Japanese Psychological Research
1978, Vol. 20, No. 3, 105–114

A MAXIMUM LIKELIHOOD METHOD FOR NONMETRIC MULTIDIMENSIONAL SCALING: I. THE CASE IN WHICH ALL EMPIRICAL PAIRWISE ORDERINGS ARE INDEPENDENT—EVALUATIONS

YOSHIO TAKANE[1]

Department of Psychology, McGill University

A maximum likelihood estimation procedure for nonmetric multidimensional scaling (MAXSCAL-1) described in the previous paper (Takane, 1978) is evaluated using both Monte Carlo and real data. Two Monte Carlo studies are designed each with specific objectives. In the first study various aspects (numerical quality of estimates, robustness, etc.) of solutions are examined as functions of the number of replications per tetrad, the number of observations (the number of tetrads for which observations are made) and the magnitude of discriminal dispersions. The results strongly suggest the importance of replicated observations for obtaining solutions of " good " qualities. In the second Monte Carlo study the effects of a particular type of systematic violations of distributional assumptions are inspected. The estimates of the location parameters (stimulus coordinates) are found to be less susceptible to the kind of distributional violations examined here, while the goodness of fit statistics (the chi-square, the AIC) tend to overestimate the correct dimensionality of the representation space. Two sets of real data are analyzed to demonstrate the advantages of the current procedure, namely the availability of confidence regions, the availability of the goodness of fit statistics, and the constrained optimization feature for testing hypotheses.

We present some empirical evidence in support of the maximum likelihood nonmetric multidimensional scaling procedure (MAXSCAL-1) described in the previous paper (Takane, 1978). Small Monte Carlo studies are conducted to investigate various statistical and numerical properties of the algorithm. We use sets of real data to demonstrate ways in which a researcher might go about in applying our procedure to practical data analytic situations. In particular we emphasize various advantages of the statistical inferential features of the present method in the analyses of dissimilarity data.

[1] The author wishes to express his appreciation to Drs. Forrest W. Young and Elliot M. Cramer of the University of North Carolina for their helpful comments. Portions of the work were done while the author was at the University of Tokyo and at the L. L. Thurstone Psychometric Laboratory, the University of North Carolina. Requests for reprints should be addressed to Yoshio Takane, Department of Psychology, McGill University, 1205 McGregor Avenue, Montreal, Quebec, Canada H3A 1B1.

Monte Carlo Studies

In Monte Carlo studies we hypothesize " true " underlying structures (stimulus configurations) from which we generate sets of data under various assumptions about the data generating processes. Those generated data are then submitted to the estimation procedure to see the extent to which and under what conditions it can recover the " true " underlying structures. The effects of various assumptions about the data generating processes on various aspects of the solutions are to be examined.

There are a variety of conditions which are of potential import in affecting the overall performance of the procedure. However, for the purpose of this paper we restrict our attention to only a limited subset of all conceivable variables. First of all we assume throughout that the distance function (the representation model) is correct. That is, we generate sets of data according to a specific distance formula, and we analyze them under appropriate model specifications.

We distinguish two situations, one in which there are no violations of distributional assumptions, and another involving substantial violations. Consequently, we have designed two separate Monte Carlo studies; one is designed to investigate the effects of various conditions of data (other than distributional assumptions on errors) assuming that the error distribution is correctly specified, and the other to investigate the effects of untenable distributional assumptions on derived solutions.

In both studies we have hypothesized a common " true " underlying structure, a stimulus configuration of nine stimuli in two dimensions. Eighteen uniform random numbers are generated for the stimulus coordinates. They are then centered, normalized, and rotated to principal axis to meet the identification restrictions. For the distance function we simply assume the Euclidian distance.

Construction of Data

Monte Carlo data 1. A set of data is generated under a particular combination of assumptions about the error model. The particular combination of assumptions is that errors are additive, and are independently normally distributed with zero mean and constant variances. Distances are calculated from the assumed underlying structure. Independent normal random numbers are added to the members of a pair of distances to generate discriminal processes, on which comparisons are made to obtain ordinal judgments. Independent random numbers are used for each comparison. The frequencies of judgment " larger " for one pair of stimuli over the other are recorded for prescribed numbers of replications, which serve as input data to MAXSCAL-1.

We consider that (1) the number of replications of ordinal judgments per tetrad, (2) the number and the particular subset of tetrads for which ordinal judgments are obtained (sampling conditions), and (3) the level (magnitude) of the discriminal dispersions (sigmas), are three particularly important determinants of the performance of the procedure.

For the number of replications we use equal numbers for all tetrads. Recall that this is not a requirement of the procedure. It is much simpler, however, to discuss the effects of number of replications in terms of a single parameter. We consider three cases: 1, 10, and 38 for the number of replications. The single replication case is interesting since it is the minimal condition. The 38-replication case is chosen for the other extreme to render a direct comparison with the available real data and the 10-replication case in-between.

Recall that MAXSCAL-1 does not require ordinal relations to be observed for all possible tetrads of stimuli. It is thus interesting to compare solutions derived from different sampling conditions of observations. We consider the following five conditions: (1) Total tetrads; (2) proper tetrads; (3) triads; (4) random 378; (5) random 252. Proper tetrads condition refers to a subset of total tetrads in which all four stimuli are distinct. The number of total tetrads is 630 for nine stimuli, of which 378 are proper tetrads and 252 are triads. In order to see the effects of the particular patterns of tetrads for which observations are made, the last two conditions are set up; observations are randomly sampled from a set of total tetrads with numbers of observations equal to the number of proper tetrads

(378) and to the number of triads (252).

The magnitude of the discriminal dispersions is another important factor we look into. We consider the following three levels for the magnitude of the discriminal dispersions: (1) $\sigma^{(1)}=.586$; (2) $\sigma^{(2)}=1.172$; (3) $\sigma^{(3)}=2.344$. The value set for $\sigma^{(1)}$ is the estimate of sigma obtained from Torgerson's data under the set of assumptions equivalent to this Monte Carlo study. It should be observed that $\sigma^{(2)}=2\sigma^{(1)}$ and $\sigma^{(3)}=2\sigma^{(2)}=4\sigma^{(1)}$. Note also that a discriminal process corresponding to the l'th level of discriminal dispersion for a particular comparison t is generated by

$$\lambda_t^{(l)} = d+\sigma^{(l)}e_t\,, \quad (l = 1, 2, 3)\,,$$

where $e_t \sim N(0, 1)$ independently of t. However, the same e_t is used to obtain $\lambda_t^{(l)}$ and $\lambda_t^{(l')}$ $(l \neq l')$, so that they are perfectly correlated with each other. This is for strictly isolating the effect of the magnitude of the discriminal dispersions.

Monte Carlo data 2. The second set of Monte Carlo data are generated under systematic violations of distributional assumptions on errors. There are a virtually infinite number of ways in which distributional assumptions may be violated. We examine but two likely situations. In justifying the distributional assumptions in MAXSCAL-1 we discussed an alternative distributional assumption, namely the normality assumption on stimulus coordinates, which may be intuitively more plausible than the normality-on-distance assumption, but is more difficult to realize in an algorithmic form. We are interested in seeing how various analysis options (representing different assumptions about the model) available in the current MAXSCAL-1 react when in fact the normality on-stimulus-coordinates is the true distributional assumption.

We limit our attention to the constant variance case with triads sampling with 10 and 38 replications. Two ways of generating discriminal processes are considered. The first case considers independent discriminal processes for common stimuli in triads, while the second case considers only one common process (or two processes which are completely dependent with each other) for common stimuli in triads. (In fact the choice of triads situation is to make this distinction clearer.) In the independent case the discriminal processes $\lambda_{ij}^{(t)}$ and $\lambda_{ik}^{(t)}$ at comparison t with stimulus i as the common stimulus are generated by

$$\lambda_{ij}^{(t)} = \{\sum_{a=1}^{2}(x_{ia}+e_{ia}^{(t)_1}-x_{ja}-e_{ja}^{(t)})^2\}^{1/2},$$

and

$$\lambda_{ik}^{(t)} = \{\sum_{a=1}^{2}(x_{ia}+e_{ia}^{(t)_2}-x_{ka}-e_{ka}^{(t)})^2\}^{1/2},$$

where $e_{ia}^{(t)_1}$ and $e_{ia}^{(t)_2}$ are independent (and standard normal). This is the case in which we can derive the noncentral chi-square distribution for the squared Euclidian distances. In the dependent case we use a common $e_{ia}^{(t)}$ for both $e_{ia}^{(t)_1}$ and $e_{ia}^{(t)_2}$. Note that the noncentral chi-square model is not tenable in this situation. It seems interesting to examine this case, since it may represent a more likely state of nature than the independent case.

Monte Carlo Results

We are interested in finding out the effects of the independent variables (the number of replications, the number of observations, the magnitude of discriminal dispersions and the dimensionality of solutions) on various aspects of solutions with particular attention to (1) the degree of recovery of the "true" underlying structure, (2) the behavior of the chi-square goodness of fit statistic and (3) the behavior of the AIC statistic.

Data 1. Data 1 have been analyzed under the Euclidian metric, and the constant variance additive error assumption. Solutions in one, two and three dimensions have been derived. Initial estimates are obtained by Torgerson's procedure. For

TABLE 1
The goodness of recoveries of the "true" underlying structure

Sampling condition	Configuration matching indices: Single replication, Discriminal dispersion .586	1.172	2.344	10-replication, Discriminal dispersion .586	1.172	2.344	38-replication, Discriminal dispersion .586	1.172	2.344
t-tetrads	.124†	.216	.508	.034 (.090)	.061 (.071)	.134 (.128)	.017 (.029)	.033 (.035)	.073 (.064)
p-tetrads	.125	.216	.411	.050 (.111)	.071 (.088)	.138 (.148)	.022 (.040)	.041 (.037)	.078 (.069)
random 378	.138	.224	.519	.050 (.139)	.082 (.176)	.133 (.208)	.017 (.031)	.031 (.046)	.053 (.056)
triads	.189	.382	.537	.051 (.128)	.111 (.119)	.252 (.246)	.022 (.040)	.037 (.048)	.106 (.079)
random 252	.210	.306	.571	.054 (.118)	.094 (.161)	.224 (.550)	.025 (—)††	.075 (.058)	.145 (.148)

† The smaller value indicates the better matching.
†† No solution is obtained.

the single replication cases the corresponding initial estimates obtained from the 10-replication cases have been used. The Euclidean elliptic norm of the gradient with respect to the inverse of the information matrix H, namely $\|\mathbf{g}\|_{H^{-1}}=(\mathbf{g}'H^{-1}\mathbf{g})^{1/2}<0.01$, is used for the termination criterion. Derived stimulus coordinates are normalized so that $\operatorname{tr} D^2=2n\operatorname{tr}(X'X)=4n^2$ where D is the matrix of (Euclidian) distances.

In order to see the goodness of recoveries of the "true" underlying structure as functions of the three data conditions, Schönemann and Carroll's (1970) configuration matching index has been calculated for each solution and reported in Table 1. For this we have assumed that the "true" dimensionality is known. The smaller value indicates the better matching. It can be clearly seen that the better recoveries are obtained when (1) the number of replications is large, (2) the number of observations is large, and (3) the magnitude of the discriminal dispersions is small. Values in parentheses indicate the goodness of recoveries by Torgerson's LS procedure for the same sets of data. It is interesting to note that the ML estimates are clearly better when the discriminal dispersions are small, while the superiority of the ML estimates diminishes as the dispersions get larger. (Plots of estimated stimulus coordinates against the "true" coordinate values may be helpful to obtain the intuitive sense of how good a recovery a particular value of matching index indicates. They are not reproduced here due to the space limitation, but are contained in Takane (1977).) We have failed to observe any systematic differences between different sampling schemes of observations (tetrads) with equal numbers of observations.

The chi-square goodness of fit statistic seems to serve as a reasonable measure of the general goodness of fit of a model in all cases except for the single replication cases. The statistic is more influenced by the magnitude of the discriminal dispersions when the number of replications is small, leading to an overly good fit for the small dispersion, and to an overly poor fit for the large dispersion. Note that the statistic is derived from the likelihood ratio criterion (Takane, 1978) so that the goodness of fit statistic should not be affected by discriminal dispersions.

TABLE 2
Mean standard errors of derived stimulus coordinates

Sampling condition	Mean standard errors								
	Single replication Discriminal dispersion			10-replication Discriminal dispersion			38-replication Discriminal dispersion		
	.586	1.172	2.344	.586	1.172	2.344	.586	1.172	2.344
t-tetrads	.205	.285	.430	.070	.108	.202	.036	.055	.102
p-tetrads	.246	.314	.520	.089	.137	.251	.046	.069	.124
random 378	.271	.386	.610	.092	.140	.257	.047	.071	.131
triads	.353	.448	.591	.114	.180	.332	.060	.094	.183
random 252	.321	.435	.556	.109	.182	.344	.058	.091	.169

The AIC statistic also seems useful for the detection of " true " dimensionality. It failed to predict the correct dimensionality in only two cases, which are both unreplicated cases ($\sigma=1.172$ with the triads sampling and $\sigma=2.344$ with the random 378 sampling). We can be fairly confident about what the AIC predicts when there are at least a moderate number of replications, suggesting again the importance of replicated observations.

The estimated standard errors of stimulus coordinates generally increase as the number of dimensions increases. In Table 2 the mean standard errors of estimated parameters are shown for two dimensional solutions as functions of the number of replications, the sampling conditions, and the level of the discriminal dispersions. Three trends which are readily apparent are that the standard errors are: (1) Decreasing functions of the number of replications, (2) decreasing functions of the number of observations, and (3) increasing functions of the discriminal dispersions.

Data 2. We have analyzed Data 2 under various analysis specifications, including the additive error model with $s=0$ (constant variance), $s=1$ (variance proportional to the mean), and the multiplicative error model, each crossed with varying dimensionalities (1, 2 and 3). As noted earlier, any of the distributional assumptions available in the current MAXSCAL-1 does not match in every respect the assumptions under which the data are generated. However, it is informative to see which of the available assumptions are relatively robust (and to what extent) against the kinds of distributional violations under consideration.

Taken nominally, in all cases we have attempted to analyze, the AIC has predicted that the additive error model with a constant variance fits better than any other error model irrespective of the replication conditions, and irrespective of the dependency or independency of discriminal processes. (In contrast, the multiplicative model, the log-normal model, has been found consistently the poorest among the three.) However, the AIC has predicted the correct dimensionality only in two cases (the constant variance additive error model for the 10-replication data from uncorrelated and correlated discriminal processes). In all other cases the AIC tends to overestimate the " true " dimensionality. This means that the AIC is not entirely reliable in cases where there is some doubt as to the plausibility of distributional assumptions under which a particular solution is obtained. This is probably also true for an inadequate choice of representation models. The AIC (and the chi-square goodness of fit statistic) seems more susceptible to the violations of various assumptions involved in the analyses. This is despite the fact that the recovery of the " true " stimulus configuration is remarkably good in all data conditions when the mast robust assumptions (i.e., additive errors with a constant variance) are made. Estimates of stimulus coordinates are relatively robust against the violations of distributional assumptions considered here, but not the goodness of fit statistics.

We have failed to observe any systematic effects of dependency in discriminal processes. The subtle difference is probably mixed up with the effects of inappropriate distributional assumptions.

Two propositions occur from the Monte Carlo studies. It is apparent that the more observations one obtains, the more reliable estimates one can derive. It is particularly important to obtain replicated observations per tetrad when the discriminal dispersion is large. This observation should be taken seriously because we have a control over the number of observations and replications to be made, whereas it is relatively difficult (if not totally impossible) to experimentally control the magnitude of the discriminal dispersions.

It is found that the goodness of fit statistics are not robust against violations of distributional assumptions (at least against the type of violations considered here). The interpretation afforded by these statistics may be misleading in such situations. Hence it is always recommended that the data be reanalyzed under different distributional assumptions. At the same time experiments should be carefully designed so that the assumptions on which our ML estimation procedure is based be met as much as possible.

Analyses of Real Data

In this section we report some empirical results obtained with real sets of data. The first data set has been extracted from Saito (1974), and the second from Torgerson (1958). They collected sets of dissimilarity judgments between colors to investigate the structure of psychological distances between colors. Nine colors are employed, all of which are red (7R in Munsell designations) in hue but varying in two dimensions, brightness and saturation, which correspond to value and chroma dimensions in the Munsell color code system. The stimulus configuration for the nine colors in terms of the Munsell system is displayed in Fig. 1.

Multiple-judgment sampling was used in both cases. However, the order of stimulus presentation was randomized for each subject. Analyses of the two sets of data are found to be very instructive in terms of what kinds of care should be extended for the design of experiments to meet the assumptions under which the present method is formulated.

Saito's data. Saito (1974) collected pairwise dissimilarity judgments between the nine colors for 35 tetrads with 60 replications each. His original purpose was to demonstrate his ML and LS methods for the initial scaling of " observed " distances based on the normality-on-coordinates assumption. We use his data to demonstrate our one-step ML procedure for nonmetric MDS.

Since we were not sure about the true distributional form of the error, we analyzed the data under four different distributional assumptions to see which fit best. However, analyses were all performed under the Euclidian assumption

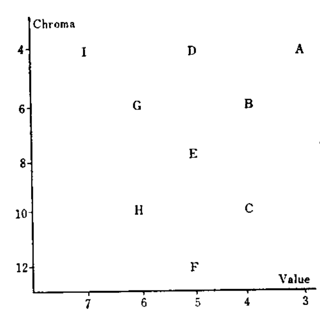

Fig. 1. Stimulus configuration of nine colors in terms of the Munsell system.

TABLE 3
A summary of results with Saito's data

Statistic	Additive model $s=0$	Additive model $s=1$	Additive model $s=2$	Multiplicative model $s=0$
Dimensionality 3				
Chi-square	15.355	16.989	17.610	17.816
Degrees of freedom	14	14	15	15
AIC+constant	2385.909	2387.542	2386.164	2386.370
Estimate of sigma	.583	.381		
Number of iterations	10	18	22	31
Dimensionality 2				
Chi-square	29.075	21.873	26.888	28.450
Degrees of freedom	20	20	21	21
AIC+constant	2387.628	2380.426	2383.442	2385.004
Estimate of sigma	.632	.439		
Number of iterations	7	5	11	13
Dimensionality 1				
Chi-square	235.280	223.999	222.911	235.717
Degrees of freedom	27	27	28	28
AIC+constant	2579.280	2568.553	2565.465	2578.271
Estimate of sigma	2.298	1.855		
Number of iterations	12	7	9	14

for the representation model. The starting configuration was obtained by Torgerson's procedure using the initial "observed" distances scaled by Saito's method (and listed in his paper). Solutions were obtained in one, two and three dimensions. A summary of results is presented in Table 3. The columns of this table represent the different distributional assumptions (the additive model with $s=0$, 1 and 2, and the multiplicative model), and the rows different dimensionalities.

It is clear that the representation requires more than one dimension (in terms of both the chi-square goodness of fit statistic and the AIC). We are somewhat undecided as to the appropriate dimensionality of the solution and as to the appropriate error model. Note that the chi-square statistics and the AIC's are remarkably "stable" across different distributional assumptions on errors. Except for the constant variance normal additive error case the minimum AIC predicts a dimensionality of two. Among all solutions obtained the minimum AIC resides in the

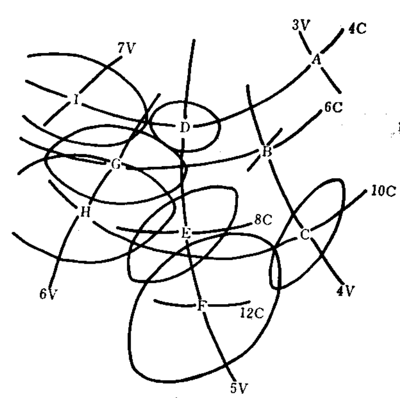

FIG. 2. Derived stimulus configuration for Saito's data: The unconstrained solution.

two dimensional solution obtained under $s=1$. This optimal solution is illustrated in Fig. 2. The 95% confidence regions are displayed around estimates of parameters along with equi-value and equi-chroma contours in terms of the Munsell system. These curves, however, should not be taken too seriously, since they rely too much on interpolations. That the method permits us to draw confidence regions is one of the favorable consequences of our method. Note that there is no confidence region for stimulus A and only a one dimensional confidence region (interval) for B. This reflects the fact that three parameters had to be fixed to identify the solution (two for translation and one for rotation).

The stimulus configuration (Fig. 2) derived from observed dissimilarity judgments between the colors, which presumably represents a cognitive map of colors (at least for those included in the analysis), looks very much like the Munsell configuration (Fig. 1) despite the presence of some topological distortion in the derived configuration relative to the Munsell configuration. How well then does the Munsell system represent the psychological distances between these colors? Our procedure permits a formal statement concerning the degree to which the Munsell system describes the cognitive map of colors.

Note first that multidimensional scaling (or the Minkowski power metric distance) is formally a special type of additive conjoint measurement on dimensionwise differences between stimuli; monotonically transformed dimensionwise differences are summed across dimensions to define overall distances, which are again monotonically related to observed dissimilarities. In view of this one aspect of what the Munsell system advocates can be translated into a statistical hypothesis, which is essentially a set of restrictions imposed on a model. We may then explicitly obtain a solution under this hypothesis and compare the goodness of fit statistic in this case against that obtained from the unconstrained solution. If the difference is significant, the hypothesis is not accepted on the statistical grounds.

We use the constrained optimization feature of the current procedure to impose equality constraints to test the hypothesis that the perceived psychological distances between colors are in fact additive functions of dimensionwise differences between colors on the Munsell dimensions. Figure 3 illustrates the stimulus configuration obtained under the equality constraints. Coordinates for stimuli D, E and F, B and C, and G and H are assumed to take equal values on dimension one (Value) as required by Munsell system. Similarly stimuli A, D and I, B and G, and C and H are constrained to take equal coordinate values on dimension two (Chroma). The chi-square goodness of fit statistic for this constrained solution is 161.363 with 27 *df*, which is significantly worse than 21.873 with 20 *df* for the unconstrained solution. Thus, the hypothesis that the Munsell system describes the psychological distances between the nine colors is rejected on statistical grounds, suggesting the necessity of some revision in the system. One reservation against drawing any definite conclusion would be the small number of observations in Saito's data. The chi-square statistic is based on the asymptotic property of the likelihood ratio criterion

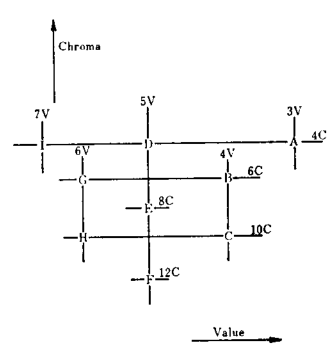

Fig. 3. Derived stimulus configuration for Saito's data: The constrained solution.

and 35 observations may or may not be sufficient to rely on this property.

Torgerson's data. Torgerson (1958) employed the complete method of triads to obtain a set of pairwise orderings of dissimilarities for 252 tetrads with 38 replications for each tetrad. Thirty-seven observations (out of 252 observations) were either 0 or 38 so that some ad hoc convention had to be employed to obtain corresponding normal deviates. The normal deviates were then used to obtain initial " observed " distances under Thurstone's Case V assumptions. The current procedure, of course, requires no such ad hoc convention.

Like Saito's data we have analyzed Torgerson's data under the four different distributional assumptions. However, we have obtained quite contrasting results to what we have observed in Saito's data; in all cases the goodness of fit is found to be extremely poor. For example, the chi-square goodness of fit statistic is 730.20 with 231 *df* in the three dimensional solution under the constant variance additive error model (which gives the minimum AIC estimates). This finding is somewhat puzzling since we have observed a moderately good fit with Saito's tetrad data employing the same set of stimuli.

In order to detect the possible sources of the poor fit we have reanalyzed Torgerson's triad data by the Thurstonian unidimensional scaling method under the Case V assumptions. We have treated the perfect discrimination data as missing observations and applied Gulliksen's inversion method (Torgerson, 1958) for incomplete pair comparison desings (which is the LS solution on normal deviates). A chi-square goodness of fit statistic based on the arcsine transformation of proportions (Mosteller, 1951) is found to be 585.407 with 179 *df*, which still indicates a poor fit. Note that for this procedure no explicit structures are assumed of distances (i.e., that the set of distances has a dimensional representation, etc.). It seems that the data radically violate the Case V assumptions. The fact that the fit has been found equally poor under the other distributional assumptions currently available in MAXSCAL-1, suggests that the problem is not only the constant variance assumption in Case V.

It should also be noted that we have never observed such a poor fit in any equivalent condition (triads, 38-replication) in either Monte Carlo study. Even when the distributional assumptions were systematically violated, the constant variance additive error model revealed some degree of robustness. We suspect that the observed poor fit in the present case may be due to the joint effects of multiple-judgment sampling and the method of triads. The independence of observations may have been difficult to attain in the multiple-judgment sampling situation in conjunction with the use of the method of triads in which pairs of dissimilarities being compared always have common stimuli.

The correlations among the discriminal processes are not necessarily damaging if they are all of the same magnitude (which is likely if the source of correlations is the common stimuli) and if the discriminal dispersions are also of the same magnitude. The Case V assumptions still hold in this case (Mosteller, 1951). If, however, discriminal dispersions are not constant over a set of dissimilarities, they will give rise to nonhomogeneous covariances which violate the Case V assumptions.

Another possible interpretation of the poor fit of Torgerson's data would be that the " comparability " (Tversky & Russo, 1969) of dissimilarities are more heterogeneous in Torgerson's case than in Saito's data. Dissimilarity pairs are more easily discriminated if they are defined on one common dimension, and are even more so if one is subsumed under the other. These situations may cause discriminal processses to covary to different extents.

Whatever the interpretation may be, however, more serious attempts should be extended to recover possibly nonzero correlational terms in the law of comparative judgment. Takane (1975) proposes several plausible ways of restricting correlational terms to avoid underdetermined models, which reduce to either one of structural assumptions on the parameters or of equality constraints.

CONCLUDING REMARKS

As we have seen in the previous sections, our ML estimation procedure for nonmetric MDS (and its Fortran implementation, MAXSCAL-1) seems to work reasonably well in both Monte Carlo and practical situations. It can recover the true underlying structure in the Monte Carlo situation, at least when the observed set of data conveys enough information to uniquely determine the set of points in a space of prescribed geometric features and dimensionality. Furthermore, it can provide additional information as to the statistical behavior of estimates as well as various statistical criteria for the appropriateness of the representation.

Note, however, that at present our ML pocedure is relatively expensive (as is usually the case for an ML procedure); the better statistical properties are not bought without cost. Furthermore, a highly specific algorithm must be constructed for each specific type of experimental operation for obtaining ordinal dissimilarities. We justify our procedure by emphasizing its qualification as the first statistical method for nonmetric MDS. The work is in process which extends the current procedure to other experimental procedures as well as to other representation models.

REFERENCES

MOSTELLER, F. 1951 Remarks on the method of paired comparisons. I. The least squares solution assuming equal standard deviations and correlations. *Psychometrika*, **16**, 203–206.

SAITO, T. 1974 Multidimensional Thurstonian scaling and its applications (I). *Nippon Univac Soken Kiyo*, **4**, 87–112.

SCHÖNEMANN, P. H., & CARROLL, R. M. 1970 Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, **35**, 245–255.

TAKANE, Y. 1975 Generalized uni- and multidimensional Thurstonian scaling: A maximum likelihood method. Unpublished manuscript. The University of North Carolina.

TAKANE, Y. 1977 Statistical procedures for nonmetric multidimensional scaling. Unpublished doctoral dissertation. The University of North Carolina.

TAKANE, Y. 1978 A maximum likelihood method for nonmetric multidimensional scaling: I. The case in which all empirical pairwise orderings are independent—theory. *Japanese Psychological Research*, **20**, 7–17.

TORGERSON, W. S. 1958 *Theory and methods of scaling*. New York: Wiley.

TVERSKY, A., & RUSSO, J. E. 1969 Substitutability and similarity in binary choices. *Journal of Mathematical Psychology*, **6**, 1–12.

(Received July 28, 1977)