Controlling the Two Kinds of Error Rate in Selecting an Appropriate Asymmetric MDS Model

Naohito CHINO*1, Shingo SABURI*2

ASYMMAXSCAL is revisited first, which is a maximum likelihood asymmetric multidimensional scaling method recently proposed by Saburi and Chino (2008). It is proven that the likelihood ratio test statistic on the quasi-symmetry hypothesis proposed by Caussinus (1965) and that of a marginal homogeneity hypothesis suggested by Andersen (1980) are mutually independent statistically. A possible application of this theorem is indicated to asymmetric relational data in the context of asymmetric multidimensional scaling.

1 Introduction


Although various asymmetric MDS methods have been proposed, these methods have remained only descriptive until recently. By contrast, Saburi and Chino (2008) have proposed a maximum likelihood method for asymmetric MDS called ASYMMAXSCAL, which extends the MAXSCAL proposed by Takane (1981) to asymmetric relational data. As with Takane’s MAXSCAL, it has three kinds of parameters pertaining to the representation model, the error model, and the response model. As for the representation model, the proximity model of object \( O_i \) to object \( O_j \), say, \( g_{ij} \), can generally be written as

\[
g_{ij} = f(x_i, y_j, c),
\]

where \( f(\bullet) \) is any asymmetric MDS model, \( x_i \) and \( y_j \), respectively, are coordinate vectors of \( O_i \) and \( O_j \), and \( c \) is the remaining parameter vector.

As regards the error model, the error-perturbed proximities are written as

\[
\tau_{ij} = y_{ij} + e_{ij}, \quad e_{ij} \sim N(0, \sigma^2),
\]

where \( \tau_{ij} \) is an error-perturbed proximity from \( O_i \) to \( O_j \), and \( e_{ij} \) is an error term.

As for the response model, we assume that subjects place error-perturbed proximities in one of the \( M \) rating scale categories, \( C_1, \ldots, C_M \). Thus, these categories are represented by a set of ordered intervals with upper and lower boundaries on a psychological continuum:

\[-\infty = b_0 \leq b_1 \leq \cdots \leq b_m \leq \cdots \leq b_{M-1} \leq b_M = \infty\]

Accordingly, the probability that the error-perturbed proximity of \( O_i \) to \( O_j \) falls in \( C_m \) is given by

\[
p_{ijm} = \text{prob} \left\{ b_{m-1} < \tau_{ij} \leq b_m \right\}.
\]

We assume that

\[
p_{ijm} = \int_{b_{m-1}}^{b_m} \phi(\tau_{ij}) d\tau_{ij},
\]

based on Torgerson’s law of categorical judgment (Torgerson, 1958). Here, \( \phi(\tau_{ij}) \) denotes the density of the standard normal distribution. For computational convenience, we approximate it by the logistic distribution.

---

*1) Department of Psychology, Faculty of Psychological & Physical Science, Aichi Gakuin University  E-mail: chino@dpc.aichi-gakuin.ac.jp
*2) Graduate School of Psychological & Physical Science, Aichi Gakuin University
We estimate all the parameters pertaining to ASYMMAXSCAL by maximizing the following joint likelihood of the total observations

\[ L = \prod_{i=1}^{n} \prod_{j=1}^{n} \prod_{m=1}^{M} Y_{ijm}^{Y_{ijm}}, \tag{5} \]

where \( Y_{ijm} \) denotes the frequency in category \( C_{im} \) in which subjects place the error-perturbed proximity of \( O_i \) to \( O_j \).

As with any model, ASYMMAXSCAL has various advantages and shortcomings. On the one hand, it enables us to determine the appropriate scaling level of the data by AIC, that is, ordinal, interval, or ratio level. It also enables us to determine the appropriate dimensionality of the model under study by AIC. Moreover, it enables us to examine whether the data are sufficiently asymmetric or not by applying some tests for symmetry prior to the scaling of objects, and by selecting a model among several candidates including some symmetry models, using AIC, on the way to the scaling.

The reason why we can apply some tests for symmetry is as follows. That is, the data obtained by the above method is a set of one-way tables, each of which corresponds to a frequency distribution of a group of subjects concerning the similarity judgment on a directional pair of objects. We call it the Type A design data, or the Type A data. If we rearrange the Type A data per rating scale category, we get \( M \) square contingency tables of order \( n \). We call them Type B (design) data. It is a bit different from the data obtained by traditional designs for the \( n \times n \times M \) table (Agresti, 2002, Bishop et al., 1975).

According to Saburi and Chino (2008), given a Type B data, under the null hypothesis,

\[ H_{0}^{(cs)} : \ p_{ijm} = p_{jim}, \quad i \neq j, \quad m = 1, \ldots, M, \tag{6} \]

the likelihood ratio test statistic,

\[ G^2 = 2 \sum_{m=1}^{M} \sum_{i=1}^{n} \sum_{j=1}^{n} Y_{ijm} \left\{ \ln Y_{ijm} - \ln \left( \frac{n_{ij}Y_{ijm} + Y_{jim}}{n_{ij} + n_{ji}} \right) \right\}, \tag{7} \]

asymptotically follows the central \( \chi^2 \)-distribution with \((M - 1)n(n - 1)/2\) degrees of freedom.

By contrast, the traditional conditional symmetry test with the special \( n \times n \times M \) contingency table, of which test statistic is given by

\[ G^2 = 2 \sum_{m=1}^{M} \sum_{i=1}^{n} \sum_{j=1}^{n} Y_{ijm} \left\{ \ln Y_{ijm} - \ln \left( \frac{Y_{ijm} + Y_{jim}}{2} \right) \right\}, \tag{8} \]

with \( Mn(n-1)/2 \) degrees of freedom under the null hypothesis.

At present, ASYMMAXSCAL enables us to select the most appropriate model among several candidate models which include some variants of symmetry models using AIC. However, such a model selection method by some information criterion does not consider the nature of the data. It will sometimes be necessary to select the representation model which reflects the nature of the data most.

To do this job, it is helpful to utilize various symmetry related tests which have been developed in the branch of mathematical statistics. Chino and Saburi (2006) attempted to administer these tests sequentially prior to the scaling step of ASYMMAXSCAL. They include tests for symmetry, quasi-symmetry, quasi-independence, independence, and some versions of marginal homogeneity. However, the relation of inclusion of these tests is rather complicated. Moreover, in performing such sequential tests, they have not taken overall statistical errors into account. It will be interesting and useful to examine whether these tests are mutually independent statistically or not. For simplicity, we shall hereafter exclusively consider a two-dimensional square contingency table.


However, there exists a small amount of literature which considers the statistical independence among test statistics on these hypotheses. For example, Goodman (1985) discusses the relationships between several symmetry and related hypotheses with respect to their implications and degrees of freedom and applied each of these models to a famous \( 4 \times 4 \) cross-classification table. However, he neither discusses the statistical independence of the test statistics corresponding to these models nor considers the problem of controlling the errors of the two kinds.

Tomizawa (1992) points out a hierarchical tree structure of some double symmetry hypotheses in addition to the symmetry hypothesis and the quasi-symmetry hypothesis, and applies each of these models to two sets of cross-classification table. However, he does not refer to the statistical independence of the test statistics, although he compares these models using Akaike’s information criterion (Akaike, 1974).

In this paper we show that the LR test statistic on the quasi-symmetry hypothesis proposed by Caussinus (1965)
and that of a marginal homogeneity hypothesis suggested by Andersen (1980) are mutually independent statistically, based on the theorems by Basu (1955) and Hogg (1961). As a result, we can control the error of the first kind when we test them sequentially. Furthermore, we can construct a more powerful test than Dunn’s and Holm’s, if we set the error rate of the quasi-symmetry test to \(\alpha\) and set that of the marginal homogeneity test under the quasi-symmetry hypothesis to \(\omega/2\), according to Hochberg (1988).

2 Statistical independence

In order to prove the statistical independence of the two statistics discussed in the previous section, we will define the following parameter spaces of some of the statistics under study according to the log-linear model (Birch, 1963). First, the total parameter space for \(\theta = (\theta_{ij}^{(12)}, \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{i,j}^{(0)})\) corresponding to an \(r \times r\) cross classification table is

\[
\Omega = \left\{ (\theta_{ij}^{(12)}, \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{i,j}^{(0)}) \middle| \begin{array}{c}
-\infty < \theta_{ij}^{(12)} = \theta_{j,i}^{(12)} < \infty, \\
i, j = 1, 2, \ldots, r
\end{array} \right\}. \tag{9}
\]

By contrast, the parameter space pertaining to the quasi-symmetry hypothesis is

\[
\omega_{QS} = \left\{ (\theta_{ij}^{(12)}, \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{ij}^{(0)}) \middle| \begin{array}{c}
-\infty < \theta_{ij}^{(12)} = \theta_{j,i}^{(12)} < \infty, \\
-\infty < \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{ij}^{(0)} < \infty, \\
i, j = 1, 2, \ldots, r
\end{array} \right\}, \tag{10}
\]

and the parameter space associated with the marginal homogeneity hypothesis proposed by Andersen (1980) is

\[
\omega_{MHo} = \left\{ (\theta_{ij}^{(12)}, \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{ij}^{(0)}) \middle| \begin{array}{c}
\theta_{ij}^{(12)} = 0, -\infty < \theta_{i}^{(1)} = \theta_{j}^{(2)} < \infty, \\
-\infty < \theta_{ij}^{(0)} < \infty, \\
i, j = 1, 2, \ldots, r
\end{array} \right\}. \tag{11}
\]

Clearly, we have

\[
\Omega = \omega_{0} \supset \omega_{QS} \supset \omega_{MHo}. \tag{12}
\]

In addition, we define two parameter spaces related to \(\omega_{QS}\) and \(\omega_{MHo}\). One is the space associated with the symmetry hypothesis in terms of the log-linear model as

\[
\omega_{S} = \left\{ (\theta_{ij}^{(12)}, \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{ij}^{(0)}) \middle| \begin{array}{c}
-\infty < \theta_{ij}^{(12)} = \theta_{j,i}^{(12)} < \infty, \\
-\infty < \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{ij}^{(0)} < \infty, \\
i, j = 1, 2, \ldots, r
\end{array} \right\}. \tag{13}
\]

The other is the parameter space for the equality of the row and column effects in the log-linear model,

\[
\omega_{ERC} = \left\{ (\theta_{ij}^{(12)}, \theta_{i}^{(1)}, \theta_{j}^{(2)}, \theta_{ij}^{(0)}) \middle| \begin{array}{c}
-\infty < \theta_{i}^{(1)} = \theta_{j}^{(2)} < \infty, \\
i, j = 1, 2, \ldots, r
\end{array} \right\}. \tag{14}
\]

Suppose now that we wish to test

\[
H_{0}^{QS} : \theta \in \omega_{QS} \quad \text{against} \quad H_{1}^{QS} : \theta \in \Omega - \omega_{QS}, \tag{15}
\]

and then

\[
H_{0}^{MHo} : \theta \in \omega_{MHo} \quad \text{against} \quad H_{1}^{MHo} : \theta \in \Omega - \omega_{MHo}. \tag{16}
\]

It is apparent that under \(H_{0}^{QS}\), \(\omega_{MHo}\) is contained in \(\omega_{S}\). In other words, \(H_{0}^{MHo}\) is considered as a special symmetry hypothesis under \(H_{0}^{QS}\). It should be noted here that the likelihood ratio statistic for testing a marginal homogeneity hypothesis

\[
H_{0}^{MHo} : \theta \in \omega_{MHo} \quad \text{against} \quad H_{1}^{MHo} : \theta \in \Omega - \omega_{MHo}, \tag{17}
\]

is written as

\[
G_{MHo}^{2} = 2 \sum_{i=1}^{r} f_{i} \left\{ \ln f_{i} - \ln \frac{f_{i} + f_{j}}{2} \right\} + 2 \sum_{j=1}^{r} f_{j} \left\{ \ln f_{j} - \ln \frac{f_{i} + f_{j}}{2} \right\}, \tag{18}
\]

and follows asymptotically the \(\chi^{2}\) distribution with \(r(r-1)\) degrees of freedom (Andersen, 1980).

Next, let us resolve the likelihood ratio statistic for testing \(H_{0}^{MHo}\) into,

\[
\lambda_{MHo} = \frac{L(\hat{\omega}_{MHo})}{L(\hat{\omega}_{0})} = \frac{L(\hat{\omega}_{QS}) L(\hat{\omega}_{MHo})}{L(\hat{\omega}_{0}) L(\hat{\omega}_{QS})} = \lambda_{QS} \lambda_{MHo}. \tag{19}
\]

Then, we have

\[
-2 \ln \lambda_{MHo} = -2 \ln \lambda_{MHo} - (-2 \ln \lambda_{QS}), \tag{20}
\]

or
\[ G_{MH_0}^2 = G_{MH_0}^2 - G_{QS}^2. \]  

(21)

It is evident that \( G_{MH_0}^2 \) is not the same as the likelihood ratio statistic \( G_{MH_0} \), for testing the marginal homogeneity symmetry hypothesis by Andersen (1980).

Moreover, note that the original marginal homogeneity hypothesis \( H_0^{MH} \) proposed by Cramér (1946), which is written as

\[ H_0^{MH} : \mu_{i\bullet} = \mu_{\bullet j}, \]  

(22)

is not necessarily equivalent to the equality hypothesis of the row effect and the column effect

\[ H_0^{ER} : \theta_i^{(1)} = \theta_j^{(2)}, \]  

(23)


It is well known that the LR statistic for testing \( H_0^{QS} \) against \( H_1^{QS} \) is

\[ G_{QS}^2 = -2 \ln \lambda_{QS} = 2 \sum_{i=1}^{r} \sum_{j=1}^{r} f_{ij} (\ln f_{ij} - \ln \hat{\mu}_{ij}), \]  

(24)

where \( \hat{\mu}_{ij} \) are the LR estimates of \( \mu_{ij} \), which satisfy

\[ f_{i\bullet} = \hat{\mu}_{i\bullet} \quad f_{\bullet j} = \hat{\mu}_{\bullet j}, \quad \text{and} \quad f_{ij} + f_{ji} = \hat{\mu}_{ij} + \hat{\mu}_{ji}. \]  

(25)

Under \( H_0^{QS} \), \( G_{QS}^2 \) follows asymptotically the \( \chi^2 \) distribution with \((r - 1)(r - 2)/2\) degrees of freedom.

Notice that the statistics, \( f_{i\bullet}, f_{\bullet j}, \) and \( f_{ij} + f_{ji} \) corresponding to the nuisance parameters \( \theta_i^{(1)}, \theta_j^{(2)}, \) and \( \theta_{ij0}^{(12)} \) under the quasi-symmetry hypothesis, give their complete sufficient statistics because of the nature of the exponential family of the joint distribution of the data under study.

Moreover, the statistic \( G_{QS}^2 \) is free of these nuisance parameters under the hypothesis of quasi-symmetry, since the term \( \hat{\mu}_{ij} \) are estimated as functions of the data, that is, \( f_{i\bullet}, f_{\bullet j}, \) and \( f_{ij} + f_{ji} \). In other words, the statistic \( G_{QS}^2 \) is a function of the complete sufficient statistics discussed above. As a result, in accordance with independence theorems due to Basu (1955) and Hogg (1961), \( G_{QS}^2 \) and \( G_{MH_0}^2 \) which is a function of the complete sufficient statistics discussed above are stochastically independent. (Q.E.D.)

3 A possible application of the theorem to asymmetric relational data

The extant asymmetric MDS models referred to in the introductory section are diverse in character. However, if we want to classify them into two, then one such classification can be made by using the concept of a quasi-symmetry-like family of asymmetric MDS (Chino & Saburi, 2009). Here, by the quasi-symmetry-like family of the asymmetric MDS models we mean members of those models composed of parameters similar to the row effect, the column effect, and the interaction effect of the log-linear model of contingency table.

These are the distance-association model (de Rooij & Heiser, 2003, 2005), the distance-density model (Krumhansl, 1978), the Okada-Imaizumi model (Okada & Imaizumi, 1987, 1997), the Saito model (Saito, 1991), the Saito-Takeda model (Saito & Takeda, 1990), the slide-vector model (Kruskal, 1973; Zielman & Heiser, 1993), the Weeks-Bentler model (Weeks & Bentler, 1982), and the wind model (Gower, 1977). De Rooij and Heiser (2003) have originally referred to them as models related to their distance-association model.

By contrast, there are some models which do not belong to the above family. They are Complex Coding (Escofier & Grorud, 1980), DEDICOM (Harshman, 1978; Harshman et al., 1982), GIPSCAL (Chino, 1978, 1990), generalized GIPSCAL (Kiers & Takane, 1994), GIPSCAL by a projected gradient approach (Trendafilov, 2000), the Gower diagram (Constantine & Gower, 1978; Gower, 1977), HFM (Chino & Shiraiwa, 1993), the Rander’s metric model (Sato, 1989). We shall call them the non-quasi-symmetry-like family of the asymmetric MDS models.

Asymmetric relational data of Type A discussed in the introductory section can be tested sequentially. That is, if \( H_0^{QS} \) is tested against \( H_1^{QS} \), and is rejected, it might be necessary and appropriate to apply some of the non-quasi-symmetry-like family of asymmetric MDS to the asymmetric relational data.

If \( H_0^{QS} \) is accepted, then we may proceed to the marginal homogeneity test under the quasi-symmetry hypothesis, that is, to test \( H_0^{MH_0} \) against \( H_1^{MH_0} \). If this test is accepted, we should apply some of the extant symmetric MDS models like MAXSCAL. However, even if it is rejected, this does not exclude the possibility that the data is symmetric. In this sense, this test is restrictive.

We have been conjectured that the likelihood ratio test statistics on the quasi-symmetry hypothesis proposed by Caussinus (1965) and that of a version of the symmetry hypothesis suggested first by him, that is,

\[ G_{S^*}^2 = G_S^2 - G_{QS}^2, \]  

(26)

are mutually statistically independent (Chino & Saburi, 2009). However, these two statistics are not statistically independent at least exactly, because \( G_{MH_0} \) is not the function of the complete sufficient statistics, \( f_{i\bullet}, f_{\bullet j} \), and \( f_{ij} + f_{ji} \). Tomizawa (2009, personal communication, December 23, 2009) conjectures that these are statistically independent asymptotically.
References
Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship...

Final version received, December 24, 2009