Neurocomputing 124 (2014) 97–104

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Data visualization for asymmetric relations

Heeyoul Choi *

Samsung Advanced Institute of Technology, Samsung Electronics, 97, Samsung2-ro, Giheung-gu, Yongin-si 446-712, Gyeonggi-do, Republic of Korea

ARTICLE INFO

Article history:
Received 28 January 2013
Received in revised form
25 April 2013
Accepted 24 July 2013
Communicated by S. Choi
Available online 2 September 2013

Keywords:
Asymmetric proximity
Data visualization
Multidimensional scaling

ABSTRACT

Most 2D visualization methods based on multidimensional scaling (MDS) and self-organizing maps (SOMs) use a symmetric distance matrix to represent and visualize object relationships in a data set. In many real-world applications, however, raw data such as a world-trade data are best captured as an asymmetric proximity matrix. Such asymmetric matrices cannot be perfectly represented by most previous methods. To handle such an intrinsic limitation, in this paper, we propose a dynamic learning for metric representations of asymmetric proximity data to better understand the data. The proposed learning generates two representations (maps) with the row vectors (sending or exporting) and column vectors (receiving or importing) of the matrix, respectively. To better present the patterns, we supplement the maps with two analysis tools: cluster analysis and distance analysis, which connect and compare the different patterns from the different maps. Experiment results using three real world data sets confirm that the proposed learning method is useful to understand asymmetric proximity data.

1. Introduction

Visualization is a procedure that helps represent the complex data (usually in a high-dimensional space) in an effective way (usually in a low-dimensional space). Dimensionality reduction algorithms have been used for this purpose. Multidimensional scaling (MDS) [1] and principal component analysis (PCA) [2] are popular linear methods for dimensionality reduction. Manifold learning is a non-linear dimensionality reduction approach, which induces a smooth nonlinear low-dimensional manifold from a set of data points drawn from the manifold. Recently, various dimensionality reduction methods (for example see [3–6]) have been developed in machine learning community drawing an attention in pattern recognition and signal processing.

In data analysis, many times, information is provided as a similarity (or dissimilarity) matrix, whose elements could be the distances between the data points. For visualization of such data, most methods based on a dimensionality reduction approach such as MDS or self-organizing maps (SOMs) [7] rely on a symmetric matrix which represents the object relationships [3]. Although sometimes those methods seem to be applied to an asymmetric matrix, they convert the asymmetric matrix into a symmetric one, and actually work on the symmetric matrix [8–10]. However, a symmetric matrix cannot represent any directional relationship between data points, while asymmetric matrix can. It means potential information loss if we use only symmetric relationship.

In many real-world applications [11], raw data can be best captured as an asymmetric proximity matrix. For example, a world-trade dataset can be represented as an asymmetric matrix with each column and each row corresponding to one country, and each cell indicating the money amount transferred from one country to another. Journal citation data (or bibliometric data) is another example for asymmetric similarity (or dissimilarity), where the numbers of citation between two journals are usually different [12]. In pattern classification, a confusion matrix as in Morse code [13] is asymmetric. Also, in social network service, the following relationship as in Twitter (www.twitter.com) can be represented as an asymmetric matrix, while the friendship as in Facebook (www.facebook.com) is symmetric.

To visualize such asymmetric proximity data, some methods have been proposed. However, most previous methods first decompose it into a symmetric and a skew-symmetric component, and then put much focus on the symmetric component using MDS-like methods [14–17]. Those methods do not present the data as 2D points with their asymmetric properties, and fail to visualize asymmetric properties geometrically on a metric space and also fail to discover regularities in such data (e.g. differences between importing and exporting within and across multiple countries). Furthermore, separating the representations of the symmetric and skew-symmetric components makes it very hard to understand nature structures and properties of the data set intuitively.

In fact, asymmetric relationships cannot be represented in any single metric space including 2D space, keeping the asymmetric properties perfectly. Once a data set is represented as points in a metric space, their geometric relations become symmetric. MDS-like methods implicitly assume that the proximity matrix was obtained

* Tel.: +82 10 9623 6578.
E-mail address: heeyoul@gmail.com

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.07.030

from the data points in a metric space, which makes the methods not capable of handling asymmetric matrices. Therefore, such 2D representations cannot present the asymmetric properties. But still visual representation, if possible, is very attractive, so we propose a dynamic learning algorithm of asymmetric relations to represent the asymmetric data on metric spaces. We plot the data onto two different maps separately, focusing on one direction in the asymmetric matrix at a time: one based on the column vectors (e.g. importing) and one based on the row vectors (e.g. exporting).

In addition, the two maps are supposed to be considered together, since they are derived from the same matrix. To enable users to better understand the data, we supplement the two maps with two analysis tools on the maps: cluster analysis and distance analysis, allowing users to connect and compare various patterns within each map and across two maps. These tools provide quantitative measures of the difference between before and after learning, or the difference between the sending map and the receiving maps. Although the proposed method does not provide a perfect view of asymmetric properties, it increases users' understanding of the asymmetric data.

Several case studies conducted with real data sets reveal the effectiveness of our approach. The cola-brand-switching data is presented to explain the maps and analysis functions, and the other data sets follow after that to confirm that the proposed method is useful to understand asymmetric proximity data.

2. Background

2.1. Asymmetric proximity data

Given N objects with a similarity matrix $\boldsymbol{S}$, where S_{ij} is the proximity between the ith and jth objects. Asymmetric relation is defined when $S_{ij} \neq S_{ji}$. If this inequality is from noise or error, we can symmetrize it by $(\boldsymbol{S}+\boldsymbol{S}^T)/2$. Otherwise, we have to deal with the asymmetric property differently from the previous methods that are based on the symmetry assumption.

In this paper, to present how our method works, we use three data sets: cola-brand switching data between 15 cola brands [14], threat display behaviours data [18], and 113 countries' trading data in 2009 from International Monetary Fund (IMF) web page, www.imf.org. These data can be summarized as an asymmetric matrix as shown in the left of Fig. 1. Note that for the world-trade data, we use logarithmic values of the trade amount since the absolute amount of trade is too much dominated by a few countries. The directional flow of the trade among the countries can be presented as in the right of Fig. 1. However, it is hard to understand the underlying pattern from the matrix or the directional flow arrows.

2.2. Previous work

Although multidimensional scaling (MDS) is not perfect for asymmetric data, it can approximately present the data points on a 2D space after the symmetrization procedure. Many previous methods are based on MDS and our proposed learning method takes MDS as an initial step to present the asymmetric properties of data.

2.2.1. Multidimensional scaling

To simply connect the asymmetric matrix to classical MDS, we need to transform the asymmetric matrix to a distance matrix. Given an asymmetric matrix, $\boldsymbol{M}$ for N objects, we make a symmetric matrix

$$\boldsymbol{M}^s = \frac{\boldsymbol{M}+\boldsymbol{M}^T}{2}. \tag{1}$$

Then, as in [1], a distance matrix, $\boldsymbol{D}$, can be given by

$$D_{ij} = 1 - M^s_{ij}. \tag{2}$$

Note that $\boldsymbol{M}$ is a normalized matrix by the maximum value of the elements of $\boldsymbol{M}$. Next, the constant adding technique is applied to make sure that each distance is the length of two points on a metric space, as in kernel Isomap [19]. In classical MDS, given a distance matrix, $\boldsymbol{D}$, a popular cost function is defined by

$$J = \sum_{ij} (d(\boldsymbol{x}_i, \boldsymbol{x}_j) - D_{ij})^2, \tag{3}$$

where $d(\cdot,\cdot)$ is a binary function to calculate the Euclidean distance between two points on a low dimensional space, and $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]$ is a representation on the low-dimensional space. Let $\boldsymbol{B}$ be the inner product matrix, where

$$\boldsymbol{B} = \boldsymbol{X}^\top \boldsymbol{X}. \tag{4}$$

Considering $\boldsymbol{B}$ as a kernel matrix as described in [6], $\boldsymbol{B}$ can be given by

$$\boldsymbol{B} = -\tfrac{1}{2}\boldsymbol{H}\boldsymbol{D}^2\boldsymbol{H}, \tag{5}$$

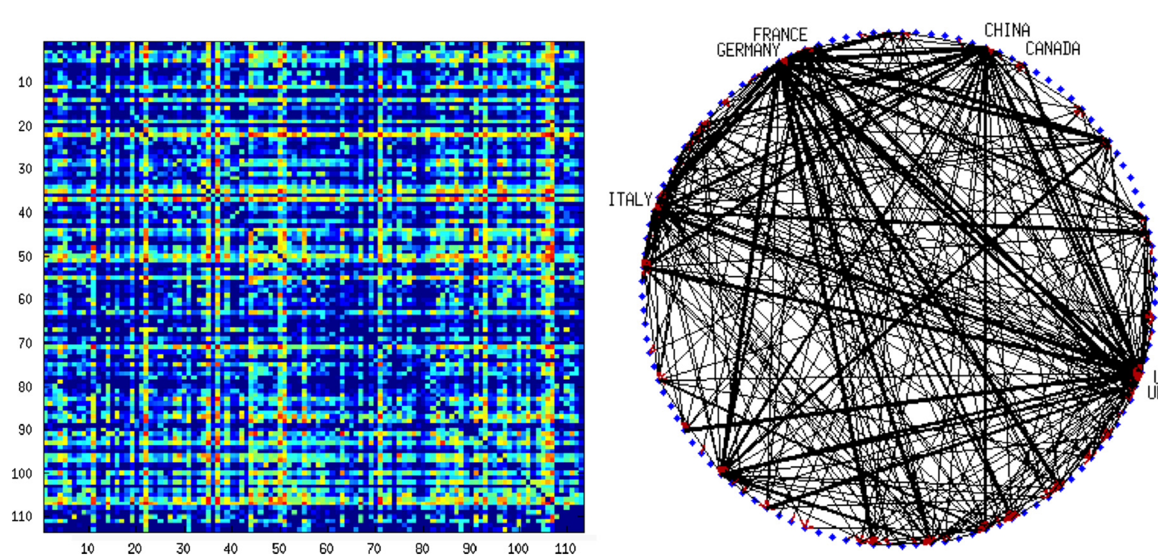

Fig. 1. Asymmetric world-trade data. (Left) 113 countries' trade data matrix, (Right) the directional flow of the trade which shows only the flows with higher values than twice the standard deviation of the matrix elements from the mean. The G7 countries' names are shown.