Journal of Classification 6:271-272 (1989)

ERRATUM

In Mirkin (1987), the additive clustering model of Shepard and Arabie (1979) is considered, given a proximity matrix $\mathbf{B} = (b_{ij})$ (where i, j are objects from the set I). The model consists of the equation

$$b_{ij} = \sum_m \lambda_m f_{im} f_{jm} + \varepsilon_{ij}$$

(where the clusters $f_m = (f_{im})$ and the weights λ_m are unknown, λ_m are real numbers, and f_{im} are ones or zeros) and of the criterion

$$\sum_{i,j} \varepsilon_{ij}^2$$

which has to be minimized by all real λ_m and zero-one f_{im}.

The consecutive exhaustion method for fitting the model was developed by Mirkin (1987) using the following steps:

1. Define $m = 1$, $\mathbf{B}_m = \mathbf{B}$.
2. Resolve, at least locally, the problem of minimizing the criterion

   $$\sum_{i,j} (b_{ij} - \lambda f_i f_j)^2$$

   by arbitrary (or perhaps only) positive real λ and zero-one f_i, f_j. Then define $f_{im} = f_i$, $\lambda_m = \lambda$.
3. Examine some stopping rules for the factors f_m's construction, using the values $\lambda_m^2 (f_m, f_m)^2$ of the contributions of the factors m to the sum of squares (not to the variance accounted for) in the data. If one of

---

Author's address: Boris G. Mirkin, Central Economics-Mathematics Institute, Krasikova str. 32, Moscow W-418, U.S.S.R. 117418.

the rules is satisfied, the algorithm terminates; otherwise, we compute residual proximities

$$\mathbf{B}_{m+1} = \mathbf{B}_m - \lambda_m f_m f_m T \ ,$$

increase m by 1, and return to step 1.

Unfortunately, the author's paper (1987) contains an error: the last column in Table 3 of the paper (which accounts for the contributions of the clusters in the proximities' variance for the data taken from Shepard and Arabie) is based on inappropriate computations.

Mirkin (1987) proved that the procedure 1-3 (in the form of QFA-O method) finds the values λ_m, f_{im} which generate the following decomposition of the sum of squared values in the proximity matrix:

$$\sum_{i,j} b_{ij}^2 = \sum_m \lambda_m^2 \, r(S_m) + \sum_{i,j} \varepsilon_{ij}^2 \qquad (1)$$

where $r(S) = (\#S)^2$ or $\#S(\#S - 1)$, depending on whether the diagonal elements b_{ii} are given. The method of finding the clusters S does not influence the form of (1) if λ_m are optimal ones (for fixed S_m). But according to (1) we may analyze the contributions of the clusters S, found by QFA-O method, to the variance (rather than to the sum of squared data values) of the proximities b_{ij} only in the case when $\mathbf{B}$ is centered (that is, the mean of all b_{ij} equals 0). The last column in Table 3 of Mirkin (1987) containing the contributions to the variance is indeed not appropriate because the clusters in Table 3 were found using the initial proximity matrix $\mathbf{B}$ (without preliminary centering of its entries). But according to (1), only the contributions of the clusters to the sum of squares (rather than to the variance) may be considered. So direct comparison of the author's results with the results of Shepard and Arabie (1979) in this concrete example is impossible, contrary to the incorrect conclusions in Mirkin (1987).

References

MIRKIN, B. G. (1987), "Additive Clustering and Qualitative Factor Analysis Methods for Similarity Matrices," *Journal of Classification, 4*, 7-31.

SHEPARD, R. N., and ARABIE, P. (1979), "Additive Clustering: Representation of Similarities as Combinations of Overlapping Properties," *Psychological Review, 86*, 87-123.