Journal of Computational Mathematics
Vol.35, No.4, 2017, 469–485.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1702-m2016-0748

AN INEXACT SMOOTHING NEWTON METHOD FOR EUCLIDEAN DISTANCE MATRIX OPTIMIZATION UNDER ORDINAL CONSTRAINTS*

Qingna Li
School of Mathematics and Statistics, Beijing Key Laboratory on MCAACI, Beijing Institute of Technology, Beijing 100081, China
Email: qnl@bit.edu.cn
Houduo Qi
School of Mathematics, University of Southampton, Highfield, Southampton SO17 1BJ, UK
Email: hdqi@soton.ac.uk

Abstract

When the coordinates of a set of points are known, the pairwise Euclidean distances among the points can be easily computed. Conversely, if the Euclidean distance matrix is given, a set of coordinates for those points can be computed through the well known classical Multi-Dimensional Scaling (MDS). In this paper, we consider the case where some of the distances are far from being accurate (containing large noises or even missing). In such a situation, the order of the known distances (i.e., some distances are larger than others) is valuable information that often yields far more accurate construction of the points than just using the magnitude of the known distances. The methods making use of the order information is collectively known as nonmetric MDS. A challenging computational issue among all existing nonmetric MDS methods is that there are often a large number of ordinal constraints. In this paper, we cast this problem as a matrix optimization problem with ordinal constraints. We then adapt an existing smoothing Newton method to our matrix problem. Extensive numerical results demonstrate the efficiency of the algorithm, which can potentially handle a very large number of ordinal constraints.

Mathematics subject classification: 90C30, 90C26, 90C90.
Key words: Nonmetric multidimensional scaling, Euclidean distance embedding, Ordinal constraints, Smoothing Newton method.

1. Introduction

Suppose that we are given the coordinates of a set of points, namely $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ with $\mathbf{x}_i \in \mathbb{R}^r$. It is straightforward to compute the pairwise Euclidean distances: $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$, $i, j = 1, \ldots, n$. The matrix $D = (d_{ij}^2)$ is known as the (squared) Euclidean Distance Matrix (EDM) of those points. However, the inverse problem is more interesting and important (and challenging). Suppose D is given. The method of the classical Multi-Dimensional Scaling (cMDS) generates a set of such coordinates that preserve the pairwise distances in D. We give a short description of it below. Let

$$J := I - \frac{1}{n}\mathbf{1}\mathbf{1}^T \qquad \text{and} \qquad B := -\frac{1}{2}JDJ,$$

* Received September 21, 2016 / Accepted February 7, 2017 /
Published online June 1, 2017 /

where I is the $n \times n$ identity matrix and $\mathbf{1}$ is the (column) vector of all ones in $\mathbb{R}^n$. In literature, J is known as the centralization matrix and B is the double-centralized matrix of D (also the Gram matrix of D because B is positive semidefinite). Suppose B admits the spectral decomposition:

$$B = [\mathbf{p}_1, \ldots, \mathbf{p}_r] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{bmatrix} \begin{bmatrix} \mathbf{p}_1^T \\ \vdots \\ \mathbf{p}_r^T \end{bmatrix}, \tag{1.1}$$

where $\lambda_1, \ldots, \lambda_r$ are positive eigenvalues of B (the rest are zero) and $\mathbf{p}_1, \ldots, \mathbf{p}_r$ are the corresponding orthonormal eigenvectors. Then the following coordinates $\mathbf{x}_1, \ldots, \mathbf{x}_n$ obtained by

$$[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] := \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_r} \end{bmatrix} \begin{bmatrix} \mathbf{p}_1^T \\ \vdots \\ \mathbf{p}_r^T \end{bmatrix} \tag{1.2}$$

preserve the known distances in the sense that $\|\mathbf{x}_i - \mathbf{x}_j\| = d_{ij}$ for all $i, j = 1, \ldots, n$. This is the well known cMDS. We refer to [16, 25, 30, 31] and [4, 7, 8] for detailed description of cMDS and its generalizations.

In almost all applications, D is not fully given or contains noise and in this case D is often denoted by Δ. When Δ is not far from a true EDM, cMDS works fairly well. This has been justified in various situations including [28] on manifold learning and [9] where the noises can be bounded. Instead of directly applying cMDS on Δ, one may also modify Δ so as for it to be a true EDM. Research on this line were mainly contributed from numerical linear algebra and optimization and include e.g., [1,3,15,21,23,29]. Those methods, in addition to those discussed in [4], belong to the category of metric MDS, which means that the magnitudes of the known distances are far more important than the rest information. Metric MDS works well when the noise in D is not at a very high level.

When D contains inaccurate distances whose errors cannot be regarded to be from random noises, the magnitudes of the erroneous distances may have a disastrous effect on the embedding constructions. Let us take a look at a simple example, which clearly demonstrates the undesired effect. Network (a) in Fig. 1.1 is a true square network with the first point being in the center. The true distance matrix is

$$D = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 4 & 2 \\ 1 & 2 & 0 & 2 & 4 \\ 1 & 4 & 2 & 0 & 2 \\ 1 & 2 & 4 & 2 & 0 \end{bmatrix}.$$

Now suppose the distances from the center to the rest points are messed up with other distances. For example, $D_{12} = D_{14} = 2$ and $D_{13} = D_{15} = 4$. Comparing with the true distance, which is 1, the errors are large. We apply several well-known methods to the corresponding Δ matrix and the reconstructed networks are shown from (b) to (f) in Fig. 1.1. It can be seen that the smoothing Newton method proposed in this paper is the only one which correctly recovers the true network topology. In our experiments, we enforced the ordinal constraints:

$$D_{13} \leq D_{12}, \ D_{13} \leq D_{14}, \ D_{15} \leq D_{12}, \ D_{15} \leq D_{14}.$$

We note that those ordinal constraints are obeyed by the true network, but are violated by the erroneous D. It is those explicitly enforced ordinal constraints that distinguish our method from the rest.