Choice of Options in PREFMAP, and other points

(a memo by Charles Jones, on applying MDS(X) PREFMAP to Occupational Cognition Data)

I. A useful procedure for using PREFMAP is:

(A) Do an initial run in Phases I through IV. Use this for estimating whether or not it is necessary in general to hypothesize Differential Rotation (and weighting) of the input configuration (all subsequent phases are done in the Average Subjects' Rotated and Weighted Space not in the space of the input configuration). Use LFITSW = 0 (metric)at this stage, otherwise the F-test is not strictly permissible. The hypothesis is best tested by:

(a) Testing the Average Subject's F-ratios for adjacent phases: (3,4),(2,3),(1,2) (The lowest acceptable phase-level is to be preferred. The test is basically whether "going up a level" significantly increases the variance explained).

(b) Estimate the dispersion in individual phase locations by deciding (in an identical way, but for each individual) on the (lowest acceptable) phase to which the individual should be located. n.b. Attention should also be paid to goodness-of-fit (individual corrections), especially for removal of cases. Where there is conflict between (a) and (b) I would choose the modal phase.

(B) Having decided on the upper limit of phases, submit a further operational run. At this stage, choose the relevant LFITSW (metric/quasi-nonmetric) option and run the phases, down to IV.

II. Before using PREFMAP on POOC data in any systematic way, I thought it best to examine empirically the effect of various options available to the user, by using the dame data set (of 16 persons) under a combination of options.

(1) As currently implemented:

(A) If 1PS=l and 1PE=4 :

The average subject's rotation of the input configuration, differentially stretched by the (evaluative) weights of this subject, form the coordinate for subsequent phases

If 1PS=2

The average subject's differentially-weighted axes of the input configuration form the coordinate axes

If 1PS=3 or 4

The input configuration remains the same for subsequent phases.

(B) In phase 3, there are no differential dimensional weights, but the sign of the weights should be noted:

- if an individual's weights are positive on all dimensions, then the subject point is an ideal point

-if an individual's weights are negative on all dimensions, then the subject point is an anti-ideal point

-if an individual's weights are mixed positive and negative, then the subject point is a "saddle point" of some form.

The effect of these constraints is that it is only possible to compare large numbers of subjects (more than 49) by sticking to Phases III and IV, unless some synthetic or "sampled" a priori configuration is input under our proposed new version of PREFMAP, which will allow the original input configuration to be replaced at each phase.

(2) The options of most importance and significance seemed to be:

- LFITSW, which controls the type of regression function (metric (0) or quasi-nonmetric (1-3) under standard tieing options)."
- NORS, which normalises the scale values, (1), or leaves them as input
- lSHAT, which either creates new fitting values (S) at each stage, or uses those of the previous phase as starting values.

The last option is easily dealt with; the only effect (rightly so) was to increase CPU time. It has no effect on the solution.

The data on certain combinations of the other two parameters are contained in the Appendix to this note.

The same set of 16 subjects' data was used for the analysis. One purpose of the experiment was to see the effect of treating RB [ratings] data under the non-metric options, and seeing whether the solution was identical to the solution from the corresponding ranked-(RA) data.

For testing the effects of the parameters, various aspects of the final solution are extracted and noted in the Appendix.

(1) The first row of the rotation matrix (the second row differs only in sign, being skew symmetric here) in Phase 1

(2) The evaluative weights in Phase I

(3) The location of ideal points (w.r.t. the original axes, for purposes of comparison)

(The above all refer to the average subject in Phase I).

(4) The RMS (goodness of fit) for each of the 4 phases

(5) The apparently most "reasonable" allocation of the average subject (see PREFMAP pt.l memo.)

The F-ratio between adjacent phases, for the average subject. n.b. This is only strictly permissible for LFITSW=O.

Certain conclusions seem justified:

(I) The major effect is undoubtedly the metric/nonmetric option. This is not unreasonable, since linear constraints are greater than monotonic constraints on the solution.

(2) Normalising seems to have very little effect

(3) Treatment of ties (LF1TSW=1 versus 2) has a predictable effect, leading to slightly more constrained and ill-fit solution compared to primary treatment of ties (cf. Kruskal, Wish "Comments on Dimensionality" and Young on "The variable monotonic transformation")

(4) Most differences centre on the Phase I parameters. The differences in weights and ideal point locations seem especially sensitive to changes of parameters, but goodness-of-fit measures do not

(5) The F-ratio is very conservative for the quasi-nonmetric option, and should be used cautiously.

(6) The difference between RB data (treated as quasi non-metric) and RA data is not identical, but gives rise to very small differences.

Overall, the results of the different analyses were a good deal more similar than I had originally anticipated. One should expect a good deal more variability if the data had been used to parameterize both the stimuli and the subject locations (internal analysis). The main differences of importance seem to be inferential and/or allocatory: beware of too rosy goodness-of-fit in the quasi non-metric case.