store models, and is expressed as follows:

$$\theta_h | \bar{\theta}, \Sigma \approx N_k(\bar{\theta}, \Sigma), \text{ where } k = \dim \theta = \dim \bar{\theta}$$

Thus, $\theta_1, \dots, \theta_H, \sigma_1^2, \dots, \sigma_H^2, \bar{\theta}, \Sigma$ are the parameters to be estimated. Assume these parameters to be the following.

$$\theta_h | \bar{\theta}, \Sigma \approx N_k(\bar{\theta}, \Sigma)$$

$$\sigma_h^2 \approx IG\left(\frac{V_0}{2}, \frac{S_{oh}}{2}\right)$$

$$\bar{\theta} \approx N_k(0, \Lambda_0)$$

$$\Sigma \approx IW(\nu_0, V_0)$$

That is, we can explicitly express full conditional distributions of each parameter, and estimate these parameters by the MCMC method using Gibbs sampling.

CALCULATION OF DYNAMIC PRICE ELASTICITIES

Data

The weekly sales data of green iced tea in a 2L PET bottle is used for parameter estimation. This data is provided by a retailer with roughly 130 supermarkets located around Tokyo. The data period is from May 24, 2010 to January 16, 2012, an 87-week data. Note that green iced tea sales are not seasonal. Setting the parameters estimated using the first 20 weeks' prior cumulative data, the parameters beginning at the 21st week (November 27, 2010) are Bayesian updated immediately using the previous week's parameters.

Model

In this study, I estimate the parameters of the Hierarchical Bayesian Regression model by the MCMC method using Gibbs sampling, and then calculate price elasticities by week and by store.

Let us examine the within-subject model. The dependent variable is Purchase Incidence (PI) y_{ht} , that is, sales volume per 1,000 shoppers. The independent variables are average shelf price x_{1ht} and number of fliers x_{2ht} per week. Suppose that external factors, like weather, competition, and other factors do not affect the in-store sales, but the number of shoppers, PI is regarded as the sales index relative only to the in-store factors. Thus, the within-subject model is expressed as follows.

$$y_{ht} = \theta_{1h} + \theta_{2h}x_{1ht} + \theta_{3h}x_{2ht} + \varepsilon_{ht} \approx N(\theta_{1h} + \theta_{2h}x_{1ht} + \theta_{3h}x_{2ht}, \sigma_h^2)$$

And suppose that the prior distribution of the between-subject model is as follows.

$$\begin{aligned}\theta_h | \bar{\theta}, \Sigma &\approx N_3(\bar{\theta}, \Sigma) \\ \bar{\theta} &\approx N_3(0, \Lambda_0) \\ \Sigma &\approx IW(5, 5I_3)\end{aligned}$$

As stated earlier, setting the parameters estimated from the first 20 weeks' data, each subsequent parameter is updated and price elasticities are calculated for the 92 stores. Using the same procedure, 67 weeks' price elasticities from 92 stores were calculated.

Calculation of Dynamic Price Elasticities

To estimate the parameters, I used the "rhierLinearModel" function with the Bayesian package of R language (Manual of Package "bayesm"). The number of iterations of the MCMC method was set at 100,000. Although most computations converge in 50 iterations (Figure 2), I set the entire burn-in period to 10,000. Posteriors were calculated using 90,000 samples. My PC's CPU is an Intel Core i7-2600 3 40GHz, and computing one price elasticity takes roughly one hour.

Three parameters (constant θ_1 , price θ_2 , and flier θ_3) were estimated for one week and one store, yielding a total number of 18,492 estimated parameters, that is, 92 stores times over 67 weeks. Accordingly, the price elasticity formula for regression is $\theta_2 x / (\theta_1 + \theta_2 x)$, and 6,164 (92 stores times 67 weeks) price elasticities were calculated.

VISUALISATION OF DYNAMIC PRICE ELASTICITIES

In this study, I visualized the calculated 6,164 price elasticities via Google Earth (Figure 3). Google Earth uses not only a Graphic User Interface but also Keyhole Markup Language (KML), which is based on Extensible Markup Language (XML). I wrote a KML program that indicates 97 price elasticities at points where each store is located on the Kanto area map (Akanemaru, Uchibe, & Morita, 2007). These price elasticities vary weekly by manipulating the slider bar in the screen's upper left corner. The size of this program is roughly 30,000 rows, 1.5 Mbytes.

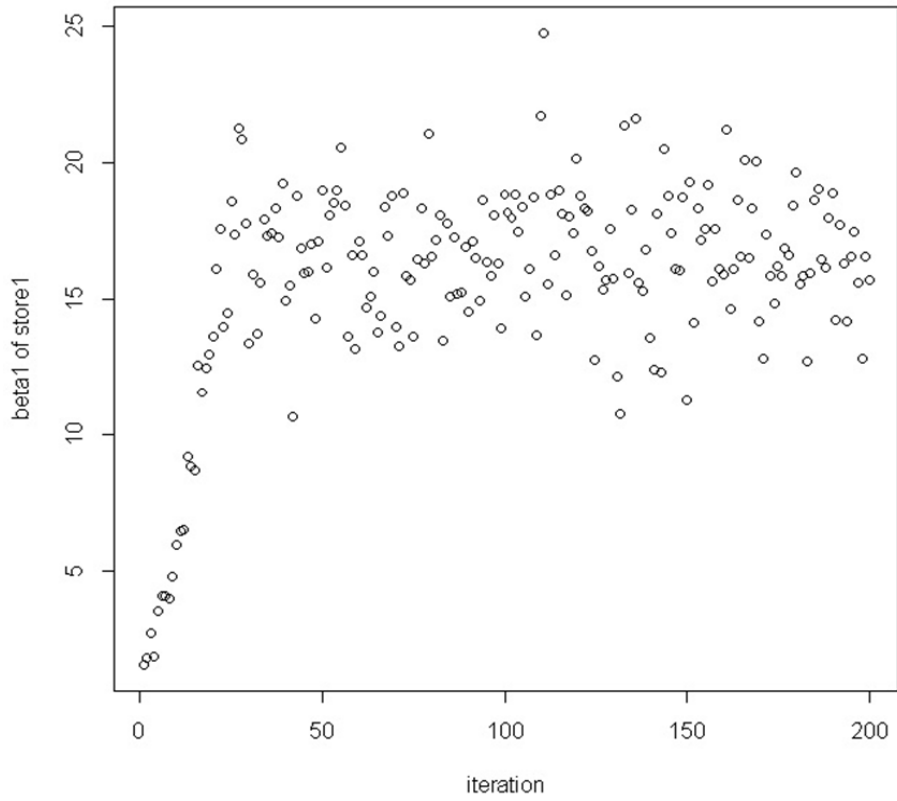


Figure 2 Drawings of a Parameter until 200 Iterations

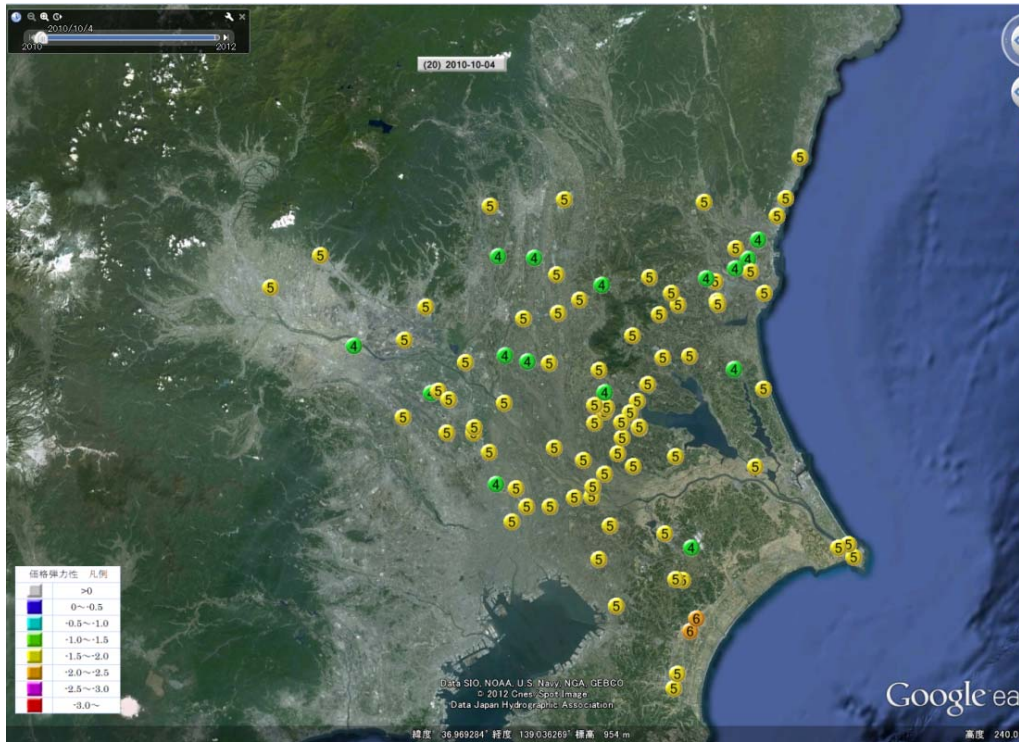


Figure 3 Price Elasticity Visualization

CONCLUSIONS

In this study, using the MCMC method with Gibbs sampling, I estimated the parameters of the Hierarchical Bayesian Regression model and calculated price elasticities by store and week. The parameters estimated are stable. As stated in Section 2.3, the Hierarchical Bayesian Regression model effectively estimates parameters.

The MCMC method was developed and improved upon for use primarily in areas of thermodynamics, statistical mechanics, and others. Because of the dramatic improvement in the performance capabilities of personal computers, the MCMC method is now also used in marketing research because of its parameter estimation power using large quantities of marketing data.

The number of price elasticities calculated in this study was 6,164, across 92 stores and 67 weeks. Google Earth was used to visualize these price elasticities. I wrote a KML program that indicates 97 price elasticities at the point of each store's location on the Kanto area map, by week. The viewer can observe these price elasticities varying weekly by manipulating the slider bar. The size of this program is roughly about 30,000 rows and 1.5 Mbytes which is accessible for most modern computers.

In this paper, dynamic price elasticities by store and week were estimated by Hierarchical Bayesian Regression model using sales history data, and were visualized using Google Earth.

These visualized dynamic price elasticities are substitutes for dynamic environmental factors including marketing activities of competitive stores, and combined with geographical information, therefore these can be useful for marketers to support decision making.

Finally the author would like to point out that because this approach using Hierarchical Bayesian Regression model with Google Earth enables us to analyze insufficient data to identify heterogeneity between subjects considering geographical information, this can be used not only to pricing strategies of retailers, but also to many fields including town planning, analyzing markets worldwide, and so on.

ACKNOWLEDGEMENT

The author would like to thank a Japanese supermarket company for providing weekly aggregated POS data. The author also would like to thank Enago (www.enago.jp) for the English language review. And the author received many valuable comments from two reviewers.

REFERENCES

- Akenemaru, Uchibe, T., & Morita, A. (2007). *Guidebook of Google Earth, KML Language 2.2, and COM API*. Gijutsu Hyoron Sha.
- Blattberg, R. C., & George, I. (1991). Shrinkage Estimation of Price and Promotional Elasticities. *Journal of the American Statistical Association*, 86 (Jun), 304–315. <http://dx.doi.org/10.2307/2290562>
- Chintagunta, P. K., Dube, J. P., & Singh, V. (2008). Balancing Profitability and Customer Welfare in a Supermarket Chain. *Quantitative Marketing and Economics*, 1, 111–147.
- Dobson, P. W., & Waterson, M. (2008). Chain-Store Competition: Customized vs. Uniform Pricing. *Working Paper, No.840*, University of Warwick, Department of Economics, Coventry.
- Levy, M., & Weitz, B. A. (2011). *Retailing Management with Connect Plus*, McGraw-Hill/Irwin.
- Rossi, P. (2012). Package 'bayesm'. Retrieved Dec 3, 2013, from <http://cran.r-project.org/web/packages/bayesm/bayesm.pdf>
- Montgomery, A. L. (1997). Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data. *Marketing Science*, 16, 315–337. <http://dx.doi.org/10.1287/mksc.16.4.315>
- Rossi, P. E., Allenby G. M., & McCulloch R. (2005). *Bayesian Statistics and Marketing*. Wiley.
- Tajima H. (2012). Some Considerations on the Bayesian Inference of Logit Models Using MCMC Algorithm. *The Journal of Tokyo Keizai University*, 274, 145-162.
- Tienery L., & Kadane J. B. (1986). Accurate Approximation for Posterior Moments and Marginal Densities. *Journal of American Statistical Association*, 81, 82–86.

