MULTIDIMENSIONAL SCALING: Using SPSS/PROX SCAL

SPSS uses Forrest Young’s ALSCAL (Alternating Least Squares Scaling) as its main MDS program. However, ALSCAL has been shown to be sub-optimal giving exaggerated importance to large data dissimilarities (Ramsay). DO NOT USE IT (or only AYOR!)

SPSS10 offers PROXSCAL (PROXimity SCALing) as an alternative to ALSCAL for multidimensional scaling: USE IT !!

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 dis/similarity measure within SPSS from a raw datafile (ANALYZE -> CORRELATE -> DISTANCES); for metric data, many use Euclidean Distance Measure.
3. OR (in PROX SCAL 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 tedious, but possible (subject of a separate handout).
SETTING UP
data-transformation-model

EQUIVALENCES IN SPSS PROXSCAL

• n.b. Use Right Click for supplementary information in SPSS10
• PRO X SCAL “proximities” = dis/similarities
• PRO X SCAL “Transformed proximities” = disparities.

1. DATA

Analyze → Scale → Multidimensional scaling (PROXSCAL)

First Window (Data Format)
• Data Format: (✓ data are proximities (or create from raw data – initiates “Create proximities from data” procedure)
• Number of sources: (✓ One for 2W1M data; Multiple for INDSCAL etc)
Second Window:

- 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-PILOTS-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 secondary approach to ties, unless ...
    ✔ Untie tied values (= primary)
  ➤ **Spline** (cf Ramsay & MULTISCALE): piece-wise polynomial transformation of the original data.
    “Pieces” and shape of transformation are specified by:
    ✔ **degree** (1= linear; 2= quadratic ...), and
    ✔ **Number** of internal knots.

• **APPLY TRANSFORMATIONS** (applies only to IN DSCAL and higher models; in effect local versus global application).
RESTRICTIONS = EXTERNAL or CONSTRAINED / CONFIRMATORY ANALYSIS

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 (Δ vs d-hat – Monotonic Fit)
- transformed proximities vs distances of solution (d-hat vs d – OLS fit)
OUTPUT

• **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 ≠ raw stress)
  ▶ Stress decomposition = point contribution to stress
  ▶ “Save to new file” is equivalent of MDSX’s “PUNCH”. Useful for graphic output.