Visualizing and Controlling the Optimization Process of Geometric Multidimensional Scaling

Martynas Sabaliauskas

E-mail: martynas.sabaliauskas@mif.vu.lt

Vilnius University, Institute of Data Science and Digital Technologies

Abstract

Geometric Multidimensional Scaling (GMDS) is an efficient iterative approach for dimensionality reduction derived from a geometric interpretation of the raw stress optimization problem. It provides analytically defined steps corresponding to the anti-gradient direction of the local stress.
This research presents a rigorous theoretical investigation into the **asymptotic properties** of the GMDS iteration formula and the **geometry of the optimization**. We formally prove three novel discoveries regarding the asymptotics near singularities, at infinity, and during configuration coalescence.
These findings significantly deepen the theoretical understanding of GMDS dynamics and the geometric characteristics of the MDS stress function near critical regions, enabling visualization and control of the optimization process.

The MDS Optimization Challenge

MDS seeks a configuration of points $Y = \{Y_1, ..., Y_m\}$ in $\mathbb{R}^n$ such that the inter-point distances closely match the original pairwise dissimilarities $D = \{d_{ij}\}$. This is achieved by minimizing the **Raw Stress** function:

$$S(Y) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} (d_{ij} - ||Y_i - Y_j||)^2. \tag{1}$$

The Problem: The landscape of $S(Y)$ is highly non-convex and characterized by numerous local minima. Traditional methods (e.g., SMACOF) often get trapped and operate as "black boxes".

Geometric MDS (GMDS) Framework

GMDS optimizes the position of a single point Y_j while others are fixed, minimizing the **local stress**:

$$S^*(Y_j) = \sum_{i \neq j}^{m} (d_{ij} - ||Y_i - Y_j||)^2. \tag{2}$$

The update is derived geometrically using **Auxiliary Points** A_{ij}, representing the ideal location for Y_j relative to Y_i:

$$A_{ij} = Y_i + d_{ij} \cdot \frac{Y_j - Y_i}{||Y_j - Y_i||}. \tag{3}$$

The new position Y_j^* is the **centroid** of these auxiliary points:

$$Y_j^* = \frac{1}{m-1} \sum_{i \neq j}^{m} A_{ij}. \tag{4}$$

Key Property (Anti-Gradient): The GMDS step corresponds exactly to the anti-gradient direction, guaranteeing descent [6]:

$$Y_j^* = Y_j - \frac{1}{2(m-1)} \nabla S^*|_{Y_j}. \tag{5}$$

The Research Focus: The GMDS formula (Eq. 4) is undefined at singularities ($||Y_j - Y_i|| \to 0$). We analyze the algorithm's dynamics near these critical regions.

Three Key Discoveries

We provide a rigorous asymptotic analysis of the GMDS iteration formula under extreme conditions. This analysis reveals the geometric structure of the optimization landscape and the basins of attraction near the boundaries of the configuration space.

Discovery 1: Asymptotics Near Singularities ($Y_j \to Y_k$)

We investigate the scenario where the point being moved, Y_j, approaches another fixed point Y_k.

Theorem

(Limit Hypersphere Near Singularity). *As $||Y_j - Y_k|| \to 0$, the limit set of the resulting GMDS positions Y_j^* forms an n-dimensional hypersphere, $S_{j,k}$.*

The limit position $Y_j^*(u)$, parameterized by the direction of approach $u \in S^{n-1}$ (the unit sphere), is:

$$Y_j^*(u) = C_{j,k} + R_{j,k} \cdot u. \tag{6}$$

The center $C_{j,k}$ is analytically derived (see paper Eq. 12), and the radius is:

$$R_{j,k} = \frac{d_{jk}}{m-1}. \tag{7}$$

Implication (Stability): The radius $R_{j,k}$ quantifies the **repulsive force** near the singularity. This reveals an inherent mechanism within GMDS that prevents the degenerate collapse of points if $d_{jk} > 0$.

Discovery 2: Asymptotics at Infinity ($||Y_j|| \to \infty$)

We analyze the dynamics when the point Y_j moves infinitely far away.

Theorem

(Limit Hypersphere at Infinity). *As $||Y_j|| \to \infty$, the limit set of Y_j^* also forms an n-dimensional hypersphere, S_∞.*

$$Y_j^*(u) = C_\infty + R_\infty \cdot u. \tag{8}$$

The center C_∞ is the **centroid** of the remaining configuration:

$$C_\infty = \frac{1}{m-1} \sum_{i \neq j} Y_i. \tag{9}$$

The radius R_∞ is the **average dissimilarity** of object j:

$$R_\infty = \frac{1}{m-1} \sum_{i \neq j} d_{ij}. \tag{10}$$

Implication (Centering): This reveals an inherent **"pull-back" mechanism**. If a point diverges, the iteration pulls it back towards the centroid, managing outliers and stabilizing the visualization.

Discovery 3: Configuration Coalescence ($Y_k \to Y_l$)

We investigate how the optimization landscape for Y_j evolves when two distinct fixed points, Y_k and Y_l, approach each other. This analyzes how basins of attraction merge.
We consider the interaction between the limit hyperspheres $S_{j,k}$ and $S_{j,l}$ (from Discovery 1).

Theorem

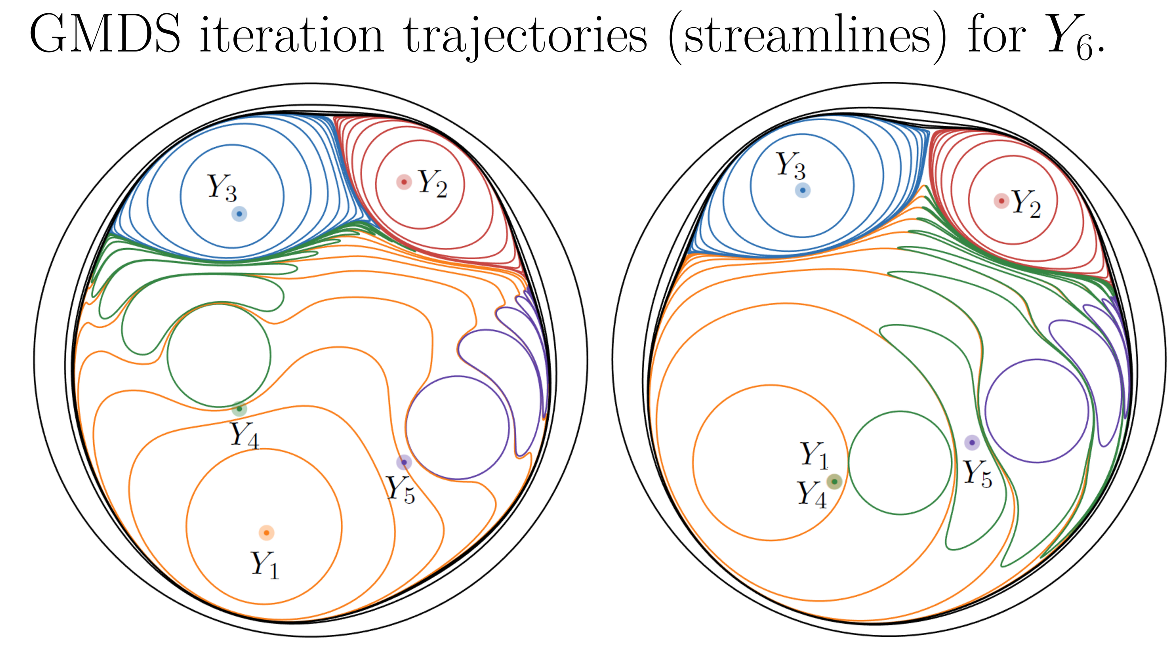
(Limit Tangency under Coalescence). *As $Y_k \to Y_l$, the distance between the centers of the limit hyperspheres $C_{j,k}$ and $C_{j,l}$ converges exactly to the sum of their radii.*

$$\lim_{Y_k \to Y_l} ||C_{j,k} - C_{j,l}|| = R_{j,k} + R_{j,l}. \tag{11}$$

Implication (Landscape Evolution): The limit hyperspheres become exactly **tangent externally** in the limit. This characterizes the precise geometric mechanism by which basins of attraction merge as the configuration approaches degeneracy.

Visualization of the Landscape

We illustrate the theoretical findings using a 2D visualization ($n = 2$) for $m = 6$ points. We visualize the GMDS iteration trajectories (streamlines) for Y_6.



(a) Well-separated configuration: Colored circles show limit sets $S_{6,k}$ (Discovery 1). The outer black circle shows the limit set at infinity (Discovery 2).
(b) Coalescence of Y_1 and Y_4: The corresponding limit circles are nearly tangent (Discovery 3). The landscape deforms, and the basins merge.

Conclusions

This study provides a rigorous mathematical framework for understanding the GMDS iteration and the geometry of the MDS optimization.

1. **Stability Mechanisms:** We proved that GMDS steps near singularities and at infinity are constrained to well-defined hyperspheres, revealing inherent mechanisms for stability and centering.
2. **Foundation for Advancement:** These theoretical insights provide a foundation for robust, visualization-guided GMDS algorithms, transforming the optimization from a "black box" to a controllable process.