THREE LEVEL HIERARCHICAL BAYESIAN ESTIMATION IN CONJOINT PROCESS

Paweł Kopciuszewski

Institute of Mathematics and Computer Science, Czestochowa University of Technology, Poland

Abstract. The conjoint process is the process which is based on some latent variables and the experts opinion is the crucial information in the estimation of unknown parameters. That is why the excellent technique applied in the conjoint process is Bayesian analysis including hierarchical Bayesian models. In the paper, the three level Bayesian model is proposed to estimate the parameters of product and consequently to choose the best product profile. The final estimation is executed with MCMC methods.

Introduction

Suppose that the considered product has some features which can be set from the least important to the most important of them. Historical observations of customers behavior gives us the information what features are more or less important to customers. But this information is not available for new products. The only way to rank the all potential product features is to collect the data having obtained in statistical survey with the participation of some respondents as potential customers. The most popular technique to solve this problem is conjoint technique. The conjoint technique was proposed in 1964 by Luce and Tukey [1] and it was introduced to the marketing research in early 1970's [2]. The most popular software to conjoint is sawtooth software. There are many kinds of conjoint technique such as Adaptive conjoint analysis (ACA), Choice based analysis (CB), Ratings based conjoint (RBC) and others.

From the statistical point of view the most popular conjoint technique is based on regression analysis and factor analysis. In this paper the Bayesian hierarchical models are adopted to conjoint process.

The conjoint technique is applied to obtain the most expected product features by customers. The crucial point of this analysis is a collection the data needed to building the statistical model. The data is collected with questionnaires fulfilled in by respondents taking part in statistical survey. The construction of all questionnaires is the most sensitivity point to the final results. It should be taken into account that a respondent can answer only a little number of question and that the sequence of all questions determines the possible answers. It is assumed that each respondent can compare no more than six different product profiles. The popular are questionnaires with two different profiles to compare. Certainly, there are many different two-profiles questionnaires given to respondents in statistical survey. Two-profiles questionnaires are simply to fill in by respondent but the collected data of them is difficult to analysis. The method of questionnaire construction is not the aim of this paper but its importance must be stressed here. The conjoint process consists of three stages: data collection, theoretical model building, part-worths estimation.

Firstly all respondents fill in the given questionnaires with the answers about the importance of all variables. The importance of the variable is called its weight. This questionnaire is only preliminary step to identify the product with the highest quality level. Notice that respondents do not take into account different relatives among all variables or levels of these variables. Certainly, the questionnaires for all respondents must include different sorts of questions, because in the case of many questions, the ones from the beginning of the questionnaire are more precisely considered than others by respondents.

Secondly, the respondents fill in the questionnaires with hypothetical partially-profiled products and rank them all. The second questionnaire step gives us the most information to build the statistical model to choose the best product profile.

1. Hierarchical Bayesian model in conjoint process

Data collection:

Let us assume that there are m features of the products represented by m different categorical variable and each of them has q_i different categories $i = 2,\ldots,m$. $\sum_{i=1}^{m} q_i = q.$ All of these variables characterize the product and our aim is to find the combination of their categories which guarantees the highest acceptance level by the customer but without the historical observations. The profile of the product is any combination of categories of these variables. Hence we have $\prod_{i=1}^{m} q_i$ possible profiles. Certainly respondents can not evaluate all of these profiles. That is why some representative profiles are chosen to estimate the most acceptable categories of variables.

Assume that there are k respondents which answer the questionnaires' questions. To the estimation of importance of all categories of variables we build three-level hierarchical Bayesian model as it is presented in Figure 1.

Null level in hierarchy - likelihood definition is defined as follows:

$$y_{ih} = IL(x_{ih}{}^T \beta) + \varepsilon_{ih},\ i = 1,\ldots,k;\ h = 1,\ldots,n_i \quad (1)$$

where:

IL - the inverse logit transformation, that is:

$$IL(z) = \frac{1}{1+\exp(-z)},\ z \in R \tag{3}$$

y_{ih} - the binary variable of the acceptance the *h*-th profile by the *i*-th respondent,

x_{ih} - the vector of the independent variables associated with the *h*-th profile considered by the *i*-th respondent,

β - the vector of part-worths with dimension $d = \sum_{i=1}^{m} q_i - m + 1$

ε_{ih} - a random error with binomial distribution.

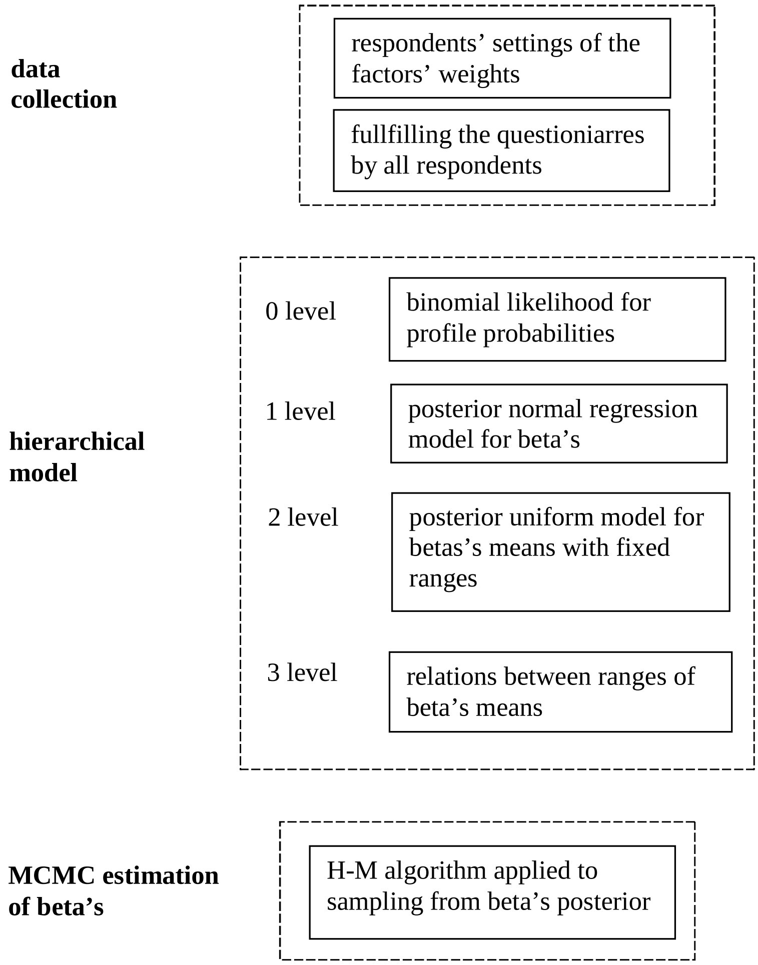

Fig. 1. Conjoint process

The first level in hierarchy is defined as follows:

$$\beta \sim N(\alpha, D) \tag{4}$$

where:

α - the mean of the vector β,

α_1 - the mean of the intercept,

α_{ij} - the mean of the part-worths for the given j-th category of the i-th variable, $i = 2,\dots,m$, $j = 1,\dots,q_i$.

D - covariance matrix of the vector β.

The second level in hierarchy is defined as follows:

$$\alpha_1 \sim U(-a_1, a_1) \tag{5}$$

$\alpha_{ij} \sim U(-a_i, a_i)$, j-th category of the i-th variable, $i = 2,\dots,m$

where: $a_i > 0$ - constraint parameter.

Notes that α_{ij} are independent random variables.

These parameters are estimated from the first stage questionnaires. If the variable is more important to the respondent then its potential influence on the target in regression model is higher than the other variables. That is it can changes on the wider range around zero and then the constant a_i is bigger. We propose the following formula to evaluate the parameters a_i.

$$a_i / a_j = w_i / w_j \tag{6}$$

where w_i, w_j are the weights related to all variables from the first stage questionnaires.

Notice that parameters α_{ij} are random variables with prior distributions.

The third level of hierarchy is defined as follows:

The parameters a_i gives us the information about the importance of all variables (not of the importance of the all related dummies). But the ranges of these parameters are unknown from the questionnaires. These ranges must be established from any model based on historical observations if they are available. In our case this model is unknown, that is why we choose noninformative proper prior for these parameters. Hence

$$a_2 \sim U(0, A) \tag{7}$$

$$a_1 \sim U(0, c)$$

where A, c are the known constants chosen with the expert's opinion.

The choice of the constants A and c is not crucial to the parth-worths estimation.

From (6) and (7) the priors for all alpha's can be obtained.

2. Estimation of the posterior parameters with MCMC algorithm

The estimation of the vector β is the main aim of our interest.
The joint posterior is as follows:

$$\pi(\beta,\alpha,a) \propto \prod_{i=1}^{k}\prod_{h=1}^{n_i} B(y_{ih};IL(x_{ih}{}^{T}\beta))\cdot N(\beta;\alpha,D)\cdot\left(\frac{1}{a_2}\right)^{q}\cdot\frac{1}{a_1} \tag{8}$$

where $B(u,p)$ is the binomial probability in the point u with acceptance probability of the positive value p. Hence the posterior marginal of beta is:

$$\iint_{D(\alpha,a)} \pi(\beta,\alpha,a)\,d\alpha\,da \tag{9}$$

where $D(\alpha,a)$ denotes the support of the density $\pi(\beta,\alpha,a)$. The above integral can be simplified as follows:

$$\pi(\beta) \propto \prod_{i=1}^{n}\prod_{h=1}^{n_i} B(y_{ih};IL(x_{ih}{}^{T}\beta))\cdot[\Phi(D^{-1}(A_{\max}-\beta))-\Phi(D^{-1}(A_{\min}-\beta))] \tag{10}$$

where: Φ is the q-dimensional normal distribution, $A_{min} = -A_{max}$, $(A_{max})_i = a_i$.

The strictly value of the posterior mode can be obtained only by applying Monte Carlo methods. Notice that the formula of the marginal posterior density for beta parameter determines the process of the posterior density generation as follows (accordingly to the H-M method):

1. Draw the vector beta from any n-dimensional density (if the historical prior estimates of α are known then it may be the normal density, otherwise the uniform density on the range of β possible values). Mark this vector with β_k.
2. Accept or reject the value β_k drawn in the previous step accordingly to the Hastings-Metropolis proposition probability (the right side of the proportion (10)).
3. If the R-ratio (introduces by Geman and Geman in 1984) [3] is near 1 then stop the algorithm, otherwise back to the step 1 and generate the next value β_{k+1}. The authors suggest values less than 1.2 to stop the algorithm.

This algorithm allows us to generate the sample from the above multivariate density with generating the sample from multivariate normal density and with calculating some values of the posterior density $\pi(\beta)$.

Certainly, the algorithm has to be executed many times to obtain the sample from the posterior density of the vector β. But two states from the drawn sample can be chosen only if they are not correlated, that is the generated Markov chain is the process with long-time memory.

References

[1] Luce R.D., Tukey J.W., Simultaneous conjoint measurement. A new type of fundamental measurement, Journal of Mathematical Psychology 1964, 1, 1-27.

[2] Green P.E., Srinivasan V., Conjoint analysis in consumer research, Issues and Outlook, Journal of Consumer Research 1978, 5, 103-123.

[3] Gelman A.B., Carlin J.S., Bayesian Data Analysis, Chapman & Hall, Harvard University, USA, 2000.