Analysis of Multicategory Purchase Incidence Decisions and Consumer Behaviors at a Department Store by Multidimensional Scaling

Atsuho Nakayama
Rikkyo (St. Paul's) University
3-34-1, Nishi-Ikebukuro, Toshima-Ku, Tokyo, Japan 171-8501
email: nakayama@lis.rikkyo.ne.jp

Abstract — **The present paper reports an analysis of purchase data at a department store. Consumers who rank in the top 20% in purchase amount were selected. The purchase data consist of 14 categories for a one-year period. The present analysis makes use of two kinds of similarities calculated from the purchase data. One is the similarities that show pattern of the joint purchase of two categories. The other is a pattern of the joint purchase of three categories. The former represents one-mode two-way data and is analyzed by one-mode two-way multidimensional scaling. The latter expresses one-mode three-way data and is analyzed by one-mode three-way multidimensional scaling. The results of the present analysis revealed the relationships among the categories on the multicategory purchase incidence decisions.**

1. Introduction

Generally, consumers make various decisions during shopping trips. The choice of one category may lead to a joint purchase of another category based on the complementary nature and similar purchase cycles of the two categories. On the other hand, the joint purchase may not occur because of the monetary reason, temporal restriction, or other observed factors. Therefore, it is necessary to make a finer analysis into how and when joint purchase decisions may or may not take place.

Multidimensional scaling (MDS) is a method for analyzing similarity data on a set of categories. It is considered that MDS is a suitable method to analyze multicategory purchase incidence decisions. MDS representation provides a common and useful criterion of multicategory. MDS is a method to map similarities into low dimensional Euclidean space.

A variety of models have been proposed and classified into different types according to the number of ways and modes (Arabie, Carroll, & DeSarubo, 1987[1]; Carroll & Arabie, 1980[2]). For example, the INDSCAL model (Carroll & Chang, 1970[3]) deals with two-mode three-way data. MDS has been widely used in many fields. In psychology, we have used these methods to understand the perception and evaluation of auditory stimuli such as speech and musical tones, visual stimuli such as colors and faces, and social entities such as personality traits and social situations. For further details, the readers should refer to Rosenberg and Kim (1975) [16], Rosenberg, Nelson, and Vivekananthan (1968) [17], Rothkopf (1957) [18], and Wish (1976) [24]. Sociology has used these methods to determine the structure of groups and organizations, based on member's perceptions of one another and their interaction patterns (see Coombs, 1964[5]; McGee, 1968[14]; Torgason, 1958[22]; Tucker & Messick, 1963[23]; Wish & Carroll, 1974[25]). Anthropology has used these methods for comparing different cultural groups based on their beliefs, language, and artifacts. Economists and marketing researchers have been using these methods for investigating the reactions of consumers to a wide variety of product classes (Green, Carmone, & Smith, 1989[8]; Green & Rao, 1971[9]). Educational researchers or educational psychologists have used these methods to study the structure of intelligence, different test batteries, and classroom environments.

The present paper reports the MDS analysis of purchase data provided from a department store. The interpretation of the configuration reveals the similarities and differences between the multicategory purchase incidence decisions by representing each category in multidimensional space. The present analysis analyzed two similarities. One is relationships of the joint purchase of two categories and the other is the joint purchase of three categories. The similarities on the joint purchase of two categories are analyzed by one-mode two-way MDS and those on the joint purchase of three categories by one-mode three-way MDS. The present analyses expressed the pattern of the multicategory purchase incidence decisions among 14 categories of a department store.

2. The Data

Purchase history data from January 1 to December 31 of consumers who rank in the top 20% of purchase amount in a client club of a department store were used. The data consist of 14 categories excluding the category of Groceries that is known to show very different characteristics in purchase decisions. The 14 categories are shown in Table 1 and Kimono means traditional Japanese clothes.

Table 1. Fourteen categories.

1 Woman's dresses and suits	8 Bedding
2 Woman's Shirts and Blouses	9 Furniture
3 Woman's Accessories	10 Household Equipments
4 Men's Apparel	11 Jewelry and Watches
5 Men's Accessories	12 Hobbies and Stationery
6 Children's Apparel	13 Sports
7 Kimono	14 Restaurants

The two aforementioned similarities are calculated from the purchase data of 14 categories.

3. The Analysis

A. One-mode Two-way Data Analysis

A 14×14 one-mode two-way symmetric similarities matrix was calculated from the purchase data. The matrix expresses the count of joint purchases. The matrix was analyzed by one–mode two-way MDS. The method and procedure were proposed by Kruskal (1964a[11], b[12]).

The analysis was done by using the maximum dimensionalities of ten through six. Therefore, the first stress values were obtained in ten- through unidimensional spaces. The second stress values were obtained in nine- through unidimensional spaces. The third stress values in eight- through unidimensional spaces, the fourth stress values in seventh- through unidimensional spaces, and the fifth stress values were in six- through unidimensional spaces. The smallest stress value in each dimensional space was chosen as the minimized stress value in that dimensional space. The resulting minimized stress values in five- through unidimensional spaces are 0.084, 0.143, 0.207, 0.292, and 0.435. The minimized stress value indicates that the two-dimensional results are an appropriate solution, and two-dimensional configuration helps easy understanding of the relationships among the categories. Thus the two-dimensional configuration is now discussed in one-mode two-way data analysis.

B. One-mode Three-way Data Analysis

A 14×14×14 one-mode three-way symmetric similarities matrix was calculated from the purchase data. The matrix shows the counts that three categories were purchased simultaneously. The matrix was analyzed by one–mode three-way MDS. That was proposed by Nakayama (2003[15]). The Shepard-Kruskal approach for non-metric multidimensional scaling (Kruskal, 1964a[11], b[12]; Shepard, 1962a[20], b[21]) was modified in Nakayama's model. The Shepard-Kruskal model is a simple one that can be easily generalized from two-way distance model into three-way distance model. For other methods on one–mode three-way MDS, see, for example, Cox, Cox, and Branco (1991) [6], De Rooij and Heiser (2000) [7], Hayashi (1989) [8], and Pan and Harris (1991) [19].

Three-way similarities δ_{ijk}, which are called three-way similarities function when a set of n objects are given and i, j, and k represent any three objects, are generalizations of two-way similarities so that δ_{ijk} determines how similar the three categories are. One configuration, $\mathbf{X}$ of n points $\mathbf{x}_i$= $(x_{i1},\ldots,x_{ip})$, is assumed i = 1, …, n, in a p-dimensional Euclidean space, where a point $\mathbf{x}_i$ in the coordinates corresponds to category i. The distance among three categories d_{ijk} is meant to reflect the values of the similarities δ_{ijk}. The distance among three categories d_{ijk} is then defined as

$$d_{ijk} = \sqrt{d_{ij}^2 + d_{jk}^2 + d_{ki}^2 - \min(d_{ij}^2, d_{jk}^2, d_{ki}^2)}. \quad (1)$$

Inside of the square root of the right hand-side of Equation (1) is subtraction of the smallest distance in d_{ij}^2, d_{jk}^2, and d_{ik}^2 form the sum of squared distances. The sum of squared distances is not exact enough to explain triadic relationships since the model must be represented by nonlinear quadratic form. The Shepard-Kruskal model for non-metric MDS is applied to Nakayama's model. In order to get the best fitting configuration, the initial configuration was firstly determined and then d_{ijk} is calculated. Next, monotone regression was used to find $\hat{d}_{ijk}$ which satisfies the following conditions,

$$\delta_{ijk} > \delta_{rst} \Rightarrow \hat{d}_{ijk} \leq \hat{d}_{rst} \quad \text{(for all } i<j<k,\ r<s<t).$$

The measure of badness fit of d_{ijk} to δ_{ijk} is called stress S, which is based on the stress formula 2 (Kruskal & Carroll, 1969[13]). It is defined as

$$S = \sqrt{\sum_{i<j<k}^{n} (d_{ijk} - \hat{d}_{ijk})^2 \Big/ \sum_{i<j<k}^{n} (d_{ijk} - \bar{d}_{ijk})^2}\ . \quad (2)$$

The one-mode three-way symmetric similarities were analyzed. The configurations were derived in the ten- through unidimensional space by the same way as one-mode two-way analysis. The stress values obtained in the five- through unidimensional space were 0.513, 0.518, 0.524, 0.536, and 0.593. Examination of these five stress values suggests that the two-dimensional configuration should be adopted as a solution. Moreover, the two dimensional configuration can easily be visualized and we can understand the relationships among the categories. The two dimensional results are chosen as the solution in one-mode three-way data analysis.

4. Results

The two dimensional solution obtained from one-mode two-way data analysis is shown in Figure 1. In the right half of Figure 1, non-garment categories such as Kimono, Furniture, and Jewelry and Watches are placed. In the left half of Figure 1, garment categories such as Men's Apparel, Men's Accessories, Sports, and Ladies' are given. Horizontal Dimension 1 of the solution would represent the garment versus non-garment dimension. Kimono is regarded as traditional handicrafts rather than casual wear. Kimono is made of silk and is usually very expensive. Nowadays, Kimono is hardly worn daily life. Kimono is worn at formal or traditional occasions such as funerals, weddings or tea ceremonies. Therefore, it is thought that characteristics of Kimono are more similar to non-garment than garment. The categories associated with luxury goods such as Sports, Jewelry and Watches, and Kimono are placed in the upper half of Figure 1, and the categories associated with domestic

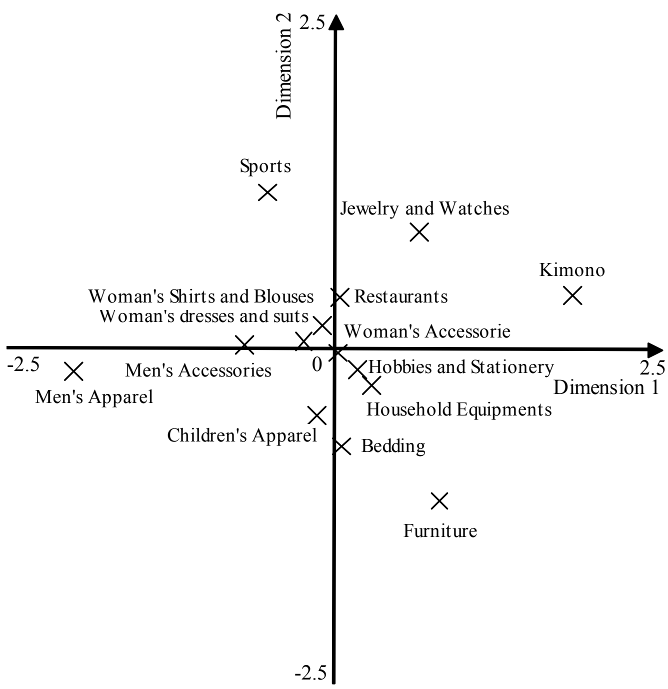

Figure 1. The two dimensional configuration obtained from the 14×14 similarities by using one-mode two-way analysis.

articles such as Furniture, Bedding, and Children's Apparel are in the lower half of Figure 1. Vertical Dimension 2 then would represent the luxury goods versus domestic articles dimension.

The categories such as Woman's dresses and suits, Woman's Shirts and Blouses, and Woman's Accessories are often purchased together with other categories because these categories are located in the central part of the configuration. Figure 1 also shows that the categories located in the edge of Figures 1 are not jointly purchased with the other categories because these categories are far from the other categories.

The two dimensional solution obtained from one-mode three-way data analysis is shown in Figure 2. In order to compare the results of one-mode three-way analysis with those of one-mode two-way analysis, the configuration for the former analysis (Figure 2) matched to the configuration for the latter analysis (Figure 1) by using Procrustes analysis (Cliff, 1966[4]) is given in Figure 3.

The dimensions of Figure 3 represent tendencies similar to those of Figure 1. In the right half of Figure 3, non-garment categories such as Kimono, Jewelry and Watches, Hobbies and Stationery, and Furniture are placed. In the left half of Figure 3, garment categories such as Men's Apparel, Sports, Men's Accessories, and Children's Apparel are placed. Horizontal Dimension 1 of the solution would reflect the difference between garment categories and non-garment categories. The categories associated with luxury goods such as Jewelry and Watches, Sports, Kimono are placed in the upper half of Figure 1, and the categories associated with domestic articles such as Furniture, Children's Apparel, and Bedding in the lower half of Figure3. Vertical Dimension 2 would represent the difference between the categories of luxury goods and the categories of domestic articles.

Figure 3 shows that the categories which are located in the central part of the configuration often purchased together with other categories. However, the categories are located in the edge of the configuration hardly are jointly purchased with the other categories because these categories are far from the other categories.

5. Discussion

The solutions of one-mode three-way data analysis and

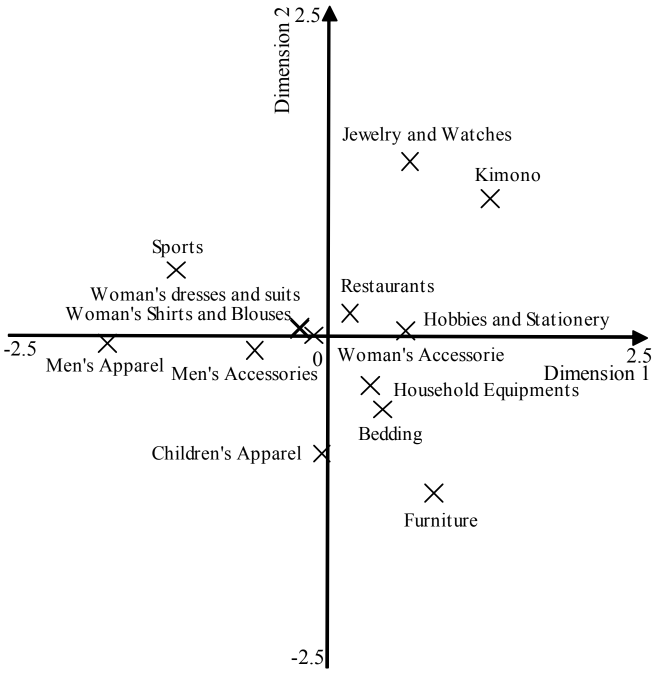

Figure 2. The two dimensional configuration obtained from the 14×14×14 similarities by using one-mode three-way analysis.

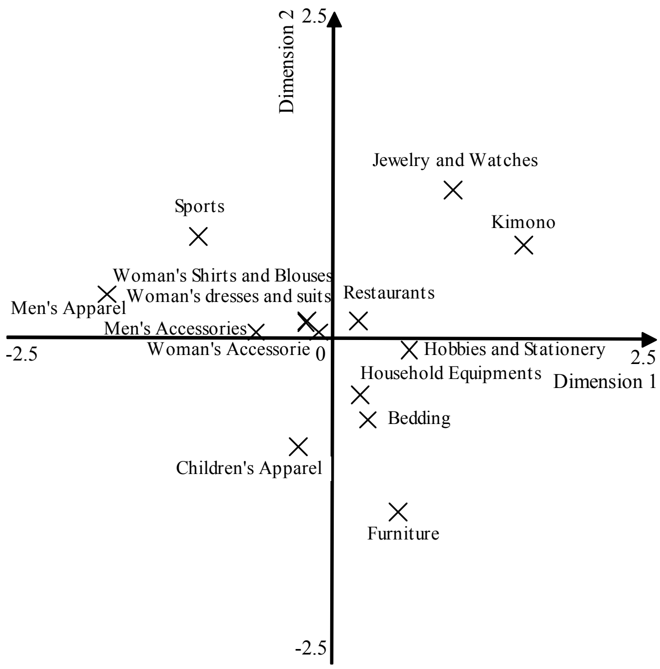

Figure 3. The configuration for one-mode three-way analysis matched to the configuration for one-mode two-way analysis by using Procrustes analysis.

one-mode two-way data analysis show that the multicategory purchase incidence decisions consist of two patterns of choices. One is the choice of garment categories or non-garment categories. The other is the choice of categories of luxury goods or categories of domestic articles. These choices represent a criterion of the common judgment for joint purchase. Consumers make a combination of choices when they make joint purchase of two categories or three categories. There exist some frameworks of the purchase of categories on the basis of these choices.

The frameworks of the purchase of two categories consist of Women-related categories, Recreation goods, Domestic articles, Men-related categories, and Luxury goods. Women-related categories refer to joint purchase of Woman's dresses and suits, Woman's Shirts and Blouses, and Woman's Accessories. Recreation goods contain Household Equipments and Hobbies and Stationery. Domestic articles include Children's Apparel and Bedding. Men-related categories show Men's Apparel and Men's Accessories. Luxury goods represent Kimono and Jewelry and Watches.

The frameworks of the purchase of three categories slightly differ from those of two categories. They have three tendencies on Men-related categories, Women-related categories, and Commodities. Men-related categories consist of the Men's Apparel, Men's Accessories, and Sports.

The three-joint purchase for Women-related categories occurs most frequently among Woman's dresses and suits, Woman's Shirts and Blouses, and Woman's Accessories. The followings are the two secondly-frequent cases. One is a combination of two-joint purchase from Woman's dresses and suits, Woman's Shirts and Blouses, and Woman's Accessories with the use of Restaurants. The other is a combination of two-joint purchase from Woman's dresses and suits, Woman's Shirts and Blouses, and Woman's Accessories with purchase of Men's Accessories.

For Commodities five combinations take place. The first is a combination of Household Equipments, Bedding, and Hobbies and Stationery. The second is a combination of Household Equipments, Bedding, and Children's Apparel. The third is a combination of Restaurants, Hobbies and Stationery, and Household Equipments, the fourth a combination of Restaurants, Hobbies and Stationery, and Bedding, and the last is a combination of Bedding, Children's Apparel, and Furniture.

The important insights into multicategory purchase incidence decisions at a department store were obtained by the results of one-mode three-way data analysis and also by the results of one-mode two-way data analysis. Consumers make a joint purchase of the categories based on two patterns of choices. Therefore, the strategic approaches which are based on these choices would be of great significance. Some frameworks of categories are also derived from these choices. Sales success could be attained by promotions based on these patterns of purchase. For example, direct mails should be sent by consideration of categories which are included in the patterns of purchase. These categories should be arranged to occupy nearer positions to each other in the department store.

Finally, the stress values of one-mode two-way data analysis cannot be directly compared with those of one-mode three-way data analysis. The stress values of one-mode two-way data analysis were converted for the purpose of comparison. The converted stress value in two-dimensional space is 0.633. The stress value in two-dimensional space for the one-mode three-way data analysis is 0.536. Consequently, it is concluded that the one-mode three-way data analysis more appropriately expresses the information of the data than one-mode two-way data analysis does.

Acknowledgements

The author would like to express their gratitude to reviewers for their constructive and helpful reviews. The heartful gratitude also goes to Professor Akinori Okada for a great deal of his valuable advices and suggestions.

References

[1] Arabie, P., & Carroll, J. D., & DeSarbo, W. S. (1987). *Threey- way scaling and Clustering.* Newbury, Park, CA: Sage.

[2] Carroll, J. D., & Arabie, P. (1980). Multidimensional scaling. In M. R. Rosenzweing & L. W. Porter (Eds.), *Annual Review of Psychology*, (Vol. 31, pp. 607-649). Palo Alto, CA: Annual Reviews.

[3] Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an *N*-way generalization of "Eckart-Young" decomposition. *Psychometrika*, *35*, 283-319.

[4] Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, *31*, 33-42.

[5] Coombs, C. H. (1964). *A theory of data*. New York: John Wiley.

[6] Cox, T. F., Cox, M. A. A., & Branco, J. A. (1991). Multidimensional scaling for *n*-tuples. *British Journal of Mathematical and Statistical Psychology*, *44*, 195-206.

[7] De Rooij, M., & Heiser, J. W. (2000). Triadic distances models for the analysis of asymmetric three-way proximity data. *British Journal of Mathematical and Statistical Psychology*, *53,* 99-119.

[8] Green, P. E., Carmone, F. J., & Smith, S. (1989). *Multidimensional scaling: Concepts and applications*. Boston, MA: Allyn and Bacon.

[9] Green, P. E., & Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, *8*, 355-363.

[10] Hayashi, C. (1989). Multiway data matrix and method of quantification of qualitative data as a strategy of data analysis. In R. Coppi and S. Bolasco (Eds.), *Multiway data analysis*, (pp. 131-142). Amsterdam: North-Holland.

[11] Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychomertrika*, *29*, 1-27.

[12] Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychomertrika*, *29*, 115-129.

[13] Kruskal, J. B., & Carroll, J. D. (1969). Geometrical models and badness-of-fit functions. In P. R. Krishnaiah (Eds.), *Multivariate analysis,* (Vol. *2*, pp. 639-671). New York: Academic Press.

[14] McGee, V. E. (1968). Multidimensional scaling of *N* sets of similarity measures: A nonmetric individual differences approach. *Multivariate Behavioral Research*, *3*, 233-248.

[15] Nakayama, A. (2003). *Three-way data analysis model in multidimensional scaling*. Manuscript submitted for publication.

[16] Rosenberg, S., & Kim, M. P. (1975). The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, *10*, 489-502.

[17] Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impression. *Journal of Personality and Social Psychology*, *9* , 283-294.

[18] Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some pairedassociate learning tasks. *Journal of Experimental Psychology*, *53*, 94-101.

[19] Pan, G., & Harris, D. P. (1991). A new multidimensional scaling technique based upon associations of triple objects - P_{ijk} and its application to the analysis of geochemical data. *Mathematical Geology*, *6*, 861-886.

[20] Shepard, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika, 27*, 125-140

[21] Shepard, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika, 27*, 219-246.

[22] Torgason, W. S. (1958). Theory and methods of scaling. New York: John Wiley.

[23] Tucker, L. R., & Messick, S. (1963). An individual differences model for multidimensional scaling. *Psycholometrika*, *37*, 3-27.

[24] Wish, M. (1976). Comparisons among multidimensional structures of interpersonal relations. *Multivariate Behavioral Research, 11*, 297-324.

[25] Wish, M., & Carroll, J. D. (1974). Applications of individual differences scaling to studies of human perception and judgment. In E. C. Carterette and M. P. Friedman (Eds.) *Handbook of Perception*, (vol. *2*. pp. 449-491). New York: Academic Press.