Equivariant Dimensionality Reduction on Stiefel Manifolds

Maddie Weinstein

Stanford University

mweinste@stanford.edu

with Andrew Lee, Harlin Lee, and Nikolas Schonsheck
project advised by Jose Perea

Outline

- general problem: dimensionality reduction
- our setting: Stiefel manifolds
- our problem
- our solution

Dimensionality Reduction

- **General Idea:** Let $n < N$. Suppose we have a data set $X \subset \mathbb{R}^N$ of intrinsic dimension n. We wish to find an n-dimensional subspace $S \subset \mathbb{R}^N$ and a subset $\tilde{X} \subset S$ that best represents X.
- **Our Setting:** We will present an algorithm for dimensionality reduction on **Stiefel manifolds** that respects their topology.

Stiefel Manifolds

- Let $0 < t < s$. The **Stiefel manifold** $V_t(\mathbb{R}^s) \subset \mathbb{R}^{s \times t}$ is the set of orthonormal t-frames in $\mathbb{R}^s$.
- The orthogonal group $O(t)$ acts on $V_t(\mathbb{R}^s)$ from the right via matrix multiplication. The quotient is the real **Grassmannian** $G_t(\mathbb{R}^s)$.

Equivariant Maps

- Let X, Y be spaces with an action of a group G, and $\pi : X \to Y$ be a map. We say that π is G-**equivariant** if for all $x \in X$, we have $\pi(g \cdot x) = g \cdot \pi(x)$.
- If a map π is equivariant, then it respects equivalence classes. That is, if $x_1 \sim x_2$, then $\pi(x_1) \sim \pi(x_2)$.
- In particular, if our dimension reduction map π from a subset of $V_k(\mathbb{R}^N)$ to a subset of $V_k(\mathbb{R}^n)$ is $O(k)$-equivariant, then frames that span the same k-dimensional subspace of $\mathbb{R}^N$ will map to frames that span the same k-dimensional subspace of $\mathbb{R}^n$. Thus π will descend to a map from a subset of $G_k(\mathbb{R}^N)$ to a subset of $G_k(\mathbb{R}^n)$.

The Problem

Let $k < n << N$. Suppose that we are given a data set $X \subset V_k(\mathbb{R}^N)$.

We seek:

- An embedding $\alpha : V_k(\mathbb{R}^n) \to V_k(\mathbb{R}^N)$ that is optimal with respect to X.
  - The set of possible embeddings α is parametrized by the Stiefel manifold $V_n(\mathbb{R}^N)$.
  - An embedding α is optimal with respect to X if it minimizes the sum of the squared distances between each data point $x_i \in X$ and its image $\pi_\alpha(x_i)$.
- An equivariant projection map $\pi_\alpha : X \to \alpha(V_k(\mathbb{R}^n))$.

The image $\tilde{X} := \pi_\alpha(X)$ is a lower-dimensional representation of X.

An Equivariant Dimension Reduction Map

Suppose we have chosen an embedding α. We now define π_α.

- Polar decomposition: Let $A \in \mathbb{R}^{n \times k}$ with $n \geq k$. There exists a matrix $U \in \mathbb{R}^{n \times k}$ with orthonormal columns and a unique self-adjoint positive semidefinite matrix $H \in \mathbb{R}^{k \times k}$ such that $A = UH$. If $\text{rank}(A) = k$, then H is positive definite, hence invertible, and U is uniquely determined by $U = AH^{-1}$.
- Fix $\alpha \in V_n(\mathbb{R}^N)$. Let $L = \{y \in V_k(\mathbb{R}^N) | \, \text{rank}(\alpha^T y) < k\}$. Define $\pi_\alpha \colon V_k(\mathbb{R}^N) \setminus L \to \alpha(V_k(\mathbb{R}^n))$ as follows. Let $\alpha^T y = UH$ be the unique polar decomposition of $\alpha^T y$. Define π_α by $\pi_\alpha(y) := \alpha U$.
- Proposition: For fixed α, π_α minimizes the sum of squared distances from $x_i \in X$ to their images in $\alpha(V_k(\mathbb{R}^n))$.
- Proposition: π_α is $O(k)$-equivariant.

Finding an Optimal Embedding

- The set of possible embeddings α is parametrized by the Stiefel manifold $V_n(\mathbb{R}^N)$.
- Under strict assumptions on the data set X, PCA supplies a critical embedding α.
- Under looser assumptions on the data set X, we use gradient descent with the α supplied by PCA as an initial point.
- The software package `manopt` implements gradient descent on Stiefel manifolds.

Thank you!