KEER2010, PARIS | MARCH 2-4 2010
INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH 2010

STUDY ON CONFLICT OF EVALUATION VALUES AND INTEREST OF RESPONDENT

Yosuke Watanabe[*], Tomohiro Yoshikawa[*], Takeshi Furuhashi[*]

[*]*Nagoya University, Japan*

ABSTRACT

Companies often employ questionnaire(s) in order to design marketing strategies or to grasp the trends. Recently, Web questionnaire survey has become popular with the spread of the Internet in order for companies to reduce cost and to get a lot of questionnaire data in a short time. However, in the Web questionnaire survey, some respondents do not respond the questions seriously because it is not done face-to-face but is done basically to get giveaway or points. This paper proposes a quantification of the conflict of response and the respondent's interest based on the evaluation values. Conflict of response is based on the contradiction by the pair(s) of questions with counter meaning. Respondent's interest is derived from the average of variances of evaluation values for each evaluation object. This paper also show an interactive grouping method based on the visualization result, in which a user makes some groups of the respondents interactively with trial and error. It is expected that appropriate analysis can be done by the proposed method.

Keywords: *Analysis of Questionnaire Data, Visualization, Interactive Grouping, Fuzzy Theory, Multi Dimensional Scaling*

1. BACKGROUND

Companies often employ questionnaire(s), which is the quantified data of respondents' Kansei, in order to design marketing strategies or to grasp the trends. For example, when companies plan a new project, they often survey the impression of people by a questionnaire. They also carry out questionnaire(s) and the analysis of them gives some marking strategy with the prediction of marketing scale or the target groups to sell or the hint of modification of the product [1][2]. The rating scale method is widely used in questionnaires [3][4],

designed to obtain impressions for evaluation objects such as products, services and brands. This method requires plural evaluation objects and questions, and respondents answer the sets of questions for each evaluation object according to their impressions with multiple grading scales. In this way, people's impressions on evaluation objects can be quantified in the form of graded data.

Questionnaire data used to be carried out by face-to-face survey or mail-in survey. Recently, Web questionnaire survey has become popular with the spread of the Internet in order for companies to reduce cost and to get a lot of questionnaire data in a short time. However, in the Web questionnaire survey, some respondents do not respond the questions seriously; for example, they might respond to them randomly or equally without reading questions, because it is not done face-to-face but is done basically to get giveaway or points. If those responses are included into questionnaire data, there is a possibility that the analysis result of the data is not accurate and shows different characteristics or trends.

This paper proposes a quantification of the conflict of response and the respondent's interest based on the evaluation values. The conflict of response is based on the contradiction by the pair(s) of questions with counter meaning. The respondent's interest is defined as the average of variances of evaluation values for each evaluation object. This paper also show an interactive grouping method based on the visualization result [5], in which a user makes some groups of the respondents interactively with trial and error [6]. It is expected that appropriate analysis can be done by the proposed method.

This paper applies the proposed method to an actual questionnaire data on a new outdoor product. Firstly, it verifies the conflict of response based on the distances between questions [7] and the respondent's interest by comparing the number of characters in free text form of the questionnaire between respondents who have answered with interest and those without it. Secondly, this paper compares the analysis result between the proposed method and the conventional method, without considering the conflict of response and the respondent's interest. It shows that important groups of respondents for the marketing strategy can be found by the proposed method, which is difficult to find by the conventional one, and an effective analysis can be done.

2. PROPOSED METHOD

2.1. Conflict of response

Conflict of response is quantified using pair(s) of questions which have counter meaning each other. In the proposed method, the contradiction of evaluation value is defined as quasi-coefficient of correlation between counter questions that employs the median value of grading scales instead of the average of evaluation values. For the graded data obtained with the rating scale method, evaluated values for the paired questions, 1 and 2, on i-th object are defined as x_{i1} and x_{i2}, and the median value is defined as M. The value of conflict of response C is defined as follows:

$$C = \frac{\sum_{i=1}^{n}(x_{i1} - M)(x_{i2} - M)}{\sqrt{\sum_{i=1}^{n}(x_{i1} - M)^2}\sqrt{\sum_{i=1}^{n}(x_{i2} - M)^2}}, \quad (1)$$

where n denotes the number of evaluation objects. C runs from -1 through 1, and C becomes large when the evaluated values for the paired questions with counter meanings are either lower or higher than the M. Thus, if a respondent's C is large, the data of the respondent concerned have a possibility that there is in a conflict in it, because he/she might have answered without reading questions. Thus, it is expected that unserious answers such as random zigzag pattern can be detected by conflict of response. Moreover, it is preferable to use plural paired questions to make possibility to detect these answers high and to acquire various values not binary one, conflict or not.

2.2. Respondent's interest

Respondent's interest [8] is derived from the average of variances of evaluation values for each evaluation object. In the rating scale method, the questions usually consist of both positive meaning and negative one, and then it can be assumed that the variance of evaluation values becomes large by being used various grades of evaluation values when a respondent has interest for the evaluated objects because he/she deeply thinks about them and has the impression from various viewpoints. If the number of evaluation objects and that of questions are defined as n and m, the evaluated value of question j of object i is defined as y_{ij}, and the average evaluated value of questions of object i is defined as $\bar{y}_i$, I can be expressed as follows:

$$I = \frac{1}{n}\frac{1}{m-1}\sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij} - \bar{y}_i)^2. \quad (2)$$

When I is large, the respondent has a possibility that he/she answered questionnaire earnestly with interest and vice versa. It is expected that answers without interest which shows same evaluation value can be detected by respondent's interest.

2.3. Weighted Multi Dimensional Scaling method

2.3.1. Multi Dimensional Scaling Method

Multi Dimensional Scaling method (MDS) [9][10] can visualize relationships among data in high-dimensional space[11] such as the evaluation values of respondents. The distance between evaluation data of respondents i and j in the original space and that in the visualized space are defined as d_{ij} and d_{ij}^*, respectively, and the sum of the error value E for the respondents is defined as follows:

$$E = \sum_{i=1}^{l}\sum_{j=1,j\neq i}^{l}(d_{ij} - d_{ij}^*)^2, \quad (3)$$

where l denotes the number of respondents. MDS is to minimize E by iterative-computing. Respondents who are near in the original space are plotted near in the visualized space, and vice versa.

2.3.2. Weighted MDS considering Conflict and Interest

In the normal MDS, the error value of each respondent is treated equally. However, if the conflicted answers or uninterested answers are mixed in the questionnaire data, it is thought that the analysis result of the data is not accurate and shows different characteristics or trends. This paper proposes a visualization method that can emphasize relationships between respondents who answered earnestly by using the conflict of response and respondent's interest shown in 2.1 and 2.2. In the proposed method, the membership functions, i.e. weights of a respondent's earnestness quantified by C and I are configured as shown in Fig. 1. The membership value of conflict of response for respondent i is defined as μ_{ci} and that of respondent's interest is defined as μ_{ii}, the earnestness k_i of respondent i is defined as follows:

$$k_i = \mu_{ci} \cdot \mu_{ii}. \tag{4}$$

By using k_i, the error value of weighted MDS, E', is defined as follows:

$$E' = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} k_i \cdot k_j (d_{ij} - d_{ij}^*). \tag{5}$$

In the past, methods involved weighted MDS like INDSCAL[12] or ALSCAL[13] have been reported. These methods consider individual difference by weighting distance in the original space. In the proposed method, respondents are visualized by weighting on error value by k_i. In addition, the proposed method focuses on the earnestness of respondents, and the weighted MDS itself is used as a tool for visualization of evaluated questionnaire data of them.

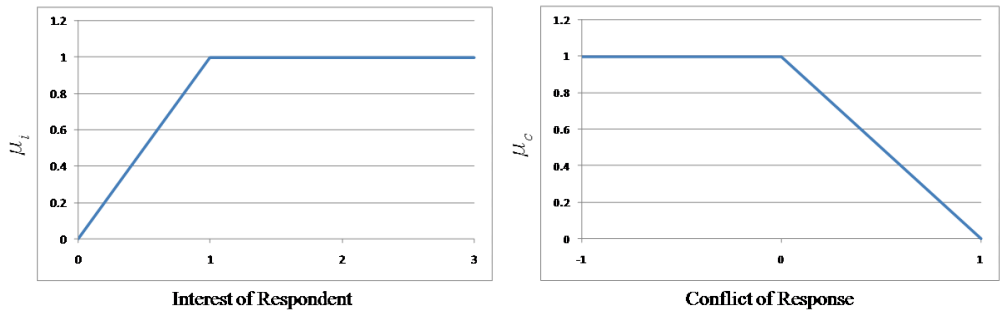


Figure 1: Membership function

2.4. Interactive search

Using the visualization map in 2.3, respondents can be clustered into some groups. Visualization makes it easier to grasp the characteristics of data with the distribution of them and interactive search supports to discover characteristics or tendencies of data and then to

get useful knowledge for marketing by grouping respondents with trial and error interactively. When respondents are visualized by using weighted MDS, respondents who answered earnestly are mapped accurate, i.e. keeping the distance with others, but respondents who did not answer earnestly have large error value. In the usual calculating, the averages of evaluated value of respondents in the groups including these not-earnest answers would be inaccurate. Thus, in this paper, the average of evaluated values in each group by interactive search is calculated considering respondent's earnestness as follows:

$$A_{Cab} = \frac{\sum_{i \in C} k_i \cdot x_{iab}}{\sum_{i \in C} k_i}, \tag{6}$$

where a, b, and k_i denote objects, questions, and the earnestness of respondent i ,respectively. And the average evaluated value of a group C is defined as A_{Cab}, the evaluated value of respondent i of question b and object a is defined as x_{iab} .

3. EXPERIMENT AND DISCUSSION

3.1. Experimental setting

This experiment involved 1,448 respondents and 6 scenes using an outdoor product α as the evaluation subjects. The experiment employed the rating scale method and the respondents were asked to choose one of five grades 1,2,3,4,5 in response to each of 10 questions. In this survey, grade 5 means "applicable" while grade 1 means "not applicable." Table 1 shows the 6 scenes (presented by moving image during the questionnaire) used as the evaluation objects and Table 2 shows the 10 questions. In addition, respondents were asked to express "why do/don't you want this product" in free-text form. In Table 2, Q3-Q8 and Q9-Q10 were the sets as pairs of questions which have counter meanings shown in 2.1.

Table 1: Evaluation objects

Scene	Contents
Scene1	Watching movie using a projector
Scene 2	Using coffee maker and refrigerator
Scene 3	Using PC and blog
Scene 4	Using shower and dryer
Scene 5	Boating with electric thruster
Scene 6	Making Pure water from river water

3.2. Verifying earnestness of respondents

First, it was verified whether pairs of questions described above actually had counter meanings for respondents by using the distance between questions [7]. Secondly,

respondents with small interest, *I*, were also verified by comparing the number of characters in the free-form text.

3.2.1. Distance between questions

The distance between questions is defined as a distance between the evaluated value of questions in the original space generated by the axes of evaluation objects. The nearer questions are, the more similar meaning the respondents considered. If the evaluated value of questions i and j of evaluation object k for respondent l ($l \in 1...N_r$, N_r: number of respondents) are defined as y_{ki} and y_{kj}, then the distance between questions of the respondent, d_{lij}, is defined as follows:

Table 2: Questions

Question	Contents
Question 1	It would make me feel superior to people around me.
Question 2	It looks as if things could be done cleanly even outside.
Question 3	I wanted to do this kind of things outdoors.
Question 4	It looks as if it would be a big help in a difficult situation, such as a natural disaster.
Question 5	It looks difficult to carry around.
Question 6	It would be a lot of trouble to assemble and set up.
Question 7	My friends and family would like it.
Question 8	I don't think I want to do this kind of thing outdoors.
Question 9	This could be done even without the α.
Question 10	It would enhance outdoor leisure activities.

$$d_{lij} = \sqrt{\sum_{k=1}^{n} (y_{ki} - y_{kj})^2}, \qquad (7)$$

where n denotes the number of evaluation objects. If the number of questions is defined as m, each respondent has the vector of distances ∂_l as follows:

$$\partial_l = (d_{l12},...,d_{lij},...,d_{l(m-1)m}) \quad (i \neq j) \qquad (8)$$

Then, the vector of all respondents $\boldsymbol{D}$ is defined as follows:

$$\boldsymbol{D} = (\partial_1, \partial_2, \ldots, \partial_{Nr})^T \qquad (9)$$

Relationships among questions based on $\boldsymbol{D}$ are visualized as a dendrogram shown in Fig. 2. In Fig. 2, the evaluated values of Q5 "It looks difficult to carry around," Q6 "It would be a lot of trouble to assemble and set up," Q8 "I don't think I want to do this kind of thing outdoors," Q9 "This could be done even without the α ." were reversed, "1" was exchanged with "5" and "2" with "4," so that evaluated value "5" had good meanings for all questions, which made the analyses easier to understand the impressions based on the average of evaluated values. As shown in Fig. 2, Q3 "I wanted to do this kind of things outdoors." and the reversed Q8 were near, so it can be thought that these questions had counter meaning each other. On the other hand, Q9 and Q10 "It would enhance outdoor leisure activities." were not so near; thus, it would appear that they did not work as a pair of question which have counter meanings for the respondents. Therefore, only the pair of Q3-Q8 was used to derive the conflict of response shown in 2.1 in the proposed method.

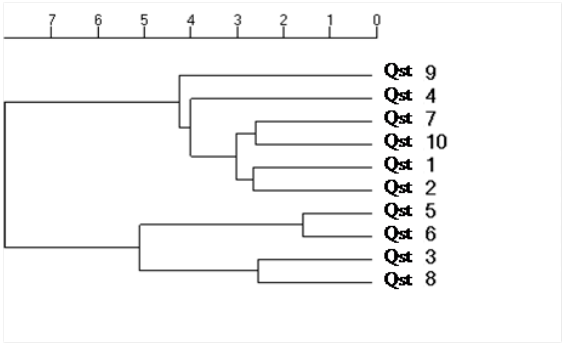


Figure 2: Dendrogram based on distance between questions

3.2.2. Comparison of the number of characters in free text

To compare the respondents who had large value of respondent's interest, *I*, with those who had small values, they were classified based on whether his/her value of interest was greater than 1.0. The respondents with high value of respondent's interest were defined as group 1, and who with low value were defined as group 2. The number of respondents of group 1 was 726 and the average number of characters in free text was 35.0. On the other hand, the number of respondent of group 2 was 722 and the average number of characters was 21.6. And it was confirmed that there was a significant difference between them at significance level of 1%.

3.3. Analysis result of conventional method

1,448 respondents' data were visualized using normal MDS shown in 2.3.1. Fig. 3 shows the visualization result. In Fig. 3, each dot represents each respondent's evaluation data and respondents plotted near have similar tendency of evaluated value. In this visualized map, two groups (group A and group B) were clustered as shown in Fig. 3.

The average of evaluated values of each group are shown in Fig. 4. In Fig. 4, the evaluated values of Q5, Q6, Q8, Q9 were reversed from the original values as described above. In Fig. 4(a), the values of Q2, Q4 and Q7 were relatively high, and those of Q5 and Q6 were low, which might lead a certain analysis. However, the values of Q3 and the reversed Q8 were both low and they were conflicted according to 3.2.1. Thus this result seemed to lack credibility. In addition, if the evaluated values of Q5, Q6, Q8, Q9 were not reversed, almost

all evaluated values of each question were centered around the value of 3 or 4. So it can be assumed that respondents who answered without interest or evaluated with evenly high values to almost all questions because they wanted to pretend favorably to analyzer, were mixed in this group. On the other hand, in group B, they highly evaluated Q5 and Q6 while they gave low evaluation to other questions, so it can be thought that they did not feel it was difficult to carry or to set up, though they did not get good impression to the outdoor product α. And if the evaluated values of Q5, Q6, Q8, Q9 were not reversed, almost all evaluated values of each question were centered around the value of 2 or 3.

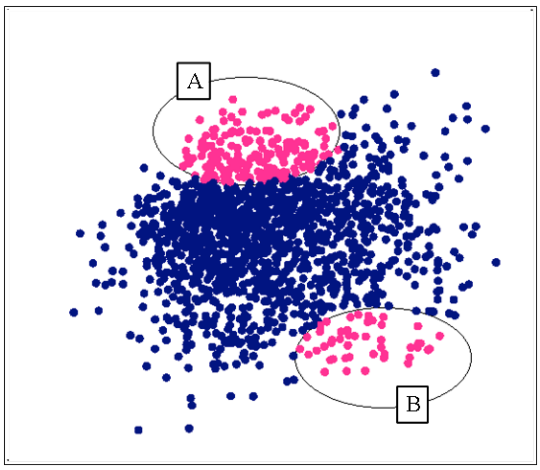
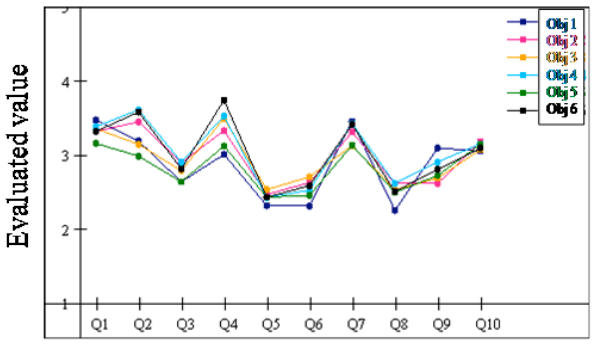
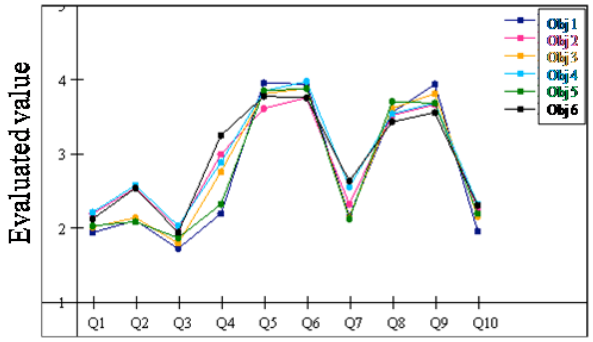


Figure 3: Visualization result (normal MDS)



(a) Group A



(b) Group B

Figure 4: Average of evaluated values

3.4. Analysis result of proposed method

The result of visualization using the weighted MDS explained in 2.3.2 is shown in Fig. 5. Here, group A' and group B', which show similar belongings with group A and group B in Fig. 3, respectively, were clustered in Fig. 5. The average of evaluated values of them are shown in Fig. 6. The tendencies of the evaluated values in group A' and group B' were different from group A and group B in Fig. 3 and they made the characteristics easier to be grasped. As shown in Fig. 6(a), the values of Q4, Q5 and Q6 were higher than those of group A. On the other hand, the values of Q3 and Q8 were both lower. So it can be assumed that they did not consider the product α as an "outdoor" product, but they thought it was easy to carry and to setup and that it would be a big help in a difficult situation such as a natural disaster. The result gives a hint of marketing strategy to sell α as a disaster supply. In Fig. 6(b), the average values were similar to those of group B, but the values of Q4 were quite high except for scene 1 and scene 5. Thus the respondents in this group also highly evaluated this product α as an emergency supply.

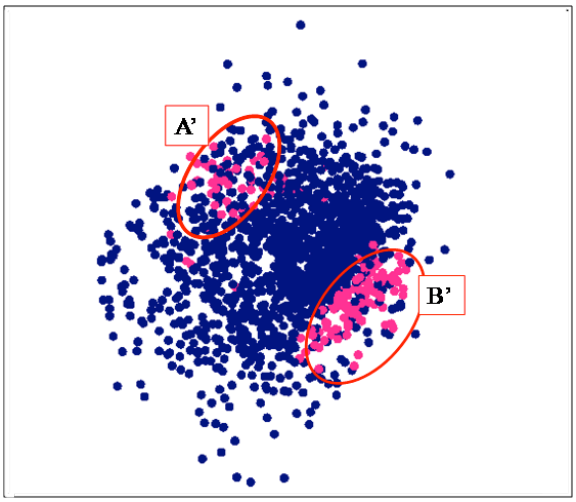
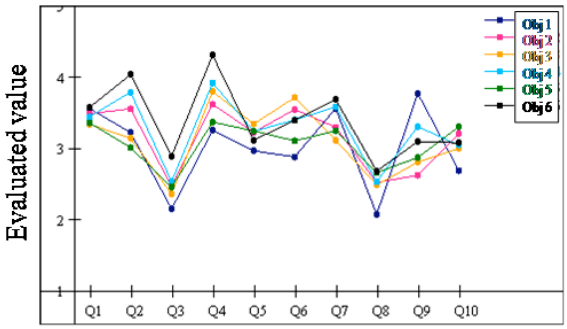
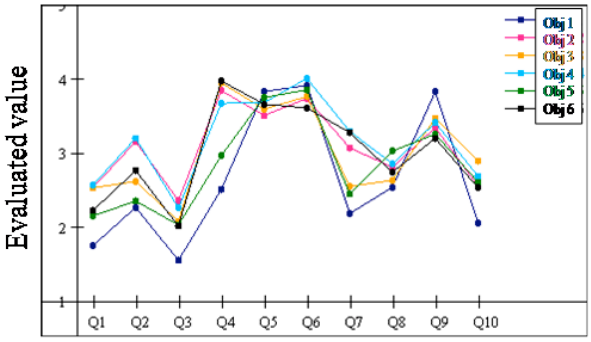


Figure 5: Visualization result (weighted MDS)



(a) Group A'



(b) Group B'

Figure 6: Average of evaluated values

4. CONCLUSION

This paper defined conflict of response and respondent's interest as a barometer to evaluate whether a respondent was earnest in answer, and proposed a visualization method that used these quantified earnestness. Applying the proposed method to a questionnaire on an outdoor product, firstly these values were shown to work appropriately. And comparing conventional method and the proposed method, it was indicated that the proposed method could avoid getting different result because of uninterested or conflicted answer. As a future work, the proposed method will be applied to other questionnaires and the designing of appropriate membership functions will be investigated.

REFERENCES

[1] Yamamoto, K., Yoshikawa, T., and Furuhashi, T.: Visualization of Similarity between Subjects' Meanings of Impression Words, Journal of Japan Society of Kansei Engineering (in Japanese), Vol.6, No2, pp.59-65, 2006

[2] Maeda, S., Futatsuka, M., Yonesaki, J., and Ikeda, M.: Relationship between questionnaire survey results of vibration complaints of wheelchair users and vibration transmissibility of manual wheelchair , Environmental Health and Preventive Medicine, Springer Japan, Vol.8, No.3, p82-89, 2003

[3] Osgood, C.E., Suck, G.J., and Tannenbaum P.H.: The Measurement of Meaning, University of Illinois Press, 1957

[4] Fukami, T., Yoshikawa, T., Furuhashi, T., Hara, I., and Mochizuki, T.: Analysis of Questionnaire Data based on Individuality of Frequency of Using Evaluation Values, Proc. of International Conference on Kansei Engineering and Emotion Research 2007 (KEER2007), L-14 (6p), 2007

[5] Takane, Y.: Ningenkagaku no Toukeigaku 1 -Tajigensyakudokouseihou- (in Japanese), Asakura Publising Co.Ltd, 1980

[6] Fukami, T., Watanabe, Y., Yoshikawa, T., Furuhashi, T., Hara, I., and Yoneda, H.: Discovering Minority Groups by Interactive Clustering in Visible Space, Proc. of International Conference on Kansei Engineering and Emotion Research 2009 (KEER2009), 11G-04 (9p), 2009

[7] Yamaga, S., Yoshikawa, T., Furuhashi, T.: Study on Difference of Respondents' Interpretation by Setting of Questions in Questionnaire (in Japanese), The 11th Annual Conference of JKSE, 3F3-2 (6p), 2009

[8] Tatematsu, H., Yoshikawa, T., Furuhashi, T., Iguchi, H., Hirao, E.: Analysis of Questionnaire Data based on Interest Degree of Objects, Proc. of The 6th International Symposium on Advanced Technology (ISAT..6th), pp.58-62, 2007

[9] Torgerson, W. S.: Theory and Methods of Scaling. New York, Wiley and Son, 1958

[10] Saito, T., Yadohisa, H.: "Kanrensei de-ta no kaisekihou" (in Japanese), KYORITSU SHUPPAN CO., LTD, 2007

[11] Kagie, M., Wezel, M., and Groenen, P, J.F.: Online shopping Using a Two Dimensional Product Map, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Vol. 4655, p89-98, 2007

[12] Carroll, D. J., Chang, J. J.: Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition, Psychometrika, 35, pp.283-319, 1970

[13] Takane, Y., Young, F.W., de Leeuw, J.: Nonmetric individual differences multidimensionalscaling: An alternative least squares method with optimal scaling features, Psychometrika, 42, pp.1-68, 1977