9. MRSCAL (MetRic SCALing)

1. OVERVIEW

Concisely: MRSCAL (MetRic SCALing) provides internal analysis of a two-way data matrix by means of a Minkowski distance model using either a linear or a logarithmic transformation of the data.

Following the categorisation developed by Carroll and Arabie (1979) MRSCAL may be described as:

Data:	One mode	Model:	Minkowski metric
	Two-way		One set of points
	Dyadic		One space
	Unconditional		Internal
	Complete		
	One replication		

1.1 ORIGIN AND VERSIONS OF MRSCAL

The MRSCAL program is the basic metric distance scaling program in Roskam's MINI series. The MRSCAL program in the MDS(X) series is based upon the 1971 and KUNST (1977) versions.

1.2 BRIEF DESCRIPTION OF MRSCAL

The MRSCAL algorithm is a metric counterpart to MINISSA. Its aim is to position a set of stimulus objects as a set of points in a space of minimum dimensionality in much the same way as MINISSA, except that the distances in this space will be a linear (or optionally a logarithmic) function of the dissimilarities between the stimuli. In this it has obvious similarities to 'classic' MDS (Richardson 1938, Young and Householder 1938) and to the linear (metric) scaling procedure developed by Messick and Abelson (1956) and made more widely known by Torgerson (1958). The MRSCAL algorithm however, utilises the iterative procedures which Guttman, Lingoes and Roskam (1971) developed and also allows the user additional options, both in the manner by which the distances in the solution space are measured (see Section 2.2.2) and

in the form of the transformation function linking data to distances in the solution (see Section 2.2.4) which make it both more general and more robust than the original procedures.

1.3 RELATION OF MRSCAL TO OTHER PROGRAMS IN THE LIBRARY

MRSCAL is an exact metric counterpart to MINISSA, differing from it in that it restricts the field of possible transformation of the data to linear (or power) ones.

Output from MRSCAL may be input to PINDIS.

2. DESCRIPTION

MRSCAL is dealt with in Chapter 5 of the Users' Guide.

2.1 DATA

MRSCAL accepts as input the lower triangle (without diagonal) of a square symmetric data matrix. Each entry of this matrix will be a measure of the (dis)similarity between the row-element and the column element. If the linear transformation option is chosen it should be borne in mind that product moment correlations and covariances are not generally acceptable in that they are only monotonically (and not linearly) related to distance.

The aim of the algorithm is to position these elements as points in a space of minimum dimensionality such that a STRESS-like measure of departure from perfect fit (the coefficient of alienation) between the (linearly) rescaled data and the distances in the solution is minimised. A perfect fit occurs if a linear (or logarithmic) transformation of the data is found which is a set of actual distances.

2.1.1 Example

Benjamin (1958) collected data on the social mobility of some 2600 subjects using thirteen occupational categories. Macdonald, who is investigating the notion of social distance, uses the index devised by Blau and Duncan (1967, p.43) to measure the dissimilarity in mobility between occupational groups. (For a fuller description of this index see section 2.3.3.4 of the Users' Guide). The measure, writes Macdonald (1972, pp. 213-14) may be interpreted as "the percentage of the sons of (group) A that would have to be reallocated jobwise for the sons of A to match the sons of B". He assembles the index values into a lower diagonal matrix, and these are included in the examples described in section 4. The scaling solution is discussed at length in Macdonald's article.

2.2 THE ALGORITHM

The program proceeds as follows.

1. An initial configuration is input (or one may be generated by the program (see 2.2.1 below)).

2. The configuration is normalised.

3. The inter-point distances are calculated according to the Minkowski metric chosen by the user (see 2.2.2 below).

4. A set of fitting quantities are computed that are

   i) a linear (or power) transformation of the data; and

   ii) a least-squares best-fit to the distances (for details see Appendix 2.)

5. The coefficient of alienation between the fitting-quantities and the distances is computed.

6. A number of tests is performed to determine whether the iterative process should continue; e.g. Is STRESS sufficiently low? Has the improvement in STRESS over the last few iterations been great enough to warrant continuing ? Has a specified maximum number of iterations been performed ?

7. If not, then the gradient is computed. This gives for each point on each dimension the direction in which that point should be moved on that dimension in order that STRESS be minimized.

8. If the gradient is zero then the configuration is output as solution.

9. If not, then the points are moved in accordance with (7) and the program returns to step 2.

2.2.1 Initial configuration

The user may provide a starting configuration by means of the control card READ CONFIG, with its associated INPUT MEDIUM and INPUT FORMAT cards. In this case a coordinate for each point on each dimension is input. This may be done either by stimuli (rows) by dimensions (columns) or dimensions(rows) by stimuli (columns). In this latter case the parameter MATFORM should be given the value 1 on the PARAMETER 1 card.

If this is not done, however, then the program constructs an initial configuration from the original data by the Lingoes-Roskam procedure which, as has often been shown, is a good initial approximation of a solution and also has certain desirable geometrical properties.

2.2.2 Distances in the configuration (UG App. 2.2)

The user may choose the way in which the distance between the points in the configuration is measured by means of the MINKOWSKI parameter. The default value 2 provides for the ordinary Euclidean metric where the distances between two points will be the length of the line joining them. The user may specify any value for the parameter. Commonly used values, however, include 1, the so-called 'city-block' or 'taxi-cab' metric where the distance between the two points is the sum of the differences between their co-ordinates on the axes of the space, and infinity (in MRSCAL approximated by a large number (>25)) the so-called 'dominance' metric when the largest difference on any one axis will eventually come to dominate all others. (Users are warned that high values of MINKOWSKI are liable to produce program failure due to overflow).

2.2.3 STRESS and the coefficient of alienation

The family of STRESS formulae for the MINI series is based on the sum of the squared differences between the fitting-values and the distances. In MRSCAL, since the fitting-values are at interval level, a product-moment form is applicable, represented by MU which is the

correlation between the distances and the fitting-values, and is hence a measure of goodness of fit. In addition, a related badness of fit measure very similar to STRESS is calculated, known as the coefficient of alienation, K. The two measures used in MRSCAL are related by:

$$K = \sqrt{(1-MU^2)}$$

2.2.3.1 Angle factor and step-size

At step 7, the algorithm computes the direction in which each point should be moved in order to reduce STRESS. This is done by calculating the partial derivative of STRESS with respect to each point - the negative gradient. It is also important, however correctly, to compute the optimal amount of movement in that direction. This is the so-called 'step-size'. This step-size may be changed at each iteration. These changes are monitored by the 'angle factor', which is in effect the cosine of the angle between successive gradients, i.e. the correlation between them. This ensures that, as the program moves towards convergence, and the gradient becomes less steep the step-size will decrease, so as to minimize the possibility of overshooting a minimum STRESS value. MRSCAL prints out at termination the final angle factor. At this stage the value ought to be very small if it is large, then more iterations should be attempted.

2.2.4 Linear and logarithmic transformations

The most common use of MRSCAL is to find a linear transformation of the data which best fits a configuration of points in the chosen dimensionality. The program will also, however, perform an analysis using logarithmic transformations of the data values. In this case the Shepard diagram will show a smooth exponential curve. The user must specify which transformation is required. If no PARAMETERS card is read and/or no specification of the transformation made, then no analysis will be performed.

2.3 FURTHER FEATURES

2.3.1 The CRITERION parameter

In step 6 of the algorithm a number of stopping tests are performed. One of these involves calculating the improvement in fit between the present and the previous iteration. If the improvement is less than the value given by CRITERION on the PARAMETERS card, then the process is terminated and the current configuration is output as solution. A large value for CRITERION will have the effect of stopping the iterative process earlier than would otherwise be the case. This allows the user to make more cheaply a number of exploratory analyses.

2.3.2 The final configuration

When the iterative process is terminated, the current configuration is output as the solution. If the metric is euclidean (i.e. MINKOWSKI (2)) then the configuration is rotated to principal axes. It should be noted that these axes are arbitrary from the point of view of interpretation, but have certain desirable geometric properties. In particular the coordinates of the points on the axes are uncorrelated. Furthermore it is often helpful in deciding on the 'correct' dimensionality of the solution to notice how much variation is associated with each axis. This variation is given in the output by the value SIGMA which is the standard deviation of the coordinates on each axis.

2.3.3 Dimensionality

As a general rule solutions should be computed in a number of dimensionalities. Since a perfect fit will be obtained in n-2 dimensions the trial dimensionalities should always be in dimensionalities less the n-3. As a guide to the choice of trial dimensionalities it is recommended that the product of stimuli × dimensions should be less than half the number of data elements.

A further method is one superficially similar to the 'scree' test of factor analysis. This involves examining the plot of stress by dimensionality. Since MU is a measure of goodness of fit the plot will show an ascending function and the elbow test for appropriate dimensionality may be performed. The 'appropriate' dimensionality, i.e. one which interpretation may be attempted, is that at which the graph shows an 'elbow', i.e. where the addition of extra dimensions is otiose.

3. INPUT PARAMETERS

3.1 LIST OF PARAMETERS

Keyword	Default Value	Function
DATA TYPE	0	0: The data are similarities 1: The data are dissimilarities.
LINEAR TRANSFORMATION	0	0: Linear transformation is not performed 1: Linear transformation is performed.
LOG TRANSFORMATION	0	0: Logarithmic transformation is not performed 1: Logarithmic transformation is performed.
CRITERION	0.00001	Sets the criterion value for terminating the iterations.
MINKOWSKI	2	Sets the Minkowski metric for the analysis.
MATFORM	0	(ONLY RELEVANT WHEN 'READ CONFIG' IS USED) 0: The input configuration is punched: stimuli (rows) by dimensions (columns). 1: The input configuration is punched: dimensions (rows) by stimuli (columns).

N.B. Either LINEAR TRANSFORMATION or LOG TRANSFORMATION must be specified

```
col 1                          col 16
RUN NAME                       OCCUPATIONAL DISSIMILARITY DATA
TASK NAME                      AS IN SEC. 2.1.1
N OF STIMULI                   13
DIMENSIONS                     5 to 1
PARAMETERS                     LINEAR(1), DATA(1)
INPUT FORMAT                   (12F5.0)
COMMENT                        THE GROUPS ARE:
                               1.  FARMERS
                               2.  AGRICULTURAL WORKERS
                               3.  HIGHER ADMIN ETC.
                               4.  OTHER ADMIN ETC.
                               5.  SHOPKEEPERS
                               6.  CLERICAL WORKERS
                               7.  SHOP ASSISTANTS
                               8.  PERSONAL SERVICE
                               9.  FOREMEN
                               10. SKILLED WORKERS
                               11. SEMI-SKILLED WORKERS
                               12. UNSKILLED WORKERS
                               13. ARMED FORCES (OR)
READ MATRIX
51.1
71.4  75.8
63.0  52.7  36.9
58.6  57.7  40.8  32.3
67.0  55.6  38.6  17.7  38.2
63.4  52.3  39.4  13.4  27.8  27.3
54.5  43.3  55.5  29.3  41.1  35.0  23.5
71.2  47.5  56.5  26.2  41.0  35.6  21.1  36.1
65.2  44.3  62.3  33.0  45.1  42.1  27.4  32.0  14.7
65.7  43.0  68.2  39.0  50.8  47.3  33.3  36.0  15.7   8.4
60.1  34.2  69.4  39.8  51.9  47.2  35.5  30.4  23.9  21.1  19.3
66.7  41.9  62.7  36.1  44.6  42.7  29.0  35.9  21.2  20.7  18.4  18.9
PLOT                                                SHEP(1)
COMPUTE
FINISH
```

3.2 NOTES

1. The card $\begin{Bmatrix}\#\\N\\NO\end{Bmatrix}$ OF SUBJECTS is not valid with MRSCAL.

2. The card $\begin{Bmatrix}\#\\N\\NO\end{Bmatrix}$ OF STIMULI may be replaced by

   the card $\begin{Bmatrix}\#\\N\\NO\end{Bmatrix}$ OF POINTS

3. a) The program expects input to be in the form of the lower triangle of a matrix of real (F-type) numbers.

   b) The INPUT FORMAT should read the longest, i.e. last, row of this matrix.

4. Maximum no. of stimuli = 80
   Maximum no. of dimensions = 8

3.3 PRINT, PLOT AND PUNCH OPTIONS

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

3.3.1 PRINT options (output to line printer)

Option	Form	Description
INITIAL	p × r matrix	Initial configuration, either generated by the program or printed by the user (p = no. of stimuli, r = no. of dimensions).
FINAL	p × r matrix	Final configuration, rotated to principal components.
DISTANCES	lower triangular, with diagonal	Solution distances between points, calculated according to MINKOWSKI parameter.
FITTING	lower triangular, with diagonal	Fitting values: the disparities (DHAT) values.
RESIDUALS	lower triangular, with diagonal	The difference between the distances and the disparities.

By default only the final configuration and the final STRESS values are printed.

3.3.2 PLOT options (output to line printer)

Option	Description
INITIAL	Up to r(r-1)/2 plots of the initial configuration. (r = no. of dimensions).
FINAL	Up to r(r-1)/2 plots of final configuration (r = no. of dimensions).
SHEPARD	The Shepard diagram of distances plotted against data. Fitting values are shown by *, actual data/distance pairs by 0.
STRESS	Plot of STRESS by iteration.
POINT	Histogram of point contributions to STRESS.
RESIDUALS	Histogram of residual values (logged).

By default, only the Shepard diagram and the final configuration will be plotted. Configuration plots are calibrated both from 0 to 100 and from 0 to the maximum coordinate value.

3.3.3 PUNCH options

Option	Description
SPSS	Outputs I (Row index), J (Column index) and corresponding DATA, DISPARITIES, DISTANCES, RESIDUALS values in the format: (2I4, 4F10.0).
FINAL	Outputs final configuration as stimulus (row) by dimension (column) matrix. Each row is prefaced by the stimulus number. Format: (I4,rF9.6) where r is the number of dimensions.
STRESS	Outputs STRESS value by iteration.

By default, none of these options is produced.

4. EXAMPLES

4.1 TEST RUNS

```
col 1                   col 16
-------------------------------------------------------
RUN NAME                8 POINT ZERO STRESS DATA
TASK NAME               AS MADE FAMOUS BY USERS' GUIDE
N OF STIMULI            8
DIMENSIONS              2
INPUT FORMAT            (7F4.0)
PARAMETERS              LINE(1), DATA(1)
READ MATRIX

  <data>

PRINT                   ALL
PLOT                    SHEP (2)
COMPUTE
FINISH
-------------------------------------------------------
```

APPENDIX 1: RELATION OF MRSCAL TO SIMILAR PROGRAMS OUTSIDE MDS(X)

The earliest work in MDS assumed that the data dissimilarities were direct estimates of Euclidean distances, and solved for the coordinates of the space that generated them. This so-called "classic MDS" thus assumes the distances are at the ratio level of measurement. Later developments (Messick and Abelson, 1956) assumed that the data were "relative" distances - i.e. a linear function of the solution distances, thus implying interval level of measurement - and therefore had to solve additionally for the "additive constant" necessary to turn the data into distance estimates (see Appendix 3). A surprisingly robust procedure for implementing such "linear" or metric scaling is described in detail in Torgerson 1958.

Similar procedures to those provided by MRSCAL are implemented in the following package and programs:

(1) KYST (the successor to the original general purpose package known as MDSCAL) provides options for specifying linear and power transformations relating data to the solution distances, and thus implement linear and logarithmic scaling respectively.

(2) ALSCAL-4 (the successor to POLYCON and TORSCA) also allows the user to specify ratio or interval levels of measurement, which also implement classical and linear scaling respectively. There is an additional facility for the user to specify a polynomial in degree 1 to 4 as the nearest equivalent to a logarithmic transformation.

APPENDIX 2: THE MRSCAL ALGORITHMS

This appendix is based on Roskam 1972 which is used with permission.

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

be a matrix of coordinates of n points in r dimensions.
The Minkowski distance between two points is

$$d_{ij} = \sqrt[u]{\sum_a | x_{ia} - x_{ja} |^u} \tag{1}$$

where u is the Minkowski parameter.

Let $\underset{\sim}{\Delta} \equiv \{\delta_{ij}\}$ be a matrix of observed dissimilarity values.
The matrix is symmetrical and may not contain missing values.
(Similarly $\underset{\sim}{\sum}$ denotes a matrix of similarities).

Let $\underset{\sim}{D}^o \equiv \{d^o_{ij}\}$ be a matrix of values which is obtained by applying any admissible transformation to $\underset{\sim}{\Delta}$

We choose $\underset{\sim}{D}^o$ to minimize

$$\sum_{ij} (d_{ij} - d^o_{ij})^2 \tag{2}$$

The minimizing d^o is denoted $\hat{d}$.

The STRESS of a configuration S1 is

$$\text{STRESS}_1 = \sqrt{\frac{\sum_{ij} (d_{ij} - \hat{d}_{ij})^2}{\sum_{ij} d^2_{ij}}} \tag{3}$$

If the admissible transformations include a free choice of unit, the $\hat{d}$ has the property

$$\sum_{ij} (d_{ij} - \hat{d}_{ij}) \hat{d}_{ij} = 0 \tag{4}$$

Hence $STRESS_1$ is equal to the coefficient of alienation

$$K = \sqrt{\frac{(\sum_{ij} d_{ij} \hat{d}_{ij})^2}{\sum_{ij} d_{ij}^2 \sum (\hat{d}_{ij})^2}} \tag{5}$$

which is the sine of the angle between $\underset{\sim}{D}$ and $\underset{\sim}{D}^o$ when considered as vectors in a $n(n-1)/2$ dimensional space.

Alternatively MU, the coefficient of monotonicity is

$$\mu = \sqrt{1 - K^2} = \sqrt{1 - S^2}$$

The MRSCAL procedure is double-phase

The first phase improves a configuration $\underset{\sim}{X}$ so as to minimize K for a given $\underset{\sim}{D}^o$.

The second phase finds, for a given $\underset{\sim}{D}$ the transformation $\underset{\sim}{D}^o$ which minimizes K.

The phases are repeated iteratively. K is evaluated after the second phase.

The first phase is itself iterative, at the end of which MU is evaluated.

Iterations are indexed superscript (s): the iterations within the first phase by t.

The first phase

In the first phase we seek to minimize S. The formula from which $X^{(s)}$ is obtained by finding the derivatives of K w.r.t. $\underset{\sim}{X}$ and setting these to zero.

Let x^t be chosen so that

$$\sum_{ij} \left(d_{ij} - \alpha^t \hat{d}_{ij}^{s-1} \right) \alpha^t \hat{d}_{ij}^{s-1} = 0 \qquad (6)$$

$$d_{ij}^{t=0} = d_{ij}^{s-1}$$

$$x_{ia}^{t=0} = x_{ia}^{s-1}$$

$$\alpha^{t=0} = 1$$

At the end of the first phase - i.e. at or short of convergence of the t iterations we have

$$d_{ij}^{s} = d^{t_{max}}$$

$$x_{ia}^{s} = x_{ia}^{t_{max}}$$

Now let

$$K^t = \sqrt{1 - \frac{(\sum_{ij} d_{ij} \hat{\alpha d_{ij}})^2}{\sum_{ij} d_{ij}^2 \sum \alpha d_{ij})^2}} \qquad (7)$$

where

$$d_{ij} = d_{ij}^{t} \text{ and } d_{ij}^{o} = \hat{d}_{ij}^{s-1}$$

Using (6) we find $K^t = S^t$

$$K^t = \text{STRESS}_1^t = \sqrt{\frac{\sum_{ij} (d_{ij} - \alpha\hat{d}_{ij})^2}{\sum_{ij} d_{ij}}} \tag{8}$$

where

$$d_{ij} = d_{ij}^t; \quad \hat{d} = \hat{d}^{s-1}$$

Note the $\text{STRESS}_1^t = \text{STRESS}$ iff $t = 0$

Since D^o is invariant under multiplication by a constant α and since $\tilde{K}^t$ is independent of α we may minimize K^t using (8). Note that $K^t = S^t$ if α is chosen so that (6) is satisfied.

Differentiating K^t (omitting the t index for simplicity)

$$\frac{\partial K}{\partial d_{ij}} = \left(\frac{1}{L} (d_{ij} - \alpha\hat{d}_{ij}) - \frac{1}{N} d_{ij} \right) \tag{9}$$

where

$$L = \sum_{ij} (d_{ij} - \alpha\hat{d}_{ij})^2$$

$$N = \sum_{ij} d_{ij}^2$$

i.e. $K = \frac{L}{N}$

Differentiating d_{ij} w.r.t. x_{ka}

$$\frac{\partial d_{ij}}{\partial x_{ka}} = \left(\delta^{ki} - \delta^{kj} \right) \left(\left| \frac{x_{ia} - x_{ja}}{d_{ij}} \right| \right)^{u-1} \text{sign } (x_{ia} - x_{ja}) \tag{10}$$

where

$\delta^{ki} = 1$ if $k = i$ and $\delta^{ki} = 0$ otherwise.

Define

$$w_{ija} \equiv \left\{ \left| \frac{x_{ia} - x_{ja}}{d_{ij}} \right| \right\}^{u-2} \qquad (i \neq j) \tag{11}$$

w_{iip} is any arbitrary value; we choose $w_{iip} = 1$

Combining (9) (10) and (11) and expanding $\delta^{ki}\ \delta^{kj}$ we obtain the negative gradient.

$$-\frac{\partial K}{\partial x_{ia}} = \frac{K}{L} \sum_j \left(1 - S^2 - \frac{\alpha \hat{d}_{ij}}{d_{ij}}\right) w_{ija} \left(x_{ia} - x_{ja}\right) \tag{12}$$

Next define

$$\gamma_{ja} = \frac{1}{\alpha} \left[\delta^{ij} \sum_k \frac{\alpha \hat{d}_{ik}}{d_{ik}} w_{ika} - \frac{\alpha \hat{d}_{ij}}{d_{ij}} w_{ija} + (1 - K^2)\, w_{ija} \right] \tag{13}$$

And

$$U = \sum_{ij} d\, d_{ij} \Big/ \sum d_{ij}^2 = \frac{1}{\alpha} (1 - K^2) \tag{14}$$

So

$$\gamma_{ija} = \delta^{ij} \sum_k \frac{\hat{d}_{ik}}{d_{ik}} w_{ika} - \frac{\hat{d}_{ij}}{d_{ij}} w_{ija} + U\, w_{ija} \tag{15}$$

Using 13 and 15 we may rewrite 12 as

$$-\frac{\partial K}{\partial x_{ia}} = \frac{\alpha K}{L} \sum_j \gamma_{ija} \left(x_{ia} - x_{ja}\right) \tag{16}$$

Note that the diagonal γ_{jja} may be freely chosen. In (16) they have no effect since they are multiplied by $(x_{ia} - x_{ja}) = 0$. Setting to zero, we obtain

$$x_{ia} = \frac{1}{\sum_j \gamma_{ija}} \sum \gamma_{ija}\, x_{ja} \tag{17}$$

which, using (11) and (15) is equal to

$$x_{ia} = \frac{1}{U \sum_j w_{ija}} \sum_j \gamma_{ija} x_{ja} \tag{18}$$

Following Guttman we solve iteratively by setting

$$x_{ia}^{t+1} \left(\frac{1}{U \sum_j w_{ija}} \sum_j \gamma_{ija} x_{ja} \right)^{s-1,t} \tag{19}$$

where s identifies $\hat{d}_{ij} = \hat{d}_{ij}^{s-1}$ in the formula for γ and U

and t identifies $x_{ia} = x_{ia}^{t}$ there and in the formulae for w_{ija}

and for d_{ij}. Note that α is a dummy coefficient which drops out

when the derivative is put to zero.

For an euclidean metric $w_{ija} = 1$ and (19) simplifies accordingly.

The process in (19) seems to converge, but for unknown reasons.

The second phase

The second phase minimizes

$$\sum (d_{ij}^{s} - \alpha \hat{d}^{s-1} - \beta)^2 \qquad s = 1,... \tag{20}$$

This is a simple regression problem solved by

$$\alpha = \alpha^s = \frac{\sum_{ij} (d_{ij} \hat{d}_{ij} - \bar{d}\,\bar{\hat{d}})}{\sum (\hat{d}_{ij} - \bar{\hat{d}}^2)} \qquad \begin{matrix} (\hat{d} = \hat{d}^{s-1} \\ d = d^s) \end{matrix} \tag{21}$$

where $\beta = \beta^s = \bar{d} - \alpha^s \bar{\hat{d}}$

where $\bar{(\ )}$ denotes the arithmetic mean over i,j.

In order to begin the process $\underset{\sim}{X}^{s=1}$ (the initial configuration) is defined following Guttman and Lingoes (1971)

$\underset{\sim}{D}^{o\ s=0}$ is set to $\underset{\sim}{\Delta}$

All d_{ij} are set to max d^{o}_{ij}

The matrix $\underset{\sim}{C} \equiv$

$$c_{ij} = \delta^{ij} \sum_{k} \left(\frac{d^{o}_{ik}}{a} - \frac{d^{o}_{ij}}{a} \right) \qquad (2^{\cdots}$$

The initial configuration consists of the first r dimensions of a principal components rotation of $\underset{\sim}{C}$.