Journal of the Institute for Psychological and Physical Science, Vol. 3, No. 1, 101–107, 2011

Review

Asymmetric Multidimensional Scaling
—1. Introduction

Naohito CHINO*

The history and scope of asymmetric multidimensional scaling (abbreviated as asymmetric MDS) are briefly discussed. First, the unidimensional sensory scaling in psychophysics is contrasted with the unidimensional psychological scaling in psychometrics. Second, the law of comparative judgment and the law of categorical judgment are introduced, which constitute the basis for asymmetric MDS. Third, several symmetric MDS' s which are extensions of the unidimensional scalings and theorems which underpin the symmetric MDS are described. Fourth, asymmetric MDS's which are extensions of the symmetric MDS and a theorem which is a basis for asymmetric MDS are discussed.

Keywords: asymmetric MDS; law of comparative judgment, law of categorical judgment; the Young-Householder theorem; the Chino-Shiraiwa theorem; likelihood ratio test; test for symmetry; Euclidian space; Minkowski space; Hilbelt space; information criteria

1 Introduction

As is frequently the case, psychological phenomena which we encounter in daily life are multidimensional in nature. For example, a famous intelligence test, WAIS, uses 11 scales in order to diagnose adult intelligence. Similarly, EPPS consists of 15 scales which diagnose personality. In these tests, the number of dimensions is presupposed. It is also frequently the case that such a number is not known a priori.

Multidimensional scaling, which we shall hereafter abbreviate as MDS, is a method for locating stimuli which vary with respect to an unknown number of dimensions, in a certain dimensional psychological continuum, given a set of dissimilarities between stimuli. Here, stimuli may be physical objects, persons, groups, nations and so on. Of course, such a set of dissimilarities between stimuli can be stacked into a matrix form, $\boldsymbol{\Delta}$ of order n, where n is the number of stimuli. We shall call the method *symmetric MDS* if $\boldsymbol{\Delta}$ is symmetric, and *asymmetric MDS* if it is asymmetric. Symmetric MDS is an extension of unidimensional scaling to be discussed below, and asymmetric MDS is an extension of symmetric MDS.

Asymmetric MDS has a long history. In fact, if we view sensory scalings in psychophysics (e.g., Fechner, 1860; Herrnstein & Boring, 1965; Stevens, 1951; Weber, 1834) as the origin of the unidimensional scaling, asymmetric MDS goes back to the 19th century. However, it may be natural and appropriate not to view sensory scalings as the root of asymmetric MDS, because it is usual that psychological scalings in psychometrics do not necessarily assume any correspondence between the scale values to be attached to stimuli and the physical quantities which belong to them.

In this sense, asymmetric MDS may be said to go back to the early 20th century, especially to the pioneering work of Thurstone (1927a). According to him :

… we have then two criteria, one for the stimuli and one for the discriminal processes of these stimuli. The stimulus continuum must of course be defined in terms of some definite stimulus attribute. The discriminal continuum is a qualitative one which does not necessarily have either magnitude or intensity (Thurstone, 1927a, p. 370).

His seminal work was a watershed of unidimenisional scalings, and it extended the applicability of scaling methods drastically, although one of the key terms in his work, *discriminal process*, which means psychological value to be attached to any object or person in general, sounds today somewhat strange.

In a companion paper (Thurstone, 1927b) he proposed the famous *law of comparative judgment* using the notion

*This paper is based on the author's works on asymmetric MDS since 1977. Requests for reprint should be sent to Naohito Chino, Dept. of Psychology, Aichi Gakuin University, Nissin-cho, Aichi, Japan, or chino@dpc.aichi-gakuin.ac.jp.

of discriminal process, which plays a central role in his theory. Let X_j and X_c be two stimuli on a discriminal continuum. According to this law, the probability that $X_j > X_c$ (e.g., X_j is superior than X_c, or X_j is heavier than X_c, and so on) is expressed as

$$P_{jc} = Prob(X_j > X_c)$$
$$= \frac{1}{\sqrt{2\pi}\sigma_{jc}} \int_0^{\infty} exp\left\{-\frac{1}{2}\left(\frac{y-\mu_{jc}}{\sigma_{jc}}\right)^2\right\} dy, \quad (1)$$

where μ_{jc} is defined as the *discriminal difference*, $\mu_j - \mu_c$. Here, μ_j and μ_c are the psychological values on a discriminal continuum corresponding to the two stimuli, and are respectively the mean values of the two discriminal processes,

$$v_j = \mu_j + \epsilon_j, \quad \epsilon_j \sim N(0, \sigma_j^2), \quad (2)$$

$$v_c = \mu_c + \epsilon_c, \quad \epsilon_c \sim N(0, \sigma_c^2). \quad (3)$$

Here, σ_j^2 and σ_c^2 are variances of the discriminal processes, v_j and v_c, respectively. Moreover, σ_{jc} in Eq. (1) is the standard deviation of the discriminal difference. Eq. (1) can be abbreviated as $P_{jc} = \Phi\left(\frac{\mu_{jc}}{\sigma_{jc}}\right)$. Bock and Jones (1968) call it the *response function*.

Later, Torgerson (1954, 1958) extended the law of comparative judgment to the case of a rating scale with a few ordered categories, and proposed the *law of categorical judgment*. According to this law, the probability that X_j is judged to fall in less than or equal to the kth category is written as

$$P_{jk} = Prob(X_j \leq \tau_k) = \frac{1}{\sqrt{2\pi}\sigma_{j-k}} \int_{-\infty}^{0}$$
$$exp\left\{-\frac{1}{2}\left(\frac{v-(\mu_j-\tau_k)}{\sigma_{j-k}}\right)^2\right\} dv = \Phi(Z_{jk}), \quad (4)$$

where τ_k is the boundary value of the kth category and k + 1th category on the discriminal continuum, and $Z_{jk} = (\tau_k - \mu_j)/\sigma_{j-k}$. Here, σ_{j-k} is the standard deviation of the variable v in Eq. (4).

Ten years after Thurstone's seminal work, referred to above, two papers were published which underpin the classical MDS to be published later. One is Young and Householder (1938) and the other Richardson (1938). The former is concerned with the necessary and sufficient condition that the set of numbers d_{jk} between stimuli constitutes the mutual distances of a real set of points in *Euclidian space*. According to the *Young-Householder theorem*, this condition is that $\boldsymbol{\Delta}$ is *positive semi-definite* (abbreviated hereafter as *p.s.d.*). The latter is concerned with a model of symmetric MDS, which might be viewed as a "preliminary work" of Torgerson's classical MDS, considering Torgerson's citation (Torgerson, 1952) to Richardson's work. The reason that we view Richardson's work as a *preliminary one* is that his work is merely a summary of a 15-minute presentation in an academic meeting and there exists no way of knowing the details of his work at this point. It is Tucker and Messick (1963) that clearly point out that MDS was developed by Richardson (1938), and extended by Torgerson (1952).

Although Young and Householder's theorem by itself does not describe the way of determining the number of dimensions of the space in which stimuli are embedded, Young and Householder (1938) refer to the paper which enables us to fit a lower dimensional set of points to a given set. This is nothing but the work of Eckart and Young (1936), which is concerned with the method for obtaining an $m \times n$ matrix $\boldsymbol{B}$ of rank r that minimizes a Frobenius norm, $\| \boldsymbol{B} - \boldsymbol{A} \|_F$, given an $m \times n$ matrix $\boldsymbol{A}$ of rank k, and a nonnegative integer $r < k$ (e.g., Lawson & Hanson, 1974). Since a solution to the problem can be obtained by utilizing the famous *singular value decomposition* (abbreviated as *SVD*) of $\boldsymbol{A}$ as in Lawson and Hanson, SVD plays an important role in the theory of MDS.

Nowadays the MDS formulated fully by Torgerson is called the *classical MDS*, because Torgerson assumed that the dissimilarities in his MDS had to fulfill a distance property in the Euclidian space, although he called it the *comparative distance* which did not necessarily satisfy the ratio scale level of measurement.

Kruskal (1964a, b) discarded the metric restriction of the classical MDS and extended MDS to the case in which the similarity is measured at an ordinal level. Therefore, his MDS is called the *nonmetric MDS*. It extended the applicability of MDS to a certain extent. Guttman (1968) developed another algorithm for nonmetric MDS called the *Smallest Space Analysis*, abbreviated as SSA, and Lingoes (1973) provided a program series composed of several versions of SSA.

The applicability of MDS was further extended by several researchers who developed the so-called *individual differences MDS* which enables us to examine differences in individuals in some senses. This method assumes in general that interstimulus distances are defined for each individual. Let such a set of data be $\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \cdots, \boldsymbol{\Delta}_m$, where m is the number of individuals.

For example, Tucker and Messick (1963) proposed a *Points of View Analysis* abbreviated as PVA. In the first stage of PVA, we first rearrange elements of each of the data matrices into a column vector of order $n(n-1)/2$. Next, we stack m vectors into an $n(n-1)/2 \times m$ matrix $\boldsymbol{X}$, decompose the inner product matrix $\boldsymbol{X}^t\boldsymbol{X}$ by the Young-Householder theorems (i.e., including the Eckart-Young theorem), and obtain a best fit matrix $\hat{\boldsymbol{X}}_p$ of order p which is less than m. The best fit matrix is further rotated into a simple structure. Each of the column vectors of the rotated matrix $\boldsymbol{F}_p$ is viewed as a *view point* of the dissimilarity judgments. In the second stage we compute some sort of compromise matrices $\boldsymbol{T}_a$, $a = 1, 2, \cdots, p$ in such a way that $\boldsymbol{T}_a = \frac{1}{m}\sum_{i=1}^{m} f_{ia}\boldsymbol{\Delta}_i$, where f_{ia} is the (i, a) element of $\boldsymbol{F}_p$. Finally we administer MDS to each of these compromise matrices, and obtain p configurations of stimuli corresponding to the p view points, respectively.

Carroll and Chang (1970) developed another type of the individual differences MDS called INDSCAL, considering shortfalls of PVA. They assumed a weighted Euclidian distance model, i.e.,

$$d_{ijk} = \sqrt{\sum_{t=1}^{r} w_{it}(x_{jt} - x_{kt})^2}, \quad i = 1, 2, \cdots, m, \quad (5)$$

where w_{it} denotes a weight corresponding to each individual i on each dimension t. It is apparent that INDSCAL assumes a *stimulus space* which is common to all individuals, and a *subject space* which represents individual differences in relative saliences or importances in the dimensions.

Escoufier and his colleagues (e.g., Escoufier, 1973; Lavit, C., Escoufier, Y., Sabatier, R., and Traissac, P., 1994) proposed a third type individual differences MDS model called the *Structuration des Tableaux A Trois Indices de la Statistique (STATIS)* (Structuralization of the three-way statistical table). STATIS consists of the three steps. Step 1 is called the *inter-structure analysis*, which uncovers a multidimensional inter-structure among m occasions. Step 2 is called the *intra-structure analysis*, which discloses a holistic multidimensional intra-structure among n stimuli. Step 3 is called the *trajectory analysis*, which estimates the trajectory of each stimulus through occasions in the holistic intra-structure (e.g., Grorud, Chino, & Yoshino, 1995).

Takane, Young, & de Leeuw (1977) proposed an *Alternating Least Squares algorithm for individual differences sCALing* method called ALSCAL, which includes not only some individual differences MDS methods but also the classical MDS as well as various MDS methods developed by that time including nonmetric MDS. They compiled the fruits of years of study on symmetric MDS developed by that time into an algorithm.

Although symmetric MDS methods developed by the late 1970's were merely *descriptive*, some psychometricians made inroads into the *inferential symmetric MDS* at that time. For example, Ramsay (1977, 1978, 1982) proposed a *maximum likelihood (ML) MDS* called *MULTISCALE*. It assumes that the observed dissimilarities δ_{jkr} obey the log-normal distribution, that is, ln $\delta_{jkr} \sim N(\ln d_{jk}, \sigma^2)$. Here, subscript r denotes the rth replication of dissimilarity judgment on the pair of stimuli, (j, k). In general, unknown parameters such as coordinates of stimuli etc. in ML MDS are estimated so as to maximize the likelihood of the data under some specified distributional assumption.

Takane (1978a, b) developed a *maximum likelihood (ML) method for nonmetric MDS* for a set of empirical orderings on a set of pairs of stimuli, which he calls *MAXSCAL-1*, utilizing Thurstone's law of comparative judgment. In his nonmetric ML MDS, coordinates of stimuli to be recovered are related to the distance d_{ij} between stimuli i and j by the Minkowski power metric, which is called the *representation model*. The d_{ij} is then assumed to error-perturbed to give rise to a psychological value $\lambda_{ij}^{(t)}$ in two ways. One is the normal error, and the other the multiplicative error. He calls such a process the *error model*. Assume here that a new variable Y_{ijkl} equals 1 whenever λ_{ij} is greater than λ_{kl}, and otherwise 0. Then, $Pr(Y_{ijkl} = 1) = Pr(\lambda_{ij} > \lambda_{kl})$ which is an important assumption implied by the law of comparative judgment. For the normal error, $Pr(\lambda_{ij} > \lambda_{kl}) = \Phi(h_{ijkl})$, where $h_{ijkl} = (d_{ij} - d_{kl})/\sigma_{ijkl}$. Likelihood of the data can thus be defined using the function Φ.

Takane (1981) developed another ML MDS called MAXSCAL-2 utilizing law of categorical judgment. As in MAXSCAL-1, this method assumes some representation models and error models. In addition, it presupposes *response models*. That is, subjects place error-perturbed proximities in one of the M rating scale categories $C_1, C_2, \cdots, C_M$. Therefore, the probability that the error-perturbed proximity of stimuli i and j, say, τ_{ij} falls in c_m is given by

$$p_{ijm} = pr(b_{m-1} < \tau_{ij} < b_m). \quad (6)$$

Utilizing the law of categorical judgment, one can then construct the likelihood of the data.

Takane & Carroll (1981) proposed a third ML MDS method in the case when dissimilarity measures are taken by ranking procedures such as the method of conditional rank orders or the method of triadic combinations. They call it *MAXSCAL-4*.

After Ramsay, Takane, and Carroll's pioneering works, some other researchers have also proposed different ML MDS methods (e.g., Groenen, 1993; Groenen, Mathar, & Heisser, 1995; MacKay, 1989). Recently, Oh & Raftery (2001) criticized ML MDS's and proposed a *Bayesian MultiDimensional Scaling* abbreviated as Bayesian MDS (BMDS). Their major criticisms of the ML MDS are (1) justification of ML relies on asymptotic theory, and the number of parameters to be optimized over typically grows as fast as the number of objects, so that the asymptotic theory may not apply in high dimensions (Cox, 1982), (2) the likelihood surface will tend to have many more local minima when there are more dimensions, and finding a good initial estimate will be correspondingly more difficult.

In their BMDS a Euclidian distance model is used and a Gaussian measurement error is assumed in the observed dissimilarity. *Markov Chain Monte Carlo* (MCMC) algorithm is used in order to obtain a Bayesian solution of the stimulus configuration. According to them, their BMDS provided a much better fit to the data than did the classical MDS and a moderately better fit than did ALSCAL in all the examples they tested. A simple Bayesian criterion called MDSIC which is based on the *Bayes factor* or ratio of integrated likelihood is used in order to choose an appropriate dimension.

All the symmetric MDS models discussed above assume that the dissimilarity data matrix $\boldsymbol{\Delta}$ is symmetric. For example, Kruskal (1964a) discusses nonsymmetry of dissimilarities in his seminal paper on nonmetric MDS. He recommends to average δ_{ij} and δ_{ji} if they are measurements on the same underlying quantity, and differ only because of statistical fluctuation. However, we shall frequently encounter asymmetric relationships which might not be considered as statistical fluctuation in daily life.

Asymmetric MDS goes back to the work of Young (1975). He proposed a *weighted Euclidian model*

$$d_{ij} = \sqrt{\sum_{a=1}^{r} w_{ia}(x_{ia} - x_{ja})^2}, \quad w_{ia} \geq 0, \tag{7}$$

in order to handle the problem of nonsymmetry of dissimilarities, and called it the *ASYMSCAL model*. We shall refer it simply as *ASYMSCAL*.

Various asymmetric MDS methods have been proposed by a body of researchers since then. However, as in the case of ASYMSCAL by Young, most of them until recently have merely been concerned with how to represent asymmetric relationships between stimuli in a multidimensional continuum. In other words, most of them have merely discussed representation models.

For example, Gower and Constantine (Constantine & Gower, 1978; Gower, 1977) discuss several possible methods for analyzing asymmetric data including SVD, canonical analysis of asymmetry, a cyclone model, and so on. We shall hereafter abbreviate the canonical analysis of asymmetry as *CASK* (e.g., Chino & Shiraiwa, 1993), although it is sometimes called the *Gower diagram*.

In CASK, the observed dissimilarity data matrix $\boldsymbol{\Delta}$ is decomposed into the symmetric part $\boldsymbol{S}_s$ and the skew-symmetric part $\boldsymbol{S}_{sk}$ first, that is, $\boldsymbol{\Delta} = \boldsymbol{S}_s + \boldsymbol{S}_{sk}$. Then, the latter part is decomposed by SVD to yield,

$$\boldsymbol{S}_{sk} = \boldsymbol{X\Gamma K X}^t, \tag{8}$$

where $\boldsymbol{\Gamma}$ is a diagonal matrix of order n such that $\boldsymbol{\Gamma} = diag(\gamma_1, \gamma_1, \gamma_2, \gamma_2, \cdots, (0))$, and $\boldsymbol{X}$ is an orthogonal matrix of order n. Moreover, $\boldsymbol{K}$ is a special matrix of order n.

As Gower (1977) points out, the elements of $\boldsymbol{S}_{sk}$ do not form a metric, and therefore, any kind of distance interpretation of such diagrams so drawn is unwarranted. To be more precise, CASK employs the standard *symplectic basis* and therefore it is different from that of a Euclidian space, as Chino & Shiraiwa (1993) point out.

By contrast, Chino (1977, 1978, 1990) proposed a similar method which is different in philosophy. The generalized version (Chino, 1990) is called the *Genenalized Inner Product multidimensional SCALing* abbreviated as *GIPSCAL*. The GIPSCAL model is expressed as follows:

$$\boldsymbol{\Delta} = a\boldsymbol{Z}\boldsymbol{Z}^t + b\boldsymbol{Z}\boldsymbol{L}_q\boldsymbol{Z}^t + c\boldsymbol{1}_n\boldsymbol{1}_n^t, \tag{9}$$

where $\boldsymbol{Z}$ is an $n \times q$ configuration matrix, and $\boldsymbol{L}_q$ is a special skew-symmetric matrix of order q. It is apparent from eq. (9) that GIPSCAL represents the symmetric part and the asymmetric part (to be precise, the skew-symmetric part) simultaneously in a configuration space.

Kiers & Takake (1994) simplified and generalized GIPSCAL in such a way that it fits the data better. Trendafilov (2002) reformulated GIPSCAL as an initial value problem for certain first order matrix ordinary differential equations, which results in a globally convergent algorithm for solving it.

Harshman (Harshman, 1978; Harshman et al., 1982) proposed a different and simpler asymmetric MDS model called the *DEcomposition into DIrectional COMponents* abbreviated as *DEDICOM*, which is expressed as

$$\boldsymbol{\Delta} = \boldsymbol{Y A Y}^t, \tag{10}$$

where $\boldsymbol{Y}$ is a configuration matrix, and $\boldsymbol{A}$ is the DEDICOM "core" matrix of order p giving the directional relationship among the basic p types or dimensions.

It was Escoufier & Grorud (1980) that proposed a similar method to GIPSCAL, but with a somewhat curious formulation which they called the *complex coding*. They decompose the observed data matrix $\boldsymbol{\Delta}$ into the symmetric part and the skew symmetric part, and compile them as

$$\boldsymbol{H} = \boldsymbol{S}_s + i\boldsymbol{S}_{sk}, \tag{11}$$

where i denotes the *imaginary number*. It is apparent that this complex matrix is a *Hermitian matrix* of order n. They solved the eigenvalue problem of the matrix, utilizing a traditional method to obtain it's eigenvalues and eigenvectors by computing a special kind of real symmetric matrix of order $2n$.

Using the first eigenvalue and the corresponding eigenvector, they approximate the symmetric part and the skew-symmetric part of the data as follows:

$$s_{jk}(s) \sim \lambda_1(u_{j1}u_{k1} + v_{j1}v_{k1}),$$
$$s_{jk}(sk) \sim \lambda_1(v_{j1}u_{k1} - u_{j1}v_{k1}), \qquad (12)$$

where (u_{j1}, v_{j1}) and (u_{k1}, v_{k1}) are, respectively, the coordinates of stimuli j and k in a real space.

By contrast, Chino & Shiraiwa (Chino, 1991a, b; Chino & Shiraiwa, 1993) proposed the same model as that of Escoufier and Grorud independently, and called it the *Hermitian Canonical Model* abbreviated as *HCM*. They applied the eigenvalue problem directly to the Hermitian matrix in Eq. (11), and obtained the following equation:

$$\boldsymbol{\Delta} = \boldsymbol{X}\boldsymbol{\Omega}_s\boldsymbol{X}^t + \boldsymbol{X}\boldsymbol{\Omega}_{sk}\boldsymbol{X}^t, \qquad (13)$$

where $\boldsymbol{\Omega}_s$ and $\boldsymbol{\Omega}_{sk}$ are special symmetric and skew-symmetric matrices, respectively.

At this point it is natural to ask whether the configuration of stimuli obtained by the complex coding or HCM has a metric property or not. Chino & Shiraiwa (1993) attacked this problem, and found the necessary and sufficient condition that the dissimilarity data is expressible in terms of a certain metric space. To be precise, the necessary and sufficient condition that the set d_{jk} gives the mutual distances of a real (true) set of points in a *finite-dimensional (complex) Hilbert space* is that $\boldsymbol{H}$ is *p.s.d.*. The *Chino-Shiraiwa theorem* is clearly an extension of the Young-Householder theorem to the complex space.

In contrast with the models discussed above, some researchers have proposed *augmented distance models* like Young's ASMSCAL (e.g., Krumhansl, 1978; Okada & Imaizumi, 1987, 1997; Saito, 1991; Saito & Takenda, 1990; Weeks & Bentler, 1982; Zielman & Heiser, 1996). In these models, the Okada and Imaizumi models abbreviated as the OI models are unique in the sense that these handle the *ordinal asymmetric data* based on Kruskal's monotone regression. For example, the one-mode two-way version of the OI models is expressed as

$$d^*_{jk} = d_{jk} - r_j + r_k, \qquad (14)$$

where d^*_{jk} is the augmented distance between stimuli j and k.

We shall briefly refer to some miscellaneous models, Sato (1988), Ten Berge (1997), and Tobler (1976–1977). Sato (1988) proposed an asymmetric MDS model of which distance function is *asymmetric*, that is, the *Minkowski metric function*. It should be noticed that the well-known *Minkowski' r-metric* which is, for example, used in Kruskal's nonmetric MDS, is a *symmetric* distance function. Ten Berge (1997) considered the definition of asymmetry implied in Gower's approach, which might be viewed as a method of reducing asymmetry by rank-one matrices, and suggested a different method. Tobler (1976–1977) proposed a unique method for analyzing the interaction between geographical areas, which is often represented by "from-to" tables.

It is interesting to note that the asymmetric MDS methods discussed above have remained to be *descriptive* until quite recently. Although Chino (1992) proposed an ML asymmetric MDS which is a natural extension of MAXSCAL by Takane (1981), it has yet remained to be completed. Saburi and Chino (2008) developed it further, and called it *ASYMMAXSCAL*. In ASYMMAXSCAL, three submodels, i.e., the representation model, the error model, and the response model are assumed as in MAXSCAL. Any asymmetric MDS models developed up to now can be basically chosen as representation models in ASMMAXSCAL, they picked up the OI model in their paper. In order to compare the goodness of fit of the several models, ASYMMAXSCAL utilizes AIC.

Another outstanding feature of ASYMMAXSCAL is that it enables us to check whether the data is *sufficiently asymmetric*. As in symmetric MDS's, asymmetry had been presupposed without examination in applying any asymmetric MDS to data. By contrast, ASYMMAXSCAL compares any asymmetric MDS model with a saturated model under the symmetry hypothesis and so on using AIC. However, such comparisons merely serve as indirect examinations of asymmetry. Moreover, information criteria such as AIC do not generally consider the nature as well as features of the data. To overcome this difficulty, ASYMMAXSCAL examines the features of the data by utilizing the traditional *test for symmetry* and related tests such as the *test for quasisymmetry* (Caussinus, 1965), and so on (Chino, 2008; Chino & Saburi, 2006, 2008).

For example, Chino and Saburi (2006) proposed to administer a sequential test for the quasi-symmetry hypothesis, the marginal homogeneity hypothesis, the quasi-independent hypothesis, the independent hypothesis, and some of the log-linear hypothesis, prior to analyze asymmetric relational data. However, some questions arise in testing these hypotheses concerning (1) how to arrange the order of testing these symmetry related hypotheses, (2) whether we can control the error of the first kind (and possibly, that of the second kind) (Chino, 2008).

Recently, Chino, Tomizawa, & Saburi (Chino & Saburi, 2009, 2010; Tomizawa, 2010, personal communication) have almost proven that the three likelihood ratio statistics pertaining to symmetry, i.e., those for testing the quasi-symmetry hypothesis, a symmetry hypothesis, and a marginal homogeneity hypothesis are *mutually independent stochastically*. This result enables us to control at least the error of the first kind in some sequential test for the symmetry and related hypothesis. In proving the above problem, theorems due to Hogg & Craig (1956) and Lehmann (1983) are utilized.

References

Basu, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhyā*, **15**, 377-380.

Bock, R. D., and Jones, L. V. (1968). *The measurement and prediction of judgment and choice*. San Francisco: Holden Day.

Carroll, D. J., and Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an *N*-Way generalization of "Eckart-Young" decomposition. *Psychometrika*, **35**, 283–319.

Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation. *Annales de la Faculté des Sciences de l'Université de Toulouse*, **29**, 77–182.

Chino, N. (1977). N-ko no taisyo-kan no hitaisyo na kannkei wo zusikika surutame no ichi-giho [A method for visualizing asymmetric relationships between N objects]. *Proceedings of the 5th annual meeting of the Behaviormetric Society of Japan*, 146–149. Okayama, Japan.

Chino, N. (1978). A graphical technique for representing the asymmetric relationships between *N* objects. *Behaviormetrika*, **5**, 23–40.

Chino, N. (1990). A generalized inner product model for the analysis of asymmetry. *Behaviormetrika*, **27**, 25–46.

Chino, N. (1991a). Fundamental problems in asymmetric scaling. *Paper presented at the symposium on "Multivariate analysis of asymmetric relational data" at the 19th annual meeting of the Behaviormetric Society of Japan*. Nagoya, Japan.

Chino, N. (1991b). A critical review for the analysis of asymmetric relational data. *Bulletin of The Faculty of Letters of Aichi Gakuin University*, **21**, 31–52.

Chino, N. (1992). Metric and nonmetric Hermitian canonical models for asymmetric MDS. *Proceedings of the 20th annual meeting of the Behaviormetric Society of Japan*. (pp. 246–249). Tokyo, Japan.

Chino, N. (2008). Tests for symmetry in asymmetric MDS. In K. Shigemasu, A. Okada, T. Imaizumi, T. Hoshino, (Eds.), *New Trends in Psychometrics* (pp. 33–40). Tokyo: Universal Academic Press.

Chino, N., and Saburi, S. (2006). Tests of symmetry in asymmetric MDS. *Paper presented at the 2nd German Japanese Symposium on Classification—Advances in data analysis and related new techniques & application*. Berlin, Germany, March 7–8.

Chino, N., & Saburi, S. (2009). Controlling the two kinds of error rate in selecting an appropriate asymmetric MDS models. *The 6th International Conference on Multiple Comparison Procedures*, Tokyo, Japan.

Chino, N., and Shiraiwa, K. (1993). Geometrical structures of some non-distance models for asymmetric MDS. *Behaviormetrika*, **20**, 35–47.

Constantine, A. G., and Gower, J. C. (1978). Graphical representation of asymmetric matrices. *Applied Statistics*, **27**, 297–304.

Cox, D. R. (1982). "Comment", *The Journal of the Royal Statistical Society*, Series A (General), **145**, 308–309.

Eckart, C., and Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, **1**, 211–218.

Escoufier, Y. (1973). Le traitement des variables vectorielles [The treatment of vector variables]. *Biometrics*, **29**, 751–760.

Escoufier, Y., and Grorud, A. (1980). Analyse factorielle des matrices carrees non symetriques [Factor analysis of square asymmetric matrices]. In: E. Diday et al. (Eds.), *Data Analysis and Informatics* (pp. 263–276), Amsterdam: North-Holland.

Fechner, G. T. (1907). *Elemente der Psychophysik*. I u. II, 3, Aufl. Leipzig: Breitkopf u. Hartel.

Gower, J. C. (1977). The analysis of asymmetry and orthogonality. In: J. R. Barra, F. Brodeau, G. Romer, B. van Cutsem (Eds.), *Recent Developments in Statistics* (pp. 109–123), Amsterdam: North-Holland.

Groenen, P. J. F. (1993). *The Majorization Approach to Multidimensional Scaling: Some Problems and Extensions*. Leiden: DSWO.

Groenen, P. J. F., Mathar, R., and Heisser, W. J. (1995). The majorization approach to multidimensional scaling for Minkowski distance. *Journal of Classification*, **12**, 3–19.

Grorud, A., Chino, N., and Yoshino, R. (1995). Complex analysis for three-way asymmetric relational data. *Proceedings of the 23th annual meeting of the Behaviometric Society of Japan*, Osaka, Japan.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, **33**, 469–506.

Harshman, R. A.(1978). Models for analysis of asymmetrical relationships among *N* objects or stimuli. *Paper presented at the First Joint Meeting of the Psychometric Society and the Society for Mathematical Psychology*. Hamilton, Canada.

Harshman, R. A., Green, P. E., Wind, Y., and Lundy, M. E. (1982). A model for the analysis of asymmetric data in marketing research. *Marketing Science*, **1**, 205–242.

Herrnstein, R. J., and Boring, E. G. (1965). *A source book in the history of experimental psychology*. Massachusetts: Harvard University Press.

Hogg, R. V., & Craig, A. T. (1956). Sufficient statistics in elementary distribution theory. *Sankhyā*, **17**, 209–216.

Kiers, H. A. L., and Takane, Y. (1994). A generalization of GIPSCAL for the analysis of nonsymmetric data. *Journal of Classification*, **11**, 79–99.

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review*, **85**, 445–463.

Kruscal, J. G. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1–27.

Kruskal, J. B. (1964b). Multidimensional scaling: A numerical method. *Psychometrika*, **29**, 115–129.

Kruskal, J. B. (1968). How to use MDSCAL, a program to do multidimensional scaling and multidimensional unfolding. Bell Telephone Laboratories.

Lavit, C., Escoufier, Y., Sabatier, R., and Traissac, P. (1994). The ACT (STATIS method). *Computational Statistics & Data Analysis*, **18**, 97–119.

Lawson, C. L., and Hanson, R. J. (1974). *Solving Least Squares Problems*. New Jersey: Prentice-Hall.

Lehmann, E. L. (1983). *Theory of Point Estimation*. New York: Wiley. Lingoes, J. C. (1973). *The Guttman-Lingoes nonmetric program series*. Ann Arbor: Mathesis Press.

MacKay, D. (1989). Probabilistic multidimensional scaling: An anisotropic model for distance judgements. *Journal of Mathematical Psychology*, **33**, 187–205.

Oh, M-S, and Raftery, A. E. (2001). Bayesian multidimensional scaling and choice of dimension. *Journal of the American Statistical Association*, **96**, 1031–1044.

Okada, A., and Imaizumi, T. (1987). Nonmetric multidimensional scaling of asymmetric proximities. *Behaviormetrika*, **21**, 81–96.

Okada, A., and Imaizumi, T. (1997). Asymmetric multidimensional scaling of two-mode three-way proximities. *Journal of Classification*, **14**, 195–224.

Ramsay, J. O. (1969). Some statistical considerations in multidimensional scaling. *Psychometrika*, **34**, 167–182.

Ramsay, J. O. (1977). Maximum likelihood estimation in multidimensional scaling. *Psychometrika*, **42**, 241–266.

Ramsay, J. O. (1978). Confidence regions for multidimensional scaling analysis. *Psychometrika*, **43**, 145–160.

Ramsay, J. O. (1982). Some statistical approaches to multidimensional scaling data. *The Journal of the Royal Statistical Society*,Series A (General), **145**, 285–312.

Richardson, M. W. (1938). Multidimensional Psychophysics. *Psychological Bulletin*, **35**, 659.

Rocci, R., and Bove, G. (2002). Rotation techniques in asymmetric multidimensional scaling. *Journal of Computational & Graphical statistics*, **11**, 405–419.

Saburi, S., and Chino, N. (2008). A maximum likelihood method for an asymmetric MDS model. *Computational Statistics and Data Analysis*, **52**, 4673–4684.

Saito, T. (1991). Analysis of asymmetric proximity matrix by a model of distance and additive terms. *Behaviormetrika*, **29**, 45–60.

Saito, T., and Takeda, S. (1990). Multidimensional scaling of asymmetric proximity: model and method. *Behaviormetrika*, **28**, 49–80.

Sato, Y. (1988). An analysis of sociometric data by MDS in Minkowski space. In: K. Matsushita (Ed.), *Statistical Theory and Data Analysis II* (pp. 385–396), Amsterdam: North-Holland.

Takane, Y. (1978a). 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.

Takane, Y. (1978b). A maximum likelihood method for nonmetric multidimensional scaling: II. The case in which all empirical pairwise orderings are independent—Evaluation. *Japanese Psychological Research*, **20**, 105–114.

Takane, Y. (1981). Multidimensional successive categories scaling: A maximum likelihood method. *Psychometrika*, **46**, 9–28.

Takane, Y., and Carroll, J. D. (1981). Nonmetric maximum likelihood multidimensional scaling from directional rankings of similarities. *Psychometrika*, **46**, 389–405.

Takane, Y., Young, F. W., and de Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: An alternative least squares method with optimal scaling features. *Psychometrika*, **42**, 1–68.

ten Berge, J.M.F. (1997). Reduction of asymmetry by rank-one matrices. *Computational Statistics & Data Analysis*, **24**, 357–366.

Thurstone, L. L. (1927a). Psychophysical analysis. *The American Journal of Psychology*, **38**, 368–389.

Thurstone, L. L. (1927b). A law of comparative judgment. *Psychological Review*, **34**, 273–286.

Tobler, W. (1976–77). Spatial interaction patterns. *Journal of Environmental Systems*, **6**, 271–301.

Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, **17**, 401–419.

Torgerson, W. S. (1954). A law of categorical judgment. *The American Psychologist*, **9**, 483.

Torgerson, W. S. (1958). *Theory and Methods of Scaling*. New York: Wiley.

Trendafilov, N. T. (2002). GIPSCAL revisited. A projected gradient approach. *Statistics and Computing*, **12**, 135–145.

Tucker, L. R. and Messick, S. (1963). An individual differences model for multidimensional scaling. *Psychometrika*, **28**, 333–367.

Weeks, D. G., and Bentler, P. M. (1982). Restricted multidimensional scaling models for asymmetric proximities. *Psychometrika*, **47**, 201–208.

Young, F. W. (1975). An asymmetric Euclidian model for multiprocess asymmetric data. *Paper presented at U.S.-Japan Seminar on MDS*, San Diego, U.S.A., 79–88.

Zielman, B., and Heiser, W. J. (1996). Models for asymmetric proximities. *British Journal of Mathematical and Statistical Psychology*, **49**, 127–146.

Final version received, December 31, 2010