Chapter 1

Association In Contingency Tables, Correspondence Analysis And (Modified) Andrews Plots

Abstract: Andrews plots (Andrews, 1972), as a tool to graphically interpret multivariate data, have recently gained considerable recognition. We present the use of Andrews plots and Modified Andrews plots which are recently introduced by the authors (Khattree and Naik, 2001), in graphically exhibiting the association in the cross classified data and in the context of correspondence analysis. A new alternative to traditional correspondence analysis introduced by C. R. Rao (1995) and implementation of our approach in this case are also discussed.

Keywords and phrases: Andrews plots, correspondence analysis, Hellinger distance, modified Andrews plots

1.1 Introduction

Andrews plots (Andrews, 1972) as a tool for the graphical representation of multivariate data have recently gained considerable popularity due to the simplicity of their plotting and many desirable and attractive mathematical properties. Khattree and Naik (2001) provide a review of Andrews and other related plots along with their merits and demerits.

For a p-dimensional multivariate observation $\mathbf{y} = (y_1, \ldots, y_p)'$, the Andrews curve in argument t is defined by

$$f_{\mathbf{y}}(t) = \frac{y_1}{\sqrt{2}} + y_2 \sin(t) + y_3 \cos(t) + y_4 \sin(2t) + y_5 \cos(2t) + \ldots \qquad -\pi \leq t \leq \pi. \tag{1.1}$$

Clearly the p-dimensional observation is then represented by a curve and hence can be plotted in two-dimensions, without any loss of information, which usually occurs in other graphical representation such as the scatter plots or the dimensionality reduction based techniques such as Bi-plots.

A plot containing Andrews curves corresponding to all the multivariate observations is called an Andrews plot. Khattree and Naik (2001) recently suggested a modified Andrews plot as a plot consisting of curves

$$\begin{aligned} g_{\mathbf{y}}(t) = \frac{1}{\sqrt{2}} \Big\{ & y_1 + y_2[\sin(t) + \cos(t)] \\ & + y_3[\sin(t) - \cos(t)] + y_4[\sin(2t) + \cos(2t)] \\ & + y_5[\sin(2t) - \cos(2t)] + \ldots \Big\}, \\ & -\pi \leq t \leq \pi, \end{aligned} \tag{1.2}$$

which has all the properties of the Andrews plots yet avoids some of the shortcomings of Andrews plots. Specifically, the odd numbered terms in the Andrews plot given in (1.1) simultaneously vanish at $t = 0$. This is a major disadvantage since human eyes tend to concentrate on and notice changes or similarities more easily and more immediately in the central regions of the graph (that is, around $t = 0$); yet, for Andrews plot around $t = 0$, these similarities or dissimilarities are mostly due to even numbered variables $y_2, y_4, \ldots$. This is not so, for the modified Andrews plots given by (1.2). It can be seen from the fact that $\frac{1}{\sqrt{2}}[\sin(kt) \pm \cos(kt)] = \sin\left(kt \pm \frac{\pi}{4}\right) = \sin\left(k\left(t \pm \frac{\pi}{4k}\right)\right)$, which has the phase angle $\pm\frac{\pi}{4k}$. Thus, none of the terms in (1.2) has the common phase angle and hence will not vanish at the same point. Therefore, for any t, the value of $f_{\mathbf{y}}(t)$ depends on at least $(p-1)$, y_i values.

The objective of this article is to use the Andrews plots to graphically represent the associations in the contingency tables. Often such associations are depicted through correspondence analysis where the rows and columns of the contingency table are represented by points in a two-or three-dimensional scatter plot. As mentioned earlier, this may not be completely satisfactory due to loss of information caused by the dimensionality reduction. Andrews and modified Andrews plots are clearly the ideal alternatives to avoid such loss. We describe the approach in the next section.

1.2 (Modified) Andrews Plots in Correspondence Analysis

For an $a \times b$ contingency table $\mathbf{N} = ((n_{ij}))$ the levels of two discrete variables A and B are represented by rows and columns, respectively. The correspondence analysis is a mathematical approach to represent these as the points in an m-dimensional space, where $m = \min(a, b) - 1$. This is done by using the singular value decomposition of the correspondence matrix $\mathbf{Q} = ((q_{ij}))$, where $q_{ij} = \frac{n_{ij}}{n}$, $n = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij}$. Clearly the correspondence matrix $\mathbf{Q}$ is the matrix of relative frequencies. The above singular value decomposition provides the representation of each row and column of $\mathbf{N}$ as a point in m-dimensional space. Projections of these points in two or three dimensions provide a way to present the information graphically. To present things more formally yet briefly, from the correspondence matrix $\mathbf{Q}$, we define

$$\mathbf{r} = \mathbf{Q1}$$

and

$$\mathbf{c} = \mathbf{Q}'\mathbf{1},$$

where $\mathbf{1} = (1, \ldots, 1)'$ is a column vector of appropriate dimension, containing 1 as the entry at each place. The vectors $\mathbf{r}$ and $\mathbf{c}$ are called, respectively, the row and column masses. We also define

$$\mathbf{D_r} = \mathrm{diag}\,(\mathbf{r}),$$
$$\mathbf{D_c} = \mathrm{diag}\,(\mathbf{c}),$$

and let

$$\mathbf{F} = \mathbf{D_r}^{-1}(\mathbf{Q} - \mathbf{rc}')\mathbf{D_c}^{-1}(\mathbf{Q} - \mathbf{rc}')'.$$

One can show with little manipulation that

$$n\mathrm{tr}\left(\mathbf{F}\right)=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\left(n_{ij}-\frac{n_{i.}n_{.j}}{n}\right)^2}{\frac{n_{i.}n_{.j}}{n}}$$

which is the Pearson's chi-squared test statistic for testing independence between A and B. The quantity $T_I=\mathrm{tr}(\mathbf{F})$ is referred to as the total inertia. It is just a multiple of Pearson's chi-squared and can be interpreted as a measure of the magnitude of the total row (or column) squared deviations. This interpretation avails us an opportunity to attempt to partition T_I into several components and allocate these components (which would essentially be the eigenvalues of $\mathbf{F}$) to various dimensions. This is done using the generalized singular value decomposition (see, Jobson, 1992) of the matrix $\mathbf{Q}-\mathbf{rc}'$. In fact, one can directly take the generalized singular value decomposition of the correspondence matrix $\mathbf{Q}$ rather than $\mathbf{Q}-\mathbf{rc}'$. The two decompositions are essentially equivalent except for one trivial (and in fact, redundant for our purpose) singular value which is always equal to 1.

In the traditional correspondence analysis only first two or three dimensions are retained and using these, one explores any associations or other striking features in the plots. However, the graphical aspect at this stage can be handled more effectively through (modified) Andrews plots. Instead of working with projections of all $m\times 1$ row and column points, we instead transform these points to Andrews or (modified) Andrews curves. Consequently, the information available in the data is used to its fullest. Row (or column) points closer to each other transform to curves nearer to each other. Any exceptional or influential categories can also be readily identified by these curves. Further, assuming a reasonable degree of symmetry in the correspondence matrix $\mathbf{Q}$, association of any row category to any column category will be reflected in the similarities between the corresponding (modified) Andrews curves. Alternatively, the asymmetric version of (modified) Andrews plots can be used (Greenacre, 1984). Independence in the contingency table should result in these plots showing no specific patterns.

It is more instructive to illustrate the techniques, approach and ways of interpretation through some examples. The next section provides these illustrations using certain real data sets.

1.3 Some Examples

Example 1. Fisher et al. (1982) present a study of reproducibility of coronary arteriography diagnostic procedure which plays an important role in deciding if a bypass surgery on a patient may be appropriate. It is thus important to monitor the quality of arteriography and ascertain the amount of agreement between the quality control site readings and the clinical site readings. The problem is essentially that of interlaboratory testing in a categorical data context. Table 1.1 presents the results for 870 randomly selected patients with two readings. For these readings the degree or intensity of disease was classified as none, zero vessel disease but some disease and one-, two-, and three-vessel disease. For identification in plots, these groups for the quality control site are denoted by 1, 2, 3, 4 and 5, respectively, and for the clinical site, the notation used are A, B, C, D and E, respectively.

Table 1.1: Agreement with Respect to Number of Diseased Vessels

Quality Control Site Reading	Clinical Site Reading				
	Normal (A)	Some (B)	One (C)	Two (D)	Three (E)
Normal (1)	13	8	1	0	0
Some (2)	6	43	19	4	5
One (3)	1	9	155	54	24
Two (4)	0	2	18	162	68
Three (5)	0	0	11	27	240

The row and column points $\mathbf{y} = (y_1, y_2, y_3, y_4)'$ for the data are obtained by performing the generalized singular value decomposition and these are reported in Table 1.2. Using these, we obtain the functions $f_{\mathbf{y}}(t)$ and $g_{\mathbf{y}}(t)$ for each of the five rows and columns. These plots are given in Figures 1.1 and 1.2. In the case of perfect agreement between the clinical site and quality control site readings, all row points should be the same as the corresponding column points, and thus the corresponding Andrews curves should be identical also. As seen in the plots, the agreement between the two readings appears to be quite good for each of the five classification groups and the (modified) Andrews curves of the corresponding points follow each other very closely in case of $f_{\mathbf{y}}(t)$ as well as $g_{\mathbf{y}}(t)$.

Table 1.2: Clinical and QC Site Evaluations: Row and Column Points

Points		Coordinates			
		y_1	y_2	y_3	y_4
Row (QC Site)	Normal	3.14513	1.72789	0.91817	1.52739
	Some	1.67179	0.25096	-0.32230	-0.97903
	One	0.24290	-0.84605	-0.32512	0.27160
	Two	-0.34541	-0.16330	0.72270	-0.15729
	Three	-0.61364	0.68014	-0.34912	0.05434
Column (Clinical)	Normal	3.12841	1.74992	0.96954	1.71600
	Some	1.94382	0.40705	-0.25836	-1.07768
	One	0.35725	-0.89110	-0.45605	0.26900
	Two	-0.26104	-0.32340	0.72042	-0.12930
	Three	-0.56821	0.59772	-0.26196	0.02836

The same argument can be made for the plots in Figures 1.3, 1.4 and 1.5, where the components of vector $\mathbf{y}$ have been permuted so that different y_i are assigned to the constant $\frac{1}{\sqrt{2}}$ in the expressions of $g_{\mathbf{y}}(t)$ (similar plots corresponding to $f_{\mathbf{y}}(t)$ are not shown here). It is not surprising, since in the present case, none of the four dimensions can be deemed redundant in that each explains a significant percent of total inertia. Specifically, these percentages are 43.80, 28.59, 16.31 and 11.30%. However, in many cases, the major portion (up to 85 or 90%) of the total inertia can be attributed to the first two dimensions. Whenever this does not happen, as in the present case, the use of only the first two dimensions to obtain a plot of points will not be very effective compared to the use of Andrews plots which utilize all dimensions.

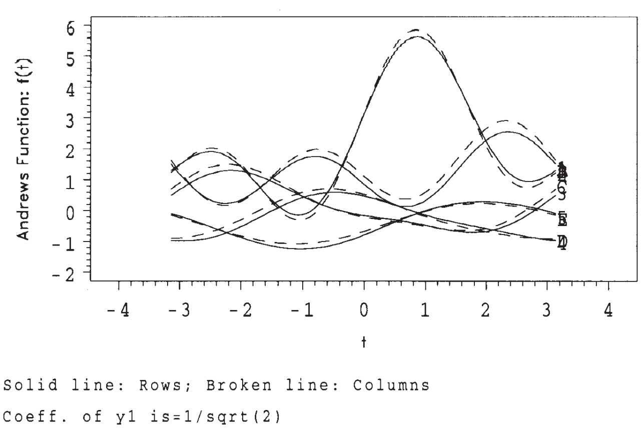


Figure 1.1: Agreement w.r.t. No. of Diseased Vessels

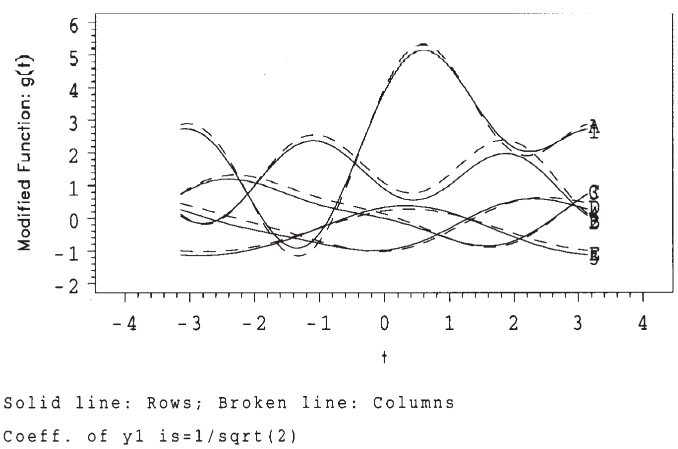


Figure 1.2: Agreement w.r.t. No. of Diseased Vessels

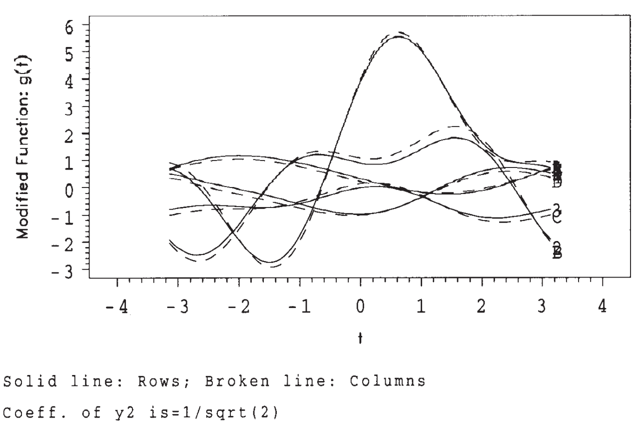


Figure 1.3: Agreement w.r.t. No. of Diseased Vessels

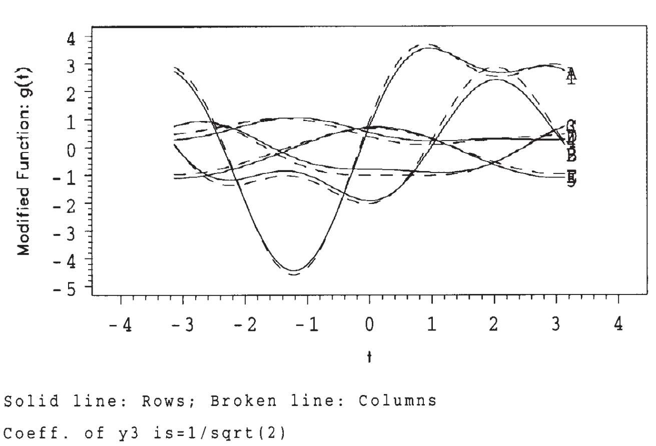


Figure 1.4: Agreement w.r.t. No. of Diseased Vessels

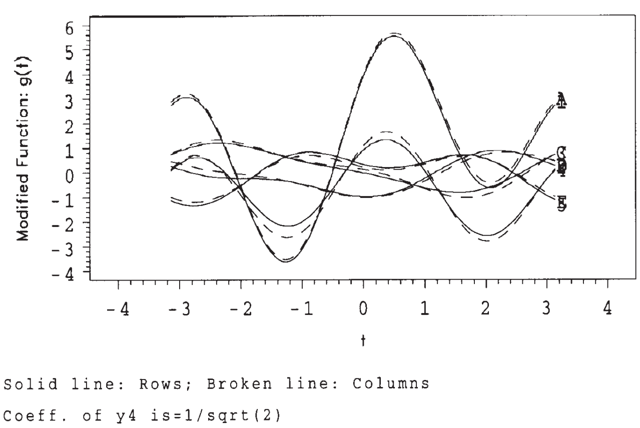


Figure 1.5: Agreement w.r.t. No. of Diseased Vessels

Example 2. Data in Table 1.3 are taken from Srole et al. (1978) and are reported in Agresti (1990) as well. The objective of the study was to examine the relationship, if any, between the mental impairment and parents' socioeconomic status. Six levels of socioeconomic status, 1 (high) to 6 (low) and four levels of mental health status, A (well), B (mild symptom formation), C (moderate symptom formation) and D (impairment) are taken. Data are obtained on a sample of 1660 residents of Manhattan. It may be pointed out that both variables are ordinal in nature.

As per the structure of the data for plotting, we have six row points and four column points in the three-dimensional space. However, the first dimension alone is able to explain 93.95% of the total inertia, and the first two dimensions combined explain almost 99% of that. Thus, the third dimension contains no significantly useful information for all practical purposes. The ten row and column points are listed in Table 1.4. The modified Andrews plots are given in Figures 1.6 and 1.7. The following observations are made from Figure 1.6.

(a) The two variables are ordered categorical in nature and this fact clearly shows up in Figure 1.6. Also, it can be observed that there is possibly further subgrouping within the socioeconomic status. Status 1 and 2 are similar to each other; 3 and 4 are similar and to some extent, there is some similarity between status 5 and 6. However, as subgroups, these three subgroups clearly separate themselves out in these plots.

(b) With respect to mental health status, the health status B (mild symptom formation) and C (moderate symptom formation) appear to be

Table 1.3: Cross-Classification of Mental Health Status and Parents' Socioeconomic Status

Parents' Socioeconomic Status	Mental Health Status (Well) A	(Mild Symptom Formation) B	(Moderate Symptom Formation) C	(Impaired) D
1 (high)	64	94	58	46
2	57	94	54	40
3	57	105	65	60
4	72	141	77	94
5	36	97	54	78
6 (low)	21	71	54	71

similar, but are clearly different from A (well) and D (impaired). Further, the ordinal nature of these health status is self evident in these plots.

(c) There appears to be positive association between the mental health status of children and parents' socioeconomic status, with mental health of children to be generally better at the higher levels of parents' socioeconomic status. Further, the upper two levels of parents' socioeconomic status groups generally correspond to a "well" mental health status; the next two levels are closely related to "mild" or "moderate symptom formation" and the lowest two levels of parents' socioeconomic status generally severely affect the mental health status of children towards impairment. It may be remarked that ideally one should use the *asymmetric plots* to assess the closeness of row and column points. However, in the present case our regular plot provides the interpretable results.

Plot in Figure 1.7 correspond to the case when y_2 has been assigned to the constant $\frac{1}{\sqrt{2}}$ in the function $g_{\mathbf{y}}(t)$. This plot reveals a few additional interesting observations.

Table 1.4: Mental Health and Parents' Socioeconomic Status: Row & Column Points

Points		Coordinates		
Row		y_1	y_2	y_3
(Socioeconomic)	1	0.180932	0.019248	0.027525
	2	0.184996	0.011625	-.027386
	3	0.059031	0.022197	-.010575
	4	-.008887	-.042080	0.011025
	5	-.165392	-.043606	-.010368
	6	-.287690	0.061994	0.004824
Column	Well	0.259536	-.012102	0.022589
(Mental Health)	Mild	0.029588	-.023651	-.019818
	Moderate	-.014210	0.069901	-.003230
	Impaired	-.237392	-.018897	0.015848

(a) With respect to parents' socioeconomic status, curves corresponding to groups 1 and 2 show the patterns just opposite to those corresponding to groups 5 and 6.

(b) With respect to mental health status, groups A (well) and D (impaired) follow the patterns opposite to each other.

(c) There, possibly, is very little difference between groups 3 and 4 of parents' socioeconomic status and between groups B and C with respect to mental health status. Conceivably, in each case, the corresponding two groups can be combined. If this is done, then there is a positive association between the combined groups (3, 4) and (B, C). Plots for this situation are not shown.

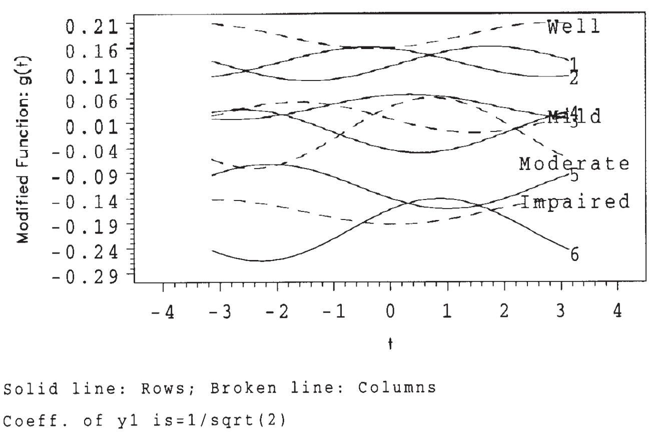


Figure 1.6: Agreement w.r.t. No. of Diseased Vessels

It may be emphasized that so much information could not have been extracted from these data if traditional two-dimensional scatter plots of correspondence analysis were constructed or if the usual Pearson's chi-squared test for independence was applied. Further, we note that that the suggested modification of Andrews plots provides curves which are easier to interpret and are more informative.

1.4 Modified Andrews Plots and Rao's Correspondence Analysis

Rao (1995) suggested the use of Hellinger distance instead of chi-squared distance for measuring the distance between two row (or column) profiles in correspondence analysis. This was done for the two main reasons (a) the chi-squared distance is not just the function of individual profiles under consideration but it also depends on the average of all profiles and (b) for the reason given in (a) and while measuring the affinities between profiles, undue emphasis may be given to categories with low frequencies. However, one important point to be emphasized here is that regardless of the choice of distance, (modified) Andrews plots can still be used in the context of correspondence analysis. To discuss the approach briefly, let $\mathbf{R} = ((r_{ij}))$ be

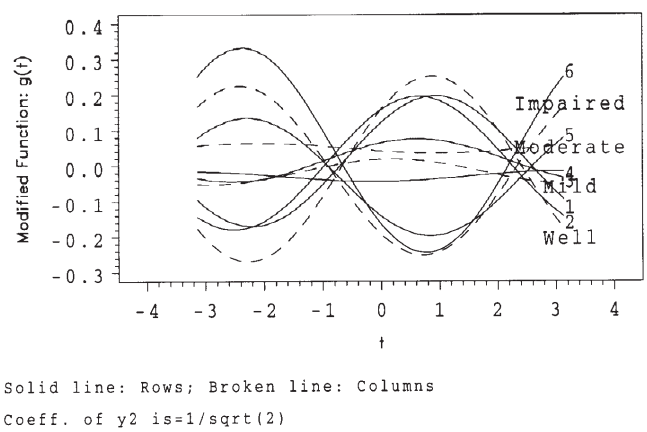


Figure 1.7: Agreement w.r.t. No. of Diseased Vessels

the $a \times b$ matrix of row profiles defined as

$$\mathbf{R} = \mathbf{D_r}^{-1}\mathbf{Q} = \begin{bmatrix} \frac{p_{11}}{p_{1.}} & \frac{p_{12}}{p_{1.}} & \cdots & \frac{p_{1b}}{p_{1.}} \\ \frac{p_{21}}{p_{2.}} & \frac{p_{22}}{p_{2.}} & \cdots & \frac{p_{2b}}{p_{2.}} \\ \vdots & & & \\ \frac{p_{a1}}{p_{a.}} & \frac{p_{a2}}{p_{a.}} & \cdots & \frac{p_{ab}}{p_{a.}} \end{bmatrix}.$$

Define

$$\mathbf{H} = ((h_{ij})) = (\sqrt{r_{ij}}) = \begin{bmatrix} \mathbf{h}_1' \\ \mathbf{h}_2' \\ \vdots \\ \mathbf{h}_a' \end{bmatrix},$$

where the vector $\mathbf{h}_i$ is defined as

$$\mathbf{h}_i = (\sqrt{r_{i1}}, \sqrt{r_{i2}}, \ldots, \sqrt{r_{ib}})', i = 1, 2, \ldots, a.$$

The Hellinger distance between the i^{th} and j^{th} profile is defined as

$$\delta_{ij}^2 = (\mathbf{h}_i - \mathbf{h}_j)'(\mathbf{h}_i - \mathbf{h}_j),$$

and as stated earlier, it depends only on the i^{th} and j^{th} profiles.

The canonical coordinates for representing row profiles are determined from the singular value decomposition of the $a \times b$ matrix $\mathbf{D}_{\mathbf{r}}^{\frac{1}{2}}(\mathbf{H} - \mathbf{1}'\boldsymbol{\eta})$, where Rao suggests the choice of vector $\boldsymbol{\eta}$ as

$$\boldsymbol{\eta} = \left(\sqrt{\frac{n_{.1}}{n}}, \sqrt{\frac{n_{.2}}{n}}, \ldots, \sqrt{\frac{n_{.b}}{n}}\right)'$$

or

$\boldsymbol{\eta}$=H'r.

It must be noted that the dimension of the space m, in this case, is the rank of $\mathbf{D}_{\mathbf{r}}^{\frac{1}{2}}(\mathbf{H} - \mathbf{1}'\boldsymbol{\eta})$ and is not necessarily equal to $\min(a, b) - 1$. If the singular value decomposition of $\mathbf{D}_{\mathbf{r}}^{\frac{1}{2}}(\mathbf{H} - \mathbf{1}'\boldsymbol{\eta})$ is

$$\mathbf{D}_{\mathbf{r}}^{\frac{1}{2}}(\mathbf{H} - \mathbf{1}'\boldsymbol{\eta}) = \lambda_1 \mathbf{u}_1 \mathbf{v}_1' + \lambda_2 \mathbf{u}_2 \mathbf{v}_2' + \cdots + \lambda_m \mathbf{u}_m \mathbf{v}_m',$$

then canonical coordinates for row profiles are given by

$$\lambda_1 \mathbf{D}_{\mathbf{r}}^{-\frac{1}{2}} \mathbf{u}_1,\ \lambda_2 \mathbf{D}_{\mathbf{r}}^{-\frac{1}{2}} \mathbf{u}_2, \ldots, \lambda_m \mathbf{D}_{\mathbf{r}}^{-\frac{1}{2}} \mathbf{u}_m,$$

and the column profiles are given by standard coordinates

$$\lambda_1 \boldsymbol{\Delta}_{\mathbf{c}}^{-1} \mathbf{v}_1,\ \lambda_2 \boldsymbol{\Delta}_{\mathbf{c}} \mathbf{v}_2, \ldots, \lambda_m \boldsymbol{\Delta}_{\mathbf{c}}^{-1} \mathbf{v}_m,$$

where

$$\boldsymbol{\Delta}_{\mathbf{c}} = \left[\mathrm{diag}[(\mathbf{H} - \mathbf{1}\boldsymbol{\eta}')'(\mathbf{H} - \mathbf{1}\boldsymbol{\eta}')]\right]^{\frac{1}{2}}.$$

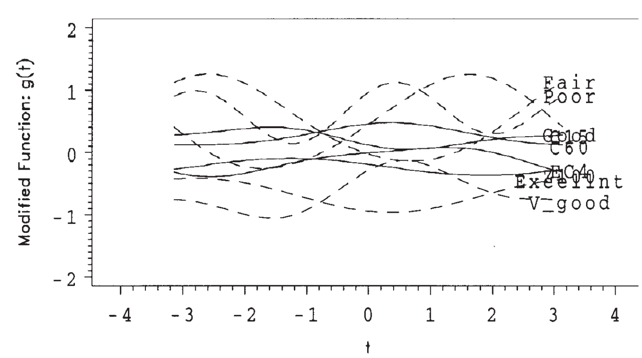
These can be plotted in the same plot. If one wishes to have the canonical coordinates for column profiles and standard coordinates for row profiles, above formulas can be suitably modified.

We will illustrate the modified Andrews plots based on Rao's correspondence analysis using a drug efficacy data of Calimlim et al. (1982).

Example 3. The data are reported in Table 1.5. The objective of the study was to examine the ratings assigned by the hospital patients to four drugs, with an objective to determine whether a particular drug or a group of drugs are favored by the patients. The four drugs are named as Z100, EC4, C60 and C15 and the five ratings on ordinal scale are from poor to excellent. We compute here the canonical coordinates for row profiles representing various drugs and standard coordinates for the column profiles representing the ratings. The coordinates are not shown here to save space. Also, we present here only the modified Andrews plots as Figures 1.8 and 1.9.

Table 1.5: Results of a Survey on Analgesic Efficacy of Drugs

	Ratings of Drug's Efficacy					
Drug	Poor	Fair	Good	Very Good	Excellent	Total
Z100	5	1	10	8	6	30
EC4	5	3	3	8	12	31
C60	10	6	12	3	0	31
C15	7	12	8	1	1	29
Total	27	22	33	20	19	121



Solid line: Rows; Broken line: Columns

Coeff. of y1 is=1/sqrt(2)

Figure 1.8: Drug vs. Efficacy Rating

Figure 1.8 clearly illustrates the closeness of the modified Andrews curves corresponding to drugs C15 and C60 and that of curves corresponding to EC4 and Z100. In the same plot, one also observes the anticipated clustering of the efficacy ratings, namely, {Excellent and Very Good} forming one cluster and the remaining three ratings forming the other cluster. The proximity of EC4 and Z100 to the higher two efficacy ratings and proximity of C15 and C60 to the lowest three ratings is also noted. This indicates that the EC5 and Z100 are perceived as the superior choices by patients.

Figure 1.9 represents modified Andrews plot when the roles of y_1 and y_2 (that is, the scores corresponding to the first and second dimensions) in (1.2) have been interchanged. This obviously changes the curves. However, in this plot the similarities and dissimilarities between various drugs and ratings and associations between the particular drugs and ratings are depicted not

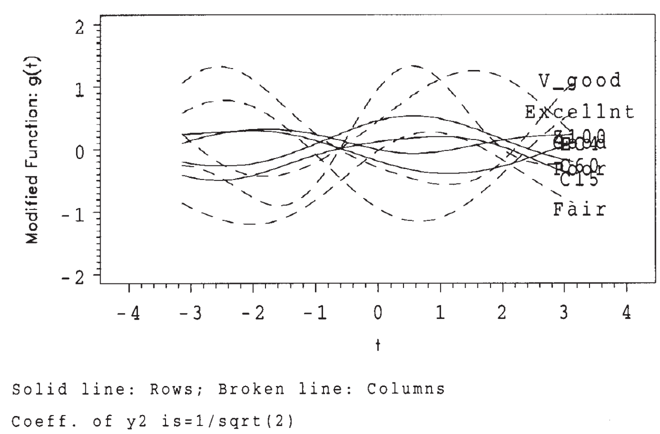


Figure 1.9: Drug vs. Efficacy Rating

by the relative closeness of corresponding curves but by their particular patterns (in terms of ups and downs). The curves corresponding to EC4 and Z100 have larger amplitudes compared to the other two drugs. A similar observation is made about the efficacy ratings. At the same time, it is also observed that the curves for EC4 and Z100 and those corresponding to C15 and C60 follow the opposite directions (when the former ones go up, the latter ones tend to go down). At the same time such changes for {EC4, Z100} and efficacy ratings {Excellent, Very Good} are in the same direction, thereby indicating the strong association in this case. A similar statement can be made for drugs {C15, C60} and efficacy ratings {Poor, Fair, Good}. It must be emphasized that all these features are not as distinct in the two-dimensional plots (not shown here) obtained by appropriate dimensionality reduction.

1.5 Conclusions.

Andrews plots as powerful graphical technique for multivariate data are not only useful in traditional clustering or outlier detection problems, but they can also be very important tools in analyzing experimental data in other discrete or descriptive multivariate analysis problems, such as contingency tables and correspondence analysis. The fact that there is no loss of dimension, and hence no approximations before the graphical displays of the data, allows these plots to be much more informative than the traditional scatter plots or the biplots. Further, the modified Andrews plots are even more informative and useful in the graphical displays of these data.

References

1. Agresti, A. (1990). *Categorical Data Analysis*, New York: Wiley.

2. Andrews, D. F. (1972). Plots of high dimensional data, *Biometrics*, 28, 125-136.

3. Calimlin, J. F., Wardell, W. M, Davis, H. T., Lasagna, L. and Gillies, A. J. (1982). Analgesic efficacy of an orally administered combination of Pentazocine and Aspirin: With observations on the use of and statistical efficacy of global subjective efficacy ratings, *Clinical Pharmacology and Therapeutics*, **21**, 34–43.

4. Fisher, L. D., Judkins, M. P. Lesperance, J., Cameron, A., Swaye, P., Ryan, T. J., Maynard, C., Bourassa, M., Kennnedy, J. W., Gosselin, A., Kemp, H., Faxon, D., Wexler, L., and Davis, K. (1982). Reproducibility of coronary arteriographic reading in the coronary artery surgery study (CASS), *Catheterization and Cardiovascular Diagnosis*, **8**, 565–575.

5. Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis*, London: Academic Press.

6. Jobson, J. D. (1992). *Applied Multivariate Data Analysis*, Vol. II, New York: Springer-Verlag.

7. Khattree, R. and Naik, D. N (2001). Andrews plots for multivariate data: Some new suggestions and applications, to appear in *Journal of Statistical Planning and Inference.*

8. Rao, C. R. (1995). A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance, *Quwstiio*, **19**, 23–63.

9. Srole, L. , Langner, T. S., Michael, S. T., Kirkpatrick, P., Opler, M. K., and Rennie, T. A. C. (1978). *Mental Health in the Metropolis: The Midtown Manhattan Study*, Revised edition, New York: NYU Press.