Behaviormetrika
1989, No. 26, 31-46

IDEAL POINT DISCRIMINANT ANALYSIS AND ORDERED RESPONSE CATEGORIES**

Yoshio Takane*

Many contingency tables have ordered response categories. This paper compares two major approaches to ordinal response categories, namely the point representation and the block (interval) representation. Specifically, two methods are compared, each representing each of the two approaches. One is ideal point discriminant analysis (IPDA) and the other the successive categories method (SCM). Similarities and distinctions between the two methods are explicated. The goodness-of-fit (GOF) is compared through AIC using several example data sets. IPDA and SCM were found to provide similar GOF, but IPDA was found to provide a slightly better fit in all the data sets examined.

1. Introduction

Many contingency tables obtained in social sciences, psychology, medicine, etc. have ordered response categories. For example, degrees of job satisfaction may be measured on a rating scale with response categories labelled as (1) very much satisfied, (2) moderately satisfied, (3) neutral, (4) not very much satisfied and (5) dissatisfied. These ratings are then related to subjects' demographic information. In a signal detection experiment in psychology subjects respond to two stimulus conditions, signal in the noise background and noise alone. The subjects may be asked to indicate confidence levels in their judgments by ordered response categories. The rating method is known to provide a more efficient way of collecting signal detection data than the conventional method which stipulates binary responses.

A variety of models have been proposed to capture the ordinal nature of response categories (Agresti, 1984 ; Goodman, 198 ; McCullagh, 1980). These methods roughly fall into one of two major approaches. The two approaches are distinguished by the mode of representing columns of a contingency table corresponding to the ordered response categories. They are called point representation and block (or interval) representation (de Leeuw, 1983), which closely parallel Bock's (1975) distinction between extremal and threshold concepts.

In the point representation approach both rows and columns of a contingency table are represented as points in a Euclidean space. Strengths of connections between rows and columns are defined in terms of relative locations of the points. The probability of a response (a column) given a row is assumed proportional to the strength of their connection relative to other connections. Association models (Agresti, 1984 ; Andersen, 1980 ; Good-

Key Words and Phrases ; point representation, block (or interval) representation, the Successive Categories Method (SCM), order restrictions, maximum likelihood estimation, AIC
* Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec H3A 1B1, Canada
** The work reported in this paper was supported by grant 410-88-1346 from the Social Sciences and Humanities Research Council of Canada. Request for reprints should be sent to Yoshio Takane, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec H3A 1B1, Canada.

man, 1979, 1981) and ideal point discriminant analysis (Takane, 1987, 1989; Takane, Bozdogan & Shibayama, 1987) are two representative methods that belong to this class of methods.

In the block representation approach columns are represented as successive intervals on a unidimensional continuum. Those intervals are characterized by their end points called upper and lower thresholds. Rows are represented as random variables with certain distributional properties along the same continuum. The probability of a response (a column) in a particular row is assumed equal to that of the row random variable falling into the interval corresponding to the column. The implementation of this approach is relatively new in statistics (Anderson & Philips, 1981; Cox, 1970; McCullagh, 1980), but it has a long tradition in psychometrics (Torgerson, 1958). It is called the successive categories scaling method.

It is of interest to compare these two approaches systematically. However, due to space limitation the comparison must be restricted between two specific models in this paper, each representing each of the two approaches described above. The methods are ideal point discriminant analysis (IPDA) and the logistic version (McCullagh, 1980) of the successive categories method (SCM). Models and parameter estimation procedure associated with the two methods are briefly described in the next section. The two models are fitted to several data sets and the goodness-of-fit is compared through Akaike's (1974) information criterion (AIC). The data sets used will be described in Section 3 and the results presented in Section 4. The paper concludes with a discussion in Section 5.

2. Methods to be Compared

2.1 Basic Models

Let $F=(f_{ij})$ denote an R by C contingency table where f_{ij} is the frequency of column j given row i. We assume that the table is arranged so that its columns correspond with ordered response categories. Rows, on the other hand, represent categories of an explanatory variable or variables. Let X denote an R by p design matrix for the rows. This matrix may contain a set of continuous variables, dummy-coded discrete variables or a mixture of both. When there is no obvious design for the rows, X may be set to an identity matrix of order R.

In IPDA the rows of F are mapped into an A dimensional Euclidean space by a linear combination of X; *i.e.*,

$$Y=XB \tag{1}$$

where Y is the R by A matrix of coordinates of the row points and the B the p by A matrix of unknown weights. $A \leq \min(R-1, C-1)$. A is also restricted to be smaller than or equal to the number of nonredundant explanatory variables. In order to remove translational indeterminacy in the Euclidean space, continuous variables in X are centered apriori and the weights for discrete variables are constrained so that they satisfy the centering restriction,

$$\sum_q f_{k(q)} \, b_{k(q)a}=0$$

where $k(q)$ indicate the q-th category in variable k. The $f_{k(q)}$ denotes the observed frequency of the category, and $x_{ik(q)}$ and $b_{k(q)a}$ and appropriate elements of X and B, respectively. Let M denote the C by A matrix of coordinates of column points. It is assumed that M is given by weighted centroids of Y; *i.e.*,

$$M = D_c^{-1}F'Y = D_c^{-1}F'XB \tag{2}$$

where D_c is the diagonal matrix with column totals of F on the diagonal. The strength of connection between row i and column j is measured by the negative exponential function of the squared Euclidean distance between the corresponding points; *i.e.*, exp $(-d_{ij}^2)$ with

$$d_{ij} = \left\{ \sum_{a=1}^{A} (y_{ia} - m_{ja})^2 \right\}^{1/2} \tag{3}$$

where y_{ia} and m_{ja} are appropriate elements of Y and M, respectively. The d_{ij} is a function of B. The conditional probability of column j given row i is proportional to w_j $exp\ (-d_{ij}^2)$, where $w_j(>0$ and $\Sigma w_k = 1)$ is the bias parameter for column j (similar to the prior probability of column j). That is,

$$p_{ij} = g_i w_j exp\ (-d_{ij}^2) \tag{4}$$

where p_{ij} is the conditional probability of column j given row i and $g_i = [\Sigma w_k\ exp\ (-d_{ik}^2)]^{-1}$, which is the scaling factor that makes p_{ij} add up to unity across columns. Some justifications behind the exact form of model (4) are given in Takane, *et al.* (1987). Relations of this model to other methods, such as the log-linear model, association models and correspondence analysis (dual scaling, quantification methods 2 & 3), are also given in Takane, *et al.* (1987) and Takane (1987).

In the logistic version of SCM the random variable corresponding to row i is assumed to follow the logistic distribution with mean $x_i'b^*$, where x_i' is the i-th row vector of X and b^* is the vector of unknown weights, analogous to B but restricted to be unidimensional. Let c_j denote the upper threshold value for the interval corresponding to column j. Successive intervals are assumed contiguous, so that c_{j-1}, the upper threshold value for the j-1st interval, coincides with the lower bound for the j-th interval. The probability of the random variable for row i not exceeding c_j is given by

$$h_{ij} = [1 + exp\ (-(c_j - x_i'b^*))]^{-1} \tag{5}$$

The probability of the random variable falling into the interval, $(c_{j-1}, c_j]$, is then given by

$$p_{ij} = h_{ij} - h_{i(j-1)} \tag{6}$$

This p_{ij} is assumed equal to the conditional probability of column j given row i.

The normal distribution was originally used (Torgerson, 1958) for the distribution of the row random variables. This follows the Thurstonian tradition (Thurstone, 1927) in psychometrics. The logistic distribution, adopted here, is getting more popular recently because of its closed form expression for the cumulative distribution function, which is rather crucial in SCM.

In both IPDA and SCM a number of interesting model specifications are possible within the basic framework of the models presented above. For example, some elements in B (and b^*) may be fixed at a certain prescribed value, or may be equated to each other. The design matrix, X may be manipulated in various ways to reflect presumed structures among row categories. For example, the rows may represent some products which are characterized by combinations of features or attributes. This information may be incorporated in the representation of the row categories. The best representation can be identified by a subset selection procedure applied to X. Similar structures may also be imposed on columns. For example, c_j may be equally spaced, or w_j may be equated across the columns. Takane (1989) presents an interesting case in which the values of w_j's vary systematically across different subsets of rows.

Whether p_{ij} is defined by (4) or (6) the conditional likelihood of F is stated as

$$L = \prod_i \prod_j (p_{ij})^{f_{ij}} \tag{7}$$

which is maximized with respect to model parameters, B and w_j in IPDA and b^* and c_j in SCM. Fisher's scoring algorithm is used for maximization, which is found to work efficiently. Once the maximum likelihood, L^*, is obtained, the AIC statistic (Akaike, 1974) is readily calculated; *i.e.*,

$$AIC = -2 \, In \; L^* + 2n_{\Pi} \tag{8}$$

where n_{Π} is the effective number of estimated parameters in model $_{\Pi}$.

2.2 Special Case of $C=2$ and Order Constraints

When the number of response categories is two ($C=2$), IPDA and SCM yield identical predictions (Takane, 1987). This is intuitively clear, since two categories can always be arbitrarily ordered. Note that the dimensionality of the representation space in IPDA is always one ($A=1$) in this case. The correspondence between parameters in IPDA and those in SCM in this case are as follows: Let

$$v_1 = (f_{i1}/f_i)x_i \tag{9}$$
$$v_2 = (f_{i2}/f_i)x_i \tag{10}$$
$$e = 2(v_2 - v_1) \tag{11}$$

and

$$s = (v_1 + v_2)/2. \tag{12}$$

Then

$$b^* = (e'b)\, b$$

(where b is the unidimensional B) or

$$b = b^*/(e'b^*)^{1/2},$$

and

$$c_1 = s'b^* - ln \; (w_1/w_2)$$

or

$$w_1 = 1 + exp \; (c_1 - s'b^*))^{-1}$$

(and $w_2 = 1 - w_1$). A proof is given in the Appendix. When $C>2$, however, there is no straightforward relationship between the two models.

Note that even if the dimensionality of the representation space is restricted to one (*i.e.*, $A=1$) in IPDA, there is no guarantee that the column points corresponding to the ordered response categories are arranged in a specific order. There is no built-in mechanism in IPDA that enforces a particular order in the arrangement of the column points. However, simply combining the columns which violate the order has the effect of equating the points corresponding to the columns. Let G be a partition of all columns whose elements are indexed by g. Then in general

$$L=\Big(\prod_{i,\, g\epsilon G}\prod (p_{ig})^{\Sigma_{j\epsilon g} f_{ij}}\Big)\Big(\prod_{ij\ \epsilon g}\prod (p_{ij}/p_{ig})\ f_{ij}\Big) \tag{13}$$

Combining columns amounts to fitting IPDA using only the first portion of the likelihood function stated above. The second factor in (13) may be calculated separately and subsequently multiplied to the first to obtain the total likelihood. All possible partitions of columns may be tried out. We examine which partitions of columns provide arrangements of the column points consistent with prescribed orders and choose the one which gives the best fit among them. This procedure is a special case of multi-sample cluster analysis (Takane, *et al.*, 1987) proposed to test the equality of column points which are not necessarily ordered. The total number of possible partitions of columns in 2^{c-1}. In creating a partition of contiguous columns boundaries for subsets of the columns may be placed at $C-1$ possible locations, where actual boundaries may be or may not be placed. We thus obtain 2^{c-1}. This number is substantially smaller than that of possible partitions of unordered columns, which is given by

$$\sum_{k=1}^{c} S(C,k) \text{ where } \quad S(C,\, k)=\sum_{g}\binom{k}{g}(-1)^{g}\,(k-g)^{c}/K\,! \text{ (Duran \& Odell, 1974).}$$

This is because a subset of ordered columns in a partition should include only adjacent columns. Note that combining columns has the effect of removing the built-in order restriction in SCM. Thus, this operation generally worsens the GOF (goodness of fit) of IPDA, while it improves the GOF in SCM.

3. Data

Extensive model comparison was conducted by analyzing several sets of actual data by the two methods, IPDA and SCM. The data sets used are presented in Tables 3.1-3.5. Brief descriptions of the data sets will follow.

(1) *Ogilvie & Creelman's (1968) data.* This data set comes from a signal detection experiment for two-point touch sensitivity using 0.5 inch separation on the forearm. "Signal" refers to presentation of two points and "noise" to presentation of only one point. In each trial either the signal or the noise is presented, and the subject is to indicate his confidence level for either signal or noise by a six-category rating method.

(2) *Guilford's (1936, p. 187) data* (see also Bock, 1975, p. 549-550). The data pertain to judgments of apparent differences in lifted weights. The subject, who is blindfolded, lifts two weights successively and states whether the second is "greater" than or "less" than the first. If the subject is in doubt, he reports "doubtful". One of the weights (B) is standard at 200 grams, and the other (A) varies from 185 grams through 215 grams in steps of 5 grams. This method of data collection is called the constant method.

(3) *Bradley, Katti & Coon's (1962) data.* This is from Example 3 in their papers.

Table 3.1
Ogilvie & Creelman's (1968) Data from a Signal Detection Experiment by the Rating Method

Presentation	Response Category						Total
	Signal			Noise			
	Sure	Medium	Unsure	Unsure	Medium	Sure	
Noise:	15	17	40	83	29	66	250
Signal:	68	37	68	46	10	21	250
Total	83	54	108	129	39	87	500

Table 3.2
Guilford's (1936) Data on Judgment of Lifted Weights by the Constant Method

Weight, g		Judgment: A is			Total
A	B	Greater	Doubtful	Less	
185	200	5	4	91	100
190	200	12	18	70	100
195	200	15	25	60	100
200	200	30	42	28	100
205	200	55	35	10	100
210	200	70	18	12	100
215	200	85	9	6	100
Total		272	151	277	700

Table 3.3
Bradley, Katti & Coons (1962) Data (Example 3): Five Treatments Rated on a Five Point Rating Scale

Treatment	Terrible	Poor	Fair	Good	Excellent	Total
I	9	5	9	13	4	40
II	7	3	10	20	4	44
III	14	13	6	7	0	40
IV	11	15	3	5	8	42
V	0	2	10	30	2	44
Total	41	38	38	75	18	210

There are five treatment conditions rated on five-point rating scales. Descriptive labels of the five response categories are: Terrible, poor, fair, good, and excellent. Unfortunately, no detailed explanations of the five treatment conditions are provided in Bradley, *et al.*'s paper.

(4) *Merit distribution data.* Members of 14 faculties at McGill University are classified according to their merit salary increase per annum in 1987. There are four categories of increase, $2,400, $ 1,650, $ 750 and $ 0. In one of the faculties, Science, the

Table 3.4
Merit Distribution Data Across Faculties of McGill University in 1987

Faculty	(Abr.)	$2400 (1)	$1650 (2)	$750 (3)	$0 (4)
1. Agriculture	(Ag)	13	27	19	15
2. Arts	(Ar)	56	81	68	13
3. Dentistry	(De)	7	9	3	1
4. Education	(Ed)	30	32	27	11
5. Engineering	(En)	36	42	32	15
6. Graduate Studies	(Gr)	13	11	11	5
7. Law	(La)	13	10	6	2
8. Management	(Ma)	20	13	13	8
9. Medicine	(Me)	24	46	44	52
10. Music	(Mu)	9	9	11	9
11. Religious Studies	(Re)	7	4	3	3
Departments in Science					
A		11	13	7	3
B		12	7	6	1
C		3	7	3	3
D		2	8	4	2
E		13	13	7	7
F		2	2	2	0
G		1	1	1	1
H		12	12	4	3
I		8	11	9	1
12. Science (total)	(Sc)	64	74	43	21
13. Libraries	(Li)	18	27	19	4
14. Others	(Ot)	9	13	13	15

Table 3.5
Williams' Data on Periodontal Condition and Average Daily Calcium Intake of Women

Periodontal condition	Calcium per day (in g.)				Total
	<0.40	0.40–0.55	0.55–0.70	>0.70	
A	5	3	10	11	29
B	4	5	8	6	23
C	26	11	3	6	46
D	23	11	1	2	37
Total	58	30	22	25	135

data are available for each of its nine departments. It will be interesting to see if the merit distribution does not differ significantly across the different departments in the Faculty of Science.

(5) *Williams' (1952) data.* This data set was analyzed previously by Williams (1952), Goodman (1981), and Tsujitani (1988). The data concern the relationship between four periodontal conditions and the amount of daily calcium intake. The amount of intake is categorized into four categories, less than .40 grams a day, between .40 and .55 grams a day, between .55 and .70 grams a day and more than .70 grams a day. There is no detailed description of the four periodontal conditions in any of the papers cited above.

4. Results

Major results are reported in Tables 4.1-4.5. Entries in these are the values of AIC's for fitted models along with the effective numbers of parameters given in parentheses. The saturated model means $p_{ij}=f_{ij}/f_i$ (where f_i is the marginal total of row i in the original data). This model gives an upper bound of the likelihood function. The null model, on the other hand, postulates $p_{ij}=w_j=f_j/f$ for all i (where f_j is the marginal total of column j and f is the grand total). These two models are used as bench-mark models.

The results will be reported data set by data set:

(1) *Ogilvie & Creelman's data.* Table 4.1 indicates that IPDA provides the minimum AIC solution. The difference, however, between IPDA and SCM is relatively minor; the latter still provides a better solution than the saturated model.

Figure 4.1a displays the optimal representation of rows and columns of Ogilvie & Creelman's data under IPDA. Notice that the column points corresponding to the six response categories are arranged in prescribed order despite the fact that no order restriction was explicitly imposed in deriving the point locations. The dimensionality of the representation space is restricted to unity in this case, since the number of rows in the data set is two. The interval representation of the same data set derived by SCM is depicted in Figure 4.1b for comparison. This, however, is not an opimal solution.

(2) *Guilford's data.* Comparisons are made among a variety of possible structures for rows of the table as well as between IPDA and SCM. Also, min $(R-1, C-1)>2$, so that the comparison between unidimensional and two-dimensional solutions is interesting in certain cases. Results are summarized in Table 4.2. In the table "Row=nominal" means that the rows are unconstrained (*i.e.*, $X=I$). "Row=linear", on the other hand, means that the linear trend was assumed over the successive rows. This specification was motivated by the fact that the rows correspond with the comparison stimulus whose weight

Table 4.1
Summary Results for Ogilvie & Creelman's Data

Saturated Model	1645.8 (10)
Null Model	1732.3 (5)
IPDA, dim=1	1638.1 (6)
SCM, dim=1	1640.5 (6)

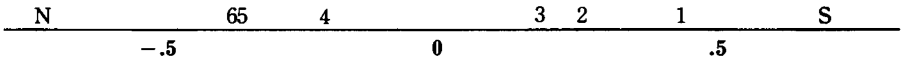

Fig. 4.1a The Best Point Representation (IPDA) of Ogilvie & Creelman's Data

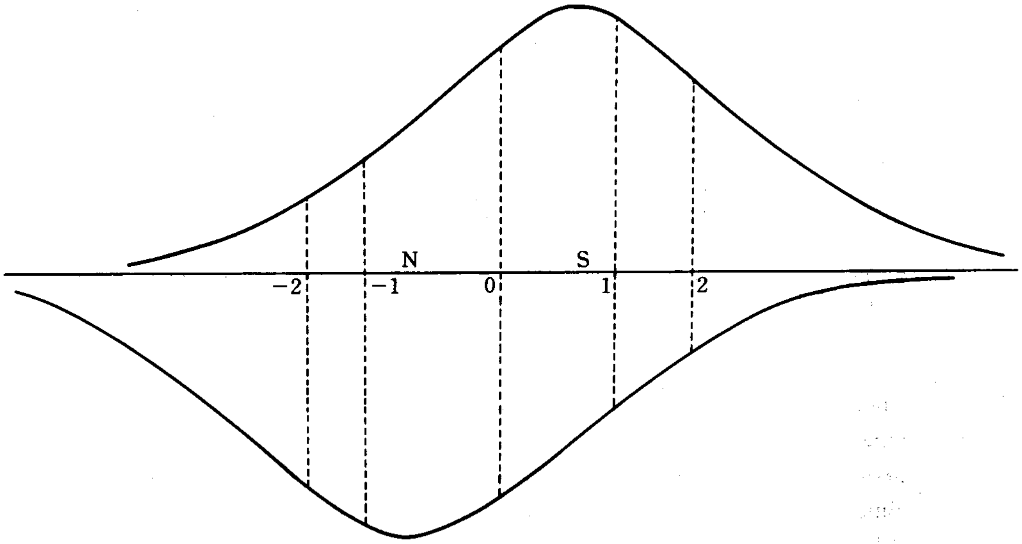

Fig. 4.1b The Corresponding Interval Representation (SCM)

was increased in equal physical unit. In "Row=linear & quadratic" the quadratic trend was added. Finally, "Row=linear in log" means that the log of physical weight of the comparison stimulus was used as the explanatory variable X. This last hypothesis was motivated by the Fechner's law (*e.g.*, Torgerson, 1958) in psychophysics, stating that the subjective weight changes as a logarithmic function of the physical weight.

Table 4.2 indicates that the two-dimensional IPDA solution with linear and quadratic trends on the rows provides the minimum AIC solution. Note that for unidimensional representations SCM generally provides better fitting solutions than IPDA. However, the representation seems to require more than one dimension; two-dimensional IPDA solu-

Table 4.2
Summary Results for Guilford's (1936) Data

Saturated Model	1119.9 (14)	
Null Model	1495.0 (2)	
	IPDA	SCM
Row=nominal, dim=2	1118.2 (13)	
dim=1	1134.8 (8)	1121.4 (8)
Row=linear, dim=1	1130.2 (3)	1116.9 (3)
Row=linear & quadratic, dim=2	1113.9 (5)	
dim=1	1131.9 (4)	1117.9 (4)
Row=linear in log, dim=1	1129.9 (3)	1116.3 (3)

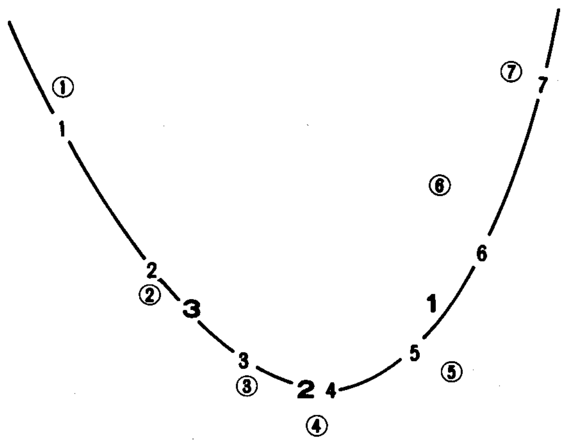

Fig. 4.2 The Derived Two-Dimensional IPDA Configurations (both unconstrained and constrained solutions superposed) Derived from Guilford's Data.

tions are generally better than the corresponding SCA solutions which are restricted to be unidimensional.

The unrestricted (Row=nominal) IPDA solutions were obtained in both one dimension and two dimensions. The two-dimensional solution turned out to be a much better solution. The row points from the unrestricted model are depicted in Figure 4.2 by encircled numerals, which show a clear quadratic configuration. This motivated the fitting of the constrained two-dimensional model by IPDA with the linear and quadratic trends imposed on the rows. This turned out to be the best fittting solution. The row points and the column points from this solution are also depicted in Figure 4.2. The row points, indicated by numerals of smaller size, are lying exactly on a quadratic curve. The column points, indicated by numerals of larger size, are lying in expected order along the curved unidimensional configuration. This type of guadratic configuration, or the curved unidimensional configuration, is quite common in multidimensional scaling and is known as the horseshoe phenomenon (*e.g.*, van Rijckevorsel, 1987). Whether this phenomenon has any substantive meaning or is just an artifact of the fitted model is yet to be investigated.

(3) *Bradley, Katti & Coon's data* (*Example 3*). The unidimensional IPDA solution indicated violations of order among the column points. They are arranged in the order of 2, 1, 5, 3, and 4. There are two possible venues to explore in such situations. One is to ignore the ordinal nature of the columns and the other is to find a constrained solution that satisfies the prescribed order of the columns. We did both. Two- and three-dimensional IPDA solutions were obtained, ignoring the order of the columns. The two-dimensional solution turned out to be optimal (AIC=582.3), and is presented in Figure 4.3b. This two-dimensional configuration may suggest that the violations of order among the column points, which looked totally unintelligible in the unidimensional configuration, may be another instance of the horseshoe phenomenon. The column points are roughly in the prescribed order on a curved unidimensional manifold within the two-dimensional space.

Multi-sample cluster analysis was also performed with both IPDA and SCM and for all possible partitions of the ordered columns. Results are presented in Table 4.3. In the

Table 4.3
Summary Results for Bradley, Katti & Coon's Data (Example 3)

Saturated Model	591.5 (20)	
Null Model	644.7 4)	
Clustering Alternatives	**IPDA**	**SCM**
1, 2, 3, 4, 5 dim=1	586.1 (8)	616.5 (8)
=2	582.3 (11)	
=3	584.4 (13)	
(1, 2), 3, 4, 5		613.4 (8)
1, (2, 3), 4,5		624.5 (8)
1, 2, (3, 4), 5		618.9 (8)
1, 2, 3, (4, 5)		607.0 (8)
(1, 2, 3), 4, 5		627.5 (8)
1, (2, 3, 4), 5		639.7 (8)
1, 2, (3, 4, 5)		600.2 (8)
(1, 2), (3, 4), 5		615.4 (8)
(1, 2), 3, (4, 5)	595.6 (8)	603.5 (8)
1, (2, 3), (4, 5)	614.8 (8)	616.2 (8)
(1, 2), (3, 4, 5)	596.1 (8)	596.1 (8)
(1, 2, 3), (4, 5)	620.1 (9)	620.1 (8)
(1, 2, 3, 4), 5	639.8 (8)	639.8 (8)
1, (2, 3, 4, 5)	626.6 (8)	626.6 (8)

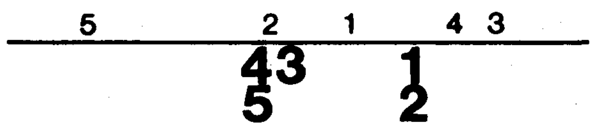

Fig. 4.3a The Order-Restricted Unidimensional IPDA Solution for Bradley, Katti, & Coons' Data

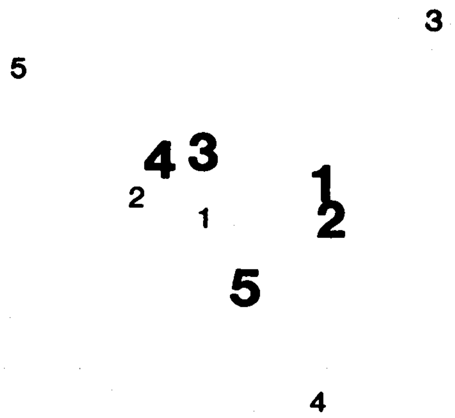

Fig. 4.3b The two-dimensional IPDA solution for Bradley, Katti, & Coons' Data

Table 4.4
Summary Results for the Merit Distribution Data

	14 Rows (Depts. in Fac. of Science Equated)	22 Rows
Saturated Model	3225.6 (32)	3255.0 (66)
Null Model	3248.3 (3)	3248.3 (3)
IPDA dim=2	3204.3 (28)	3221.5 (44)
dim=1	3196.4 (16)	3203.8 (24)
SCM dim=1	3207.7 (16)	3217.3 (24)

table column numbers enclosed in parentheses indicate those which were combined for the purpose of analysis. The AIC values from IPDA are not reported for some partitions of the columns (called "clustering alternatives"). These partitions did not yield the column points consistent with the prescribed order. The constrained IPDA solution, in which columns 1 and 2, and columns 4 and 5 were equated, is found to be the best solution (AIC= 595.6) among those which satisfied the order restriction. This solution is presented in Figure 4.3a. Which of Figure 4.3a and Figure 4.3b is more informative is difficult to judge, because of our lack of knowledge is the subject matter of research in which this data set was collected. Note the all the AIC values are identical for IPDA and SCM in the constrained solutions in which the assumed number of distinct columns in two. This is in line with our discussion in Section 2.2 that two methods are equivalent for $C=2$.

(4) *The merit distribution data.* Both IPDA and SCM solutions were obtained for the nine departments in the Faculty of Science treated separately (22 Rows) or unitarily (14 Rows). The latter assume that there are no significant differences in merit distributions across the nine departments, so that they may be treated as one. Table 4.4 indicates that in all comparable cases this assumption is adequate since solutions from the "14 Rows" are associated with the smaller values of AIC. Both one- and two-dimensional solutions were obtained by IPDA, and the one-dimensional solution was found to be better. The comparison between the unidimensional IPDA and the SCM solutions indicates that the former is a better solution, making it the best solution obtained.

The unidimensional IPDA solution is presented in Figure 4.4. The column points representing the four merit categories are arranged in the expected order. Interestingly, on average School of Dentistry received the most favorable ratings and Medical School the

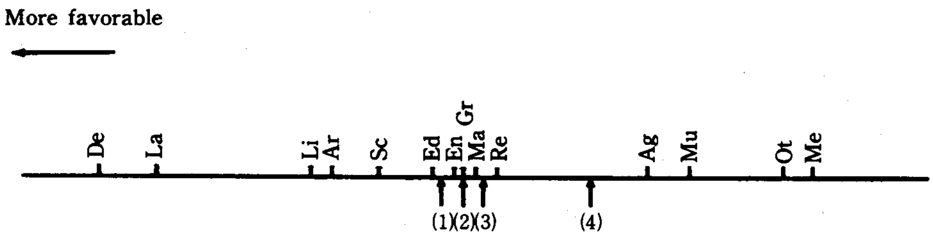

Fig. 4.4 . The Unidimensional Representation of Merit Distribution Data by IPDA

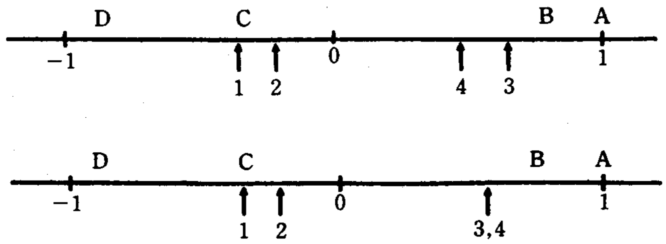

Fig. 4.5 Unconstrained and Order-Constrained Unidimensional Representations of William's Data by IPDA.

Table 4.5
Summary Results for Williams' Data

Saturated Model		329.5 (12)	
Null Model		358.4 (3)	
	Clustering Alternatives	All Distinct 1, 2, 3, 4	3 & 4 Combined 1, 2, (3, 4)
IPDA			
Row-nominal, dim=2		321.8 (8)	
dim=1		319.3 (6)	320.2 (6)
SCM			
Row=nominal, dim=1		327.4 (6)	323.4 (6)
Association Model			
Row=nominal, dim=1		322.3 (8)*	324.2 (8)**

*Obtained by Goodman, 1981
**Obtained by Tsujitani, 1988

least favorable ratings. McGill's Medical School is ranked fifth best among medical schools in North America (tenth in the world) according to some source, while School of Dentistry has no comparable reputation.

(5) *Williams' data.* IPDA solutions were obtained in one and two dimensions. The unidimensional solution turned out to be a better solution. This solution is also better than the corresponding SCM solution. However, in the unidimensional IPDA solution columns 3 and 4 violated the expected order. (See the top configuration in Figure 4.5). The two columns were combined and analyses were repeated. The difference in GOF between IPDA and SCM diminished considerably. However, IPDA still yields a slightly better fit. The last row in Table 4.5 gives the GOF of the association model. The solution in which columns are not order-constrained was obtained by Goodman (1981). The order-constrained solution, on the other hand, was recently obtained by Tsujitani (1988; see also Agresti, Chuang, & Kezouh, 1987). In both unconstrained and constrained cases, IPDA is found to fit to the data better.

5. Concluding Remarks

In this paper performance of IPDA and SCM was systematically compared in representing ordered response categories. This was done by actually fitting them to several real data sets. IPDA was generally found to fit better. Although in this paper results were reported only for five data sets, this tendency also held for numerous other data sets we tried.

When the ordered response categories violate their prescribed order, IPDA tended to give a considerably better fit than SCM. This is because points are not intrinsically order-restricted in IPDA, whereas in SCM intervals representing the ordered response categories on a unidimensional continuum always assume a certain order. It was shown that combining the response categories which violate prescribed orders had the effect of imposing the order restrictions in IPDA (by equating the point locations), while the same operation had the effect of removing the intrinsic order restrictions in SCM (by treating those categories as nominal). Fair comparisons should be either between order-constrained IPDA and order-constrained SCM, or between order-unconstrained IPDA and order-unconstrained SCM. This having been taken into account IPDA was still found to fit better, though only slightly in most cases.

Model comparisons made in this paper were all besed on the AIC statistic, which in turn was based on asymptotic properties of the maximum likelihood estimation. This criterion may not be optimal for small samples (*e.g.*, Williams' data). Methods based on randomization or permutation with possible approximations by Monte-Carlo methods may be used in such situations. The methods, however, depend on the size of contingency tables, which prevents systematic investigations. Perhaps the best strategy is to obtain an approximate randomization (permutation) distribution in each specific situation as the necessity arises.

One advantage of IPDA lies in the possibility of multidimensional representations. When the data requires a multidimensional representation (as in Guilford's data and Bradley, Katti, & Coon's data), the two-dimensional IPDA solution was substantially better than the SCM solution which is restricted to be unidimensional. One way to extend the SCM model is to incorporate the row-specific dispersion parameters, a_i. That is,

$$h_{ij}=[1+exp(-a_i(c_j-x_i'b))]^{-1}$$

The comparison between IPDA and this version of SCM for the data for which IPDA required two-dimensional representations would undoubtely be of interest.

Appendix

Takane (1987) shows that p_{i1} in IPDA can be written as $p_{i1}=(1+\exp(-q_i))^{-1}$, since $d_{ij}^2=(y_i-m_j)^2$, $j=1, 2$. Here

$$q_i=-2(m_2-m_1)x_i'b+(m_2-m_1)(m_1+m_2)-ln(w_1/w_2).$$

The first term on the right hand side has to be equal to $-x_i'b^*$ and the remaining terms to c_1 in SCM. Note first $m_1=v_1'b$ and $m_2=v_2'b$ where v_1 and v_2 were defined in (9) and (10). We thus obtain

$$-2(m_2-m_1)x_i'b=-2(v_2-v_1)'bx_i'b$$
$$=-e'bx_i'b=-x_i'(e'b)b$$

which is equal to $-x_i'b^*$ when $b^*=(e'b)b$ with e defined in (11). If we solve the last equation for b by premultiplying both sides by e', we obtain $b=b^*/(e'b^*)^{1/2}$. Also with s defined in (12),

$$(m_1-m_2)(m_1+m_2)-ln(w_1/w_2)=e'bs'b-ln(w_1/w_2)$$
$$=s'(e'b)b-ln(w_1/w_2),$$

which should be equal to c_1. From this we obtain

$$w_2/w_1=\exp(c_1-s'b^*)=r$$

or $w_2=rw_1$, but since $w_1+w_2=1$,

$$1=w_1+rw_1=(1+r)\ w_1,$$

or $w_1=(1+r)^{-1}$.

References

Agresti, A. (1984). *Analysis of ordinal categorical data.* New York: Wiley.

Agresti, A, Chuang, C., & Kezouh, A. (1987). Order-restricted score parameters in association models for contingency tables. *Journal of the American Statistical Association,* **82**, 619-623.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control.* **19**, 716-723.

Andersen, E.B. (1980). *Discrete statistical models with social science applications.* Amsterdam: North-Holland.

Anderson, J.A., & Philips, P.R. (1981). Regression, discrimination and measurement models for ordered categorical variables. *Applied Statistics,* **30**, 22-31.

Bock, R.D. (1975). *Multivariate statistical methods in behavioral research.* New York: McGraw Hill.

Bradley, R.A., Katti, S.K., & Coons, I.J. (1962). Optimal scaling for ordered categorical data. *Psychometrika,* **27**, 355-374.

Cox, D.R. (1970). *Analysis of binary data.* London: Methuen.

de Leeuw, J. (1983). Models and methods for the analysis of correlation coefficients. *Journal of Econometrics,* **22**, 113-137.

Duran, B.S., & Odell, P.L. (1974). *Cluster analysis : A survey.* New York: Springer-Verlag.

Goodman, L.A. (1979). Simple models for the analsis of cross-classifications having ordered categories. *Journal of the American Statistical Association,* **74**, 537-552.

Goodman, L.A. (1981). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association,* **76**, 320-334.

Goodman, L.A. (1984). *The analysis of cross-classified data having ordered categories.* Cambridge, MA: Harvard University Press.

Guilford, J.P. (1936). *Psychometric methods.* 1st ed., New York: McGraw Hill.

McCullagh, P. (1980). Regression models for ordinal data (with discussion). *Journal of Royal Statistical Society, Series B,* **42**, 109-142.

Ogilvie, J.C., & Creelman, C.D. (1968). Maximum likelihood estimation of receiver operating characteristic curve parameters. *Journal of Mathematical Psychology,* **5**, 377-391.

Takane, Y. (1987). Analysis of contingescy tables by ideal point discriminant analysis. *Psychometrika,* **52**, 493-513.

Takane, Y. (1989). Ideal point discriminant analysis: Implications for multiway data analysis. In

R. Coppi, *et al.* (Eds.), *Analysis of multiway data matrices.* Amsterdam : North Holland, 287-299.

Takane, Y., Bozdogan, H., & Shibayama, T. (1987). Ideal point discriminant analysis. *Psychometrika,* **52**, 371-392.

Thurstone, L.L. (1927). A law of comparative judgments. *Psychological Review,* **34**, 273-286.

Torgerson, W.S. (1958). *Theory and methods of scaling.* New York : Wiley.

Tsujitani, M. (1988). Optimal scaling for association models when category scores have a natural ordering. *Statistics & Probability Letters,* **6**, 175-180.

van Rijckevorsel, J.L.A. (1987). *The implication of horseshoes and fussy coding in multiple correspondence analysis.* Leiden : DSWO Press.

Williams, E.J. (1952). Use of scores for the analysis of association in contingency tables. *Biometrika,* **39**, 274-289.

(Received September, 1989.)