

Bayesian multidimensional scaling for the estimation of a Minkowski exponent

KENSUKE OKADA

Senshu University, Kanagawa, Japan

AND

KAZUO SHIGEMASU

Teikyo University, Tokyo, Japan

The Minkowski property of psychological space has long been of interest to researchers. A common strategy has been calculating the stress in multidimensional scaling for many Minkowski exponent values and choosing the one that results in the lowest stress. However, this strategy has an arbitrariness problem—that is, a loss function. Although a recently proposed Bayesian approach could solve this problem, the method was intended for individual subject data. It is unknown whether this method is directly applicable to averaged or single data, which are common in psychology and behavioral science. Therefore, we first conducted a simulation study to evaluate the applicability of the method to the averaged data problem and found that it failed to recover the true Minkowski exponent. Therefore, a new method is proposed that is a simple extension of the existing Euclidean Bayesian multidimensional scaling to the Minkowski metric. Another simulation study revealed that the proposed method could successfully recover the true Minkowski exponent. BUGS codes used in this study are given in the Appendix.

Since its first development by Torgerson (1952), multidimensional scaling (MDS) has been used as a method for modeling psychological phenomena in low-dimensional space. In many such applications, the geometry of MDS serves as a model of psychological space. The distance function defined on psychological space then serves as a model of mental arithmetic (Borg & Groenen, 2005, chap. 17). The Minkowski family of distance functions has been considered by many previous studies as a prospective candidate for a psychological distance function. For any value $r \geq 0$, the Minkowski distance between two points \mathbf{x}_i and \mathbf{x}_j in \mathbb{R}^p is defined by

$$d_{ij} = \left[\sum_{k=1}^p |x_{ik} - x_{jk}|^r \right]^{1/r}, \quad (1)$$

where the parameter r is called the *Minkowski exponent* and k represents the k th dimension. In a special case of the Minkowski distance, in which $r = 2$, Equation 1 yields the well-known *Euclidean distance*. When $r = 1$, it yields the *city block distance*, which is also known as the *Manhattan distance* or *taxicab distance*.

Kruskal (1964) first presented a method to find the best Minkowski metric when analyzing dissimilarity data by MDS. His approach was simple: First, compute MDS solutions for several different Minkowski exponents r and then choose the one that results in the lowest stress, which

is a loss function used in MDS analysis (typical examples of stress are shown in the next paragraph). Arnold (1971) essentially reformulated the same procedure, with a detailed description. In the original article, Kruskal applied this method to the experimental data presented by Ekman (1954), who obtained directly judged similarities data of pure spectral colors. Since that study, there have been many investigations that compared the stress value of different Minkowski metrics to study the property of the psychological space for specific subject areas. Arabie (1991), Borg and Groenen (2005, chap. 17), and Townes and Abbott (1974) provided reviews of such studies.

However, from a modern psychometric point of view, several drawbacks for this popular method are identified. A primary problem can be attributed to the loss function that is minimized in MDS analysis. In Kruskal's (1964) approach, the so-called *Kruskal's stress-1*,

$$\sigma_1 = \sqrt{\frac{\sum_{i>j} (\delta_{ij} - d_{ij})^2}{\sum_{i>j} d_{ij}^2}}, \quad (2)$$

where δ_{ij} is the observed dissimilarity of objects i and j , was minimized. This approach was taken in many other application studies. However, we have no definite reason to exclude other possible loss functions. In fact, many other measures of fit for MDS analysis have been proposed. For a review

K. Okada, ken@psy.senshu-u.ac.jp



of such MDS measures of fit, see Borg and Groenen (2005, chap. 11). For example, when \bar{d} is the average distance, other candidates for the loss functions include *stress-2*,

$$\sigma_2 = \sqrt{\frac{\sum_{i>j} (\delta_{ij} - d_{ij})^2}{\sum_{i>j} (d_{ij} - \bar{d})^2}}, \tag{3}$$

normalized stress,

$$\sigma_n = \frac{\sum_{i>j} (\delta_{ij} - d_{ij})^2}{\sum_{i>j} \delta_{ij}^2}, \tag{4}$$

s-stress,

$$\sigma_s = \sum_{i>j} (\delta_{ij}^2 - d_{ij}^2)^2, \tag{5}$$

and more.

Choosing a suitable loss function is very important, because different loss functions result in different values of the optimized Minkowski exponent. However, from the above list of the proposed stress measures, it is difficult to choose a single, universal measure.

Furthermore, although in all of the stress functions listed above, the degree of deviation between the observed dissimilarity and proximity is evaluated in terms of the sum of squared Euclidean distance—that is, $\sum_{i>j} (\delta_{ij} - d_{ij})^2$ —this is just a common practice. Instead, a loss that is based on Minkowski distance could be (and maybe should be) used for MDS with Minkowski distances. For example, when we are to calculate the loss for the MDS with Minkowski distance of order r , we could calculate the Minkowski version of *stress-1* as

$$\sigma_1^{(r)} = \left[\frac{\sum_{i>j} |\delta_{ij} - d_{ij}|^r}{\sum_{i>j} d_{ij}^r} \right]^{1/r}, \tag{6}$$

instead of the Euclidean version (Equation 2). Again, the choice of such a loss function will affect the optimized value of r that minimizes the loss function.

This kind of arbitrariness problem in the loss function occurs because of a common strategy of dealing with Minkowski distance in MDS. As described previously, the common strategy is to fix the Minkowski exponent r , to perform MDS and calculate the stress value, and to then repeat these processes for several r values to find the optimal r that results in the lowest stress value. However, in the present article, we treat the Minkowski exponent r as a parameter and directly estimate it and other parameters by using the Bayesian MDS technique (Oh & Raftery, 2001; Okada & Shigemasa, 2009). Note that almost all existing Minkowski MDS applications have used a least-squares type MDS. On the other hand, by using Bayesian estimation, we can treat the Minkowski exponent r as a parameter and directly evaluate its marginal posterior distribution. This solves the arbitrariness problem because we no longer have to choose a loss function. Of course, the benefits of Bayesian inference, such as the availability of a point estimate, the whole posterior distribution of parameters, and the ability to incorporate prior uncertainty,

are gained (see, e.g., Box & Tiao, 1992; Gelman, Carlin, Stern, & Rubin, 2004).

In fact, Lee (2008) used such a method as one of the three case studies in which he applied Bayesian analysis to psychological models. Lee’s model and method were proposed for individual subject data, which are an $n \times n \times m$ array representing the (dis)similarity of objects i and j ($i, j = 1, \dots, n$) judged by a subject t ($t = 1, \dots, m$). However, in many studies dealing with the Minkowski exponent, this kind of data is often averaged to obtain an $n \times n$ matrix, which is then used for MDS analysis (for reviews of such studies, see Arabie, 1991; Borg & Groenen, 2005). Also, MDS is often applied to the data of a single $n \times n$ (dis)similarity matrix. Although Lee’s method is theoretically useful and seems to work well for individual subject data, it is less certain whether this method is directly applicable to the averaged or single data.

The present article is organized as follows: First, the natural averaged data version of Lee’s (2008) individual subject model is described. Then, we present a numerical simulation study to evaluate this method and show that the result is not very satisfactory. Next, we propose a new Bayesian MDS methodology to estimate the Minkowski exponent. The evaluation of this proposed method is based on a numerical simulation study similar to the previous one. Finally, the conclusions are presented.

The implementation of Bayesian inference has become easier with the advent of MCMC software, such as WinBUGS (Lunn, Thomas, Best, & Spiegelhalter, 2000) and OpenBUGS (Spiegelhalter, Thomas, Best, & Lunn, 2007). The Appendix provides BUGS codes used in this study. Using these code sequences, readers can repeat the present study as well as apply the proposed method to their own data.

AVERAGED DATA VERSION OF LEE’S (2008) METHOD

As mentioned in the previous section, Lee (2008) proposed a new approach for finding the best Minkowski metric. The novelty of his method lies in treating the Minkowski exponential r as a random variable and estimating it using the Bayesian estimation technique. In this section, we describe a straightforward extension of Lee’s method to the averaged data problem.

In Lee’s (2008) model, the observed similarity is assumed to be generated as the exponential decay of the distance of points in multidimensional space; however, it is subject to noise. The noise process is assumed to be normal with zero means. Therefore, the model distribution for the observed similarity s_{ij} of objects i and j is

$$s_{ij} \sim N \left[\exp(-d_{ij}), \omega^2 \right], \tag{7}$$

where d_{ij} is the Minkowski distance between objects i and j (Equation 1).

For prior distributions, Lee (2008) modeled that the coordinate x_{ik} has an equally prior probability of being anywhere in a sufficiently large interval with bounds $(-\iota, +\iota)$,

$$x_{ik} \sim U(-\iota, +\iota), \tag{8}$$

where $\iota > 0$, and $U(\alpha, \beta)$ denotes uniform distribution between α and β . The term *sufficiently large* was used to identify an interval large enough that increasing the value of ι almost does not change the posterior distribution over the coordinate point parameters. The prior distribution for Minkowski exponent r was a uniform distribution with explicit end points,

$$r \sim U(0, 2). \quad (9)$$

The prior distribution for the variance parameter ω^2 was an inverse gamma distribution,

$$\omega^2 \sim \text{IG}(\xi, \zeta), \quad (10)$$

where ζ was set to near zero. With these settings, the posterior distribution of r (as well as other parameters) was estimated using the MCMC algorithm.

The above method proposed by Lee (2008) is important, because it introduced a new concept to the problem of finding the best Minkowski exponent: the Minkowski exponent as a random variable, which is to be estimated in the analysis as well as other MDS parameters. His article confirmed that this method worked for individual subject data. However, as we noted previously, it is not known whether this method is directly applicable to averaged or single data, which are common in psychology and behavioral science. We emphasize this partly because, from the knowledge of former studies in statistics, it could be troublesome to set a lognormal distribution on the observed similarity with the mean Minkowski distance when the objective is to estimate the Minkowski exponent value (r in Equation 1, where d_{ij} satisfies Equation 7). It is known that a lognormal distribution is very heavily tailed, and sometimes the moment of higher orders makes parameter estimation unstable (Aitchinson & Brown, 1957; Schmoyer, Beauchamp, Brandt, & Hoffman, 1996). Additionally, in this setting, the objective is the estimation of the parameters of the distance function, which is, simultaneously, a location parameter of a distribution function. This situation is uncommon in standard statistical parameter estimation, and, therefore, little knowledge about when the estimation succeeds is available. Therefore, it is our view that a simulation study is needed for this method to be applicable. Consequently, in the next section, we provide a simulation study that investigates an averaged data version of Lee's method.

NUMERICAL EXPERIMENT 1

As noted above, the aim of the first simulation is to assess the performance of the natural extension of the existing Minkowski exponent estimation proposed by Lee (2008). For this purpose, we generated random data from the true model (that is, the model explained in the previous section) and then analyzed it with Lee's method.

The simulation data were generated as follows. The true configuration matrix \mathbf{X} was set to be a 15×2 matrix. This means the number of dimensions p was set to 2. Each element of \mathbf{X} was generated from a uniform distribution,

$$x_{ik} \sim U(-\iota_0, \iota_0), \quad (11)$$

which corresponds to the true model structure of Equation 8. Two conditions—namely, $\iota_0 = 1.5$ and $\iota_0 = 0.15$ —were considered for the values of ι_0 . The generated configuration \mathbf{X} was transformed to the true distance $\mathbf{D} = \{d_{ij}\}$ by Equation 1. In doing so, three conditions of the true Minkowski exponents were used: $r = 1$ (city block metric), 1.5 (in between), and 2 (Euclidean metric). Then, it was transformed to the observed similarity matrix, $\mathbf{S} = \{s_{ij}\}$, by

$$s_{ij} = \exp(-d_{ij}), \quad (12)$$

which corresponds to Equation 7, when $\omega^2 = 0$ (i.e., when no error is introduced).

Fifty independent data sets of the artificially “observed” similarity matrix \mathbf{S} were generated as above for each of the three true Minkowski exponent conditions and for each of the two true uniform distribution parameter (ι_0) conditions. The artificial data were then analyzed by Lee's (2008) method. The analysis was performed using MATLAB (The MathWorks, Inc., Natick, MA) and OpenBUGS (Spiegelhalter et al., 2007) via the use of the MATBUGS program (Murphy & Mahdavian, 2005). We modified Lee's OpenBUGS script¹ for our data and used it in the analysis, using 10 as a “sufficiently large” value of ι in Equation 8.² The first 20,000 MCMC iterations were discarded as burn-in, and the following 40,000 iterations were used for posterior estimation.

Because in Lee's (2008) method, the Minkowski exponent r is restricted to the interval $[0, 2]$ (Equation 9), the posterior distribution of r is typically not symmetric. Therefore, rather than the posterior means, the posterior mode is appropriate for the point estimate of the Minkowski exponent. To find a posterior mode of Minkowski exponent r , the kernel density estimation with a normal (Gaussian) kernel is used to fit a continuous density. Then, the point that gives the mode of the fitted continuous density is used as a posterior point estimate. Furthermore, we recorded the 95% credibility intervals of r for use in interval estimation.

Result

The resultant posterior mode estimates of the Minkowski parameter r and its 95% credibility intervals are plotted for all trials in Figure 1. From the result, it is clear that the estimated value of the Minkowski parameter is a little smaller than, but close to, 2, regardless of the true value. Therefore, it was determined that the estimate from the natural extension of Lee's (2008) method to the averaged data problem is strongly biased. Moreover, it would not be appropriate for investigating the value of the true Minkowski exponent in the averaged data problem. Some possible reasons why this method did not work are presented in the Conclusions section.

PROPOSED METHOD

The simulation study in the previous section showed that, although the idea of treating the Minkowski parameter r as a random variable seems promising, the “naive” extension of Lee's (2008) method to the averaged data problem may not be adequate for analyzing the true

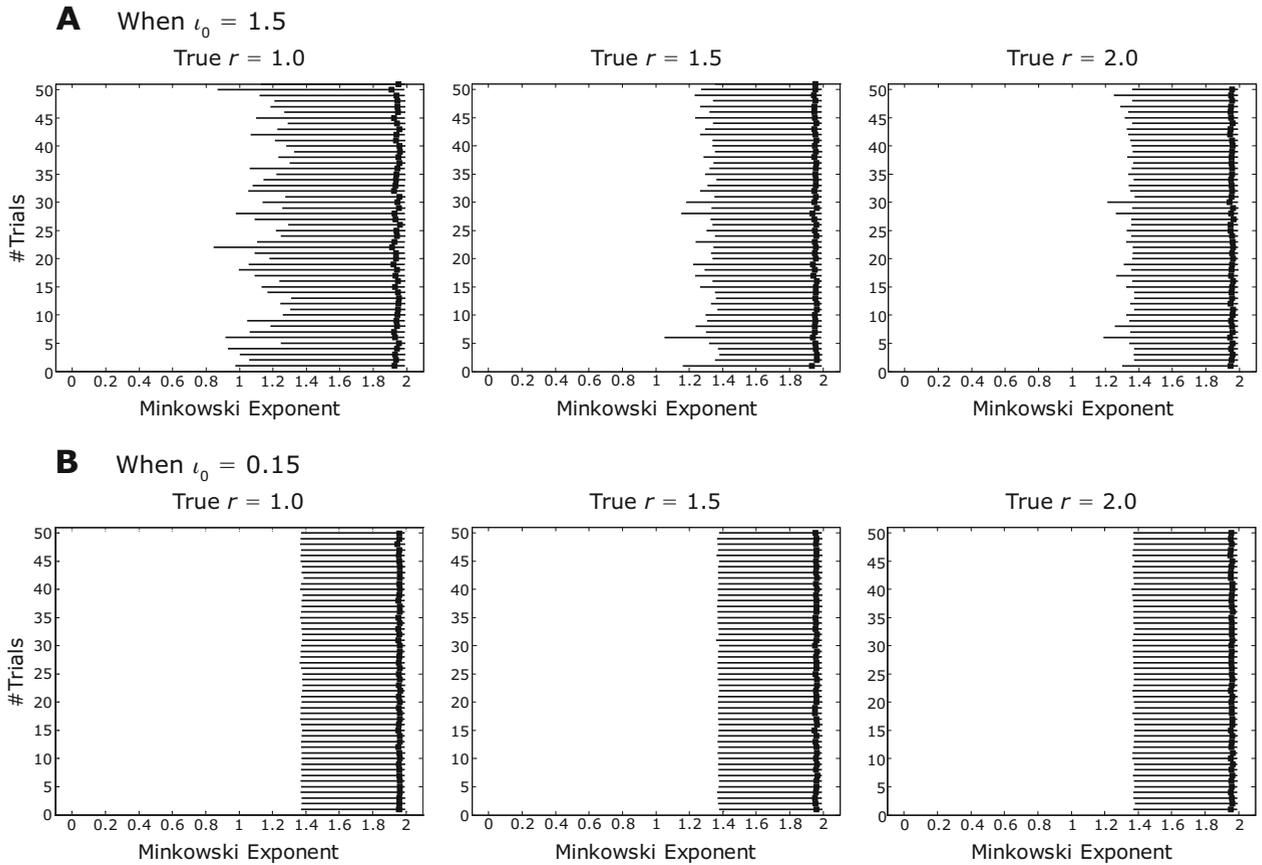


Figure 1. Posterior modes (squares) and 95% credibility intervals (lines) of Minkowski exponent r by an averaged data version of Lee’s (2008) method, when the true r equals 1 (left), 1.5 (middle), and 2 (right). (A) When $t_0 = 1.5$. (B) When $t_0 = 0.15$. One box contains the results of 50 trials.

Minkowski exponent; therefore, another method is required. This section introduces a new approach to Bayesian MDS with Minkowski distance.

The idea is simply to extend the Euclidean MDS model of Oh and Raftery (2001) to the general Minkowski distance model. Namely, the observed dissimilarity δ_{ij} between objects i and j is modeled using the normal distribution,

$$\delta_{ij} \sim N(d_{ij}, \phi^2), \tag{13}$$

where d_{ij} is the Minkowski distance among the object configurations (Equation 1). For the prior distribution of the Minkowski exponent r , we follow the setting of Lee’s (2008) model,

$$r \sim U(0, 2). \tag{14}$$

For other parameters, we used the prior settings of Oh and Raftery (2001). Specifically, for the prior distribution of \mathbf{x}_i , we used a multivariate normal distribution,

$$\mathbf{x}_i \sim N(\mathbf{0}, \Lambda), \tag{15}$$

independently for $i = 1, \dots, n$ objects, where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ is a diagonal covariance matrix. For each element of Λ , an inverse gamma hyperprior,

$$\lambda_k \sim \text{IG}(\alpha, \beta_k), \tag{16}$$

is set independently for $k = 1, \dots, p$. Also, for the prior of the error variance ϕ^2 , an inverse gamma distribution

$$\phi^2 \sim \text{IG}(a, b) \tag{17}$$

is used. With these settings, the posterior distribution of r (as well as other parameters) can be estimated using the MCMC algorithm.

NUMERICAL EXPERIMENT 2

In order to investigate the applicability of the proposed model, we conducted a new simulation study.

As was the case with Numerical Experiment 1, the artificial data were generated from the true model (i.e., the proposed model). Again, the three conditions of the true Minkowski parameters were set as $r = 1, 1.5, \text{ and } 2$. For each condition, we generated artificial data as follows: Each element of the configuration \mathbf{X} was generated from a normal distribution,

$$x_{ik} \sim N(0, \lambda_0), \tag{18}$$

which corresponds to the true model structure (Equation 15) with $\Lambda = \text{diag}(\lambda_0, \dots, \lambda_0)$. We considered two conditions for the variance parameter λ_0 —namely, $\lambda_0 = 1$ and $\lambda_0 = 0.32$. The number of dimensions p was set to 2. Thus generated,

\mathbf{X} is transformed to the distance $\mathbf{D} = \{d_{ij}\}$ by Equation 1 with the three conditions of the true Minkowski exponents: $r = 1, 1.5,$ and 2 . Then, this \mathbf{D} is used as a dissimilarity data matrix $\mathbf{\Delta} = \{\delta_{ij}\}$. This operation corresponds to generating δ_{ij} from Equation 13, with $\phi^2 = 0$; that is, just as in numerical Experiment 1, no error is introduced. We generated 50 independent sets of artificially “observed” dissimilarity $\mathbf{\Delta}$ per one condition in this manner and then analyzed them using the proposed method. As for the values of hyperparameters in Equations 16 and 17, we used the same settings used in Oh and Raftery (2001). Again, we performed the analysis using MATLAB and OpenBUGS via the use of the MATBUGS program. The first 20,000 MCMC iterations were discarded as burn-in; we used the subsequent 40,000 iterations for posterior estimation.

Result

The resultant posterior estimates (posterior means) of the Minkowski parameter and its 95% credibility intervals are plotted for all 50 trials in Figure 2. It can be seen that, although in some estimations the credibility intervals are somewhat large, the posterior point estimate (squares) is close to the true value in all conditions. Comparing this result with that of the previous model (Figure 1), it is clear that the model proposed here recovers the true Minkowski exponent better. Also, comparing Figure 2A with Fig-

ure 2B shows that the credibility intervals are larger when λ_0 is small. This can be because, when λ_0 in Equation 18 is small, every object becomes closer to one another and, therefore, the difference between them becomes smaller, making the estimation of the Minkowski exponent harder than it would be otherwise.

REAL DATA ANALYSIS

Because the proposed method seems to work well, we applied it to real data. We analyzed the experimental data presented by Helm (1959). In this experiment, 10 examinees were asked to judge the dissimilarity of 10 color chips that were different in hues but constant in both brightness and saturation. The colors ranged over the entire spectrum from red (long wavelength) to purple (short wavelength). Data from 10 participants are used for the analysis.³

The proposed model is used for analysis. Prior distributions and hyperparameters are the same as those used in the previous simulation study (Numerical Experiment 1). The first 10,000 MCMC iterations were discarded as burn-in; we used runs of the subsequent 200,000 iterations for posterior inference.

The resultant posterior density of the Minkowski exponent r is shown in Figure 3. The gray bars indicate the histogram of 200,000 posterior MCMC samples of the

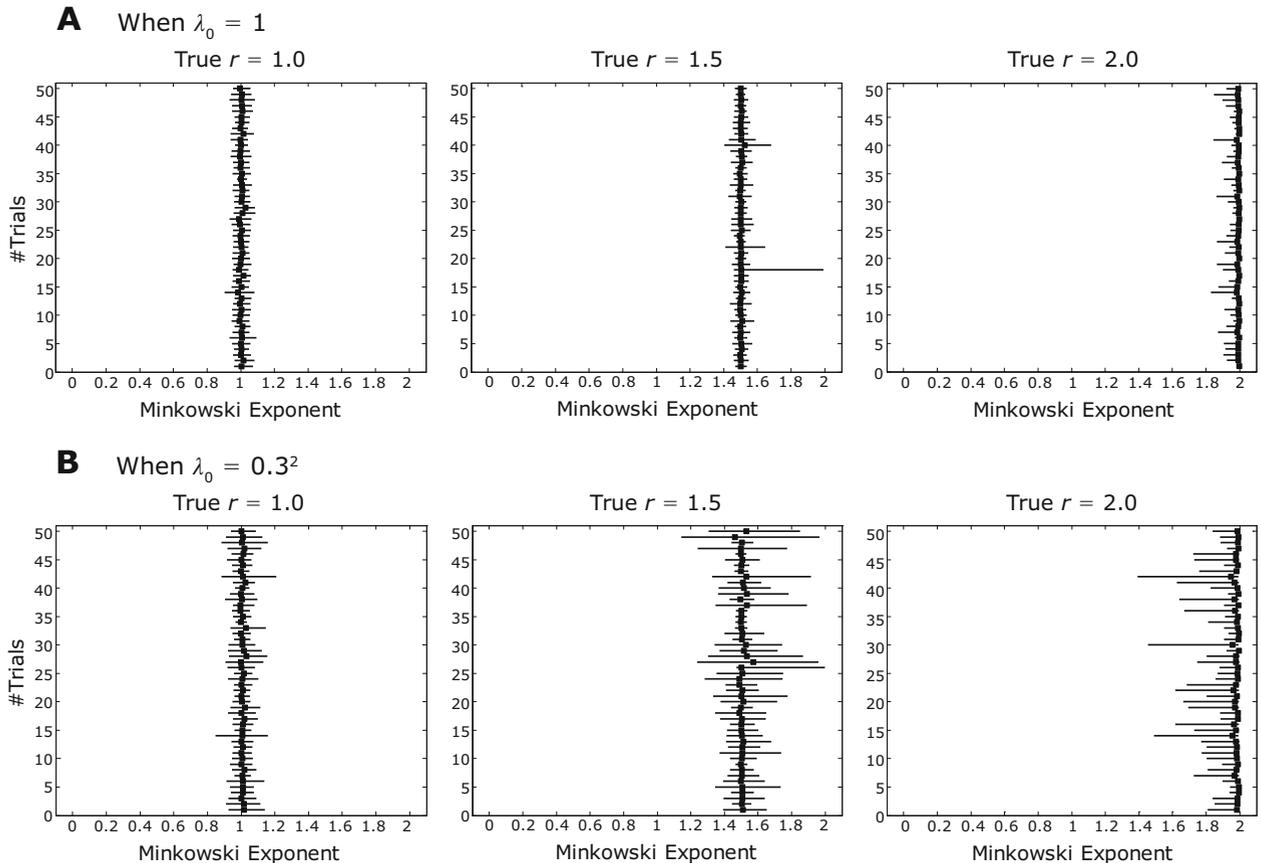


Figure 2. Means (black squares) and 95% credibility intervals (lines) of Minkowski exponent r by the proposed method, when the true r equals 1 (left), 1.5 (middle), and 2 (right). (A) When $\lambda_0 = 1$. (B) When $\lambda_0 = 0.3^2$. One box contains the results of 50 trials.

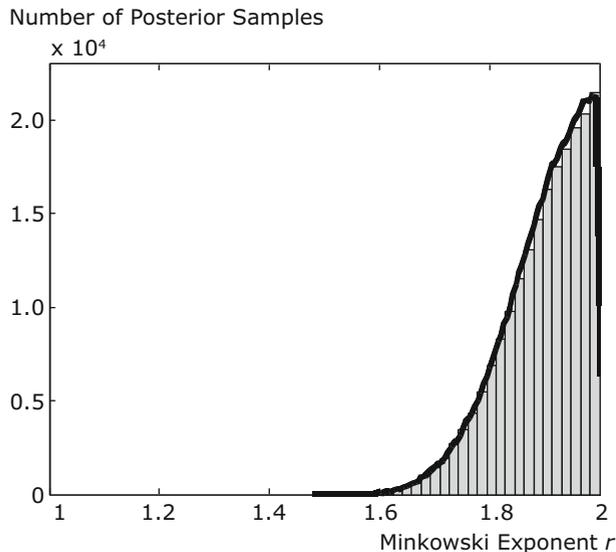


Figure 3. Empirical (gray histogram) and smooth fitted (black line) posterior distribution of Minkowski exponent r for Helm's (1959) color data.

Minkowski exponent r , and the black line indicates the fitted empirical posterior density using the kernel estimation method with a normal (Gaussian) kernel. As shown in Figure 3, the posterior mode is very close to 2, indicating the fit of Euclidean distance. This result seems to be in agreement with prior studies, which state that the human color space is almost Euclidean (Backhaus, Menzel, & Kreißl, 1987; Borg & Groenen, 2005, chap. 17; Kruskal, 1964).

CONCLUSIONS

The contribution of the present article is summarized in the following four points:

In the first simulation study, recovery of the true Minkowski exponent parameter reveals a difficulty in the natural extension of Lee's (2008) existing method to averaged- or single-data problems (Numerical Experiment 1). Therefore, although the idea of estimating the Minkowski exponent as a parameter is attractive, this method may not be recommended for averaged- or single-data analysis.

A new Bayesian method for the estimation of the Minkowski exponent in the averaged- (or single-) data MDS problem is proposed (see the Proposed Method section). The proposed model is a generalization of the existing Euclidean Bayesian MDS model for the Minkowski metric.

The proposed method is evaluated by another simulation study. It is found that the proposed method is able to recover the true Minkowski exponent (Numerical Experiment 2).

The proposed method is applied to analyze the experimental data of Helm (1959). The posterior distribution of the Minkowski exponent r has its mode close to 2, thereby indicating the fit of Euclidean distance. This result agrees with former studies of human color space (see the Real Data Analysis section).

With regard to Lee's (2008) method, it would be desirable to determine why this method could not recover the value of the true Minkowski exponent in its extension to the averaged data problem, even though it worked for individual subject data (Lee, 2008). One of the possibilities would be that the exponential transformation in Equation 12 causes a problem in the estimation of the Minkowski exponent, because, as described previously, former studies have found difficulties in the estimation using a lognormal distribution. Another plausible explanation would be that fitting individual subject data helps make the recovery of r more robust.

Note that in our method—as well as in Lee's (2008) method—the possible range of the Minkowski exponent is restricted to be within $[0, 2]$. There are four reasons for this restriction, of which the first three were presented in Lee's article. First, the common interest of psychological researchers is usually in whether r is 1 (city block) or 2 (Euclidean), both of which are within this range. Second, metrics with $r < 1$ have been given a psychological justification in terms of modeling stimuli with component dimensions that compete for attention. Third, it is difficult to give a psychological meaning to metrics with $r > 2$. Fourth, on the basis of our experience, if one allows r to assume larger values, numerical trouble frequently occurs. This may happen because the actual value of the Minkowski distance almost does not change when the value of the Minkowski exponent becomes large, which makes it difficult to distinguish the true value.

Our final remark is about application studies. As noted in the introduction, thus far, most empirical studies that deal with a Minkowski property of a psychological space have been conducted using a stress-minimization strategy. It would be of interest to reanalyze the existing results by using the proposed method.

AUTHOR NOTE

This research was supported in part by a Grant-in-Aid for JSPS Fellows (21-5613). The authors express their gratitude to Michael D. Lee, Man-Suk Oh, and Adrian E. Raftery for releasing their code online. The authors thank Junko Kato for helpful discussions and feedback. The authors also thank the reviewers for their comments, which have greatly improved the article. Address correspondence concerning this article to K. Okada, Department of Psychology, School of Human Sciences, Senshu University, 2-1-1, Higashimita, Tama-ku, Kawasaki-shi, Kanagawa 214-8580, Japan (e-mail: ken@psy.senshu-u.ac.jp).

REFERENCES

- ATTCHINSON, J., & BROWN, J. A. C. (1957). *The lognormal distribution*. Cambridge: Cambridge University Press.
- ARABIE, P. (1991). Was Euclid an unnecessarily sophisticated psychologist? *Psychometrika*, *56*, 567-587. doi:10.1007/BF02294491
- ARNOLD, J. B. (1971). A multidimensional scaling study of semantic distance. *Journal of Experimental Psychology*, *90*, 349-372. doi:10.1037/h0031565
- BACKHAUS, W., MENZEL, R., & KREIßL, S. (1987). Multidimensional scaling of color similarity in bees. *Biological Cybernetics*, *56*, 293-304.
- BORG, I., & GROENEN, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications* (2nd ed.). New York: Springer.
- BOX, G. E. P., & TIAO, G. C. (1992). *Bayesian inference in statistical analysis*. New York: Wiley.
- EKMAN, G. (1954). Dimensions of color vision. *Journal of Psychology*, *38*, 467-474.

- GELMAN, A., CARLIN, J. B., STERN, H. S., & RUBIN, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton: Chapman & Hall/CRC.
- HELM, C. E. (1959). *A multidimensional ratio scaling analysis of color relations*. Princeton, NJ: Princeton University & Educational Testing Service.
- KRUSKAL, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1-27. doi:10.1007/BF02289565
- LEE, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, **15**, 1-15. doi:10.3758/PBR.15.1.1
- LUNN, D. J., THOMAS, A., BEST, N., & SPIEGELHALTER, D. (2000). WinBUGS—A Bayesian modeling framework: Concepts, structure, and extensibility. *Statistics & Computing*, **10**, 325-337. doi:10.1023/A:1008929526011
- MURPHY, K., & MAHDAVIANI, M. (2005). MATBUGS: A MATLAB interface to WinBUGS [Computer software and manual]. Retrieved from <http://people.cs.ubc.ca/~murphyk/Software/MATBUGS/matbugs.html>.
- OH, M.-S., & RAFTERY, A. E. (2001). Bayesian multidimensional scaling and choice of dimension. *Journal of the American Statistical Association*, **96**, 1031-1044. doi:10.1198/016214501753208690
- OKADA, K., & SHIGEMASU, K. (2009). BMDS: A collection of R functions for Bayesian multidimensional scaling. *Applied Psychological Measurement*, **33**, 570-571. doi:10.1177/0146621608321761
- SCHMOYER, R. L., BEAUCHAMP, J. J., BRANDT, C. C., & HOFFMAN, F. O., JR. (1996). Difficulties with the lognormal model in mean estimation and testing. *Environmental & Ecological Statistics*, **3**, 81-97. doi:10.1007/BF00577325
- SPIEGELHALTER, D. J., THOMAS, A., BEST, N., & LUNN, D. (2007). *OpenBUGS user manual* (Version 3.0.2). Cambridge: MRC Biostatistics Unit.
- TORGERSON, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, **17**, 401-419. doi:10.1007/BF02288916
- TOWNES, B., & ABBOTT, R. D. (1974). The psychological interpretation of metrics in ordinal multidimensional scaling: A review. *JSAS, Catalog of Selected Documents in Psychology*, **4**, 36.

NOTES

1. Retrieved December 11, 2008, from www.socsci.uci.edu/~mdlee/bgm.html.
2. This value was set to 5 in the original code. To be on the safe side, we set it to 10, which is more sufficiently large.
3. Although the original data contained 10 color-normal participants and 4 color-deficient participants, we used only the data from the color-normal participants, because the objective here was to analyze the Minkowski property of color-normal participants in their recognition of colors.

APPENDIX BUGS Code Sequences

The BUGS code for the natural extension of Lee's (2008) model to averaged (or single-subject) data, as used in Numerical Experiment 1, is as follows:

```
model {
  for (i in 1:n) {
    for (j in 1:n) {
      S[i,j] ~ dnorm(mu[i,j], invomega2)
      mu[i,j] <- exp(-d[i,j])
      d[i,j] <- pow(pow(abs(X[i,1]-X[j,1]), r) + pow(abs(X[i,2]-X[j,2]), r), 1/r)
    }
  }
  for (i in 1:n) {
    for (k in 1:2) {
      X[i,k] ~ dunif(-10,10)
    }
  }
  invomega2 ~ dgamma(zeta, zeta)
  r ~ dunif(0,2)
  zeta <- 0.001
}
```

The BUGS code for the proposed model used in Numerical Experiment 2 is as follows:

```
model{
  for(i in 2 : n) {
    for(j in 1 : i-1) {
      Delta[i,j] ~ dnorm(d[i,j], invphi2)
      d[i,j] <- pow(pow(abs(X[i,1]-X[j,1]), r) + pow(abs(X[i,2]-X[j,2]), r), 1/r)
    }
  }
  invphi2 ~ dgamma(a,b)
  for(k in 1 : 2) {
    for(i in 1 : n) {
      X[i,k] ~ dnorm(0, invlambda[k])
    }
  }
  invlambda[k] ~ dgamma(alpha, beta[k])
}
r ~ dunif(0,2)
}
```