Advanced Topics In Machine Learning Project Report : Low Dimensional Embedding of a Pose Collection Fabian Prada

## 1 Introduction

In this project we present an overview of (1) low dimensional embedding, and (2) K-Most Dissimilar Subset detection for a set of samples for which only the matrix M of pairwise dissimilarity is provided.

For evaluation (and visual validation) our dissimilarity matrix is computed on a collection of shapes, but the analysis and techniques discussed can be applied to any set that admits a dissimilarity measure.

Following the approach of [2], we compute for each shape  $x_i$  its characteristic function  $\chi_i$  and its euclidean distance transform  $s_i$ . We define,

$$M_{ij} = \min_{q \in G} \langle \chi_i, s_j \circ g \rangle + \langle s_i, \chi_j \circ g \rangle \tag{1}$$

where G is the set of rigid transformations in  $\mathbb{R}^3$ .

We must notice that an ideal dissimilarity matrix M should satisfy metric properties<sup>1</sup>, i.e, $M_{ij} = d(x_i, x_j)$  for some metric  $d : X \times X \to \mathbb{R}$ . Our shape dissimilarity 1 satisfy positivity, identity, and symmetry properties but not necessarily triangle inequality<sup>2</sup>. Methods based in Lp distance of rotational invariant feature vectors satisfy triangular inequality but tend to have less discriminative power.

Given a dissimilarity matrix M we will derive two types of affinity matrices that will be used in the sections below. We define the *product-affine* matrix A as:

$$A = -\frac{1}{2}J(M \odot M)J \tag{2}$$

<sup>&</sup>lt;sup>1</sup>Metric properties are necessary for M to admit a perfect embedding.

<sup>&</sup>lt;sup>2</sup>In our evaluation less than 8% of all triplets dont satisfy triangular inequality

where  $J = I - \frac{1}{N} \mathbf{1} \mathbf{1}^{\top}$  is a centering operator. We also define the *kernel-affine* matrix W as:

$$W_{ij} = \exp\{-\frac{M_{ij}^2}{2\sigma^2}\}$$
 (3)

In contrast to dissimilarity matrices, affinity matrices have large values at entries associated to similar samples.

Finally, we denote by K to the k-nearest neighbor mask associated to M. The k-nearest neighbor affinity matrix correspond to  $W \odot K$ . Abusing of notation, we will keep denoting those affinity matrices as W.

## 2 Low Dimensional Embedding

**Problem 1:** Given dissimilarity M compute  $Y = \{y_1, \ldots, y_N\} \subset \mathbb{R}^d$  that minimizes,

$$\operatorname{argmin}_{Y} \sum_{i,j} (||y_i - y_j|| - M_{ij})^2$$
(4)

The problem above is non-convex and not admit a close form solution.

### 2.1 MultiDimension Scaling(MDS)

MDS states an alternative problem in terms of inner-product affinities,

$$\operatorname{argmin}_{Y} \sum_{i,j} (\langle y_i, y_j \rangle - a_{ij})^2 = \operatorname{argmin}_{Y} ||Y^\top Y - A||_F^2$$
(5)

Close form solution to this problem is given by  $Y^{\top} = R \Sigma_d^{1/2} U_d$ , i.e., an optimal *d*-rank approxiantion of *A*.

In equation 2 we formulate the relation between dissimilarity matrix and inner-product affinity matrix. Lets formalize this derivation: **Proposition 1.** Given  $x_1, \ldots, x_N \subset \mathbb{R}^D$ , let  $x_i = Uy_i + \mu$  be its principal componet decomposition. Define  $M_{ij} = ||x_i - x_j||$  and  $A_{ij} = \langle y_i, y_j \rangle$ . Then  $A = -\frac{1}{2}J(M \odot M)J$ 

Proof.

$$(M \odot M)_{ij} = M_{ij}^2 = ||x_i - x_j||^2 = ||y_i - y_j||^2 = ||y_i||^2 - 2\langle y_i, y_j \rangle + ||y_j||^2$$

For any matrix T,  $(JTJ)_{ij} = T_{ij} - \overline{T}_i - \overline{T^j} + \overline{T}$ . Since  $\overline{y} = 0$ , it can be checked that  $\overline{(M \odot M)_i} = ||y_i||^2 + \overline{||y||}^2, \overline{(M \odot M)^j} = ||y_j||^2 + \overline{||y||}^2$ , and  $\overline{(M \odot M)} = 2\overline{||y||}^2$ . Thus  $-\frac{1}{2}(J(M \odot M)J)_{ij} = \langle y_i, y_j \rangle$ .

From the previous proposition, we obtain the next corollary,

**Corollary 2.** If M is the dissimilarity matrix of points in a d-dimensional affine linear space, and  $A = -\frac{1}{2}J(M \odot M)J$ , then MDS provides a perfect embedding (i.e. energy 4 is zero).

### 2.2 Laplacian Eigenmaps(LE)

LE minimize an energy that also admits close form solution,

$$\operatorname{argmin}_{Y} \sum_{i,j} w_{ij} ||y_i - y_j||^2 \text{ s.t. } Y \mathcal{D} \mathbf{1} = \mathbf{0} \text{ and } Y \mathcal{D} Y^\top = I$$
(6)

The embedding in this case is given by the second to (d+1)-st eigevectors of the generalized eigenvalue problem,  $\mathcal{L}Y^{\top} = \Delta \mathcal{D}Y^{\top}$  where  $\mathcal{L} = \mathcal{D} - W$ , and  $\mathcal{D}_{ii} = \sum_{j} w_{ij}$ 

Large values of  $w_{ij}$  enforce proximity between  $y_i$  and  $y_j$ , while the constraints avoid a trivial solution. In fact, the following condition holds.

**Proposition 3.** Let  $\rho_i = \frac{D_{ii}}{\sum_i D_{ii}}$ , and  $\{y_i\}_i$  the embedding positions. Then the weighted point distribution  $(y_i, \rho_i)_i$  has zero mean and covariance  $\frac{1}{\sum_i D_{ii}} I_d$ .

For our evaluation we set  $w_{ij} = \exp\{-\frac{M_{ij}^2}{2\sigma^2}\}$  where  $\sigma^2 = \overline{M_{ij}^2}$ . We evaluate LE with both normalized and unnormalized constraints and with different *k*-nearest neighbour masks.

### 2.3 Low Dimensional Embedding Evaluation

All the tests in this section were done on a sequence of 2283 frames captured in a continuous performance.

#### Test 1 : MDS Approximation Ratio

In this experiment we plot the approximation ratio  $1 - \frac{||Y^{\top}Y - A||_{F}^{2}}{||A||_{F}^{2}}$  as a function of the embedding dimension d. For d = 3 we get an approximation ratio of 0.72. Since A is not positive definite it can not represented as  $Y^{\top}Y$ , however for this example we get an optimal approximation error of 0.9809.



Figure 1: Approximation Ratio.

#### Test 2 : Scaled Embedding Error

Given an embedding  $Y = \{y_1, \ldots, y_N\}$  we define the optimal scaling  $\lambda$  as

$$\operatorname{argmin}_{\lambda} \sum_{i,j} (||\lambda y_i - \lambda y_j|| - M_{ij})^2$$
(7)

Solution to equation above is given by  $\lambda = \frac{\sum ||y_i - y_j|| M_{ij}}{\sum ||y_i - y_j||^2}$ . The scaled embedding error is the value of the energy in 7 for the optimal scaling.

The following were the results of this test for 3 dimensional embedding on the evaluation sequence

Zero Emb.	Rand. Emb.	MDS	LE	LE-10NN
5.11e+14	8.78e + 13	$4.46e{+}13$	4.80e + 13	2.37e + 14

As expected MDS provided the best embedding error<sup>3</sup>.

#### Test 3 : Embedding Qualitative Results

In Figure 3 we observe the embedding provided by LE, LE-10NN and MDS. The embedding for LE-10NN looks like a self intersecting curve. This is particularly nice for a pose collection obtained from a continuous motion since consecutive frames are mapped to consecutive points in the curve. However, since the embedding only consider local similarity, very different poses can be (and are indeed ) mapped to close points. On the other hand LE and MDS provided suprisingly similar embeddings. Both embeddings perform a satisfactory job on preserving similar poses close and separating dissimilars.



Figure 2: 3D Embedding Results.

## 3 K-Most Dissimilar Subset

The second problem we will consider is identifying a subset of samples that are the most dissimilar among them. We will be considering the following formulation for this problem:

**Problem 2:** Given dissimilarly M, identify a subset of size k that maximize the minimal pairwise distance:

<sup>&</sup>lt;sup>3</sup>By Corollary 2, MDS attains zero error of equation 7 when the dissimilarity matrix is the euclidean distance samples in a d-dimensional subspace.

$$\max_{S \subset [n], |S|=k} \min_{i,j \in S} M_{ij} \tag{8}$$

Observe that this problem is combinatorial, and a brute force solution would require the evaluation of  $\binom{N}{K}$  subsets. This is prohibitively expensive for large values of N and K.

Rather than considering the entire set of N samples we propouse to consider only a small subset of samples that make the combinatorial search tractable. The set we will be considering is the convex hull of a low dimensional embedding for d = 2, 3.

For d = 2 and d = 3 the convex hull of a set of N samples can be computed in  $O(N \log h)$  where h is the size of the output [1]. For larger dimensions the complexity is bounded by  $O(N^{d/2})$ .

In Figure 3 we show the percentage of samples (N = 2283) that belong to the convex hull as a function of the embedding dimension. For d = 2, 3, 4the number of points in the hull for MDS and LE is roghly 1%, 3% and 10% of the input size.



Figure 3: Convex Hull. Left: plot of percentage of vertices in the hull as function of d. Right: convex hull of the 3D embedding for MDS(middle) and LE(right).

#### Test 4: Most dissimilar pair and triplet

For our evaluation set, the most dissimilar pair is (963, 1678) and the most dissimilar triplet is (645, 1679, 2283). The convex hull of the 2-dimensional embeddings provided by LE and MSD (with 19 and 21 samples, resp.) contains the pair (963, 1678). These convex hulls does not contain the triplet (645, 1679, 2283) but a very close one : (642, 1679, 2283) (see Figure 3).

Since the *d*-dimensional embedding provided by LE and MSD is a projection of the d + 1-dimensional embedding (a projection of the first *d* coordinates), all the samples belonging to the hull of the *d*-dimensional embedding are still in the hull of the higher one. Thus (963, 1678) is keep in the hull of the 3-dimensional embedding for both methods. Even more, the most dissimilar triplet (645, 1679, 2283) is perfectly contained in the hull of the 3D embedding for both methods.



Figure 4: Most dissimilar pair (top) and triplet (bottom). Both are within the vertices of the convex hull of the 2D and 3D embedding provided by LE (depicted in the images) and MDS.

#### Test 5: Greedy K-Most Dissimilar

A natural approach to generate k dissimilar poses is following a greedy approach: start by finding the most dissimilar and at each iteration add to set the element that maximize the minimum distance to the already selected poses. It is easy to check that such greedy algorithm has complexity  $O(Nk^2)$ .

If we use the Greedy K-Most Dissimilar algorithm to identify the 3 most dissimilar triplet we get the result (963, 1678, 803) which has dissimilarity score (i.e., the minima among pairwise similarity) of  $2.18e10^8$ . Instead, using exhaustive search in the convex hull of the 2-dimensional embedding of LE (which evaluates  $\binom{19}{3}$  triplets), we get the triplet (642, 1680, 2283) which has dissimilarity score  $2.76e10^8$ .

#### Conclusions

MDS and LE produced a satisfactory embedding of the pose collection: similar poses where mapped to close points in the embedding, while the most dissimilar poses were well separated. Despite LE and MDS minimize significantly different energy functions, there was a surprisingly strong similarity of the embedding points distributions and samples belonging to the convex hull. This was an unexpected result since there is no direct relation between the LE energy 6 and the optimal embedding energy 4. Further experimental and theoretical analysis is required to understand the extent of this relation. In this project we propoused a characterization of the most dissimilar subset. Our testing case reinforced the hypothesis that motivated this project: the most dissimilar poses (or a very good approximation) can be directly obtained from the convex hull of a low dimensional embedding. Additional evaluation and analysis is required to understand under what conditions this is the case. Since the convex hull can be efficiently computed ( $O(N \log h)$ ) for d = 2 and outputs a small subset of vertices (1% for d = 2 in our testing case) it would be a very useful tool to get a set vertices where the most dissimilar subset computation can be tractable.

# References

- C. Bradford Barber, David P. Dobkin, and Hannu Huhdanpaa. The quickhull algorithm for convex hulls. ACM Trans. Math. Softw., 22(4):469–483, December 1996.
- [2] Thomas Funkhouser, Michael Kazhdan, Philip Shilane, Patrick Min, William Kiefer, Ayellet Tal, Szymon Rusinkiewicz, and David Dobkin. Modeling by example. ACM Trans. Graph., 23(3), 2004.