Journal of the Korean
Statistical Society
Vol. 25, No. 3, 1996

A Graphical Approach to Paired Rankings †

Sang-Tae Han [1] and Myung-Hoe Huh [2]

Abstract

Paired rankings data comes to us in two situations. One situation is when pairs of subjects, say husbands and wives, are asked to rank a group of objects. Another situation is when subjects are asked to rank a group of same objects at two time points, say, before and after the treatment. In this study, we show how biplot techniques can be applied to represent graphically such paired rankings.

Key Words : Simultaneously paired rankings; Sequentially paired rankings; Quantification; Biplot.

1. INTRODUCTION

Ranked data are obtained by asking subjects(or judges) to rank a set of objects(or items). For example, potential buyers order different brands of personal computers in terms of which one he likes best, second best, and so on. Statistical analysis of such data were dealt in Mallows(1957), , Critchlow(1985), Baba(1986), and Diaconis(1988) among others.

---

† This paper is partially supported by Institute of Statistics, Korea University.
[1] Institute of Statistics, Korea University, Seoul 136-701, Korea.
[2] Department of Statistics, Korea University, Seoul 136-701, Korea.

Recently, Critchlow and Verducci(1992) introduced methods for analyzing paired rankings data. Their paired rankings are generated when subjects in a study independently rank a set of objects before and after the treatment. Such sequentially paired rankings need conditional approach. Thus, they used the conditional distribution of subject's post-treatment rankings given their pre-treatment rankings, under the hypothesis that each subject's post-ranking is symmetrically distributed about its pre-ranking.

Another situation is when a pair of rankings arrives simultaneously, for example, when couples of a husband and a wife are interviewed for their opinions on a number of objects in a form of ranking, it might not be natural to condition on one part of paired rankings or on the other part of paired rankings. Such simultaneously paired rankings need symmetric approach. Thus, Huh et al.(1995) have considered a symmetric extension of paired rankings. Their method analyzing paired rankings is based on probability ranking models, such as a class of Mallows' models.

In this study, we investigate the three problems as follows. First, we will classify paired rankings data into two categories. Second, we will propose biplots for both sequentially and simultaneously paired rankings data, based on quantification. The biplot, originated from Gabriel(1971), is such a nice graphical tool that subjects(observations) as well as objects(variables) can be plotted to show off mutual relationships. Third, we will also propose procedures for detecting the trend in sequentially paired rankings and the within-pair difference in simultaneously paired rankings. We have used "data analytic approach" based on description of multivariate data without any specific probability assumptions. The method we proposed is entirely different point of view from the tradition of modeling approach based on probability ranking models, such as Mallows' models.

2. QUANTIFICATION OF SIMULTANEOUSLY PAIRED RANKINGS

Suppose that n pairs of subjects are asked to rank a group of p objects. Let $r_{ij}^{[1]}$ be the rank given to the object $j(=1,\cdots,p)$ by the first subject of the $i(=1,\cdots,n)$-th pair, and $r_{ij}^{[2]}$ be the rank given by the second subject. As a row centering process, let

$$s_{ij}^{[1]} = r_{ij}^{[1]} - (p+1)/2, \quad s_{ij}^{[2]} = r_{ij}^{[2]} - (p+1)/2,$$

and write

$$S^{[1]} = (s_{ij}^{[1]}), \quad S^{[2]} = (s_{ij}^{[2]}), \quad S = \begin{pmatrix} S^{[1]} \\ S^{[2]} \end{pmatrix}.$$

In Spearman's sense, the squared rank distance between two rows i and $i'(i, i' = 1, \cdots, n, i \neq i')$ is defined by

$$d_S^2(i, i') = \sum_{j=1}^{p} (s_{ij} - s_{i'j})^2.$$

We can write the Spearman rank correlation between two rows i and i' as

$$1 - 6 \ \ d_S^2(i, i')/p(p^2 - 1).$$

Then, one type of biplot can be obtained from singular value decomposition(SVD) of $2n \times p$ matrix S(Lebart et al., 1984). Let

$$S = UD_\lambda V',$$

where U is a $2n \times p$ matrix with orthonormal columns, $V = (v_1, v_2, \cdots, v_p)$ is a $p \times p$ orthonormal matrix, and D_λ is a $p \times p$ diagonal matrix with singular values $\lambda_1 \geq \cdots \geq \lambda_{p-1} > (\lambda_p = 0$, since $\text{rank}(S) = p - 1)$ as its diagonal elements. Then two-dimensional row plot points are given by the rows of

$$G^* = S(v_1 : v_2) = UD_\lambda V'(v_1 : v_2) = U_{(2)} D_{\lambda(2)},$$

where $U_{(2)}$ is an $2n \times 2$ submatrix of U and $D_{\lambda(2)}$ is a 2×2 diagonal submatrix of D_λ.

One particular merit in the row plot is that Spearman's distances between rankings given by subjects(rows of S) are approximately preserved, since

$$SS' = (UD_\lambda V')(UD_\lambda V')' = UD_\lambda^2 U' \cong \text{G}^*\text{G}^{*'}.$$

Therefore, the goodness-of-approximation of the two-dimensional row plot is given by

$$GOA_{row(2)} = 1 - ||SV - (SV_{(2)} : O_{2n \times (p-2)})||^2 / ||SV||^2$$

$$= (\lambda_1^2 + \lambda_2^2)/(\lambda_1^2 + \cdots + \lambda_p^2).$$

where $V_{(2)} = (v_1, v_2)$ is the $n \times 2$ submatrix of V.

For the plot of columns(or objects), we consider a group of hypothetical supplementary rows(or subjects), rows of I_p, and use the same PC axis vector

v. Then, two-dimensional column plot points for objects are given by the rows of

$$H^* = V_{(2)}.$$

For further details of this procedure, see Han and Huh(1995). The goodness-of-approximation of the two-dimensional column plot is given by

$$GOA_{col(2)} = 1 - ||V_{(p-1)} - (H^* : O_{p\times(p-3)})||^2/||V_{(p-1)}||^2$$
$$= 1 - (p-3)/(p-1) = 2/(p-1),$$

where $V_{(p-1)}$ is the $p \times (p-1)$ submatrix of V.

The biplot, produced by combining row and column plots, can be interpretable as follows. Note that S can be expressed as

$$S = (UD_\lambda)V' \cong (U_{(2)}D_{\lambda(2)})V'_{(2)} = G^*H^{*\prime}.$$

Thus the raw data elements are recovered by the inner products of row and column plot vectors.

Next, we may ask a question whether there exists a systematic within-pair difference. To answer this question, we use the plot for the differences between first group rankings and corresponding second group rankings. Partition G^* into two equal-size $n \times 2$ matrices, $G^{[1]*}$ and $G^{[2]*}$, then two-dimensional row-differences plot in which plotting points are given by the rows of

$$G^{[2]*} - G^{[1]*} = \{g_i^{[2]*} - g_i^{[1]*}, \quad i = 1, \cdots, n\},$$

where $g_i^{[2]*}$ is the row plot points for the second subject of the i-th pair and $g_i^{[1]*}$ is the row plot points for the first subject of the i-th pair. An example will be given in Section 4. We may try, another type of quantification by changing the sequence of plotting and differencing. See Chapter 3 of Han's dissertation(1995) for details.

3. QUANTIFICATION OF SEQUENTIALLY PAIRED RANKINGS

Consider a situation in which subjects are asked to rank a group of same objects at two time points, say, before and after the treatment. An example, literary criticism data(Critchlow and Verducci, 1992), which we shall analyze in detail, came from a study where $n = 38$ secondary school children rank

$p = 4$ styles of criticism, 'authorial', 'comparative', 'personal' and 'textual', as applied to a certain story. The students then undergo a course in writing and literary criticism and afterwards rank the same four styles of criticism applied to a similar story. The question of interest includes whether or not the post-treatment rankings have moved in the direction of a particular idealized ranking α, e.g. the teacher's own ranking of the four styles.

Suppose that two sets of rankings are sequentially paired, yielding $r_{ij}^{[1]}$ as the first measurement rankings(pre-rankings) and $r_{ij}^{[2]}$ as the second(post-rankings) given to the object $j(=1,\cdots,p)$ by the subject $i(=1,\cdots,n)$.

Consider singular value decomposition(SVD) of $n \times p$ matrix $S^{[1]}$. Specifically, write

$$S^{[1]} = U^{[1]} D_{\lambda}^{[1]} V^{[1]\prime},$$

where $U^{[1]}$ is a $n \times p$ matrix with orthonormal columns, $V^{[1]} = (v_1^{[1]}, v_2^{[1]}, \cdots, v_p^{[1]})$ is a $p \times p$ orthonormal matrix, and $D_{\lambda}^{[1]}$ is a $p \times p$ diagonal matrix with singular values, $\lambda_1 \geq \cdots \geq \lambda_{p-1} > \lambda_p(=0)$ as its diagonal elements. Then, for the first set of rankings, two-dimensional row plot points are given by the rows of

$$G^{[1]*} = S^{[1]}(v_1^{[1]} : v_2^{[1]}) = U^{[1]} D_{\lambda}^{(1)} V^{[1]\prime}(v_1^{[1]} : v_2^{[1]})$$

$$= U_{(2)}^{[1]} D_{\lambda(2)}^{[1]} = (r_1^{[1]} : r_2^{[1]}),$$

where $U_{(2)}^{[1]}$ is a $n \times 2$ submatrix of $U^{[1]}$, $D_{\lambda(2)}^{[1]}$ is a 2×2 diagonal submatrix of $D_{\lambda}^{(1)}$ and $(r_1^{[1]} : r_2^{[1]})$ are first two PC score vectors.

For the second set of rankings, we treat them as a group of hypothetical supplementary rows(or subjects), and use the same PC axis vector v. Then, for the second set of rankings, two-dimensional row plot points are given by the rows of

$$G^{[2]*} = S^{[2]} V_{(2)}^{[1]} = S^{[2]}(v_1^{[1]} : v_2^{[1]}),$$

where $V_{(2)}^{[1]}$ is a $n \times 2$ submatrix of $V^{[1]}$. Then, write $G^{[2]*} = (r_1^{[2]} : r_2^{[2]})$. And, two-dimensional column plot points are positioned at

$$H^{**} = V_{(2)}^{[1]} = (v_1^{[1]} : v_2^{[1]}).$$

We also can define the two dimensional goodness-of approximation for the pre-ranking row plot as

$$GOA_{row(2)}^{[1]} = 1 - ||S^{[1]} V^{[1]} - (G^{[1]*} : O_{n \times (p-2)})||^2 / ||S^{[1]} V^{[1]}||^2$$

$$= (\lambda_1^2 + \lambda_2^2)/(\lambda_1^2 + \cdots + \lambda_p^2).$$

Similarly, for the post-ranking row plot, it is defined as

$$GOA^{[2]}_{row(2)} = 1 - ||S^{[2]}V^{[1]} - (G^{[2]*} : O_{n\times(p-2)})||^2/||S^{[2]}V^{[1]}||^2.$$

For the column plot, we may define as follows

$$GOA_{col(2)} = 1 - ||V^{[1]}_{(p-1)} - (H^{**} : O_{p\times(p-3)})||^2/||V^{[1]}_{(p-1)}||^2 = 2/(p-1),$$

where $V^{[1]}_{(p-1)}$ is the $p \times (p-1)$ submatrix of $V^{[1]}$.

One particular merit in row plot is that Spearman's rank distance between rankings of the first set are approximately preserved, while the second set of rankings are positioned in perceptional space derived by the first set of rankings.

By superimposing the pre- and post-ranking row plots, we may check whether there exists a trend from pre- to post-rankings. But, these plots look very complex for moderately large n and it seems to be difficult to extract any meaningful informations. To make things easy, we propose the following procedure for trend analysis:

First, plot the first principal component score vectors(= $(r_1^{[1]} : r_1^{[2]})$) of the pre- and post-rankings. We call it "the trend plot" on first principal component axis.

Second, plot the second principal component score vectors(= $(r_2^{[1]} : r_2^{[2]})$) of the pre- and post-rankings. We call it "the trend plot" on second principal component axis.

When, "ideal ranking" exists, we may apply the above procedures for trend analysis by plotting the principal scores obtained by quantification centered at the plot point of ideal ranking. Then, evaluate how close post-ranking is to the idealized ranking.

4. NUMERICAL ILLUSTRATIONS

4.1 Stress Patterns of Newlyweds

We will illustrate the above simultaneously paired rankings procedure using the data set given in Huh et al.(1995). Here, we analyze blood concentration of the hormonal prolactin, which is known to be positively correlated with the amount of stress that an individual is experiencing. The data, collected by Dr. Malarkey and his colleagues at Ohio State University Hospital, is described by Huh et al.(1995) as follows:

> Blood samples were taken from pairs of wife and husband during a 24 hour interval. Of particular interest are the relative levels of prolactin at five periods of time during this day: baseline (early morning), conflict session (during the hour after baseline), intermission (15 minutes after the end of the conflict session), post-session (one hour after the intermission), and the ensuing night. The conflict session centered around a wife-husband negotiation on an issue known to be especially stressful for that particular couple. For each individual, the five time periods are ranked according to the concentration level of prolactin in that person's blood sample. The time periods are listed in their temporal order: B = Baseline, C = Conflict, I = Intermission, P = Post-session, and N = Night. Each person's ranking is represented by an ordered 5-tuple that lists the relative ranks of the five time periods, ranked according to increasing order of prolactin concentration.

The row plots are shown in Figure 1 with the goodness-of-approximation 78%. Thus almost eighty percent of the total variation among the subjects is explained by the first two principal component axes. Also, the column plot is shown in Figure 2 with the goodness-of-approximation 50%. By superimposing the row and column plots, we obtain biplot of simultaneously paired rankings. In Figure 1, the largest one of both wives' and husbands' cluster are found in the right-side region along the first axis, of which corresponding position in Figure 2 is assigned to put N as "high prolactin concentration". So, we may interpret that prolactin concentration peaks during the night for both wives and husbands.

Next, row-differences plot is shown in Figure 3. By superimposing the row-differences plot (Figure 3) and column plot (Figure 2), we see that No.7, No.8, No.16, No.21, No.26, No.28, and No.31 husbands have relatively higher prolactin concentration during "P" compared to their wives, while No.30 husband has relatively higher prolactin concentration during "I". On the other hand, No.12, No.14, No.15, No.18, No.22, No.32, and No.33 husbands have higher prolactin concentration during "B". Finally, No.1, No.2, and No.9 husbands have relatively higher prolactin concentration during "N". However, in Figure 3, we observe that the existence of systematic within-pair difference is not apparent since the row-differences plot points are distributed evenly around the origin. This result is quite similar to the significance testing result obtained by Huh et al.(1995).

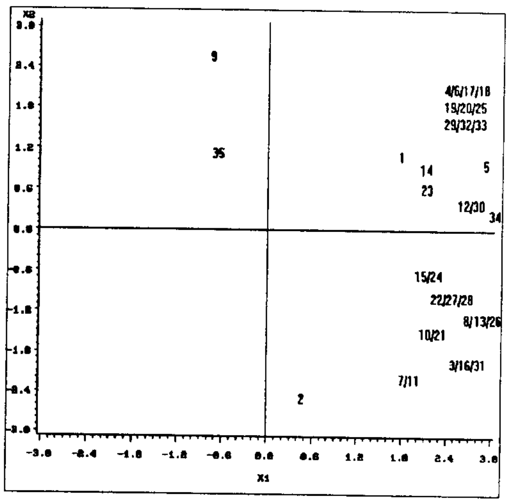

Figure 1. Row plots for newlywed couples data ($GOA_{row(2)} = 78\%$)
(a) Wives' ranking (b) Husbands' ranking

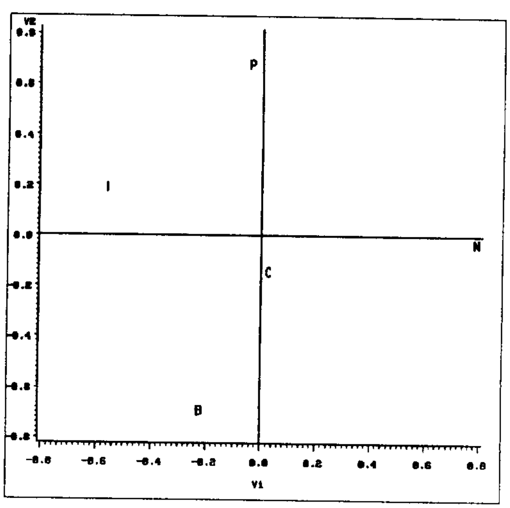

Figure 2. Column plot for newlywed couples data ($GOA_{col(2)} = 50\%$)

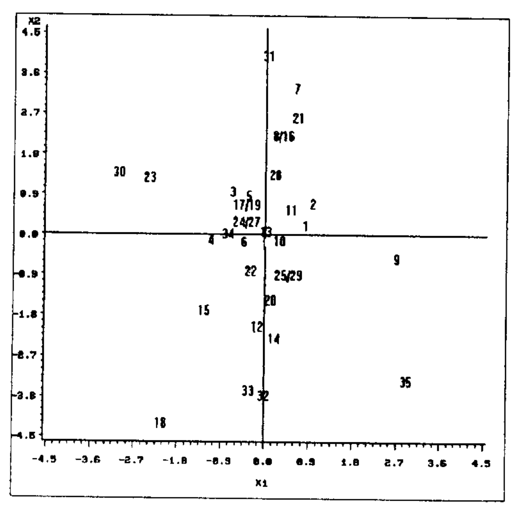

Figure 3. Row-differences plot for newlywed couples data ($GOA_{row(2)} = 78\%$)

4.2 Pre-and Post-course Rankings of Four Critical Styles

We will illustrate the sequentially paired rankings procedure using the data set given in Critchlow and Verducci(1992). The data was collected by Rogers(1990) under the following contexts:

> Thirty-eight ninth grade students are participants in a course in writing and literary criticism. At the beginning of the course, they read the story "A rose for Emily" by William Faulkner, and are asked to rank four critical paragraphs according to their personal preferences. The selected paragraphs typify four distinct critical styles: A = Authorial, C = Comparative, P = Personal, and T = Textual. At the end of the course, the students read another story by William Faulkner, "That Evening Sun," and again rank four critical paragraphs written in the same four styles.

One special feature in this data is that there exists the ideal ranking α which is also called the attractor in Critchlow and Verducci(1992). The ideal ranking is specified as α =(P, C, A, T), where object P is ranked as the best, object C is ranked as the second best, object A is ranked as the third best, and object T is ranked as the worst.

Pre-and post-rankings row plots(including ideal ranking α) are shown in Figure 4 and Figure 5 with the goodness-of-approximation 75.8% and 59.8%, respectively. Also, column plot is shown in Figure 6 with the goodness-of-approximation 66.7%. By superimposing the row and column plots, we obtain biplot of sequentially paired rankings. Next, we evaluate how close his post-ranking is to the idealized ranking through the trend plots. Trend plots on principal component axes are shown in Figure 7 and Figure 8 respectively. In this particular case, for easy perception of the trend, the principal component scores obtained by quantification were subtracted from the plot point of ideal ranking α. In Figure 7 and Figure 8, Region 1 is the improved area towards the ideal ranking and Region 2 is the area for the disimprovement. In this particular example, about half of the subjects have trend towards the ideal ranking, while the remainders have trend with the opposite direction. Specifically, in Figure 7 of the trend in the first PC axis, No.8, No.16, No.24, and No.26 students made significant improvements towards the ideal ranking, while No.7, No.10, and No.14 students got worse. Also, in Figure 8 of the trend in the second PC axis, we see that No.23, and No.19 students made significant improvements towards the ideal ranking, while No.12, No.14, and No.6 students got worse a lot. (This result is not fully consistent with the significance testing result obtained by Critchlow and Verducci(1992). But, Thompson(1993) reached a similar conclusion to ours by plotting the frequencies of the pre- and post-ranking on a truncated octahedron).

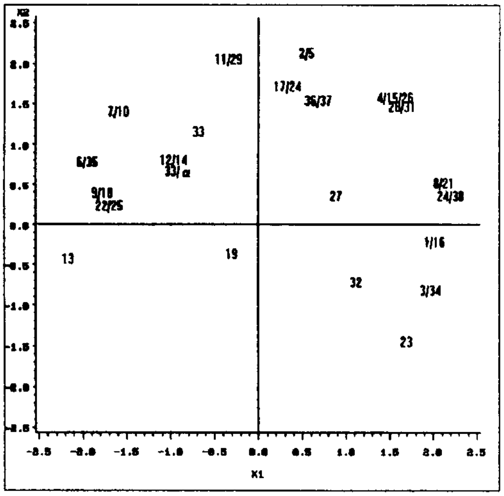

Figure 4. Row plot for pre-rankings of literary critism data ($GOA^{(1)}_{row(2)} = 75.8\%$)

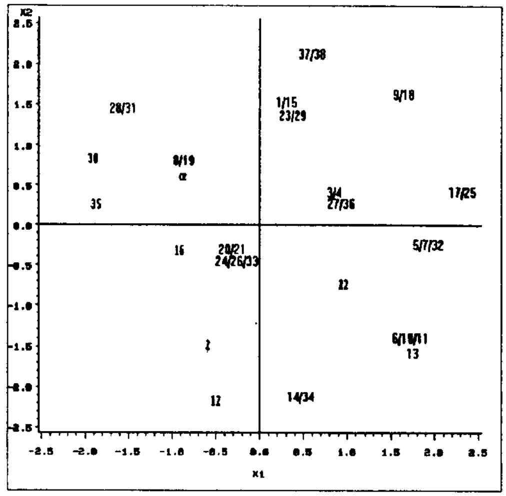

Figure 5. Row plot for post-rankings of of literary critism data ($GOA^{(2)}_{row(2)} = 59.8\%$)

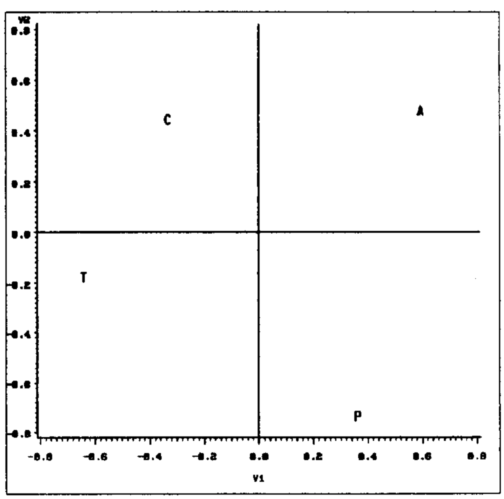

Figure 6. Column plot for literary criticism data ($GOA_{col(2)} = 66.7\%$)

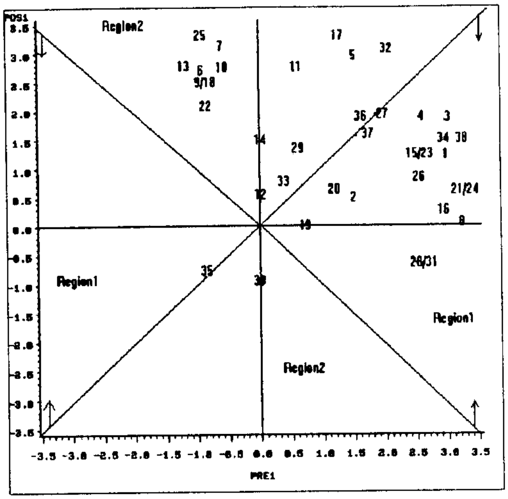

Figure 7. Trend plot for literary criticism data on the first principal component axis

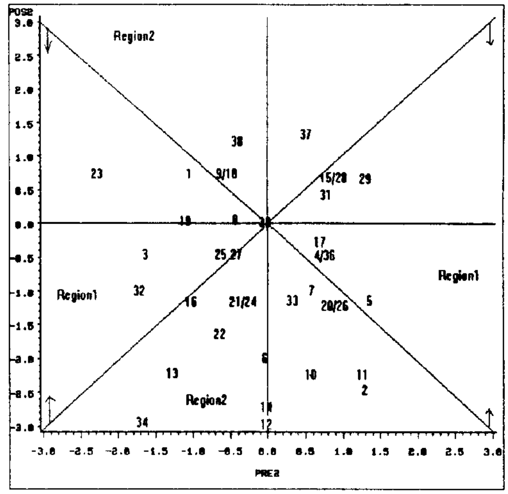

Figure 8. Trend plot for literary criticism data on the second principal component axis

5. CONCLUDING REMARK

In this study, we derived biplots for two types of paired rankings by quantifying objects and judges. Specifically, we proposed a row plot, in which the interpoint distances between plot points can be interpreted as an approximation of the (squared) rank distance between rankings given by corresponding subjects in Spearman's sense. Similarly, we also proposed a column plot for objects having a sensible relationship to the row plot for subjects. Specially, these graphs give us good information for the clustering of judges. Using biplots for paired rankings, we are able to visually detect the within-pair difference and the trend between two sets of rankings . And we also find clusters of subjects showing similar patterns of statistical change. Also, in each biplot, we define and compute "goodness-of-approximation" measures to guard ourselves against over-interpretation. We would like to add one remark.

The same idea in Section 2 and Section 3, can be also applied to another type coding scheme(Kendall-type coding scheme) for paired rankings data in a similar way. The Kendall-type coding scheme that is derived directly from pairwise comparison of objects is convenient for computing rank distances in Kendall's sense. Then, we may produce biplots in which Kendall's rank distances instead of Spearman's are preserved approximately. For more details, see Han(1995).

REFERENCES

(1) Baba, Y. (1986). Graphical analysis of rank data. *Behaviormetrika*, **19**, 1-15.

(2) Critchlow, D. E. (1985). *Metric Methods for Analyzing Partially Ranked Data. Lecture Notes in Statistics*, No. **34**. Springer, New York.

(3) Critchlow, D. E. and Verducci, J. S. (1992). Detecting a trend in paired rankings. *Journal of the Royal Statistical Society, Series* C **41**, 17-29.

(4) Diaconis, P. (1988). *Group Representations in Probability and Statistics.* Lecture Notes-Monograph Series, Vol. **11**. Institute of Mathematical Statistics. CA: Hayward.

(5) Gabriel, K. R. (1971). The biplot graphics display of matrices with applications to principal component analysis. *Biometrika*, **58**, 453-467.

(6) Han, S. T. (1995). *Quantification Approach to Ranked Data Analysis.* Unpublished Ph. D. Dissertation, Dept. of Statistics, Korea University.

(7) Han, S. T. and Huh, M. H. (1995). Biplot of ranked data. *Journal of the Korean Statistical Society*, **24**, 439-451.

(8) Huh, M. H., Critchlow, D.E., Verducci, J. S., Kiecolt-Glaser, J., Glaser, R. and Malarkey, W. B.(1995). A symmetric analysis of paired rankings with application to temporal patterns of hormonal concentration. *Biometrics*, **51**, 1361-1371.

(9) Lebart, L., Morineau, A. and Warwick, K. (1984). *Multivariate Descript- ive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices.* Wiley, new York.

(10) Mallows, C. L. (1957). Non null ranking models I. *Biometrica*, **44**, 114-130.

(11) Rogers, T. (1990). *Students as literary critics: A case study of the interpretative theories, processes, and communities of ninth grade students.* Unpublished Technical Report.

(12) Thompson, G. L. (1993). Generalized permutation polytopes and exploratory graphical methods for ranked data. *The Annals of Statistics*, **21**, 1401-1430.