Nonlinear Label Spreading on Hypergraphs

Konstantin Prokopchik

Gran Sasso Science Institute, L'Aquila (Italy)

`konstantin.prokopchik@gssi.it`

Semi-supervised learning is the problem of finding clusters in a graph or a point-clould dataset where we are given "few" initial input labels. Label Spreading (LS) is a standard technique for this problem, which can be interpreted as a random walk diffusion process of the labels on the graph [1]. Recent work in network science has shown that in many graph algorithms a great advantage can be obtained when accounting directly for higher-order interactions modeled as hypergraphs, where, for example, a hyperedge directly connects all the authors of a paper in a co-authorship network [2]. In this work we propose a variation of LS for hypergraphs where the standard random walk Laplacian matrix is replaced by a nonlinear Laplacian-inspired map which is defined in terms of the hypergraph incidence matrix [3]. We show the convergence of the new nonlinear diffusion process to the global minimum of a regularized objective function that aims at reducing the variance across the hyperedge nodes and we demonstrate the efficiency of our approach on a variety of point cloud and network datasets.

Joint work with A. Benson *and* F. Tudisco

References

[1] D. Zhou and O. Bousquet and T. N Lal, and J. Weston and B.Schölkopf , *Learning with local and global consistency*, NeurIPS, 2004, pp. 321–328.

[2] A. R Benson and D. F Gleich and J Lescovec, *Higher-order organization of complex networks*, Science, (353) 2016, pp. 163–166.

[3] F. Tudisco and A. R Benson and K. Prokopchik, *Nonlinear higher-order label spreading*, Proceedings of the Web Conference, 2021.