Sparse Subspace Clustering for the Popularity Adjusted Block Model

John Koo, Minh Tang, Michael Trosset

Abstract

TODO

1 Introduction

1.1 Notation

P denotes the edge probability matrix for the PABM. $A_{ij} \overset{\text{indep}}{\sim} Bernoulli(P_{ij})$ for $i > j$, and $A_{ji} = A_{ij}, A_{ii} = 0 \; \forall i, j \leq n$ to make A the edge weight matrix for a hollow, unweighted, and undirected graph. X is an ASE of A while Y is constructed using the popularity vectors $\{\lambda^{(kl)}\}_K$ and the projection matrix Π, which is an ASE of P. $Z = XQ - Y$ for $Q = \arg \min_{Q \in \mathbb{O}(p,q)} ||XQ - Y||_F$. Let $x_i^\top, y_i^\top, z_i^\top$ be the rows of X, Y, Z. $X^{(n)}$ represents the full X matrix for a sample of size n. $X^{(n,k)}$ represents the k^{th} block of $X^{(n)}$.

2 Main Results

Theorem 1. The subspace detection property holds for Y with probability at least $1 - \sum_k^K n_k e^{-\sqrt{K(n_k - 1}} - K^{-2} \sum_{k \neq l} \frac{4e^{-2t}}{(n_k+1)n_l}$.

This falls out of Theorem 2.8 from Soltanolkotabi and Candés [2]. The subspaces in Y are orthogonal, so $\text{aff}(S_k, S_l) = 0 \; \forall k, l \leq K$.

Property 2. By Rubin-Delanchy et al. [1], $\max_i ||Q_n x_i^{(n)} - y_i^{(n)}|| = O_P\Big(\frac{(\log n)^c}{n^{1/2}}\Big)$. Then $Z^{(n)} \to 0$, $\delta^{(n)} \to 0$, and $r(X^{(n,l)} Q^{(n,l)}) \to r(Y^{(n,l)})$. Here we assume $r(Y^{(n,l)}) > 0 \; \forall n > 3$ and $l \leq K$.

Property 3. $P(\mu(Y^{(k)}) = 0) \geq 1 - \frac{1}{K} \sum_{k \neq l} \frac{4e^{-2t}}{(n_k+1)n_l}$ [2].

Theorem 4. ***TODO*** Find property 3 for X or XQ (?).

Theorem 5. The subspace detection property holds for $X^{(n)} Q^{(n)}$ as $n \to \infty$.

Theorem 6. $\exists M \in \mathbb{N}$ such that $\mu(X^{(n,l)}) < r(X^{(n,l)}) \; \forall n \geq M$.

Corollary. If $n \geq M$, $\exists \lambda > 0$ such that the LASSO subspace detection property holds for $X^{(n)}$ with parameter λ.

This falls out of Theorem 6 of Wang and Xu [3] and Theorem 6 of this paper.

Theorem 7. ***TODO*** Find the probability for which the subspace detection property holds for $X^{(n)}$ or $X^{(n)}Q^{(n)}$.

References

[1] Patrick Rubin-Delanchy, Joshua Cape, Minh Tang, and Carey E. Priebe. A statistical interpretation of spectral embedding: the generalised random dot product graph, 2017.

[2] Mahdi Soltanolkotabi and Emmanuel J. Candés. A geometric analysis of subspace clustering with outliers. *Ann. Statist.*, 40(4):2195–2238, 08 2012. doi: 10.1214/12-AOS1034. URL https://doi.org/10.1214/12-AOS1034.

[3] Yu-Xiang Wang and Huan Xu. Noisy sparse subspace clustering. In Sanjoy Dasgupta and David McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 89–97, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR. URL http://proceedings.mlr.press/v28/wang13.html.