7. MINIRSA (MINI Rectangular Smallest space Analysis)

1. OVERVIEW

Concisely: MINIRSA (MINI Rectangular Smallest Space Analysis) provides internal analysis of two-way data in a row-conditional format of a (dis)similarity measure by a Euclidean distance model using a monotonic transformation of the data.

Following the terminology developed by Carroll and Arabie (1979) MINIRSA may be described as:

Data:	Model:
Two-mode	Euclidean distance incorporating
Two-way	Two sets of points in
Ordinal	One space
Row conditional	The solution is internal
Complete or incomplete	
One replication	

1.1 ORIGIN, VERSIONS AND ACRONYMS

The MINIRSA program included in the MDS(X) series is adapted from Roskam's 1973 release.

1.2 BRIEF DESCRIPTION OF MINIRSA

MINIRSA performs a non-metric multidimensional unfolding analysis. Consider a set of subjects and a set of stimuli where the subjects indicate their preferences for the stimuli (the judgements need not be of preference; any asymmetric relation is acceptable). The aim of the program is to position both stimuli of subjects as points in a space of minimum dimensionality so that, for each subject, the rank order of the distances from his or her point of maximum preference in the space (the "ideal point") to the stimuli matches the subject's preference ordering as closely as possible.

1.3 RELATION OF MINIRSA TO OTHER PROGRAMS IN MDS(X)

MINIRSA analyses preference data by means of an 'ideal point' or 'point-point' model. That is to say that each subject, or "judge" is represented in the solution space as a point positioned at his(her) point of maximum preference. The stimuli are also positioned as points in the same space so that the nearer a point lies to a given subject's ideal point the greater is that subject's preference for it.

By contrast the MDPREF program implements a 'point-vector' model, where the subjects are represented in the solution space as vectors: i.e. directions of increasing preference (which is formally equivalent to having an ideal point at infinity).

MINIRSA is also equivalent to the third phase of PREFMAP except in so far as MINIRSA provides an internal analysis, that is to say that both subject and stimulus points are simultaneously positioned to satisfy the data, whereas in PREFMAP phase 3 the subject points are inserted into a pre-existing configuration of stimulus points. (Note, however, that PREFMAP also provides for a quasi-internal analysis q.v.).

MINICPA is similar to MINIRSA except that it regards the row and column points (the 'subjects' and the 'stimuli') as identical and position only one set of points in the space.

2. DESCRIPTION OF THE PROGRAM

2.1 DATA

MINIRSA takes data in a 'row-conditional' format. In the simplest case, a group of N subjects might be asked to rank in order of preference a set of p stimuli. The judgement may, of course, be a ranking (or rating) in terms of any suitable criterion of which preference is the intuitively most obvious example.

The data matrix, then, consists of N rows each of which reflects a particular subject's order of preference for the stimuli. There are p columns. The various p ways in which these may be presented are detailed below (2.1.1).

MINIRSA does not accept paired-comparisons data as such but will take the row sums of such matrices (see MDPREF, Section 2.1.2).

2.1.1 Ranks or Scores

Preference judgements may be represented for MINIRSA (as in MDPREF and other programs) in four distinct ways. The major distinction is that between a rank and a score. If a subject is asked to write down in his order of preference for five stimuli, he might respond with:

ACDEB

If these letters (or stimulus names) are given numeric values this becomes:

13452

This is the rank-ordering method (analogous to Coombs's I-scales) and means that stimulus 1 is preferred to 3 which is preferred to 4 etc.

Data may be input to MINIRSA in this form by specifying DATA TYPE(1). In various data-collection techniques it may be that the ordering obtained begins with the least-preferred stimulus so that the previous example would in this case be written as: BEDCA, signifying that B is least preferred, followed by E, and so forth. If this is the case then the data should be specified as: DATA TYPE(2).

A different way of representing such data is by the 'score' method. In this method each column represents a particular stimulus and the entry in that column gives the score or rating of that stimulus (for that subject) in his 'scale of preference'. Thus, in our original example the I-scale ACDEB (where A is preferred to C, which is preferred to D etc.) would in this method be represented as follows:

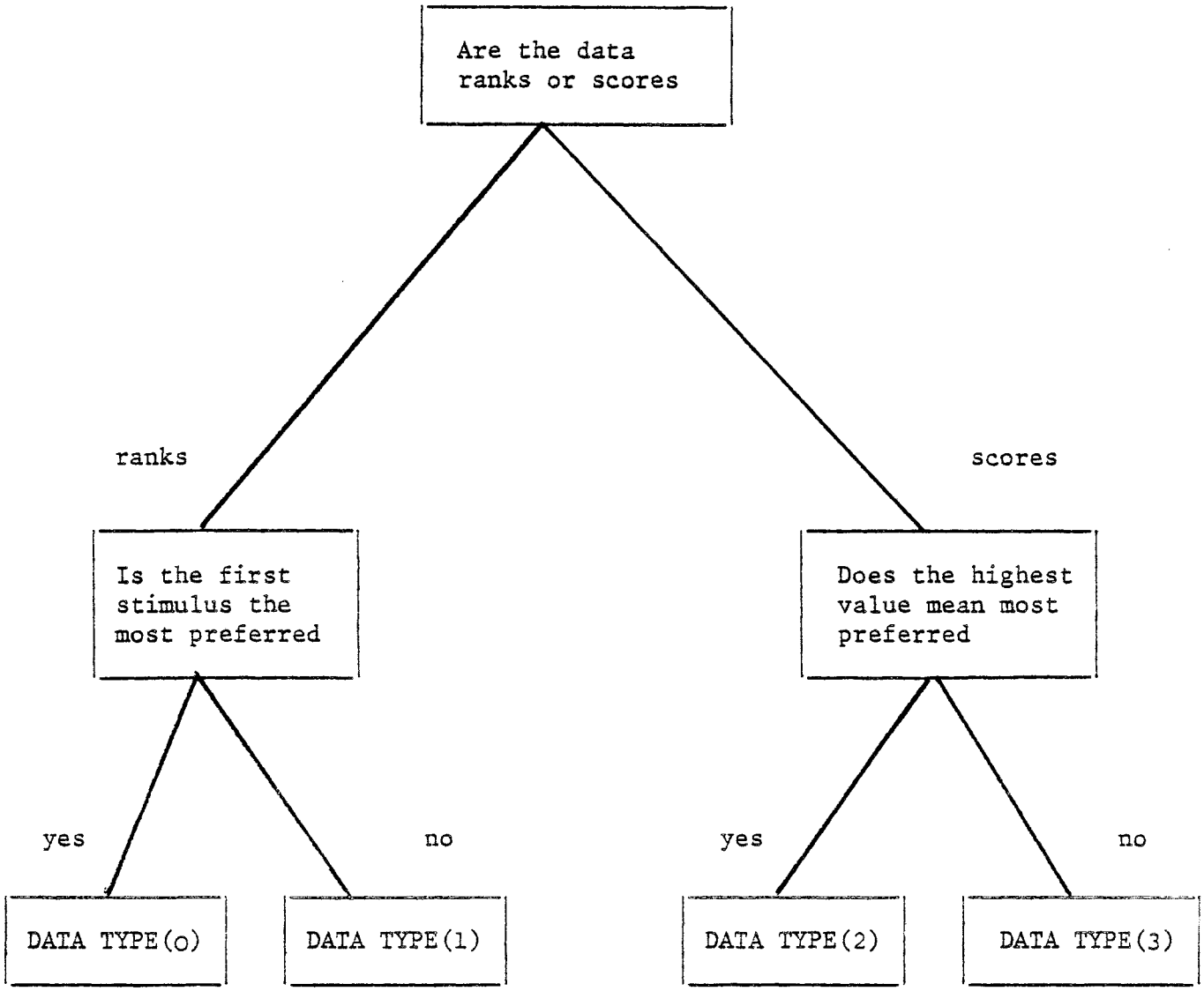
		A	B	C	D	E
subject	i	1	5	2	3	4

In this instance, the lowest number ('1') is used to denote the most preferred stimulus and the highest ('5') to represent the least preferred. This option is chosen by: DATA TYPE(3). Alternatively, the highest number might have been used to represent the most preferred stimulus and if this is so, DATA TYPE(4) should be specified.

(Although in illustrating the score method we have used the number 1 to 5, the data might equally well have been numerical ratings).

Figure 1 provides a simple means of identifying the appropriate DATA TYPE value.

Figure 1


Are the data ranks or scores
ranks
scores
Is the first stimulus the most preferred
Does the highest value mean most preferred
yes
no
yes
no
DATA TYPE(0)
DATA TYPE(1)
DATA TYPE(2)
DATA TYPE(3)

2.2 THE MODEL

Coombs (1964) developed the notion of unidimensional unfolding in which a set of stimuli were so placed along the continuum (the "I-scale") that a subject might be thought of as being located at one point (our 'ideal point') in such a way that his or her preference for the stimuli decreased the further away from the ideal point a given stimulus is situated. In this simple 1-space the fact that the distance from the subject's ideal point to stimulus a was greater than the distance from the ideal point to stimulus b implied that the subject preferred stimulus b to stimulus a. (For a more detailed overview see Appendix 3). The generalisation to spaces of higher dimensionality is intuitively obvious though computationally complex. MINIRSA is the program which performs non-metric multidimensional unfolding in the MDS(X) library.

MINIRSA takes data of the form described and seeks to position both sets of objects - subjects and stimuli - as points in a space of minimum dimensionality. The subjects are positioned at their points of maximum preference: their 'ideal points'. For each subject the distances to the stimuli will reflect the order of preference as revealed by the data: the most preferred stimulus will be the nearest stimulus point to a subject's ideal point, the least-preferred, the farthest away.

Strictly speaking, this will hold only if the data are 'perfect' (i.e. fit the given dimensionality) and for all but minimal STRESS values, some inversions will occur.

It is instructive to consider the contours enclosing areas of equal preference. In MINIRSA these will describe circles around each of the subject points (as contrasted, for instance, with PREFMAP phases I, II, where the contours are ellipses and MDPREF and PREFMAP IV where the "contours" are straight lines perpendicular to the subject's vector).

2.2.1 The Algorithm

1. If the user does not provide one, the program generates an initial stimulus configuration (see Appendix 2.5) in which the subjects are initially placed between their two most preferred stimuli.

2. The configuration is normalised.

3. The distances in the configuration (between each subject and the stimuli) are calculated.

4. The fitting values are next calculated following Kruskal's method of monotone regression.

5. STRESS is calculated.

6. If STRESS has reached zero or an acceptable minimum then the configuration is output as solution. If not, then

7. For each point on each dimension both the direction in which it should move so that STRESS is minimized and the optimal size of that move (the 'step-size') are calculated.

8. The configuration is moved in accordance with (7) and the program returns to step 2.

9. The solution is rotated to principal axes. (A translation of the origin is also allowed).

2.2.1.1 MINIRSA and MINISSA

The MINIRSA algorithm differs from the basic MINISSA algorithm on two major counts.

2.2.1.1.1 The monotonicity requirement

Since at step 5 Kruskal's method of calculating the fitting values is used, the program only enforces the requirement of weak monotonicity on the fitting value. Specifically, this means that different data values may be fit by the same fitting values.

2.2.1.1.2 STRESS

The input data to MINIRSA is considered to be 'row-conditional' (i.e. no comparability is assumed between subjects' rankings). Thus it is inappropriate to calculate STRESS according to the simple $STRESS_1$ formula, but rather a form of $STRESS_2$ is calculated. For each distinct ranking ("I-scale"), the $STRESS_2$ value is first calculated: ($STRESS_2$ is used in preference to $STRESS_1$ in order to prevent the occurrence of degenerate solutions, with fitting values all having the same value). The overall $STRESS_2$ value is then defined as a weighted average of the individual STRESS values.

2.3 FURTHER FEATURES

2.3.1 Missing Data

MINIRSA allows for missing data. The value to be regarded as indicating a missing value should be specified on the PARAMETERS card by means of the MISSING parameter: e.g. if 9 is the code for a missing datum then MISSING(9) is appropriate.

2.3.2 Repeated Orderings

If, as often happens, a number of subjects have provided the same rank-ordering, MINIRSA requires that these be presented in a particular way.

At the end of a card presenting the representative data a frequency count is appended. The INPUT FORMAT card must allow for this. If there are no repeated orderings then this frequency need not be given but the program should be instructed to read a blank field, which it will interpret as indicating an unrepeated ordering. This means that if the data fill the card, a blank card must be inserted between each subject and the frequency read from it.

2.3.2.1 Example

Suppose that five subjects have returned the ranking 'ABCDEF' on a set of six stimuli, whereas only one has indicated 'FEDCBA'. This would be presented to the program thus:

```
col 1                col 16
|-------------------------------------------|
|                                           |
|  .                                        |
|  .                                        |
|  .                                        |
|  INPUT FORMAT        (5X, 6I2, 2X, I1)    |
|  READ DATA                                |
|  .                                        |
|  .                                        |
|  .                                        |
|  _ _ _ _ _ _ 1_2_3_4_5_6_ _ 5             |
|  _ _ _ _ _ _ 6_5_4_3_2_1_ _ _             |
|  .                                        |
|  .                                        |
|  .                                        |
|-------------------------------------------|
```

3. INPUT PARAMETERS

3.1 LIST OF PARAMETERS

Keyword	Default Value	
DATA TYPE	0	0: Data are ranks (I-scales) of column indices in decreasing order of preference. 1: As 0 but in increasing order of preference. 2: Data are scores in order of column indices - high score means low preference. 3: As 2 but high scores mean high preference.
MINIMUM ITERATIONS	6	Sets the minimum number of iterations to be to be performed before convergence test.
MISSING DATA	0	Sets the data value which is to be regarded as missing data.
MATFORM	0	NOTE: only relevant when 'READ CONFIG' is used. 0: The input configuration is punched subjects and stimuli (rows) by dimensions (columns). Subjects are punched before stimuli. 1: The input configuration is punched dimensions (rows) by subjects and stimuli (columns).

3.2 NOTES

1. The {#, N, No} OF SUBJECTS card may be replaced by {#, N, No} OF ROWS.
2. The {#, N, No} OF STIMULI card may be replaced by {#, N, No} OF COLUMNS
3. See section 2.3.2 for details of frequency counts.

3.3 PROGRAM LIMITATIONS

Maximum number of subjects	=	100
Maximum number of stimuli	=	60
Maximum number of dimensions	=	5

3.4 PRINT, PLOT AND PUNCH OPTIONS

The general format for printing, plotting and punching output is described in the Overview. In the case of MINIRSA the particular options are as follows.

3.4.1 PRINT option (output to line printer

Keyword	Form	Description
INITIAL	N × r p × r	Two matrices are produced being the coordinates of the subject points and the stimulus points in the required dimensions.
FINAL	N × r p × r	Similarly, two solution matrices are printed.
DISTANCES	N × N p × p N × p	Three matrices are printed, being: 1. The distances between the subject points. 2. The distances between the stimulus points. 3. The distances between the subjects and the stimuli.
FITTING	N × p	The matrix of disparities (DHAT's).
RESIDUALS	N × p	The matrix of residuals is printed.
HISTORY		This keyword generates an extremely detailed history of the iterative process. Users are warned that this option generates a large amount of output.

By default only the final configurations and the final STRESS value are printed.

3.4.2 PLOT options (output to line printer)

Keyword	Description
SUBJECTS	A plot of the subject points only is produced.
STIMULI	A plot of the stimulus points only is produced.
JOINT	The configuration of subject and stimulus points is plotted.
SHEPARD	The Shepard diagram is produced.
STRESS	A histogram of STRESS values at each iteration is produced.
POINT	The contribution of each subject to the overall STRESS value is plotted.
RESIDUALS	A histogram of residual values is produced.

By default a Shepard diagram and the joint space only are plotted.

3.4.3 PUNCH options

Keyword	Description
SPSS	A file suitable for input to SPSS is produc The following values appear: I : the subject index no. IFR : no. of repeat orderings. J : the stimulus index no. INPUT : the datum corresponding to I,J. FITTING: the corresponding DHAT value. DIST : the solution distance between I & J. RESID : the corresponding residual value. The format of the file is 4I4, 3F10.4.
STRESS	The STRESS values at each iteration are output in a fixed and specified format.

Keyword	Description
FINAL	A card-image file of the final configuration is produced.

4. EXAMPLES

4.1 TEST RUN

```
col 1                    col 16
RUN NAME                 MINIRSA TEST DATA
TASK NAME                46 I-SCALES FROM 5 CONVEX STIMULI
ITERATIONS               80
DIMENSIONS               2
N OF SUBJECTS            46
N OF STIMULI             5
PRINT                    DISTANCES, RESIDUALS
PLOT                     POINT
INPUT FORMAT             (3X,6I1)
READ MATRIX
  <data follow here>
COMPUTE
FINISH
```

BIBLIOGRAPHY

Carroll, J.D. and P. Arabie (1979) Multidimensional scaling, in M.R. Rozenweig and L.W. Porter (eds.) Annual Review of Psychology, Palo Alto, Ca., Annual Reviews.

Coombs, C.H. (1969) A Theory of Data, New York: Wiley.

Coxon, A.P.M. (1974) The mapping of family-composition preferences: a scaling analysis, Social Science Research, 3, 191-210.

Davidson, J.A. (1972) A geometric analysis of the unfolding model: nondegenerate solutions, Psychometrika, 3, 193-216.

Delbeke, L. (1968) Construction of preference spaces, Louvain: Publications of the University of Louvain.

Gleason, T.C. (1969) Multidimensional scaling of sociometric data, Ph.D. thesis, University of Michigan, Ann Arbor, Michigan.

Goldberg, D. and G.H. Coombs (1964) Some applications of unfolding theory to fertility analysis, in Emerging Techniques in Population Research, Proceedings of the 1962 Annual Conference of the Milbank Memorial Fund, New York.

Green, P.E. and F.J. Carmone (1969) Multidimensional scaling: an introduction and comparison of nonmetric unfolding techniques, Journal of Marketing Research, 6, 330-341.

Green, P.E. and V.R. Rao (1972) Applied multidimensional scaling, New York: Holt, Rinehart and Winston.

Niemöller, B. and C. Sprenger (1974) Program MINIRSA: unfolding of preference data and comparable choice data according to Roskam, Program Bulletin No.42, Technisch Centrum F.S.W., Universiteit van Amsterdam.

Roskam, E.E. (1969) Data theory and algorithms for nonmetric scaling, I,II, (stencil) Psychology Laboratory, Mathematische Psychologie, University of Nijmegen, Nijmegen, The Netherlands.

——— (1970) Data theorie en metrische analyse, Ned. Tijdschrift Voor Psychologie, 25, 15-54 and 66-82.

——— (1975) Nonmetric data analysis: general methodology and technique with brief descriptions of miniprograms. Report No.75-MA-13. Nijmegen, The Netherlands: Psychology Laboratory, Mathematische Psychologie, University of Nijmegen.

Singson, R.L. (1973) A multidimensional scaling and unfolding analysis of store image and shopping behavior, Ph.D. thesis, University of Washington, Seattle, Washington.

Young, F.W. and R. Lewyckyj (1979) ALSCAL-4 Users' Guide, Carrboro, N.C.: Data Analysis and Theory Associates.

APPENDIX 1: RELATION OF MINIRSA TO OTHER PROGRAMS NOT IN MDS(X)

Internal multidimensional unfolding analysis, implemented by MINI-RSA, is also implemented by the SSAR-II program in the Guttman-Lingoes series and in Young and Lewyckyj's ALSCAL-4 package (with parameters set so that the measurement level is ordinal and the data type is rectangular and row-conditional).

More general variants are also possible in these packages. The Guttman-Lingoes programs permit other types of conditionality (see Lingoes 1972, pp 57-59) and ALSCAL-4 allows other levels of measurement (see Young and Lewyckyj 1979, p 23).

APPENDIX 2: THE MINIRSA ALGORITHM

A2.1 Data

Let the stimuli be denoted $j,k,\ldots,p$
and the subjects $i,\ldots,N$

$$p^i\,(j,k)$$

is interpreted to mean that subject i prefers j to k.

A2.2 The configuration

Let $\underset{\sim}{X} = \{x_{ia}\}$ $\quad i = 1,\ldots,N$, $a = 1,\ldots,r$

be a configuration of N
subject points in an r-space. Similarly,

$$\underset{\sim}{Y} = \{y_{ja}\} \qquad j = 1,\ldots,p,\ a = 1,\ldots,r$$

be a configuration of stimulus points.

A2.3 The model

In MINIRSA the measurement of the distances in the configuration is restricted to the euclidean metric (see Niemoller and Sprenger, 1974).

Specifically

$$d_{ij} = \sqrt{\sum_a (x_{ia} - y_{ja})^2} \qquad (1)$$

We further define fitting quantities $\hat{d}$ such that

$$\hat{d}_{ij} \leq \hat{d}_{ik} \quad <=> \quad p^i \ (j,k) \tag{2}$$

The STRESS of a given configuration is measured by

$$\text{STRESS}_2 = S_2 = \sqrt{\frac{\sum_j (d_{ij} - \hat{d}_{ij})^2}{\frac{1}{N}\sum_j \sum_j (d_{ij} - \bar{d}_i)^2}} \tag{3}$$

in which

$$\bar{d}_i = \frac{\sum_j d_{ij}}{p}$$

being the mean distance for each subject.

If (3) is zero then

$$d_{ij} \leq d_{ik} \quad <=> \quad p^i \ (j,k)$$

A2.4 The minimization procedure

The procedure is iterative and monitored by a superscript s in parenthesis. The method of steepest descent is used.
At each iteration

$$\underset{\sim}{X}^{(s+1)} = \underset{\sim}{X}^{(s)} - \alpha^{(s)} \underset{\sim}{G}^{(s)} \tag{4}$$

and

$$\underset{\sim}{Y}^{(s+1)} = \underset{\sim}{Y}^{(s)} - \alpha^{(s)} \underset{\sim}{H}^{(s)} \tag{5}$$

The quantity α is the step-size. $\underset{\sim}{G}$ and $\underset{\sim}{H}$ are matrices of gradients. These are determined by evaluating the partial derivatives of S_2 w.r.t. $\underset{\sim}{X}$ and $\underset{\sim}{Y}$.

First we find the partial derivative of S_2 w.r.t. d

$$\frac{\partial S_2}{\partial d} = \sum_{ij} \left[\frac{S_2^i \, (d_{ij} - \hat{d}_{ij})}{N. \;\; S_2 \;\; . \; S_o} - \frac{S_2^i \, . \, (d_{ij} - \bar{d}_i)}{N. \;\;\; S_2 \; . \; NF_2^i} \right] \qquad (6)$$

in which

$$S_o = \sum_j (d_{ij} - \hat{d}_{ij})^2$$

$$NF_2^i = \sum_j (d_{ij} - \bar{d}_i)^2$$

$$S_2^i = \frac{S_o}{NF_2^i}$$

We call the part between the square brackets in (6)

$$q_{ij}$$

The gradients then become:

$$g_{ia} = \frac{\partial S_2}{\partial x_{ia}} = \sum_j q_{ij} \cdot \frac{|x_{ia} - y_{ja}|}{d_{ij}} \cdot \delta^{ij} \, (x_{ia} - y_{ja}) \qquad (7)$$

and

$$h_{ja} = \frac{\partial S_2}{\partial y_{ja}} = \sum_i q_{ij} \cdot \frac{|x_{ia} - y_{ja}|}{d_{ij}} \cdot \delta^{ij} \, (x_{ia} - y_{ja}) \qquad (8)$$

where δ^{ij} is Kronecker delta

A2.5 The initial configuration

The Guttman-Lingoes initial configuration is used in MINIRSA in a form specifically suited to row-conditional data (see Roskam 1969; Lingoes and Roskam 1970).

Let ρ_{ij} be the rank number of the i,j 'th element

ρ_{max} be the maximum rank number

and $\bar{\rho}_{i.}$ be the average rank number over row i

We define a matrix $\underset{\sim}{C}$ in which

$$\underset{\sim}{V}: \quad c_{ij} = c_{ji} = 1 - \frac{\rho_{ij}}{\rho_{max}} \qquad \begin{matrix} i = 1,\ldots,N \\ j = 1,\ldots,p \end{matrix} \tag{9a}$$

$$\underset{\sim N}{D}: \quad c_{ii} = \sum_k \frac{\rho_{ik}}{\rho_{max}} \qquad \begin{matrix} i = 1,\ldots,N \\ k = 1,\ldots,p \end{matrix} \tag{9b}$$

$$\underset{\sim p}{D}: \quad c_{jj} = \sum \frac{\rho_{kj}}{\rho_{max}} \qquad \begin{matrix} j = 1,\ldots,p \\ k = 1,\ldots,N \end{matrix} \tag{9c}$$

Row- and column-sums ($\underset{\sim}{Q}$) of this matrix will be (N+p).

We will use the principal components of $\underset{\sim}{Q}^{-1}\underset{\sim}{C}$ to define an initial configuration. The use of $\underset{\sim}{Q}^{-1}$ is convenient and does not affect the principal components (except for a scale factor).

To avoid a matrix of order (N+p) we write

$\underset{\sim}{Q}^{-1}\underset{\sim}{C}\ \underset{\sim r}{y} = \underset{\sim r}{y}\ \lambda_r$ as:

$$\underset{\sim}{Q}^{-1}\ \underset{\sim N}{D}\ \underset{\sim Nr}{y} + \underset{\sim N}{Q}^{-1}\ \underset{\sim}{V}\ \underset{\sim pr}{y} = \underset{\sim Nr}{y}\ \lambda_r \tag{10a}$$

and

$$\underset{\sim p}{Q}^{-1}\ \underset{\sim}{V}^{1}\ \underset{\sim Nr}{y} + \underset{\sim p}{Q}^{-1}\ \underset{\sim p}{D}\ \underset{\sim pr}{y} = \underset{\sim pr}{y}\ \lambda_r \tag{10b}$$

which are taken from:

$$\underset{\sim}{Q}^{-1}\,\underset{\sim}{Y}_r = \left(\begin{array}{c|c} \underset{\sim}{Q}_N^{-1}\,\underset{\sim}{D}_N & \underset{\sim}{Q}_N^{-1}\,\underset{\sim}{V} \\ \hline \underset{\sim}{Q}_p^{-1}\,\underset{\sim}{V}' & \underset{\sim}{Q}_p^{-1}\,\underset{\sim}{D}_p \end{array}\right) \left(\begin{array}{c} \underset{\sim}{y}_{Nr} \\ \hline \underset{\sim}{y}_{pr} \end{array}\right) \tag{10c}$$

Let $\underset{\sim}{E} = (\underset{\sim}{Q}-\underset{\sim}{D})^{-1}$.

By the definition of ρ_{max}, $\underset{\sim}{D}$ is the sum of N+p values, none of which is >1, and $\underset{\sim}{Q}$ = N+p, so $\underset{\sim}{E}$ is non-negative. Thus we may use $\sqrt{\underset{\sim}{E}}$ to construct a symmetric matrix of order p (p<N).

Using $\underset{\sim}{E}$, we may rewrite 10a and 10b as

$$\underset{\sim}{E}_N\,\underset{\sim}{V}\,\underset{\sim}{y}_{pr} = \underset{\sim}{y}_{Nr}\,\lambda_r \tag{11a}$$

and

$$\underset{\sim}{E}_p\,\underset{\sim}{V}'\,\underset{\sim}{y}_{pr} = \underset{\sim}{y}_{pr}\,\lambda_r \tag{11b}$$

Premultiplying 11a by $\underset{\sim}{E}_p\,\underset{\sim}{V}^1$ we get

$$\underset{\sim}{E}_p\,\underset{\sim}{V}'\,\underset{\sim}{E}_N\,\underset{\sim}{V}\,\underset{\sim}{y}_{pr} = \underset{\sim}{y}_{pr}\,\lambda_r^2 \tag{12}$$

Using the Eckart-Young theorem that for any matrix $\underset{\sim}{A}$

$\underset{\sim}{A} = \underset{\sim}{U}\,\underset{\sim}{\Delta}^{\frac{1}{2}}\,\underset{\sim}{W}^1$ where $\underset{\sim}{U}$ and $\underset{\sim}{W}$ are orthogonal and $\underset{\sim}{\Delta}$ is diagonal.

$\underset{\sim}{A}'\,\underset{\sim}{A} = \underset{\sim}{W}\,\underset{\sim}{\Delta}\,\underset{\sim}{W}^1$

and $\underset{\sim}{A}\,\underset{\sim}{W}\,\underset{\sim}{\Delta}^{\frac{1}{2}} = \underset{\sim}{U}$

Let

$$\underset{\sim}{A} = \underset{\sim n}{E}^{\frac{1}{2}} \underset{\sim}{V} \underset{\sim m}{E}$$

$$\underset{\sim}{W} = \underset{\sim m}{E}^{\frac{1}{2}} \underset{\sim m}{Y}$$

$$\underset{\sim}{U} = \underset{\sim n}{E}^{-\frac{1}{2}}$$

and

$$\Delta^{\frac{1}{2}} = \Delta$$

then (12) may be rewritten

$$\underset{\sim}{A}' \underset{\sim}{A} \underset{\sim}{W} = \underset{\sim}{W} \underset{\sim}{\Delta} \tag{13}$$

Having found $\underset{\sim p}{Y} = \underset{\sim p}{E}^{\frac{1}{2}} \underset{\sim}{W}$ for the eigenvectors $\underset{\sim}{W}$ of $\underset{\sim}{A}'\underset{\sim}{A}$ we find $\underset{\sim n}{Y}$ from $\underset{\sim}{A} \underset{\sim}{W} \underset{\sim}{\Delta}^{-\frac{1}{2}} = \underset{\sim}{U}$ or:

$$\underset{\sim N}{E} \underset{\sim}{V} \underset{\sim p}{E}^{-1} \underset{\sim p}{Y} \underset{\sim}{\Delta}^{-\frac{1}{2}} = \underset{\sim N}{Y}$$

The $\underset{\sim}{C}$ matrix is positive semi-definite. It has one latent root equal to one, (which is also the largest) and a constant associated eigenvector.

We may replace ρ_{max} in (9) by $\bar{\rho}_i$. This makes the row-sums of $\underset{\sim}{V}$ all equal to zero. The rank of $\underset{\sim}{V}$ is $\leq p+1$. With $\bar{\rho}_i$ in (9), $\underset{\sim N}{D} = \underset{\sim N}{Q}$ and $\underset{\sim p}{E}$ may become negative. We may then simplify by assuming $\underset{\sim p}{D} \simeq$ constant and replace by $N+p = \underset{\sim p}{Q}$. Under these assumptions (10) simplifies to

$$\underset{\sim}{V} \underset{\sim pr}{y} = \underset{\sim Nr}{y} (\lambda-1) (p+N)$$

$$\underset{\sim}{V}' \underset{\sim Nr}{y} = \underset{\sim pr}{y} (\lambda-1) (N+p)$$

$$\underset{\sim}{V}'\underset{\sim}{V} \underset{\sim pr}{y} = \underset{\sim pr}{y} (\lambda-1)^2 (N+p)^2$$

Since (N+p) is a scaling factor,

$$\underset{\sim}{V}'\underset{\sim}{V}\ \underset{\sim}{y}_{pr} = \underset{\sim}{y}_{pr}\ \mu^2$$

where

$$\mu^2 = (\lambda-N-p)^2 \quad : \quad \lambda = \mu+N+p$$

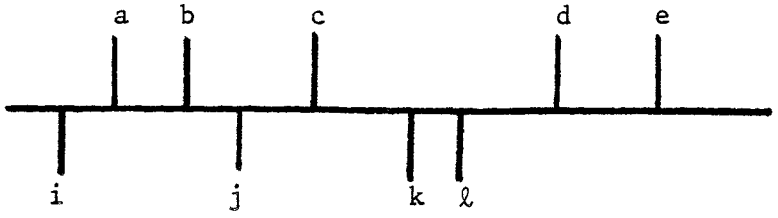
We find the eigenvectors of $\underset{\sim}{V}'\underset{\sim}{V}$ and obtain $\underset{\sim}{y}_{Nr}$ from

$$\underset{\sim}{V}\ \underset{\sim}{y}_{p}\ \underset{\sim}{\Delta}^{-1} = \underset{\sim}{y}_{N}$$

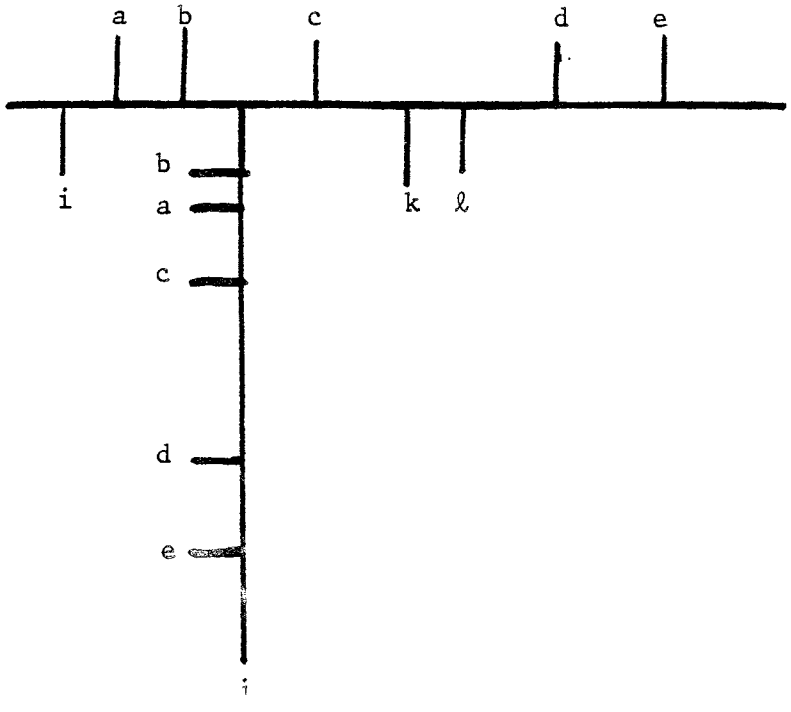
Normalizing the y_r vectors to unit length and multiplying by $\lambda^{\frac{1}{2}}$ gives the desired initial configuration.

APPENDIX 3: UNIDIMENSIONAL UNFOLDING

One well-known method of analysing preference data is based on that of Coombs (1964) known as unfolding. We assume a number of objects (stimuli) with scale values on a one-dimensional scale. On this scale each subject is represented by a single point of maximum preference. This point is called the subjective ideal point. We suppose that the subject's preference for an object is greater the nearer it is situated to his ideal point. The preference rank order is then assumed to be a monotone function of the distance between the object point and the ideal point of the subject, and is termed an "I-scale". The figure below gives a number of objects (a, b, c, d, e) and a number of subjects - their ideal points - (i, j, k, ℓ):



Such a one-dimensional scale is called a J ("joint") scale. An I scale is obtained from a J scale by folding the J scale at the subjective ideal point. The J scale given above generates for subject j the following I scale by using the unfolding technique.



If the J scale is also folded at three other subjective ideal points, the following preference rank orders for the four subjects are found:

subject	preference rank order
i	a b c d e
j	b c a d e
k	c d b e a
ℓ	d c e b a

It is clear that not all preference rank orders can be accommodated on a given J-scale. For instance, it is not possible for a subject to produce ordering e a b c d if the scale values of b, c, and d are between a and e (as in our example).

Application of this model is not restricted to preference orderings of different subjects for different objects. In general, this model may be applied if there is a conditionally weak order

$$p^i(j,k)$$

which means: $j \underset{e}{\geq} k$ for a given i. Symbol $\underset{e}{\geq}$ expresses* an observed weak order relationship for which:

$p^i(j,j)$ (reflexivity)

if

$p^i(j,k)$ and $p^i(k,j)$ then: j is equivalent to k (antisymmetry)

if

$p^i(j,k)$ and $p^i(k,h)$ then $p^i(j,h)$ (transitivity)

---

* $\underset{e}{\geq}$ is used to express an empirical relation, as opposed to the arithmetical $\geq$

Possible interpretations of $p^i(j,k)$ are: the observation that j occurs more often than k, under condition i, or that j is more attractive than k, under condition i, in which 'condition i' could mean 'the judgment of subject i' as well as 'the presentation of stimulus i'. The interpretation of $p^i(j,k)$ could also be: response j in situation i is felt to be more compatible than response k in situation i, etc. (This is not to say that in all these cases the Coombsian distance model gives a correct representation of the observations concerned; it is quite possible that for certain observations another model would be more adequate).

It can easily be seen that it is impossible to reproduce all different rank orders with a single one-dimensional scale. The possible number of orderings of n stimuli is:

$$n\,! = n \times (n-1) \times (n-2) \times \ldots \times 1$$

The number that are compatible with a particular one-dimensional J-scale is much smaller:

$$\binom{n}{2} + 1 = \tfrac{1}{2} \times n \times (n-1) + 1$$

which means that for 5 stimuli only 11 out of 120 possible orderings occur.

If such 'impossible' orderings are found, a one-dimensional unfolding is not possible.