Future developments in asymmetric MDS

◯ Naohito Chino
Faculty of Psychological & Physical Science
Aichi Gakuin University

1 Introduction

There has been a body of work on *asymmetric MDS* since Young (1975) proposed the ASYMSCAL (see for example, Chino, 1997; Cox & Cox, 2001). However, there are *a number of problems yet to be solved* on this and other related topics. In this paper we shall raise six such problems.

2 Technical problems to be solved

Problem 1: Asymmetric ML-MDS

There has been much formal literature on (symmetric) maximum likelihood MDS (ML-MDS) (Ramsay, 1969, 1977, 1978, 1982; Suppes & Zinnes, 1963; Takane, 1978a, 1978b, 1981; Takane & Carroll, 1981). However, there has been no published paper on asymmetric maximum likelihood MDS (asymmetric ML-MDS), although Chino (1992) has already proposed an algorithm which is a simple extension of Takane's algorithm (Takane, 1981) to asymmetric data.

In Chino (1992), the representation model is assumed to be

$$g_{jk} = \sum_{t=1}^{n} \lambda_t (r_{jt} r_{kt} + c_{jt} c_{kt}) + \sum_{t=1}^{n} \lambda_t (r_{kt} c_{jt} - r_{jt} c_{kt}) + a_c, \quad (1)$$

and is further assumed that the g_{jk} is error-perturbed by some psychological process in such a manner that

$$\tau_{ijkl} = g_{jk} + e_{ijk}, \quad (2)$$

where τ_{ijkl} is a psychological value (or discriminal process) for subject i corresponding to the proximity from object j to object k at replication l. As the error models, he assumes *the normal distribution model* and *the Beta distribution of the first kind model*.

Asymmetric ML-MDS enables not only the estimation of the *confidence regions* of coordinates of objects but also *tests of dimensionality hypotheses* as is often done in symmetric ML-MDS.

As will be introduced briefly in a later section, these confidence regions might be utilized if we intend to *predict means and variances of coordinates of objects* using the *Kalman filter*, given a set of longitudinal asymmetric proximity matrices (Chino, 2003b).

Problem 2: Nonmetric MDS's for various asymmetric MDS models

Although there is a body of literature on asymmetric MDS, nonmetric MDS algorithms for these MDS models have been restricted to the Okada-Imaizumi models (Okada & Imaizumi, 1987, 1997). It may be necessary and appropriate to develop nonmetric MDS algorithms to be applicable to the other asymmetric MDS models. Such versions might enable us to *choose among these asymmetric MDS models* based on the badness of fit indices such as Kruscal's STRESS formulae. Furthermore, if we develop the Takane-type asymmetric ML-MDS discussed in Problem 1 which includes nonmetric, asymmetric ML-MDS, we might be able to *select the best model based on some test statistic* or *a certain information criterion*.

Problem 3: Various asymmetric MDS approaches to contingency tables

A closely related approach to asymmetric MDS is the analysis of contingency tables (for example, Agresti, 2002; Bishop, Fienberg, & Holland, 1975). Recently, Rooij and Heiser (De Rooij, 2003; De Rooij & Heiser, 2003) have proposed a series of distance models specifically designed for the quasi-symmetry model. Their approach may be contrasted with the mere applications of extant descriptive asymmetric MDS methods to square contingency tables. For example, Rooij and Heiser (2003) approximates observed frequencies not with direct distance models but with *expected frequencies*. As is well known, linear models concerning the expected frequencies are standard and rational approaches to frequency data. Similar approaches may be taken to non-distance models for the asymmetric MDS.

Problem 4: Dynamic asymmetric MDS

There has been an increasing attention to the analysis of longitudinal asymmetric relational data matrices (for example, Ambrosi & Hansohm, 1987; Chino & Nakagawa, 1983, 1990; Grorud, Chino, & Yoshino, 1995; Okada & Imaizumi, 1997). In these models, features of location changes of objects which underlie proximity changes among objects over time are conjectured *a posteriori*, examing the results of analysis, as pointed out by Chino (2003b).

Recently, Imaizumi (2003) has proposed a vector model in which transition frequencies are assumed to be certain functions of the distance between states. By contrast, Chino has proposed an *axiomatic approach* to a set of longitudinal asymmetric proximity matrices, in which some *complex difference system models* are assumed (Chino, 2000, 2002, 2003a, 2003b). In this approach, we suppose that in pairwise relations the following three axioms hold true;

1. *Asymmetric proximities between members force them to change their proximities.*
2. *If one has a positive sentiment to the other, then one will move toward the other in a sociopsychological space.*
3. *If one has a negative sentiment to the other, then one will move away from the other in the space.*

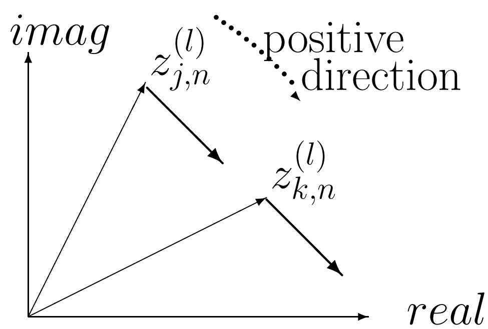

Figure 1: Motion of two members when the positive direction of the lth complex plane is assumed to be clockwise.

Especially, in Chino (2003b) *three types of dynamic MDS models* are proposed as candidates for these longitudinal data matrices. All of these models assume that the theoretical space in which members of a group interact with each other is either a *Hilbert space* or an *indefinite metric space*.

One is *a simple LS method* in which *measurement errors* are assumed for observed proximities, and *deterministic latent processes* are also assumed.

1. *measurement equation*

$$s_{jkt} = \tilde{s}_{jkt} + \varepsilon_{jkt}, \tag{3}$$

where

$$\tilde{s}_{jkt} = \sum_{d=1}^{p} \lambda^{(d)} s_{jkt}^{(d)} + \gamma, \tag{4}$$

$$s_{jkt}^{(d)} = (\rho_{jt}^{(d)}\rho_{kt}^{(d)} + \sigma_{jt}^{(d)}\sigma_{kt}^{(d)}) - (\rho_{jt}^{(d)}\sigma_{kt}^{(d)} + \sigma_{jt}^{(d)}\rho_{kt}^{(d)}),$$

and ε_{jkt}'s are mutually independent.

2. *latent deterministic process equation*

$$\boldsymbol{z}_{j,t+1} = \boldsymbol{z}_{jt} + \sum_{k \neq j}^{N} \boldsymbol{D}_{jk,t} (\boldsymbol{z}_{kt} - \boldsymbol{z}_{jt}),$$

or in scalar form,

$$\rho_{j,t+1}^{(d)} = \rho_{jt}^{(d)} + \sum_{v \neq j}^{N} \alpha_t^{(d)} (\rho_{jt}^{(d)}\sigma_{vt}^{(d)} - \sigma_{jt}^{(d)}\rho_{vt}^{(d)})(\rho_{vt}^{(d)} - \rho_{jt}^{(d)}), \tag{5}$$

$$\sigma_{j,t+1}^{(d)} = \sigma_{jt}^{(d)} + \sum_{v \neq j}^{N} \alpha_t^{(d)} (\rho_{jt}^{(d)}\sigma_{vt}^{(d)} - \sigma_{jt}^{(d)}\rho_{vt}^{(d)})(\sigma_{vt}^{(d)} - \sigma_{jt}^{(d)}), \tag{6}$$

where $\alpha_t^{(d)} \geq 0$.

Another is *a nonlinear regression model* with serially correlated error terms (Gallant, 1987). Let us now consider the model in which errors in the measurement equation discussed above are assumed to be *serially correlated.* In this case it is convenient to stack the longitudinal asymmetric matrices into a T by N^2 matrix

$\boldsymbol{Y}$ as follows:

$$\boldsymbol{Y} = \begin{pmatrix} \left(vec\ \boldsymbol{S}_1^t\right)^t \\ \left(vec\ \boldsymbol{S}_2^t\right)^t \\ \vdots \\ \left(vec\ \boldsymbol{S}_T^t\right)^t \end{pmatrix} = \begin{pmatrix} s_{111}, & s_{121}, & \cdots, & s_{NN1} \\ s_{112}, & s_{122}, & \cdots, & s_{NN2} \\ \cdots & \cdots & \cdots & \cdots \\ s_{11T}, & s_{12T}, & \cdots, & s_{NNT} \end{pmatrix}. \tag{7}$$

Furthermore, let us define

$$\boldsymbol{y}_{jk} = \begin{pmatrix} s_{jk1} \\ s_{jk2} \\ \vdots \\ s_{jkT} \end{pmatrix}, \quad \boldsymbol{\varepsilon}_{jk} = \begin{pmatrix} \varepsilon_{jk1} \\ \varepsilon_{jk2} \\ \vdots \\ \varepsilon_{jkT} \end{pmatrix}.$$

Let us first consider *a univariate nonlinear regression approach*. The simple linear model can be rewritten as

$$\boldsymbol{y}_{jk} = \boldsymbol{f}_{jk}(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_{jk}, \tag{8}$$

where $\boldsymbol{\theta}$ is the parameter vector of the model. Here we assume that the errors ($\varepsilon_{jkt},\ t = 1, 2, \ldots, T$) are *serially correlated*. If the process $\{\varepsilon_{jkt}\}_{t=-\infty}^{\infty}$ is covariance *stationary*, we can estimate *the autocovariance matrix* $\boldsymbol{\Gamma}_T$ of the disturbance vector $\boldsymbol{\varepsilon}_{jk}$, given a sufficient number of sample points. Let it be $\hat{\boldsymbol{\Gamma}}_T$. Suppose then that $\hat{\boldsymbol{\Gamma}}_T^{-1}$ can be factored as $\hat{\boldsymbol{P}}^t\ \hat{\boldsymbol{P}}$. Then we can estimate parameter vector $\boldsymbol{\theta}$ via a usual LS method by replacing $\boldsymbol{y}_{jk}$ and $\boldsymbol{f}_{jk}(\boldsymbol{\theta})$ by $\hat{\boldsymbol{P}}\boldsymbol{y}_{jk}$ and $\hat{\boldsymbol{P}}\boldsymbol{f}_{jk}(\boldsymbol{\theta})$, respectively (for example, Gallant, 1987).

In the case of *multivariate nonlinear regression*, we must estimate covariance matrix $\boldsymbol{\Sigma}_{N^2}$ of N^2 error terms $\varepsilon_{11t},\ \varepsilon_{12t},\ \ldots,\ \varepsilon_{NNt}$.

In any case, this model requires the estimations of the autocovariance matrix $\boldsymbol{\Gamma}_T$ as well as the covariance matrix $\boldsymbol{\Sigma}_{N^2}$ or $\boldsymbol{\Sigma}_N$ based on data. Such requirements seem to be very tough to meet from a practical point of view.

The third is *a nonlinear state space model* in which an observation equation and a state equation are assumed. In the first model, we can examine the qualitative features of the locomotions of objects by utilizing dynamical system theories in mathematics.

In this model we shall assume *two kinds of errors* in both the measurement equation and the latent equation. As a result, such a set of equations can be said to be a nonlinear state space model. In such a case the former equation and the latter equation can be called *the observation equation* and *the state equation*, respectively (Durbin & Koopman, 2001):

1. *the observation equation*

$$\boldsymbol{y}_t = \boldsymbol{X}_t(\boldsymbol{\theta}_t) + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0},\ \boldsymbol{H}_t), \tag{9}$$

2. *the state equation*

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{T}_t(\boldsymbol{\theta}_t) + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0},\ \boldsymbol{Q}_t), \tag{10}$$

for $t = 1, 2, \ldots, T$, where $\boldsymbol{\theta}_1 \sim N(\boldsymbol{a}_1,\ \boldsymbol{P}_1)$, and $\boldsymbol{X}_t(\boldsymbol{\theta}_t)$ and $\boldsymbol{T}_t(\boldsymbol{\theta}_t)$ are differentiable nonlinear functions of $\boldsymbol{\theta}_t$. Moreover, vector $\boldsymbol{y}_t$ is defined as follows:

$$\boldsymbol{y}_t = (s_{11t},\ s_{12t},\ \ldots, s_{1Nt},\ \ldots, s_{NNt})^t \tag{11}$$

According to Durbin and Koopman (2001), after linearizing the $\boldsymbol{X}_t(\boldsymbol{\theta}_t)$ and $\boldsymbol{T}_t(\boldsymbol{\theta}_t)$ at a trial value $\tilde{\boldsymbol{\theta}}_t$, we may apply the standard Kalman filter to the above equations with an appropriate change of notation.

The major purpose of this model is to *forecast the means and variances of coordinates of objects* at time *t+1* based on those at time t utilizing *the Kalman filter*(Durbin & Koopman, 2001).

3 Theoretical problems to be solved

Problem 5: Distribution of the Hermitian matrix in HFM

In the complex coding proposed by Escoufier and Grorud (1980) as well as the Hermitian Form Model (HFM) proposed by Chino and Shiraiwa (1993), observed proximity matrix $\boldsymbol{S}$ of order N, which equals the number of objects, is transformed first into a Hermitian matrix such that $\boldsymbol{H} = \boldsymbol{S}_s + i\ \boldsymbol{S}_{sk}$, where $\boldsymbol{S}_s = (\boldsymbol{S} + \boldsymbol{S}^t)/2$ and $\boldsymbol{S}_{sk} = (\boldsymbol{S} - \boldsymbol{S}^t)/2$.

Although Chino and Shiraiwa examined the existence of the complex metric structures embedded in this Hermitian matrix, they have not considered *the distribution of the latent roots of the matrix.* If we can write down their distribution, we may test the dimensionality hypothesis based on data. Fortunately, we have had various results on *the distribution of complex normal variates* since Wooding (1956), which might throw light on the study under consideration. There seem to have recently been some papers which are *directly* concerned with this problem. These are

1. Camarda, Hs. (1992). Statistical behavior of eigenvalues of real-symmetrical and complex Hermitian band matrices - Comparison with random-matrix theory. *Physical Review A*, **45**, 579-582.

2. Feinberg, J.; Zee A. (1997). Non-hermitian random matrix theory: method of hermitian reduction. *Nuclear Physics B*, **504**, 579-608.

3. Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in

multivariate analysis. *Communications in Statistics*, **1**, 393-417.

Problem 6: Qualitative features of complex difference systems

If the latent process of the complex difference system models are assumed to be *deterministic*, then we can examine the qualitative features of the systems using dynamical system theories in mathematics, especially a multivariate complex dynamical system theory.

We are now interested in examining the mathematical features of the following *complex nonlinear difference system models* (Chino, 2000, 2002, 2003a, 2003b):

$$\boldsymbol{z}_{j,n+1} = \boldsymbol{z}_{j,n} + \sum_{m=1}^{r} \sum_{k \neq j}^{N} \boldsymbol{D}_{jk,n}^{(m)} \boldsymbol{f}^{(m)}(\boldsymbol{z}_{k,n} - \boldsymbol{z}_{j,n}), \quad j = 1, 2, \cdots, N, \tag{12}$$

where,

$$\boldsymbol{f}^{(m)}(\boldsymbol{z}_{k,n} - \boldsymbol{z}_{j,n}) = \begin{pmatrix} \left(z_{k,n}^{(1)} - z_{j,n}^{(1)}\right)^m \\ \left(z_{k,n}^{(2)} - z_{j,n}^{(2)}\right)^m \\ \vdots \\ \left(z_{k,n}^{(p)} - z_{j,n}^{(p)}\right)^m \end{pmatrix}. \tag{13}$$

Moreover, $\boldsymbol{D}_{jk,n}^{(m)} = diag\left\{w_{jk,n}^{(1,m)}, \cdots, w_{jk,n}^{(p,m)}\right\}$, and

$$w_{jk,n}^{(l,m)} = a_n^{(l,m)} r_{j,n}^{(l,m)} r_{k,n}^{(l,m)} \sin\left(\theta_{k,n}^{(l,m)} - \theta_{j,n}^{(l,m)}\right), \tag{14}$$

where $l = 1, 2, \cdots, p,\ m = 1, 2, \cdots, r$.

The complex dynamical system theory is still an active area of research since the innovative works by Cayley, Fatou, and Julia (for example, Peitgen & Richter, 1986; Uno et al., 2003). The following is a list of references which seem to be useful in examining the qualitative features of the behaviors of real or complex nonlinear difference systems:

1. Carleson, L., and Gamelin, T. W. (1993). *Complex Dynamics*. New York: Springer.
2. Chan K-S., and Tong, H. (2001). *Chaos: A Statistical Perspective*. New York: Springer.
3. Elaydi, S. N. (1999). *An Introduction to Difference Equations*. New York: Springer.
4. Peitgen, H.-O., and Richter, P. H. (1986). *The Beauty of Fractals*. New York: Springer.
5. Sullivan, D. (1985). Quasiconformal homeomorphisms and dynamics I: Solution of the Fatou-Julia problem on wandering domains. *Annals of Mathematics*, **122**, 401-418.
6. Sullivan, D. (1985). Quasiconformal homeomorphisms and dynamics II: Structural stability implies hyperbolicity for Kleinian groups. *Acta mathematica*, **155**, 243-260.
7. Ueda, T., Taniguchi, M., Morosawa, S. (1995). Fukuso Rikigakukei Josetsu [Introduction to Complex Dynamical System]. Tokyo: Baifu-kan.