Objectives

The objectives of this research project include:

- Understanding classical MDS on finitely many points of the circle
- Defining a notion of MDS on infinite metric measure spaces
- Studying its optimal properties and goodness of fit
- Testing convergence of MDS on metric measure spaces

Introduction

Multidimensional scaling (MDS) is concerned with the problem of constructing a configuration of \( n \) points in Euclidean space using information about the dissimilarities between the objects. The distances need not be based on Euclidean distances; they can represent many types of dissimilarities between objects. The goal of MDS is to map the objects \( x_1, \ldots, x_n \) to configuration or embedding points \( f(x_1), \ldots, f(x_n) \) in \( \mathbb{R}^m \) in such a way that the given dissimilarities \( d(x_i, x_j) \) are well-approximated by the distances \( ||f(x_i) - f(x_j)||_2 \) [1].

MDS of evenly spaced points on a Circle

Let \( (S^1, d) \) be the unit circle, equipped with the geodesic metric \( d \). Figure 1 shows the MDS embedding in \( \mathbb{R}^3 \) of 1000 evenly spaced points on \( S^1 \).

Proposition. The Classical MDS embedding in \( \mathbb{R}^m \) of \( n \) evenly spaced points on \( S^1 \) lies, up to a rigid motion of \( \mathbb{R}^m \), on the curve
\[
\gamma_n: S^1 \to \mathbb{R}^m, \quad \gamma_n(e^{i\theta}) = (a_1(n) \cos(\theta), a_1(n) \sin(\theta), a_2(n) \cos(3\theta), a_2(n) \sin(3\theta), \ldots, a_n(n) \cos(n\theta), a_n(n) \sin(n\theta)) \in \mathbb{R}^m,
\]
where \( \lim_{n \to \infty} a_j(n) = \frac{\sqrt{2}}{j} \) (with \( j \) odd).


MDS on Infinite Metric Measure Spaces

Multidimensional scaling (MDS) is a popular technique for mapping a finite metric space \((X, d_X)\) into a low-dimensional Euclidean space \( \mathbb{R}^m \) in a way that best preserves pairwise distances. We are working on constructing an extension of MDS to infinite metric measure spaces \((X, d_X, \mu)\).

Suppose we start with a measure metric space \((X, d_X, \mu)\), where \( d_X \) is a real-valued \( L^2 \)-function on \( X \times X \) with respect to the measure \( \mu \). We hope that the metric \( d_X \) can be approximately represented by a Euclidean metric \( d_{\hat{X}} : \hat{X} \times \hat{X} \to \mathbb{R} \) on a space \( \hat{X} \) in a Euclidean space \( \mathbb{R}^m \), perhaps of low dimension (often \( m = 2 \) or \( 3 \)).

Goal: Convergence of MDS

Let \( X_n = ((X_n, d_{X_n}, \mu_n))_{n \in \mathbb{N}} \) be a sequence of metric measure spaces and let \( \hat{X}_n = ((\hat{X}_n, d_{\hat{X}_n}, \mu_n))_{n \in \mathbb{N}} \) be a sequence of the corresponding MDS embeddings of \( X_n \) such that \( d_{\hat{X}_n} \) is Euclidean distance on \( \hat{X}_n \subseteq \mathbb{R}^m \).

There are various notions of convergence of \( X_n \) to \( X = (X, d_X, \mu_X) \) as \( n \to \infty \), one of which can be defined using the Gromov-Wasserstein distance \( D_p \) [2]. We hope to show that if we have convergence \( X_n \overset{D_p}{\to} X \), then we also have convergence \( \hat{X}_n \overset{D_p}{\to} \hat{X} \).

In the particular case when the \( X_n \) are all finite with the same number of points, this convergence follows from [3].

Motivating questions

- If a finite sample \( X_n \subseteq X \) converges to \( X \) as we sample more points, then in what sense do the MDS embeddings of these finite samples converge to the MDS embedding of \( X \)?
- More generally, if a sequence of metric measure spaces converges to \( X \), then in what sense do the MDS embeddings of these spaces converge to the MDS embedding of \( X \)?

References


Contact Information

- Web: https://larakassab.weebly.com
- Email: lara.kassab@colorado.edu