MULTIDIMENSIONAL SCALING: Using SPSS/PROXSCAL

• SPSS 10 offers PROXSCAL (PROXimity SCALing) as an alternative to ALSCAL for multidimensional scaling: USE IT !! ALSCAL has been shown to be sub-optimal (Ramsay).

• PROXSCAL performs most Distance Model scaling (for scalar products/vector models, see SPSS Categories). (Pre-SPSS PROXSCAL .pdf Documentation by Busing is available).

• Data for basic MDS in SPSS10 can be either

1. input directly as a full SSM (square symmetric matrix) of proximities=dis/similarities into SPSS editor.
2. calculate measure within SPSS from a raw datafile (separate procedure)
3. OR (in PROXSCAL only) matrix can be imported as a SSM. SPSS will not accept LT matrices directly, (n.b. universal use of LT matrices in other programs; confusing SPSS documentation suggesting otherwise).
4. procedures for changing a LT into a SSM for SPSS are contorted but possible
(subject of a separate handout).

**SETTING UP**

Data transformation model

EQUIVALENCE IN SPSS PROXSCAL

• n.b. Use Right Click for supplementary information in SPSS10

• PROXSCAL “proximities” = dis/similarities

• PROXSCAL “Transformed proximities” = disparities.

1. DATA

Analyze Y Scale Y Multidimensional scaling (PROXSCAL)

**First Window (Data Format)**

• Data Format: (T data are proximities (or create from raw data – initiates “Create proximities from data” procedure)

• Number of sources: (T One for 2W1M data; Multiple for INDSCAL etc)
Transfer Matrix (v1 - v16) across.
(Best to keep variables name short at this stage, to avoid over-printing at graphing stage; they can be selectively lengthened later in plotting routine)

Note the crucial buttons at the foot of the Transfer Window:

**MODEL- RESTRICTIONS-OPTIONS-PLOTS-OUTPUTS**

these are the ones which set the detail of the analysis and run parameters.
• **SCALING MODEL:**
  < Identity means simple Euclidean
  < Weighted Euclidean means INDSCAL
  < Generalised Euclidean means IDIOSCAL (individual rotation and then weighting)
  < Reduced Rank means IDIOSCAL with minimal rank of matrix
• **SHAPE:** does not refer to input matrix … which must be SSM
  < Lower triangular means “only the lower triangular data are analysed” i.e. symmetric
  < Upper triangular: ditto, upper
  < Full matrix: data may be asymmetric, but are symmetrised
• **PROXIMITIES (data)**
  < Similarity data (hi means more similar)
  < Dissimilarity (hi means more dissimilar)
  n.b. this option is default: beware!
• **DIMENSIONS**
  MDS solutions proceed from max (-1) min
PROXSCAL MODEL PARAMETERS (cont.)

- **Proximity TRANSFORMATIONS**
  
  - **Ratio** (LoM) – implies metric analysis
  - **Interval** (LoM) – metric
  - **Ordinal** (LoM) – non-metric. Default is a secondary approach to ties, unless ...
    
    T Untie tied values (=primary)
  
  - **Spline** (cf Ramsay & MULTISCALE): piece-wise polynomial transformation of the original data.
    
    “Pieces” and shape of transformation are specified by:
    
    T degree (1=linear; 2= quadratic ...), and
    
    T Number of internal knots.

- **APPLY TRANSFORMATIONS** (applies only to INDSCAL and higher models; in effect local versus global application).
RESTRICTIONS = EXTERNAL or CONSTRAINED / CONFIRMATORY ANALYSIS (cf Borg and Groenen 1997, pp181-199).

These options allow for:

- Fixing some (known?) points in a configuration, and estimating the others (Some co-ordinates fixed)
- Fitting (regressing) external properties (PRO-FIT) (Linear combinations ...)

Additional information is contained in an separate (or integral) SPSS file
PLOTS

- “Common Space” = Group or Stimulus Configuration.

n.b. The Shepard Diagram is not immediately available in PROXSCAL; it combines (and can in principle be reconstructed from)

- original vs transformed proximities = disparities (* vs d-hat – Monotonic Fit)
- transformed proximities vs distances of solution (d-hat vs d – OLS fit)
• **DISPLAY**

  < Common Space (=Stimulus Configuration)
  < Distances (of solution)
  < Transformed proximities (=disparities)
  < Input data (ALWAYS recommended, to ensure the program is working on the data YOU think it is ... )
  < Iteration history (for diagnosis of stress minimization)
  < Multiple stress measures (use Stress1 for comparison with other solutions; **not** S-STRESS. Note: normalised raw stress $\bar{\alpha}$ raw stress)
  < Stress decomposition = point contribution to stress
  < “Save to new file” is equivalent of MDSX’s “PUNCH”. Useful for graphic output.