MULTIDIMENSIONAL SCALING FOR DISSIMILARITY FUNCTIONS WITH CONTINUOUS ARGUMENT

Masahiro Mizuta*

ABSTRACT

In this paper, a method of Multidimensional Scaling (MDS) for dissimilarity functions with continuous argument is discussed. MDS is one of the important methods for data analysis. Most conventional MDS methods suppose that dissimilarities are real values. Nowadays, the types of data set dealt with in data analysis are extended. Ramsay and Silverman proposed the concept of Functional Data Analysis (FDA). FDA deals with functional data or with data as functional data. When dissimilarity data among n objects are given dependent on a variable t, we would like to use methods of MDS of functional version; the aim of the method is to derive functional configuration $X(t)$ that represents the dissimilarity functional data. A method of MDS for dissimilarity functions with discrete argument is also discussed, because most dissimilarity functions are given by discrete values in view of implementation on computer.

1. Introduction

Ramsay and Silverman (1997) proposed a concept of Functional Data Analysis (FDA). FDA deals with data represented by functions. They developed several methods for analyzing functional data, including functional regression analysis, functional principal component analysis, and functional discriminant analysis. We have also proposed an extension of functional regression analysis (Shimokawa et al., 2000).

In this paper, we discuss functional multidimensional scaling (Functional MDS). Functional MDS analyzes dissimilarity functional data, which means dissimilarities between objects depending on variables, typically time or location. When we investigate two configurations, which are two results of a MDS method for different situations or conditions, we can use (generalized) procrustes analysis. In this paper, we extend the idea to functional dissimilarity data. At first, we will propose a Functional MDS for dissimilarity functional data with continuous argument. Then we describe a Functional MDS for dissimilarity functional data with discrete argument for approximation of dissimilarity functional data with continuous argument.

2. Functional data analysis (FDA) and multidimensional scaling (MDS)

We give a brief introduction and notations on functional data analysis and multidimensional scaling for preparations.

---

*Center for Information and Multimedia Studies, Hokkaido University, N.11, W.5, Kita-ku, Sapporo 060-0811, Japan E-mail: mizuta@cims.hokudai.ac.jp

Key words: Computer oriented method; Functional data analysis
2.1. Functional data analysis

Most conventional methods for data analysis assume that data are sets of numbers with some structure, i.e. a set of vectors or a set of matrices. However, we must often analyze more complex data. One of the complex types of data is functional data; data themselves are represented as functions. Functional data can be considered as infinity-dimensional data. Most methods in the book (Ramsay and Silverman, 1997) for functional data are based on an approximation with finite expansions of the functions with basis functions. Once the functional data can be thought as finite linear combinations of the basis functions, functional data analysis methods for the functional data are almost the same as those of conventional data analysis methods.

2.2. Multidimensional scaling

Conventional MDS methods are developed for finding structures among n objects with configurations of n points on the p dimensional Euclidean space. It is very easy to calculate distances among n points on p dimensional Euclidean space. But, conversely, it is hard to construct n points that represent distances among objects.

The dissimilarities among n objects are denoted by $S = \{s_{ij}\} (i, j = 1, 2, \ldots, n)$. We assume $s_{ij} \geq 0$, $s_{ij} = s_{ji}$ and $s_{ij} = 0$. The aim of MDS is to construct the configuration $X = (x_1, \ldots, x_n)$ that represent the relations among n objects. The configuration is a set of n points on the p dimensional Euclidean space. The Euclidean distances between $x_i$ and $x_j$ are denoted by $d_{ij}; d_{ij} := \| x_i - x_j \|$. Methods for MDS find a configuration $X$ such as $d_{ij} \simeq s_{ij}$. The degree of coincidence of $d_{ij}$ and $s_{ij}$ is defined for each MDS method with the values of them or the order of them. So, for most MDS methods, we are not interested in the coordinate system of the configuration itself.

One simple MDS method is Torgerson’s method (Mardia et al., 1979; etc.) The method is described here. For dissimilarity data $S = \{s_{ij}\}$, we calculate

$$b_{ij} := \frac{1}{2}(s_{ij}^2 - s_i^2 - s_j^2 + s^2),$$

where $s_{ij} = \frac{1}{n} \sum_{j=1}^{n} s_{ij}^2$, $s_i^2 = \frac{1}{n} \sum_{i=1}^{n} s_{ij}^2$, and $s^2 = \frac{1}{n^2} \sum_{i,j} s_{ij}^2$.

The matrix $B = (b_{ij})$ is applied to singular value decomposition; $B = T \Lambda T^T$, where $\Lambda = \text{diag}(l_1, l_2, \ldots, l_n)$ and $T = (t_1, \ldots, t_n)$. The configuration with Torgerson’s method is $X = (t_1, \ldots, t_n) \text{diag}(\sqrt{l_1}, \ldots, \sqrt{l_n})$. The solutions (results) are indefinite if $l_i = l_{i+1}$.

3. Functional MDS

In this section, we propose a method of functional multidimensional scaling. In the first subsection, Functional MDS for dissimilarity functional data with continuous argument is discussed. In the second subsection, we propose Functional MDS for dissimilarity functional data with discrete argument. Then we discuss the relations between the two methods.

3.1. Functional MDS for dissimilarity functions with continuous argument

We assume that we have dissimilarity functions with continuous argument between n objects. Dissimilarities among n objects depending on a variable t are denoted by $S(t) = \{s_{ij}(t)\} (i, j = 1, 2, \ldots, n), t \in [a, b]$. We restrict ourselves to two dimensional configurations.

We propose a Functional MDS method for $S(t)$. The proposed method needs a conventional MDS method for the first step, for example, Torgerson’s method. A conventional MDS method is applied to the dissimilarity data for each t. We can get two dimensional
Multidimensional Scaling for Dissimilarity Functions with Continuous Argument

configurations of $n$ objects for each $t$. They are denoted by $X(t) = (x_1(t), \ldots, x_n(t))$. If $S_{ij}(t)$ are differentiable in $t$, we can choose $A$ and $T$ such that $X(t)$ is differentiable. Then we assume that they are differentiable here. The key idea of the method is to find out the orthogonal matrix function $Q(t)$ that adjusts $X(t)$ to $Q(t)X(t)$ because the distances between points are invariant under orthogonal transformations.

The length of the curve $x_i(t)$ on the two dimensional space is

$$l = \int_a^b \sqrt{\left(\frac{dx_i(t)}{dt}\right)^2} \, dt. \quad (1)$$

Then, we would like to find the orthogonal matrix function $Q(t)$ that minimizes

$$l(Q) = \int_a^b \sum_{i=1}^n \left\| \frac{dQ(t)x_i(t)}{dt} \right\|^2 \, dt.$$

$Q(t)$ can be represented by $Q(t) = \begin{pmatrix} \cos \phi(t) & \sin \phi(t) \\ -\sin \phi(t) & \cos \phi(t) \end{pmatrix}$, where $\phi(t)$ is a function of $t$. Then

$$l(Q) = \int_a^b \left( \sum_{i=1}^n \left\| \frac{dx_i(t)}{dt} \right\|^2 \right) + 2 \left( \sum_{i=1}^n \frac{dx_i(t)T}{dt} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x_i(t) \right) \phi'(t) + \left( \sum_{i=1}^n \|x_i(t)\|^2 \right) \phi'(t)^2 \, dt.$$

So when

$$\phi'(t) = -\sum_{i=1}^n \frac{dx_i(t)}{dt} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x_i(t) \left/ \sum_{i=1}^n \|x_i(t)\|^2 \right., \quad \quad (2)$$

$l(Q)$ takes the minimum value.

The functional configuration $Q(t)X(t)$ is the result of the functional multidimensional scaling. The functional configuration can be shown with motions of $n$ points by dynamic graphics. The $n$ trajectories of $X(t)$ are also a good representation on a sheet.

3.2. Functional MDS for dissimilarity functions with discrete argument

In the previous subsection, we assumed that the dissimilarity functions have continuous argument. However, in most cases, we can only get dissimilarity functions as functions with discrete argument: we assume that the dissimilarity data among $n$ objects are given several times, $t_1, t_2, \ldots, t_m \in [a, b]: \{S(t_k)\}(k = 1, \ldots, m)$. For each $t_k$, the configurations of $n$ objects $X(t_k) = (x_1(t_k), \ldots, x_n(t_k))$ are derived with a conventional MDS method. Even if $S(t)$ is continuous, $X(t)$ is not always continuous. So, we cannot get the relations among the configurations.

Same as the previous arguments, our criterion is to minimize

$$\sum_{i=1}^n \sum_{k=2}^m \left\| Q(t_{k-1})x_i(t_{k-1}) - Q(t_k)x_i(t_k) \right\|^2 \quad (3)$$
with respect to orthogonal matrices $Q(t_k)(k=1, \ldots, m)$. Because

$$
\sum_{i=1}^{n} \| Q(t_{k-1})x_i(t_{k-1}) - Q(t_k)x_i(t_k) \|^2
$$

$$
= \sum_{i=1}^{n} \| x_i(t_{k-1}) - Q(t_{k-1})^T Q(t_k)x_i(t_k) \|^2
$$

$$
= \sum_{i=1}^{n} \| x_i(t_{k-1}) \|^2 + \| x_i(t_k) \|^2 - 2x_i(t_{k-1})^T Q(t_{k-1})^T Q(t_k)x_i(t_k),
$$

we derive the orthogonal matrix $Q(t_k)$ that minimizes $\sum_{i=1}^{n} x_i(t_{k-1})^T Q(t_{k-1})^T Q(t_k)x_i(t_k) (k=1, \ldots, m)$. $Q_k$ denote $Q(t_{k-1})^T Q(t_k)$, hereafter.

We can get $Q_k$ with

$$
Q_k = (X(t_k)X(t_{k-1})^T X(t_{k-1})X(t_k)^T)^{-1/2}X(t_k)X(t_{k-1})^T.
$$

The functional configuration $Q(t)X(t)$ is the result of the functional multidimensional scaling.

In the case of two dimensional configurations, we can explain the method more primitives. Orthogonal matrix $Q(t)$ can be represented as

$$
Q(t) = \begin{pmatrix}
\cos \phi(t) & \sin \phi(t) \\
-\sin \phi(t) & \cos \phi(t)
\end{pmatrix},
$$

where $\phi(t)$ is a function of $t$. So

$$
\sum_{i=1}^{n} x_i(t_{k-1})^T Q_k x_i(t_k) = \sum_{i=1}^{n} \left( x_i^{(1)}(t_{k-1}), x_i^{(2)}(t_{k-1}) \right) \begin{pmatrix}
\cos \phi(t) & \sin \phi(t) \\
-\sin \phi(t) & \cos \phi(t)
\end{pmatrix} \begin{pmatrix}
x_i^{(1)}(t_k) \\
x_i^{(2)}(t_k)
\end{pmatrix}
$$

$$
= \sum_{i=1}^{n} \left( x_i^{(1)}(t_{k-1})x_i^{(2)}(t_k) - x_i^{(2)}(t_{k-1})x_i^{(1)}(t_k) \right) \sin \phi(t)
$$

$$
+ \sum_{i=1}^{n} \left( x_i^{(1)}(t_{k-1})x_i^{(1)}(t_k) + x_i^{(2)}(t_{k-1})x_i^{(2)}(t_k) \right) \cos \phi(t),
$$

we choose the function $\phi(t)$ that $\sin \phi(t)$ is equal to

$$
\frac{\sum_{i=1}^{n} \left( x_i^{(1)}(t_{k-1})x_i^{(2)}(t_k) - x_i^{(2)}(t_{k-1})x_i^{(1)}(t_k) \right)}{\sqrt{\left( \sum_{i=1}^{n} \left( x_i^{(1)}(t_{k-1})x_i^{(2)}(t_k) - x_i^{(2)}(t_{k-1})x_i^{(1)}(t_k) \right)^2 + \left( \sum_{i=1}^{n} \left( x_i^{(1)}(t_{k-1})x_i^{(1)}(t_k) + x_i^{(2)}(t_{k-1})x_i^{(2)}(t_k) \right) \right)^2}},
$$

where $x_i(t_k) = (x_i^{(1)}(t_k), x_i^{(2)}(t_k))^T$.

We can get functional configuration with discrete argument $Q(t_k)X(t_k)(k=1, \ldots, m)$.

### 3.3. Relations between two methods

We have described two Functional MDS methods; the version of dissimilarity function with continuous argument and the version of dissimilarity function with discrete argument. We will show that the discrete version is an approximation of the continuous version.

Suppose $m \to \infty$, $t_k \in [a, b]$, then the formula (3) $\to$ (1).
If we divide the formula (4) by $t_k - t_{k-1}$ and put $t_k = t, t_{k-1} = t - \varepsilon$, then

$$
\sin \phi(t) \approx \frac{\phi(t)}{\varepsilon} = \sum_{i=1}^{n} \left( x_i^{(1)}(t) - x_i^{(2)}(t) \right) \times \frac{1}{\sqrt{N}}
$$

$$
= \frac{\varepsilon}{\varepsilon} \sum_{i=1}^{n} \left( x_i^{(1)}(t) - x_i^{(1)}(t - \varepsilon), x_i^{(2)}(t) - x_i^{(2)}(t - \varepsilon) \right) \left( \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right) \left( \begin{array}{c} x_i^{(1)}(t) \\ x_i^{(2)}(t) \end{array} \right) \times \frac{1}{\sqrt{N}},
$$

where

$$
N = \left( \sum_{i=1}^{n} \left( x_i^{(1)}(t) - x_i^{(2)}(t) \right) \right)^2 + \left( \sum_{i=1}^{n} \left( x_i^{(1)}(t) + x_i^{(2)}(t - \varepsilon) \right) \right)^2.
$$

Suppose $\varepsilon \to 0$, then $N \to \sum_{i=1}^{n} \left\| x_i(t) \right\|^2$. We can get (2).

4. Numerical example

We will show numerical examples of the proposed method: discrete version. For the first example, dissimilarity functional data $S(t) = \{s_{ij}(t) = s_{ij} + td_{ij}\}, (i, j = 1, \ldots, 10)$ are used ($t = 0, 1, \ldots, 200$). $s_{ij}$ are sample values from normal random generator $N(0, 1)$, and $d_{ij}$ are from $N(0, 0.01^2)$. The data have a simple structure.

Figure 1(a) shows two dimensional configurations of $S(0) = \{s_{ij}\}$ with ordinary MDS. Figure 1(b)-(d) show also configurations of $S(t); t = 1, 2, 200$. The circles represent the original configuration (i.e. $t = 0$).

The configuration of $t = 3$ is almost same as that of $t = 0$. The configuration of $t = 1$ takes the y-coordinate that is the opposite direction of that of $t = 0$. The configuration of $t = 2$ takes both coordinates that are the opposite directions of them. Figure 2 shows the trajectories of the data from $t = 0$ to 200. We can investigate the data with the trajectories.

The second example data $S(t) = \{s_{ij}(t)\} (t = 0, 1, \ldots, 50)$ are made by the rule: $S(0) = \{s_{ij}\}, s_{ij}(t + 1) = s_{ij}(t) + \varepsilon_{ijt}$, where $\varepsilon_{ijt}$ are i.i.d. from $N(0, \delta^2)$ and $\delta = 0.02, 0.1$. Figure 3 is the result of the proposed method. When $\delta = 0.02$, the lengths of the trajectories are short and the function $\phi(t)$ is low. When $\delta = 0.1$, the trajectories are in utter disorder and the function $\phi(t)$ takes high values.

5. Concluding remarks

Needless to say, there are several MDS methods for time dependent similarity data. Our goal is to develop MDS methods for general functional data, not one-dimensional functional data. The proposed methods can be extended to two-dimensional functional data. The proposed methods uses conventional MDS method, for example, Torgerson’s method. We can use most MDS methods for the proposed methods.
Fig. 1: Configurations of Example 1

(a) \( t = 0 \)  
(b) \( t = 1 \)  
(c) \( t = 2 \)  
(d) \( t = 200 \)

Fig. 2: Trajectories of Example 1 (Proposed Method; \( t = 0, 1, \ldots, 200 \))
Acknowledgements

The author would like to thank the editor and the anonymous reviewers for suggesting several good ways to improve the original manuscript.

REFERENCES

