

Evaluating Goodness of Fit in Nonmetric Multidimensional Scaling by ALSCAL

Robert MacCallum
The Ohio State University

Two types of information are provided to aid users of ALSCAL in evaluating goodness of fit in nonmetric two-way and three-way multidimensional scaling analyses. First, equations are developed for estimating the expected values of SSTRESS and

STRESS for random data. Second, a table is provided giving mean values of SSTRESS and STRESS for structured artificial data. This information provides the empirical investigator with a second comparative basis for evaluating values of these indices.

A common problem in the empirical use of nonmetric multidimensional scaling (MDS) is the evaluation of goodness of fit. In the context of two-way MDS, which involves the analysis of one symmetric matrix of proximity measures among p stimuli, goodness of fit is commonly measured by the STRESS index, first proposed by Kruskal (1964a, 1964b). To attempt to aid the investigator in interpreting an observed value of STRESS, Kruskal proposed rough guidelines for values indicating that fit is excellent, good, and so forth. However, these guidelines have been found to be of limited utility because of the lack of comparability of STRESS values for different numbers of stimuli and dimensions.

Another approach to evaluating STRESS was provided by a series of studies on the distribution of STRESS values for completely random data. Early monte carlo work by Klahr (1969) and by Stenson and Knoll (1969) demonstrated the effect of number of stimuli and number of dimensions on STRESS, provided information on the likelihood of occurrence of very low STRESS values in random data, and presented graphs allowing the estimation of the expected value of STRESS in random data. Later, Spence and Ogilvie (1973) developed a table of expected STRESS values for random data as a function of number of stimuli and number of dimensions. These studies provide the empirical investigator with sufficient information to aid in some aspects of research design (e.g., choosing the number of stimuli), as well as to determine whether an observed value of STRESS differs significantly from a value expected from random data.

All of these studies are relevant only to two-way nonmetric MDS. With the development of ALSCAL (Takane, Young, & de Leeuw, 1977), researchers now have available a viable method for three-way nonmetric MDS, which involves the analysis of an $N \times p \times p$ matrix of proximities, con-

taining a $p \times p$ matrix of proximities for each of N individuals. (Although ALSCAL does offer a metric option, only the nonmetric approach is of interest here.) ALSCAL measures fit using an index called SSTRESS, which is defined conceptually the same way as STRESS, except that SSTRESS measures correspondence between *squared* distances and *squared* optimally scaled data as opposed to *unsquared* distances and *unsquared* optimally scaled data as treated by STRESS. ALSCAL optimizes SSTRESS, but it also computes and prints STRESS. Mathematically, these measures are defined as follows:

$$SSTRESS = \left[\frac{1}{N} \sum_{k=1}^N \left[\frac{\sum_i \sum_j (d_{ijk}^2 - d_{ijk}^{*2})^2}{\sum_i \sum_j d_{ijk}^{*4}} \right] \right]^{\frac{1}{2}} \quad [1]$$

$$STRESS = \left[\frac{1}{N} \sum_{k=1}^N \left[\frac{\sum_i \sum_j (d_{ijk} - d_{ijk}^*)^2}{\sum_i \sum_j d_{ijk}^{*2}} \right] \right], \quad [2]$$

where d_{ijk} is the distance between stimuli i and j for individual k , and d_{ijk}^* is the corresponding optimally scaled (monotonically transformed) proximity judgment.

Since ALSCAL optimizes SSTRESS rather than STRESS, and since there is another variable (N) present that may affect the behavior of the fit measure, users of ALSCAL should not rely on the literature on STRESS for two-way MDS to evaluate goodness of fit of their solutions. The purpose of the present study is to provide sufficient information to allow investigators doing nonmetric two-way and three-way MDS via ALSCAL to achieve a sound evaluation of goodness of fit. This will be achieved through (1) the development of simple method for determining the expected value of SSTRESS and STRESS for random data as a function of N , p , and r , the number of dimensions; and (2) the presentation of approximate values of SSTRESS and STRESS, which can be expected for data that do in fact fit the weighted Euclidean model (also called the INDSCAL model; Carroll & Chang, 1970), which is the model fit by ALSCAL. Though results for both measures of fit will be presented, users are encouraged to focus on SSTRESS, since that is the measure being optimized by ALSCAL.

Method

A monte carlo study was carried out to determine expected values of SSTRESS and STRESS for random data. For selected values of N and p , data matrices of order $N \times p \times p$ were created by generating each element independently from a uniform distribution in the interval zero to one. (Actually, since each of the N matrices should be symmetric, only the lower triangular section of each matrix was generated.) Table 1 shows the actual values of N and p that were used, and the number of data sets generated for each combination of N and p . For the largest values of N and p , only a few data sets were needed, since obtained SSTRESS and STRESS values were very stable across replications. Though not a factorial design, the conditions represented in Table 1 essentially cover sample sizes from 1 to 40 and numbers of stimuli from 12 to 30, thus encompassing much of the range covered by empirical two-way ($N = 1$) and three-way MDS studies. The entire process produced 163 data sets representing 18 different combinations of N and p .

Table 1
Conditions Included in Generation
of Random Data

Number of Individuals	Number of Stimuli	Number of Data Sets Generated
1	12	10
1	20	10
3	12	10
3	20	10
5	12	10
5	20	10
10	12	10
10	16	10
10	20	10
20	12	10
20	16	10
20	20	10
20	30	5
30	12	10
30	16	10
30	20	10
40	16	5
40	30	3

Each data set was then analyzed using ALSICAL, with scale of measurement specified as ordinal, indicating a nonmetric analysis, and data specified as conditional, indicating a different monotonic transformation for each individual. Solutions were obtained in two through five dimensions. Values of SSTRESS and STRESS were recorded for each solution, and means of these measures were obtained for each combination of N , p , and r . The next objective was to develop simple equations for determining an expected value of SSTRESS or STRESS as a function of N , p , and r . Since it is known that the relationship of SSTRESS and STRESS to dimensionality is not linear (Klahr, 1969; MacCallum & Cornelius, 1977; Stenson & Knoll, 1969), it was clear that some type of nonlinear regression method would be needed to allow accurate prediction of SSTRESS and STRESS from N , p , and r .

To accomplish this, the nonmetric regression procedure called MORALS, developed by Young, de Leeuw, and Takane (1976) was employed. The number of individuals (N), number of stimuli (p), and number of dimensions (r) were used as independent variables. The number of dimensions was declared to be an ordinal variable, so that the optimal scaling phase of MORALS would rescale the values of this variable in such a way as to allow an optimal linear equation employing the rescaled variable. Since this rescaling is monotonic, and thus potentially nonlinear, this approach offers a simple way of fitting the nonlinear relationship between the number of dimensions and each of the fit measures. It was also suspected that N might show a nonlinear relationship to SSTRESS and STRESS, a suspicion that was borne out to be true. To fit this, a quadratic term N^2 was included in the model. There were, in effect, 72 observations in the regression analysis, representing the 18 combinations of N and p , with four different values for r for each combination.

Results

The regression analyses provided simple equations achieving very accurate estimation of the expected value of SSTRESS or STRESS for random data. To obtain the expected value of SSTRESS, given N , p , and r , the user must first obtain the optimally rescaled value of r from Table 2. Let this new value be r^* . Then, the following equation is applied:

$$E(\text{SSTRESS}) \doteq .0083(N) - .0002(N^2) + .0051(p) - .069(r^*) + .4688 \quad [3]$$

To illustrate this procedure, the expected value of SSTRESS for random data when $N = 10$, $p = 20$, and $r = 3$ would be found as follows:

$$\begin{aligned} E(\text{SSTRESS}) \doteq .0083(10) - .0002(100) + .0051(20) - .069(3.193) \quad [4] \\ + .4688 \doteq .414 \end{aligned}$$

By a similar procedure, the right-hand section of Table 2 could be employed and the following equation used to estimate expected values of STRESS:

$$E(\text{STRESS}) \doteq .0053(N) - .0001(N^2) + .0038(p) - .069(r^*) + .3869 \quad [5]$$

Table 2
Rescaled Values of Number of Dimensions
for Estimating $E(\text{SSTRESS})$ and $E(\text{STRESS})$

Number of Dimensions (r)	Rescaled Value of r for estimating $E(\text{SSTRESS})$	Rescaled Value of r for estimating $E(\text{STRESS})$
2	1.846	1.774
3	3.193	3.312
4	4.130	4.179
5	4.831	4.735

Multiple correlations for both of these equations were greater than .97. It should also be noted that standard deviations of SSTRESS and STRESS values within each of the 72 conditions represented in this study were consistently small, usually less than .01. Thus, goodness-of-fit values showed little sampling fluctuation, indicating that confidence intervals around the estimated values produced by the above equations would be fairly narrow.

Discussion and Conclusions

The developments above provide an investigator a simple, accurate method for determining the expected value of SSTRESS and STRESS from an ALSICAL analysis of two-way or three-way MDS random data. Users will find that expected values of SSTRESS and STRESS may be quite different, SSTRESS being higher; this indicates that ALSICAL users must be careful not to transfer their old frame of reference for evaluating STRESS values from two-way MDS.

To employ these developments in practice, an investigator should compare observed SSTRESS values from real data to expected values produced from Equation 3 above. If an observed SSTRESS value is not substantially lower than the expected value for random data, then there is present either very high measurement error or severe model violation, or both.

These developments cannot be used directly to determine dimensionality for a given set of data. It would be tempting to compare observed SSTRESS values from a sequence of solutions of increasing dimensionalities to corresponding expected values from Equation 3, in the hope that the two series of SSTRESS values would cross at some point and thus indicate the true dimensionality of the data. This would be analogous to the parallel analysis technique in factor analysis (Montanelli & Humphreys, 1976) using observed and expected eigenvalues. However, such an approach could not work in MDS because SSTRESS is measuring overall fit in a given number of dimensions. For instance, a truly two-dimensional data set would yield a very low SSTRESS value in two dimensions and also in three or more dimensions. Thus, an investigator should not attempt to determine dimensionality by simply comparing observed SSTRESS values to expected random values.

Although expected values of SSTRESS and STRESS for random data will be of some value to empirical investigators, there would be even more value in having some idea about what values of SSTRESS and STRESS would be likely to occur in nonrandom data, i.e., data that fit, to some degree, the weighted Euclidean model for three-way MDS. At least two monte carlo studies (Hahn, Widaman, & MacCallum, 1978; MacCallum & Cornelius, 1977) have used ALSCAL to analyze such data. Mean values of SSTRESS and STRESS for some of the conditions in these studies were obtained and combined (since the methodology was the same in both studies), as presented in Table 3.

Table 3
Mean Values of SSTRESS and STRESS Obtained
for Structured Artificial Data

Number of Stimuli	Number of Dimensions	Low Random Error		Moderate Random Error	
		Mean SSTRESS	Mean STRESS	Mean SSTRESS	Mean STRESS
12	2	.058	.053	.269	.214
	3	.047	.040	.230	.172
	4	.047	.038	.195	.140
16	2	.069	.059	.284	.225
	3	.058	.046	.234	.182
	4	.050	.037	.213	.153
20	2	.072	.064	.287	.229
	3	.062	.051	.255	.190
	4	.054	.042	.229	.164

Because these two studies constructed data that deviated from the weighted Euclidean model only due to random error, values in Table 3 can be interpreted as expected levels of fit for real data that fit the model except for error of measurement. Due to the design of the studies cited above, the results in Table 3 are limited to the three-way MDS case with N at least 10. When N exceeds 10, fit measures in MDS tend to be essentially independent of N ; thus, there is no variable N in Table 3.

Using the information presented in this paper, an empirical investigator will have a sound basis for evaluating goodness of fit of ALSCAL to three-way MDS data. By comparing observed values of SSTRESS and STRESS to expected values of these indices for random data and for data that fit the model, the researcher can more confidently draw inferences regarding the degree of some types of structure in his/her data as well as the empirical validity of an ALSCAL solution.

Of course, the user must be aware of the limitation inherent in the developments reported here. First, these results are strictly valid only for the specified ranges of N , p , and r , though they do provide some useful comparative information for empirical problems whose parameters fall outside of the given ranges. Second, these results are applicable only to ALSCAL's two-way and three-way non-metric MDS analysis; ALSCAL can do a wide range of other types of analyses, though the case treated here is probably the most commonly used.

References

- Carroll, J. D., & Chang, J. J. Analysis of individual differences in multidimensional scaling via an N -way generalization of "Ekart-Young" decomposition. *Psychometrika*, 1970, 35, 238-319.
- Hahn, J., Widaman, K. F., & MacCallum, R. C. *Robustness of INDSCAL and ALSCAL with respect to violations of metric assumptions*. Paper presented at the meetings of the Psychometric Society, Hamilton, Ontario, 1978.
- Klahr, D. A monte carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure. *Psychometrika*, 1969, 34, 319-330.
- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, 29, 1-27. (a)
- Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, 29, 28-42. (b)
- MacCallum, R. C., & Cornelius, E. T., III. A monte carlo investigation of recovery of structure by ALSCAL. *Psychometrika*, 1977, 42, 401-428.
- Montanelli, R. G., Jr., & Humphreys, L. G. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika*, 1976, 41, 341-348.
- Spence, I., & Ogilvie, J. C. A table of expected STRESS values for random rankings in nonmetric multidimensional scaling. *Multivariate Behavioral Research*, 1973, 8, 511-517.
- Stenson, H. H., & Knoll, R. L. Goodness of fit for random rankings in Kruskal's nonmetric scaling procedure. *Psychological Bulletin*, 1969, 71, 122-126.
- Takane, Y., Young, F. W., & de Leeuw, J. Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika*, 1977, 42, 7-67.
- Young, F. W., de Leeuw, J., & Takane, Y. Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features. *Psychometrika*, 1976, 41, 505-529.

Author's Address

Send requests for reprints or further information to Robert MacCallum, Department of Psychology, 404C W. 17th Avenue, Ohio State University, Columbus OH 43210.