Improved Graph Laplacian via Geometric Consistency

Yu-Chia Chen, Dominique Perrault-Joncas, Marina Meilă, James McQueen. University of Washington

Original paper: (on NIPS 2017)
Improved Graph Laplacian via Geometric Consistency
Dominique Perrault-Joncas, Marina Meilă, James McQueen. University of Washington

The task

Problem: **Estimate the radius** r of heat kernel in manifold embedding

Formally: Optimize Laplacian w.r.t. parameters (e.g. radius r)

Previous work:

- asymptotic rates depending on the (unknown) manifold [4]
- Embedding dependent neighborhood reconstruction [6]

Challenge: it's an unsupervised problem! What "target" to choose?

The radius r affects…

- Quality of manifold embedding via neighborhood selection
- Laplacian-based embedding and clustering via the kernel for computing similarities
- Estimation of other geometric quantities that depend on the Laplacian (e.g Riemannian metric) or not (e.g intrinsic dimension).
- Regression on manifolds via Gaussian Processes or Laplacian regularization.

Heat Kernels, Laplacians, and Geometry

- Heat Kernel

$$W_{ij} = \exp\left(\frac{\|x_i - x_j\|^2}{r^2}\right)$$

- Radius parameter: r
- Compute the Graph Laplacian:

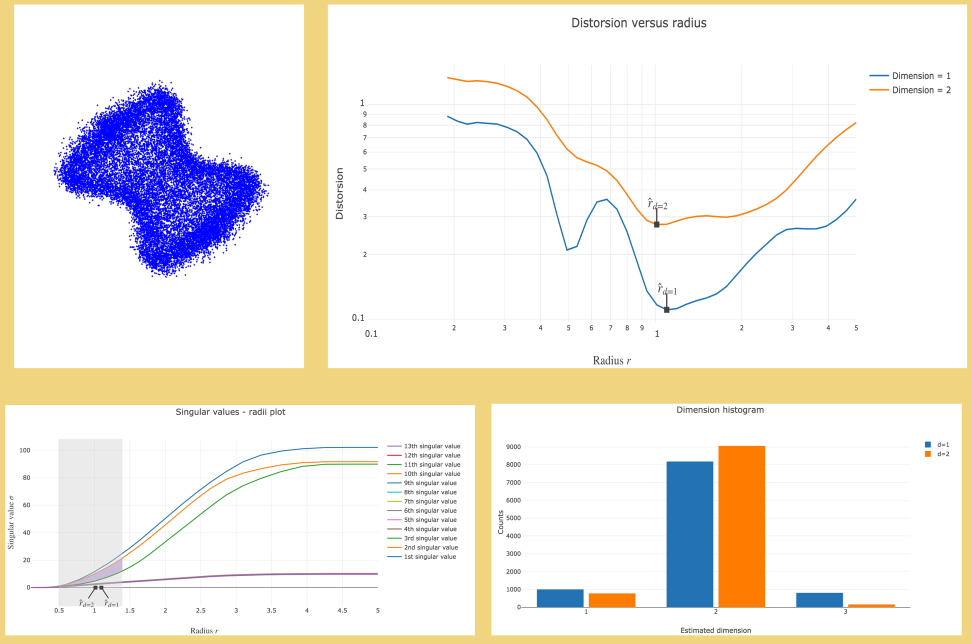
$$t_i = \sum_j W_{ij}, \qquad W'_{ij} = \frac{W_{ij}}{t_i t_j}, \qquad t'_i = \sum_j W'_{ij}$$

- Then

$$L = \sum_j \frac{W'_{ij}}{t'_j}$$

- Assume a Riemannian Manifold $(\mathcal{M}, g)$
- Riemannian Metric, g, encodes geometry e.g. volume element is $\sqrt{\det G(X)}$
- $\left(H(p)\right)_{kj} = \frac{1}{2}\Delta_{\mathcal{M}}\left(x^k - x^k(p)\right)\left(x^i - x^i(p)\right)|_{x=x(p)}$
- Optimize r for geometric consistency

Using $\hat{r}$ for dimension estimation (with [5])



- Intrinsic dimension is estimated by the eigengap of Local SVD with radius $\hat{r}$.
- Upper left: hourglass data
- Upper right: $\hat{r}$ estimates as the minimizer of distorsion.
- Lower left: avg. singular values versus radii.
- Lower right: histogram of estimated dimensions on each points.

GC Algorithm: Optimizing the Laplacian

Input Data $\{x_1, x_2, \dots x_N\}$, dimension d', pow=$1, -1$
For each ϵ
1. Estimate the Laplacian induced by r
2. For each data point x_i (in a subsample)
   1. Weights $w_j = K_r(x_i, x_j)$ for all x_j
   2. Project neighbors of x_i on tangent subspace

Algorithm 2 Tangent Subspace Projection(X, w, d')
Input: $N \times r$ design matrix X, weight vector w, working dimension d'
Compute Z using (6)
$[V, \Lambda] \leftarrow \text{eig}(Z^T Z, d')$ (i.e.d'-SVD of Z)
Center X around $\bar{x}$ from (6)
$Y \leftarrow XV_{:,1:d'}$ (Project X on d' principal subspace)
return Y

3. Treat Y as an embedding of X. Estimate the R. metric for Y

Algorithm 1 Riemannian Metric($X, i, L, pow \in \{-1, 1\}$)
Input: $N \times d$ design matrix X, i index in data set, Laplacian L, binary variable pow
for $k = 1 \to d, l = 1 \to d$ **do**
$H_{k,l} \leftarrow \sum_{j=1}^{N} L_{ij}(X_{jk} - X_{ik})(X_{jl} - X_{il})$
end for
return H^{pow} (i.e. H if $pow = 1$ and H^{-1} if $pow = -1$)

4. But Y should be isometric to X. Hence H should be the identity matrix. Penalize the difference.

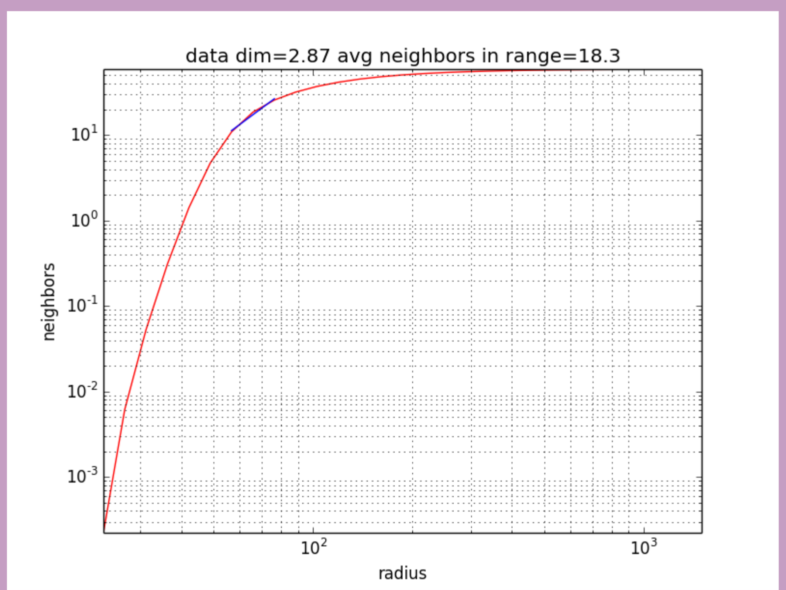
Algorithm 3 Compute Distortion(X, ϵ, d')
Input: $N \times r$ design matrix X, ϵ, working dimension d', index set $\mathcal{I} \subseteq \{1, \dots, N\}$
Compute the heat kernel W by (2) for each pair of points in X
Compute the graph Laplacian L from W by (3)
$D \leftarrow 0$
for $i \in \mathcal{I}$ **do**
$Y \leftarrow$ TangentSubspaceProjection($X, W_{i,:}, d'$)
$H \leftarrow$ RiemannianMetric($Y, L, pow = 1$)
$D \leftarrow D + \|H - I_{d'}\|^2 / |\mathcal{I}|$
end for
return D

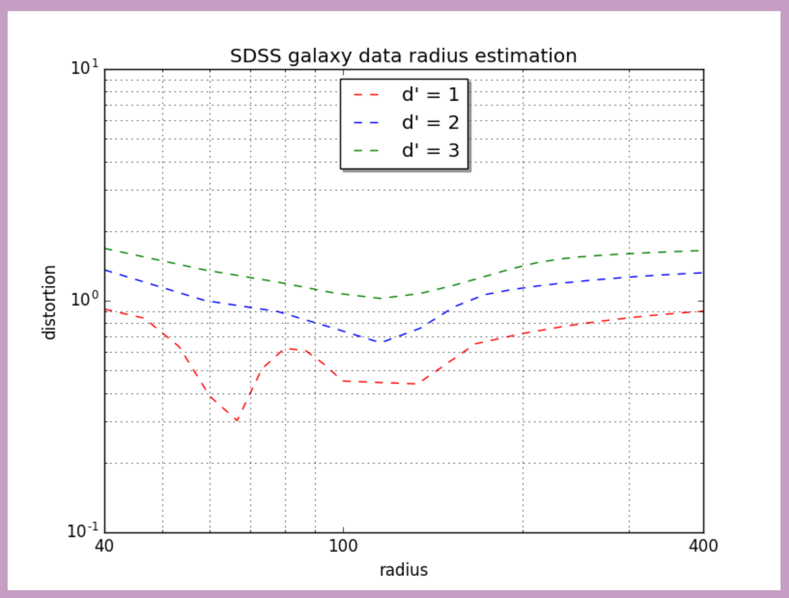
Measures departure from isometry, i.e. geometric consistency

Output $\hat{r}$ that minimizes distortion D

$\hat{r}$ for embedding Spectra of galaxies

N=670,000, r=3750 dimensions (www.sdss.org)





Radius Estimate for Galaxy Spectra.
Left: GC results for d' = 1, 2, 3; $3 \cdot r^{\text{opt}} = 66$
Right: log-log plot of radius vs avg. # nbrs;
Indicates d=3 at $r^{\text{opt}} = 22$

Conclusions: Geometry Consistency (GC) is…

- Choosing the correct radius/bound/scale is important in any non-linear dimension reduction task1.
- The GC Algorithm required minimal knowledge:
  - maximum radius, minimum radius,
  - (optionally: dimension d of the manifold.)
- The chosen radius can be used in
  - any embedding algorithm
  - semi-supervised learning with Laplacian Regularizer (see our NIPS 2017 paper)
  - estimating dimension d (as shown here)

References

[1] R. R. Coifman, S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 2006.
[2] D. Perrault-Joncas, M. Meila, Metric learning and manifolds: Preserving the intrinsic geometry , arXiv1305.7255
[3]X. Zhou, M. Belkin. Semi-supervised learning by higher order regularization. *AISTAT*, 2011
[4] A. Singer. From graph to manifold laplacian: the convergence rate. *Applied and Computational Harmonic Analysis*, 2006.
[5] G. Chen, A. Little, M. Maggioni, L. Rosasco. Some recent advances in multiscale geometric analysis of point clouds. *Wavelets and multiscale analysis*. Springer, 2011.
[6] L. ChenA. Buja. Local Multidimensional Scaling for nonlinear dimension reduction, graph drawing and proximity analysis, JASA,2009.