An Iterated Graph Laplacian Approach for Ranking on Manifolds

Xueyuan Zhou
Department of Computer Science
University of Chicago
zhouxy@cs.uchicago.edu

Mikhail Belkin
Department of Computer Science
and Engineering
Ohio State University
mbelkin@cse.ohio-state.edu

Nathan Srebro
Toyota Technological Institute
at Chicago
nati@uchicago.edu

Graph Laplacian has been a key element in several machine learning areas, for instance, semi-supervised learning (SSL) (Zhu, 2005; Chapelle et al., 2006), dimensionality reduction (Belkin & Niyogi, 2003), and ranking in information retrieval (Zhou et al., 2004b). Several graph Laplacian SSL (Zhu et al., 2003; Zhou et al., 2004a; Belkin et al., 2004) and ranking algorithms (Zhou et al., 2004b) estimate a real-valued function on the whole data set in a transductive setting, which can be used as a density adaptive ranking score for all data points. One key advantage of this family of "diffusion" based ranking algorithms is that it is density adaptive, combining the local information encoded in the local neighborhood graph representation and the global information from the graph Laplacian into the final ranking score function.

Given the success of this family of methods in machine learning, particularly SSL and ranking, however, there is little functional analysis aspect study of these algorithms. One example is that in (Nadler et al., 2009), a striking finding for a family of popular diffusion based SSL methods is that the estimator on unlabeled points degenerate to constants in the limit of infinite unlabeled points while fixing labeled ones. The essential cause is that the solution space is so rich that solutions are guaranteed to be overfitting. This problem motivates the study of diffusion based ranking algorithms on manifolds from function analysis point of view. For instance, we can ask the following questions in ranking: what the ranking function space is, what the related null space is if it exists and how it might change the ranking function, and, what the ranking function will be in the limit of infinite sample points. These are not only interesting theoretical questions, but also closely related to the applications of these algorithms in practice. Running these algorithms on a huge amount of samples is approximately the same as working with infinite random samples.

In this paper, we will try to answer these questions. First, we study an diffusion based ranking method on manifolds (Zhou et al., 2004b) using functional analysis views. We show that due to the use of a symmetric normalized graph Laplacian, the method is sensitive to a commonly used parameter, which does not happen to the unnormalized graph Laplacian. This is different than the case in spectral clustering (von Luxburg et al., 2008), where normalized graph Laplacian is preferred. Second, we propose an iterated graph Laplacian approach for ranking on manifolds using unnormalized graph Laplacian. Compared to the existing method, our method enjoy several superior properties: more robust, more density adaptive, and well defined in the limit of infinite samples. We discuss a relation between the Green's function and the reproducing kernel based on graph Laplacian which connects the diffusion based ranking algorithms to kernel methods. At last, we test two interesting empirical graph Laplacian, from a family of two-step normalized graph Laplacian. To the best of knowledge, this is the first time they are used in ranking.

1 Topic: learning algorithms

2 Preference: poster

3 X. Zhou will present the paper

References

Belkin, M., Matveeva, I., & Niyogi, P. (2004). Regularization and semi-supervised learning on large graphs. *COLT* (pp. 624–638). Springer.

Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comp*, *15*, 1373–1396.

Chapelle, O., Schölkopf, B., & Zien, A. (Eds.). (2006). *Semi-supervised learning*. Cambridge, MA: MIT Press.

Nadler, B., Srebro, N., & Zhou, X. (2009). Semi-supervised learning with the graph laplacian: The limit of infinite unlabelled data. *Twenty-Third Annual Conference on Neural Information Processing Systems*.

von Luxburg, U., Belkin, M., & Bousquet, O. (2008). Consistency of spectral clustering. *Ann. Statist.*, *36*, 555–586.

Zhou, D., Bousquet, O., Lal, T. N., Weston, J., & Schölkopf, B. (2004a). Learning with local and global consistency. *Advances in Neural Information Processing Systems 16* (pp. 321–328). MIT Press.

Zhou, D., Weston, J., Gretton, A., Bousquet, O., & Schölkopf, B. (2004b). Ranking on data manifolds. *Advances in Neural Information Processing Systems (NIPS) 16* (pp. 169–176). MIT Press.

Zhu, X. (2005). *Semi-supervised learning with graphs*. Doctoral dissertation, CMU. CMU-LTI-05-192.

Zhu, X., Lafferty, J., & Ghahramani, Z. (2003). Semi-supervised learning using gaussian fields and harmonic function. *The Twentieth International Conference on Machine Learning*.