N. CONPAR (CONcordance PARTitioning: Points-of-View analysis)

1. OVERVIEW

Concisely: CONPAR is a modern implementation of Tucker & Messick's "Points of View" (PoV) Analysis, analysing a 3-way, 2-mode set of dis/similarity (symmetric proximity) matrices (such as those submitted to INDSCAL) in a 2-step process: (i) identifying groups of individual subject matrices with concordant proximity structures, and (ii) scaling each group's averaged data separately, using MINISSA.

DATA: 3-way, 2-mode dissimilarities
TRANSFORMATION: Nonmetric, invariant under monotone transformations
MODEL: (i) Minimizing partition diameter for a specified number of groups (BBDIAM)
(ii) Euclidean distance (MINISSA)

1.1 BRIEF DESCRIPTION OF CONPAR and "POINTS-OF-VIEW" ANALYSIS

The Tucker-Messick (1963) "Points of View" model was the first viable method for representing "individual differences" in scaling. In the original Tucker-Messick model, the 3W2M data were first factor analysed to give an empirical clustering of the subjects, and the averaged dissimilarity matrix for each cluster ("viewpoint") was then separately scaled. Note that no allowance is made in this model for an overall Group Stimulus Space and the axes of the scalings are not uniquely oriented, as in the INDSCAL model (qv) (Carroll & Chang 1970) which succeeded PoV, although 3-mode Factor Analysis variants of PoV have subsequently been developed (Tzeng and Landis 1978) which do allow for a Group Space.

In this program there are two stages

(i) Creating an efficient clustering of subjects The BBDIAM procedure (see Section XX and Program Notes) is used to create an efficient set of (user-specified number of) clusters of individuals. The program takes as input a set of N matrices each of which is a lower triangular matrix (of order p) of (dis)similarity judgments/measures between the p stimuli and looks for a set of best-fitting proximity matrices, by partitioning respondents into groups who share comparable proximity matrix structures. (It is therefore best restricted to the exploration of data in which a number of distinct groups are likely to exist).

(ii) Scaling each group ("point of view) separately The option is given within the program for each of the pooled group matrices generated at stage (i) to be scaled using MINISSA, thus offering a series of representations of the stimulus space, one for each of the groups identified. There is thus no assumption that the subjects' data refer to a single, common Group Space, and the group scalings may be of different dimensionality and have no dimensional uniqueness properties (unlike INDSCAL).

PoV is therefore primarily an exploratory process used to detect homogeneous groups whose members share a common viewpoint.

1.2 RELATION TO OTHER PROCEDURES IN NewMDSX

1.2.1 INDSCAL

CONPAR is most similar to INDSCAL, and accepts the same 3W2M dis/similarity data. PoV/CONPAR explains a subject's dis/similarity judgments as arising from his or her membership of a group (cluster) all of whom share a common perspective, which can be represented as a configuration of stimulus points. Individual differences are thus thought of as arising from different group membership. But PoV makes no assumption whatever about how these different group configurations relate to each other, and there is therefore no single, common "Group Stimulus Space" The INDSCAL model, by contrast, explains differences between subjects' judgments by a variant of the distance model, where the stimuli are located in a single 'master' Group Stimulus

Space common to all subjects, but whose dimensions are weighted differentially by each subject. Thus in INDSCAL, individual differences are explained by different patters of dimensional weighting.

CONPAR offers an alternative approach by making no assumptions about either the dimensionality of the MDS solutions, or even the appropriate metric (Euclidean, city block, etc.), and the option in the program to scale using MINISSA is simply a useful expedient for representing each “point of view”. The User may decide to scale any of the CONPAR pooled group matrices by whatever model is deemed appropriate.It is offered as a preprocessing tool to be applied prior to fitting MDS models to potentially different groups of subjects.
In addition, the content of the various partitions of the subject matrices produced by CONPAR may be identifiable in the graphic displays of the Subject Space resulting from applying INDSCAL to the same set of input matrices.

1.2.2 PINDIS

If the user wishes to compare and contrast the separate group configurations produced by CONPAR’s MINISSA scalings, this can be done using PINDIS (q.v.). In effect, this allows PoV analysis to be supplemented by a Group (Centroid) Space, and an examination of the ways in which the separate group configurations relate to each other (e.g. to answer the question: Does the Configuration for group X resemble that of group Y more than it does group z?).

1.2.3 OTHER PROGRAMS

Meulmann (1996) describes a similar procedure to CONPAR

2 DESCRIPTION OF THE PROGRAM

2.1 DATA

As for INDSCAL, the usual data-collection scenario is that a group of subjects has been asked to assess the (dis)similarity between a set of objects. To explore the relationships between the various sets of judgments. CONPAR assumes that subjects can divided into subgroups, each sharing a similar view of the space represented by the stimuli with which they have been presented in arriving at their judgments and it seeks to recover the set of subgroups with the most in common in terms of their perceived spaces. Again, as with INDSCAL, there is no reason why the judgments of (dis)similarity should come from "real" individuals. They may be different occasions, methods, places, groups etc., in which case they are often referred to as 'pseudo-subjects'.

2.2 THE ALGORITHM

The objective is to identify groups of subject dis/similarity matrices that have comparable/similar structural properties. This is achieved by first collapsing three-way dissimilarity data into a subject x subject pairwise dissimilarity matrix, using a non-metric concordance measure based on gradient information. This index (Brusco, 2004; Hubert, 1987) is invariant under monotone transformations of the original matrices, and is able to capture their internal structural properties.

The resulting matrix is subsequently partitioned, identifying an upper bound by biased-sampling complete link cluster analysis, and then using a branch-and-bound algorithm to minimize the 'partition diameter', that is, the maximum of the cluster 'diameters' (pairwise dissimilarity values for the groups). The value of this process is that the individual subject matrices can then more safely be pooled for further analysis, and the groups to be considered are not constrained at all in their relative sizes. The same partitioning process can also be directly applied to a single matrix using BBDIAM (Brusco, 2003; Brusco and Cradit, 2005).

When running CONPAR in NewMDSX, once partioning into a specified number of groups is complete, an option is automatically offered to apply MINISSA to the pooled group matrices. The pooled group matrices may be separately saved for further analysis.

3. USING CONPAR

It is necessary to specify the number of SUBJECTS and the number of STIMULI represented in the individual subject matrices.

Crucially, the number of CLUSTERS to be identified by the partitioning process must be given by the CLUSTERS command. Only one partitioning at a time is permitted.

The N dis/similarity matrices of order p must follow the READ MATRIX command sequentially, according to the DATA TYPE parameter specified. These will automatically be converted to dissimilarities if necessary, for partitioning by CONPAR .

DIMENSIONS

For input to CONPAR, the number of DIMENSIONS specified is employed for the optional application of MINISSA to the pooled matrices for the clusters identified. Usually between two- and four-dimensional solutions will be adequate for any reasonable data set.

INPUT COMMANDS

Keyword	Function
N OF STIMULI [number]	Number of stimuli for analysis
N OF SUBJECTS [number]	Number of subjects for which data are to be input
DIMENSIONS [number] [number list] [number] TO [number]	A single integer or list, specifying the dimensions for an optional MINISSA analysis of the groups identified
CLUSTERS [number]	A single integer, giving the number of clusters to be defined in the partitioning process (<N OF STIMULI)
LABELS [followed by a series of labels (<= 65 chars) each on a separate line]	Optionally identify the stimuli, followed by the subjects, as required. All labels should be entered, without omissions.

PARAMETERS

Keyword	Default Value	Function
DATA TYPE	1	0: Lower triangle similarity matrix (without diagonals). 1: Lower triangle dissimilarity matrix (without diagonals). 2: Lower triangle euclidean distances (without diagonals). 3: Lower triangle correlation matrix

Apart from PRINT DATA, there are no PRINT or PLOT options for CONPAR.

PROGRAM LIMITS
Maximum no. of dimensions = 5
Maximum no. of stimuli = 30
Maximum no. of subjects = 65

5. EXAMPLE (Test Data)

```
RUN NAME       COLA DISSIMILARITY DATA
COMMENT        From Schiffman et al (1981) Intro to MDS
NO OF STIMULI   10
NO OF SUBJECTS  10
CLUSTERS        3
DIMENSIONS      2
PARAMETERS     DATA TYPE(1)
LABELS  DIET PEPSI
RC COLA
YUKON
DR PEPPER
SHASTA
COCA COLA
DIET DR PEPPER
TAB
PEPSI COLA
DIET RITE
SUBJECT 1
SUBJECT 2
SUBJECT 3
SUBJECT 4
SUBJECT 5
SUBJECT 6
SUBJECT 7
SUBJECT 8
SUBJECT 9
SUBJECT 10
READ MATRIX
16
81 47
56 32 71
87 68 44 71
60 35 21 98 34
84 94 98 57 99 99
50 87 79 73 19 92 45
99 25 53 98 52 17 99 84
16 92 90 83 79 44 24 18 98
09
90 70
87 65 06
87 77 83 83
33 79 25 89 39
86 86 99 22 90 40
81 30 57 88 69 39 97
74 20 94 78 05 81 92 88
23 26 72 94 02 76 81 20 05
49
96 96
97 92 94
68 12 90 93
77 44 88 90 26
97 93 94 25 93 49
54 76 92 94 20 24 93
47 48 92 94 35 18 94 23
21 47 90 92 68 67 87 55 15
23
99 51
99 23 78
90 16 22 49
74 55 50 99 13
14 88 77 75 50 70
```

```
25 95 48 99 99 79 99
60 36 69 24 21 53 99 99
00 89 72 81 77 71 74 51 71
62
77 16
98 14 55
76 22 40 47
84 16 16 81 07
17 80 36 93 60 90
76 93 86 80 94 36 19
74 20 16 38 05 18 06 71
10 72 78 92 92 86 16 02 99
85
82 15
97 28 56
51 31 36 43
79 27 07 82 07
13 84 38 87 76 82
82 99 73 68 80 40 20
69 24 30 27 16 12 28 80
15 80 78 90 72 66 17 05 95
10
53 75
99 99 99
87 27 65 99
60 66 72 99 99
96 99 90 10 90 75
98 99 91 98 88 34 99
73 15 90 99 09 56 95 75
54 62 84 99 95 53 85 91 49
14
61 47
79 96 77
72 21 12 73
66 12 28 81 13
66 64 75 41 71 82
51 67 32 93 49 66 86
07 20 67 71 15 56 76 69
19 51 06 88 25 81 50 08 83
11
90 69
72 26 90
93 17 69 24
39 34 36 98 80
26 82 77 85 53 99
80 74 75 99 93 87 13
73 08 91 35 17 17 99 91
24 62 90 76 85 64 77 24 65
69
63 58
76 85 79
52 14 51 81
61 39 35 83 36
80 90 93 06 78 85
28 87 83 94 64 44 90
80 20 92 98 51 23 80 33
78 28 40 99 36 71 82 62 13
COMPUTE
FINISH
```

OUTPUT:

The following clusters are identified:

```
CONPAR - INDIVIDUAL DIFFERENCES DATA IN  3 CLUSTERS.

 COMPUTATION TIME FOR  3 GROUPS         0.02
```

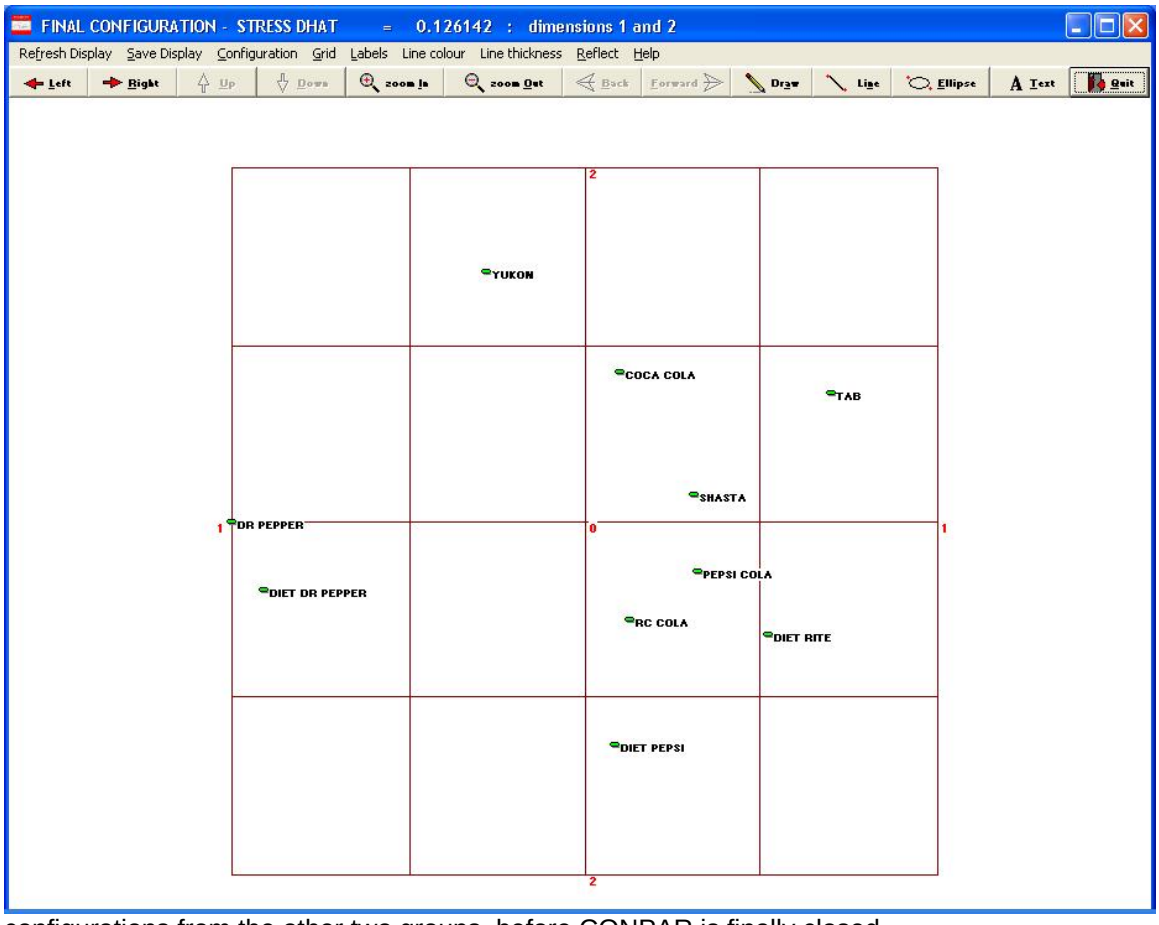
```
THE FOLLOWING CLUSTERS OF SUBJECTS ARE SUGGESTED :

CLUSTER   1
===============
 SUBJECT 1
 SUBJECT 9

 MEMBER(S)  2

CLUSTER   2
===============
 SUBJECT 2
 SUBJECT 3
 SUBJECT 7
 SUBJECT 8
 SUBJECT 10

 MEMBER(S)  5

CLUSTER   3
===============
 SUBJECT 4
 SUBJECT 5
 SUBJECT 6

 MEMBER(S)  3
```

When the results from CONPAR are listed in the NewMDSX output window, you are immediately offered the option of applying MINISSA to the pooled matrices for each of the groups identified. The largest group of subjects is seen to be distinguishing types of cola according to brand name, irrespective of whether or not they are termed "light":

```
POOLED CONPAR CLUSTER MATRICES FOR MINISSA ANALYSIS               CLUSTER  2

COMMENT   CLUSTER 2 CONSISTS OF SUBJECTS: 2  3  7  8 10

SOLUTION IN  2 DIMENSIONS:
 *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

FIT= DHAT ; ALGORITHM= HARD SQUEEZE; (SEE PROGRAM DESCRIPTION)

RAW STRESS DHAT     =     1.591177
STRESS DHAT         =     0.126142
RAW STRESS DSTAR    =     3.215417
COEF. ALIEN. DSTAR  =     0.178594


STRESS BASED ON APPROXIMATION TO RANDOM DATA =     0.154917
(I.Spence, Multivariate Behavioral Research,14,1979)

OPTIMAL SOLUTION USING DSTAR WAS REACHED AFTER ITERATION  18
OPTIMAL SOLUTION USING DHAT  WAS REACHED AFTER ITERATION  30

F I N A L   C O N F I G U R A T I O N

  1  DIET PEPSI
         0.1334     -0.9552
  2  RC COLA
         0.2026     -0.4284
  3  YUKON
        -0.4199      1.0743
  4  DR PEPPER
        -1.5175     -0.0081
  5  SHASTA
         0.4713      0.1217
  6  COCA COLA
```

```
          0.1627       0.6416
 7  DIET DR PEPPER
         -1.3721      -0.2944
 8  TAB
          1.0531       0.5502
 9  PEPSI COLA
          0.4906      -0.2174
10  DIET RITE
          0.7957      -0.4844

 MEAN    -0.0000      -0.0000

 SIGMA    0.8156       0.5786
```

Clicking in the usual way on these configuration values produces the following graphic display, which can be saved and manipulated as described in the chapter headed HOW TO USE NewMDSX FOR WINDOWS, or retained on the screen for comparision with the equivalent



configurations from the other two groups, before CONPAR is finally closed.

BIBLIOGRAPHY

Brusco, M.J. (2003) An enhanced branch-and-bound algorithm for a partitioning problem, British Journal of Mathematical and Statistical Psychology, 56, 83-92.

Brusco,M.J. (2004) On the concordance of empirical confusion matrices for visual and tactual letter recognition, Perception and Psychophysics, 66, 392-397.

Brusco,M.J. and Cradit,J.D. (2005) ConPar: a method for identifying groups of concordant subject proximity matrices for subsequent multidimensional scaling analyses, Journal of Mathematical Psychology, 49, 142-154.

Carroll J.D. & J.-J. Chang, (1970): Analysis of Individual Differences in Multidimensional scaling via an N-way generalization of 'Eckart-Young' Decomposition, Psychometrika 35, 283-319.

Hubert,L.J. (1987) Assignment Methods in Combinatorial Data Analysis, New York, Marcel Dekker.

Meulmann, J.J (1996) Fitting a distance model to homogeneous subsets of variables: Points of view analysis of categorical data, Journal of Classification, 13, 249-266

Tucker, L.R & S. Messick (1963) An individual differences model for multidimensional scaling , Psychometrika, 28, 333-367

Tzeng, O.C.S & D. Landis (1978) Three-mode multidimensional scaling with points of view analysis, Multivariate Behavioral Research, 13, 181-213

APB draft 1, amended APMC 13 Jan 2006