On the Theoretical Properties of the Network Jackknife

Qiaohui Lin, Robert Lunde, and Purnamrita Sarkar

University of Texas at Austin

ICML 2020

The Problem

- Let $A^{(n)}$ be a $n \times n$ adjacency matrix
- Consider some network functional $Z_n = g(A^{(n)})$.
- Is there a way to quantify uncertainty for general g?
- We show that the network jackknife has favorable properties under the sparse graphon model
  1. Always conservative in expectation
  2. Consistent for counts and smooth functions of counts

Sparse Graphon Model

Consider $\{A_n\}_{n\in\mathbb{N}}$ generated by the following model (Bickel and Chen, 2009):

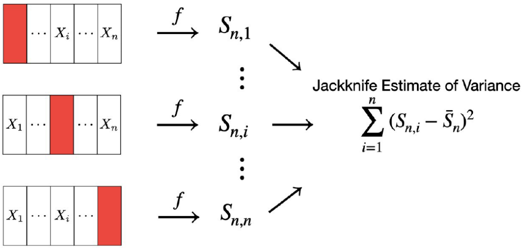
$$\begin{aligned} A_{ij}^{(n)} = A_{ji}^{(n)} &= \mathbb{1}(\eta_{ij} \leq \rho_n w(\xi_i, \xi_j) \wedge 1) \\ &\stackrel{d}{=} \text{Bernoulli}(\rho_n w(\xi_i, \xi_j) \wedge 1) \end{aligned}$$

where $\eta_{ij} \sim \text{Unif}[0,1]$ for $1 \leq i < j \leq n$, $\xi_1, \ldots, \xi_n \sim \text{Unif}[0,1]$ $\rho_n \to 0$, $E[w(\xi_i, \xi_j)] = 1$.

- Natural model for graphs that exhibit vertex exchangeability
- Subsumes many other common network models (i.e. Stochastic block models, random dot product graphs)

The (I.I.D.) Jackknife

- Developed by Quenouille (1956) and Tukey (1958)
- Let $X_1, \ldots X_n \sim P$, $S_n = f(X_1, \ldots, X_n)$
- Let $S_{n,i}$ denote functional with X_i left out, $\bar{S}_n = \frac{1}{n} \sum_{i=1}^{n} S_{n,i}$



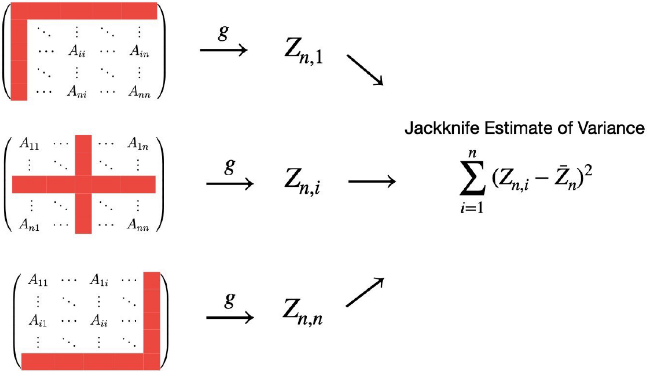
- Jackknife estimate of the variance:

$$\widehat{\mathrm{Var}}_{JACK}\ S_{n-1} = \sum_{i=1}^{n} (S_{n,i} - \bar{S}_n)^2$$

- Jackknife consistent for smooth functions (see e.g. Shao and Tu (1995)), but generally regularity conditions stronger than bootstrap
- Efron-Stein inequality (1979):

$$\mathrm{Var}\ S_{n-1} \leq E(\widehat{\mathrm{Var}}_{JACK}\ S_{n-1})$$

- Under the sparse graphon model, analog of "leave-one-out" is "leave-node out." Let $Z_{n,i}$ denote network functional with node i left out.
- This suggests the following network jackknife, first proposed by Frank and Snijders (1994):



Network Efron-Stein Inequality

Theorem 1 (Network Efron-Stein Inequality)

For any functional invariant to node-permutation,

$$\mathrm{Var}\ Z_{n-1} \leq E(\widehat{\mathrm{Var}}_{JACK}\ Z_{n-1})$$

- Intuition: Z_{n-1} may be viewed as function of independent random variables $\{\eta_{ij}\}_{1\leq i<j\leq n}$ and $\{\xi_i\}_{1\leq i\leq n}$.
- Use martingale difference techniques (Rhee and Talagrand, 1986) with appropriate filtration (Borgs et al., 2008)

Consistency for Counts

Let R be subgraph with p nodes and e edges. Let $G[S]$ denote subgraph induced by nodes of S. Consider the following functional, introduced in Bickel et al. (2011):

$$\hat{P}(R) = \rho_n^{-e} \frac{1}{\binom{n}{p}|Iso(R)|} \sum_{S \sim R} \mathbb{1}(S = G_n[S])$$

Theorem 2 (Consistency for Counts)

Suppose R is acyclic or a p-cycle. Then if $n\rho_n \to \infty$,

$$n \cdot \widehat{\mathrm{Var}}_{JACK}\ \hat{P}(R) \xrightarrow{P} \sigma^2$$

where $\sigma^2 = \lim_{n\to\infty} n \cdot \mathrm{Var}\ \hat{P}(R)$.

Consistency for Counts

Theorem 2 (Consistency for Counts)

Suppose R is acyclic or a p-cycle. Then if $n\rho_n \to \infty$,

$$n \cdot \widehat{\mathrm{Var}}_{JACK}\ \hat{P}(R) \xrightarrow{P} \sigma^2$$

where $\sigma^2 = \lim_{n\to\infty} n \cdot \mathrm{Var}\ \hat{P}(R)$.

- Intuition: $Z_{n,i} = \underbrace{Z_{n,i} - E(Z_{n,i} \mid \boldsymbol{\xi}_n)}_{A_i \text{ Bernoulli Noise}} + \underbrace{E(Z_{n,i} \mid \boldsymbol{\xi}_n)}_{B_i \text{ U-statistic}}$
- We show that:

$$n \cdot \widehat{\mathrm{Var}}_{JACK}\ Z_{n-1} = n \cdot \sum_{i=1}^{n} (B_i - \bar{B}_n)^2 + o_P(1)$$

- Jackknife is consistent for U-statistics (Arvesen, 1969).

Let $(R_1, \ldots, R_d)$ are p-cycles or acyclic graphs, $f(G_n)$ be a function of the vector $(\hat{P}(R_1), \ldots, \hat{P}(R_d))$. Let ∇f denote the gradient of f and $\mu \in \mathbb{R}^d$ the limit of $(\tilde{P}(R_1), \ldots, \tilde{P}(R_d))$ as $n \to \infty$; it turns out that μ corresponds to an integral parameter of the graphon related to the edge structure of the subgraph. We have the following result.

Theorem 3 (Consistency for Smooth Functions of Counts)

Suppose that $(R_1, \ldots, R_d)$ are p-cycles or acyclic graphs and $n\rho_n \to \infty$. Let $e^* = \max\{|E(R_1), \ldots E(R_d)\}$ and suppose that $\int_0^1 \int_0^1 w^{2e^*}(u,v)\, du\, dv < \infty$. Furthermore, suppose that ∇f exists in a neighborhood of μ, $\nabla f(\mu) \neq 0$, and that ∇f is continuous at μ. Let σ_f^2 denote the asymptotic variance of $\sqrt{n}[f(G_n) - f(E(G_n))]$. Then,

$$n \cdot \widehat{\mathrm{Var}}_{JACK}\ f(G_n) \xrightarrow{P} \sigma_f^2$$

- Network jackknife gives a variance bound for general functions of sparse graphons.
- Network bootstraps (Green and Shalizi, 2017; Levin and Levina, 2019) have only been shown to be consistent for counts.
- Subsampling (Bhattacharyya and Bickel, 2015; Lunde and Sarkar, 2019) valid under more general conditions than bootstrap, but requires weak convergence.

Simulation Results

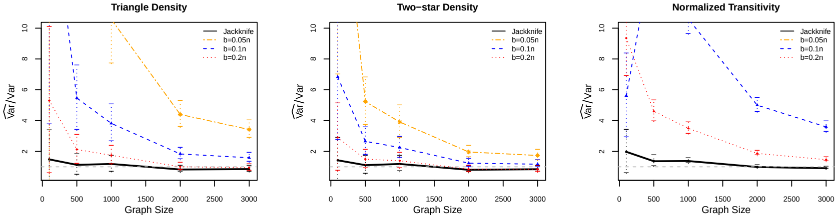


Figure: Results for a stochastic block model

Simulation Results

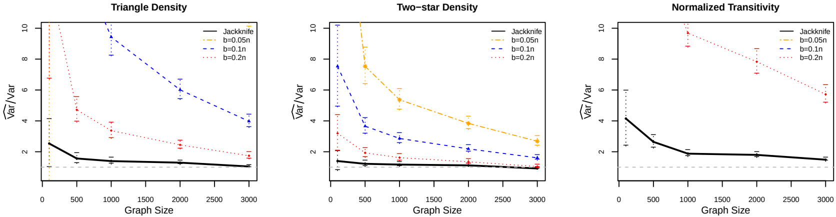


Figure: Results for a continuous graphon model

Real Data Results

We look at three college pairs: Berkeley and Stanford, Yale and Princeton, Harvard and MIT from Facebook Network Data Rossi and Ahmed (2015).

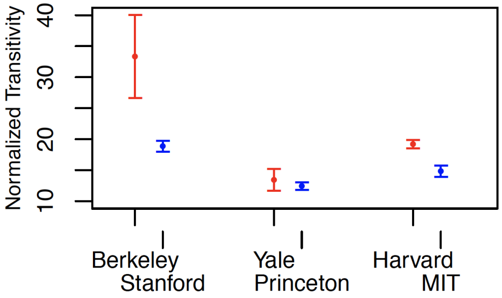


Figure: Results for Facebook college network

Figure 4 presents 97.5% CI's for normalized transitivity for each college. Thus, two disjoint CI's are equivalent to rejecting a level 0.05 test. Figure above shows that transitivity can in fact separate Berkeley and Stanford Facebook networks, as well as Harvard and MIT Facebook networks, giving us interesting information about the inherent differences between the network structures of these colleges.

- We establish a Network Efron-Stein inequality.
- We show consistency for counts and smooth functions of counts.
- In our simulation study, the jackknife exhibits better finite sample performance than subsampling.

Arvesen, J. N. (1969). Jackknifing U-statistics. *The Annals of Mathematical Statistics*, 40(6):2076–2100.

Bhattacharyya, S. and Bickel, P. J. (2015). Subsampling bootstrap of count features of networks. *Annals of Statistics*, 43:2384–2411.

Bickel, P. J. and Chen, A. (2009). A nonparametric view of network models and Newman-Girvan and other modularities. *Proceedings of the National Academy of Sciences (USA)*, 106:21068–21073.

Bickel, P. J., Chen, A., and Levina, E. (2011). The method of moments and degree distributions for network models. *Annals of Statistics*, 39:38–59.

Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T., and Vesztergombi, K. (2008). Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801 – 1851.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7(3):1–26.

Frank, O. and Snijders, T. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, 10(1):53 – 67.

Green, A. and Shalizi, C. (2017). Bootstrapping Exchangeable Random Graphs. *arXiv e-prints*, page arXiv:1711.00813.

Levin, K. and Levina, E. (2019). Bootstrapping networks with latent space structure. *ArXiv e-prints*.

Lunde, R. and Sarkar, P. (2019). Subsampling sparse graphons under minimal assumptions. *ArXiv e-prints*.

References III

Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika*, pages 353–360.

Rhee, W. T. and Talagrand, M. (1986). Martingale inequalities and the jackknife estimate of variance. *Statistics & Probability Letters*, 4(1):5– 6.

Rossi, R. A. and Ahmed, N. K. (2015). The network data repository with interactive graph analytics and visualization. In *AAAI*.

Shao, J. and Tu, D. (1995). *The Jackknife and the Bootstrap*. Springer.

Tukey, J. W. (1958). Bias and confidence in not quite large samples (abstract). *The Annals of Mathematical Statistics*, 29:614.