Spectral Clustering Survey

HU, Pili*

May 14, 2012†

Abstract

Abstract. Sources can be found in [6].

*hupili [at] ie [dot] cuhk [dot] edu [dot] hk

†Last compile:May 14, 2012

Contents

1 Introduction

Spectral Clustering(SC) was used in several disciplines long ago. For example, computer vision[12], load balancing [4], electronics design [3], etc. Spectral Embeding(SE) was also widely discussed in the community[2]. Outside spectral community, the machine learning community also developed many linear or non-linear Dimensionality Reduction(DR) methods, like Principal Component Anslysis (PCA), Kernel PCA (KPCA)[11], Locally Linear Embedding (LLE)[8], etc. Other technique like Multi-Dimensional Scaling(MDS) was successfully used in computational psychology for a very long time[1], which can be viewed as both "embedding" or "dimensionality reduction".

According to our survey, although those methods target at different problems and are derived from different assumptions, they do share a lot in common. The most significant sign is that, the core procedure involves eigenvalue decomposition or singular value decomposition, aka "spectral". They all involve an intermidiate step of embedding high-dimensional / non-Euclidean / non-metric points into a low-dimensional Euclidean space (although some do not embed explicitly). In this case, we categorize all these algorithms as Spectral Embedding Technique(SET).

1.1 A Sample Spectral Clustering Algorithm

There are many variations of SC. They all work under certain conditions and researchers don't have a rule of thumb so far. Before we analyze their procedure and justification, we present a simple but workable sample algorithm(**Alg 1**).

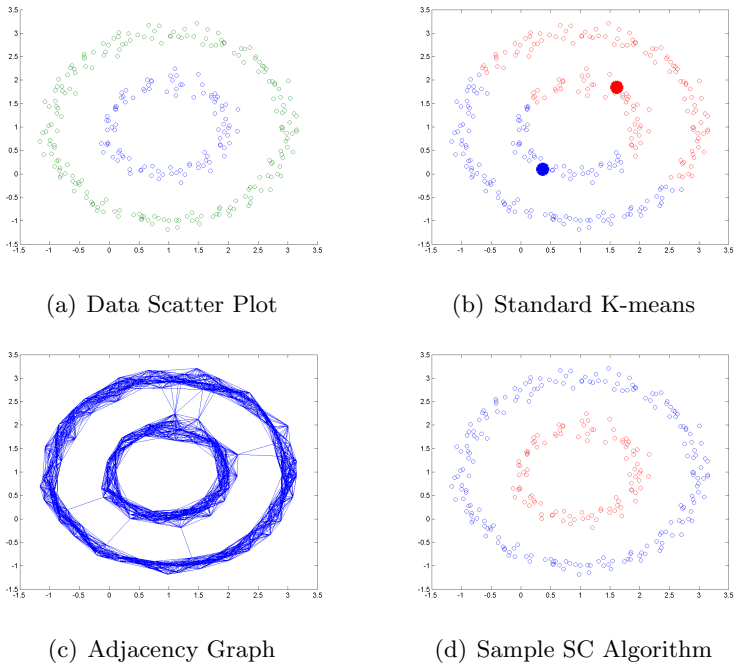
Algorithm 1 Sample Spectral Clustering

Input: Data matrix $X = [x_1, x_2, \ldots, x_N]$; Number of Clusters K.
Output: Clustering $\{C_i\}$: $C_i \in V$ and $\cap_i C_i = \emptyset$ and $\cup_i C_i = V$.

1: Form adjacency matrix A within ϵ-ball.
2: Solve $A = U\Lambda U^{\mathrm{T}}$, indexed according the eigenvalue's magnitude.
3: $Y \leftarrow$ first K columns of U.
4: Cluster Y's rows by K-means.

In **Alg 1**, the ϵ-ball adjacency graph is constructed as follows. First create one vertex for each data point. If for two points i, j satisfy $||x_i - x_j|| < \epsilon$, connect them with an edge. In this simple demonstration, we consider an unweighted graph, i.e. all entries of A are 0(disconnected) or 1(connected).

Fig 1 demonstrates the result of our sample SC algorithm, compared with standard K-means algorithm. **Fig 1(a)** shows the scatter plot of data. It is composed of one radius 1 circle and another radius 2 circle, both centered at (1,1). **Fig 1(b)** shows the result of standard K-means working on



(a) Data Scatter Plot (b) Standard K-means

(c) Adjacency Graph (d) Sample SC Algorithm

Figure 1: Demonstration of Sample SC Algorithm

Euclidean distance. **Fig 1(c)** shows the graph representation, where the adjacency graph is formed by taking a ϵ-ball and $\epsilon = 0.7$ in the example. **Fig 1(d)** shows the output of **Alg 1**. It's obvious that standard K-means algorithm can not correctly cluster the two circles. This is a known major weakness of K-means(in Euclidean): When clusters are not well separated spheres, it has difficulty recovering the underlying clusters. Although K-means works for this case if we transform the points into polar coordinate system(see [6] for code), the solution is not universal. However, in this example, our sample SC algorithm can separate the two clusters, probably because the eigenvectors of adjacency matrix convey adequate information.

A precaution is that **Alg 1** does not always work even in this simple case. Nor have we seen this algorithm from formally published works (so far), let alone justifications. This algorithm is only to show the flavour of spectral clustering and it contains those important steps in other more sophisticated algorithms. Readers are recommended to learn von Luxburg's tutorial[13] before reading the following sections. Since that paper is very detailed, we'll present overlapping topics concisely.

1.2 Linear Algebraic Properties

2 Spectral Clustering Framework

2.1 Metric Formulation

2.2 Spectral Embeding

2.3 Clustering

3 Spectral Clustering Justification

3.1 Random Walk

[von]

3.2 Normalized Cut

[shi]

3.3 Ratio Cut

[von]

3.4 Conductance

[von]

3.5 Matrix Perturbation

[andrew ng]

3.6 Low Rank Approximation

[matthew brand]

3.7 Density Estimation View

[mo chen, 2010]

3.8 Commute Time

[jihun ham, kernel]. view pseudo inverse of graph Laplacian by commute times on graphs.

3.9 Polarization

[m. brand] unifying view...

4 Other Spectral Like Embedding

4.1 MDS

4.2 isomap

4.3 Laplacian Eigenmap

4.4 Hessian Eigenmap

4.5 PCA

4.6 LLE

4.7 Kernel PCA

4.8 Kernel Framework

4.9 Graph Framework

5 Conclusion

Acknowledgements

References

[1] I. Borg and P.J.F. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer Verlag, 2005.

[2] M. Brand and K. Huang. A unifying theorem for spectral embedding and clustering. In *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, 2003.

[3] S.W. Hadley, B.L. Mark, and A. Vannelli. An efficient eigenvector approach for finding netlist partitions. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 11(7):885–892, 1992.

[4] B. Hendrickson and R. Leland. Multidimensional spectral load balancing. *Report SAND93-0074, Sandia National Laboratories, Albuquerque, NM*, ¡¡missing¿¿:¡¡missing¿¿, 1993.

[5] Pili Hu. Matrix calculus. GitHub, https://github.com/hupili/tutorial/tree/master/matrix-calculus, 3 2012. HU, Pili's tutorial collection.

[6] Pili Hu. Spectral techniques for community detection on 2-hop topology. GitHub, https://github.com/hupili/Spectral-2Hop, 4 2012. course project of CUHK/CSCI5160.

[7] Pili Hu. Tutorial collection. GitHub, https://github.com/hupili/tutorial, 3 2012. HU, Pili's tutorial collection.

[8] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[9] L.K. Saul and S.T. Roweis. An introduction to locally linear embedding. Technical report, NYU, 2000.

[10] L.K. Saul and S.T. Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *The Journal of Machine Learning Research*, 4:119–155, 2003.

[11] B. Schölkopf, A. Smola, and K.R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5):1299–1319, 1998.

[12] J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.

[13] U. Von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.

Appendix