An Iterative Step-Function Estimator for Graphons

Diana Cai[†,‡], Nathanael Ackerman[‡], Cameron Freer[†,⋆]

[†]Gamalon Labs, [‡]Harvard University, [⋆]Massachusetts Institute of Technology

http://tinyurl.com/graphon

Abstract

Exchangeable graphs arise via a sampling procedure from the measurable functions known as *graphons*. A natural estimation problem is how well we can recover a graphon given a single graph sampled from it. One general framework for estimating a graphon uses step-functions obtained by partitioning the nodes of the graph according to some clustering algorithm. We propose an *iterative step-function estimator* (ISFE) that, given an initial partition, iteratively clusters nodes based on their edge densities with respect to the previous iteration's partition. Each round of the algorithm can be seen as attempting to reduce the cut metric between the original graphon and the step-function induced by a given partition. We demonstrate ISFE's performance with respect to various clustering algorithms, and in comparison with other graphon estimation techniques.

Background

A random graph G is **exchangeable** when its distribution is invariant under arbitrary permutations of its vertices.

Graphons are symmetric, measurable functions

$$W(x,y) : [0,1]^2 \to [0,1]. \tag{1}$$

Exchangeable random graphs can be sampled from graphons as follows:

$$\begin{aligned} U_i &\overset{\text{iid}}{\sim} \text{Uniform}[0,1] \\ G_{ij} \mid U_i, U_j &\overset{\text{ind}}{\sim} \text{Bernoulli}(W(U_i,U_j)) \text{ for } i<j \end{aligned} \tag{2}$$

Every exchangeable graph arises from a mixture of such sampling procedures (Aldous, 1981; Hoover, 1979), and this mixture is unique up to measure preserving transformations.

Graphon Estimation

The "graphon value estimation problem" aims to invert the second step of the sampling procedure, and hence can be thought of as finding the local underlying structure of a graph sampled from a graphon (without concluding anything about the graphon at any location that wasn't involved in the sample). Suppose we have sampled the W-random graph $\mathbb{G}(n,W)$ using $\{U_i\}_{i\in[n]}$ as in Equation (2). Graphon value estimation consists of giving an estimator $\widehat{M} := (\widehat{M}_{ij})_{i,j\in[n]}$ for the matrix $M := (M_{ij})_{i,j\in[n]}$ where each $M_{ij} := W(U_i,U_j)$. One measure of success for the graphon value estimation problem is given by the mean squared error:

$$\text{MSE}(\widehat{M}) := \mathbb{E}\left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(M_{ij}-\widehat{M}_{ij})^2\right). \tag{3}$$

Step-Function Approximation

A graphon U is called a **step-function** when there is a partition $\mathcal{S} = \{S_0,\dots,S_{k-1}\}$ of $[0,1]$ into finitely many pieces, such that U is constant on each set $S_i \times S_j$.

For a vertex-weighted, edge-weighted graph H with vertex set $[n]$, with vertex weights α_i and edge-weights β_{ij} for $i,j\in[n]$. The **step-function graphon** W_H associated with H is $W_H(x,y) = \beta_{ij}$ for $x\in J_i$ and $y\in J_j$, where the steps $J_0,\dots,J_{n-1}$ are a partition of $[0,1]$ into consecutive intervals of size $\frac{\alpha_i}{\sum_{j\in[n]}\alpha_j}$ for $i\in[n]$. For an unweighted finite graph G, use the weighted graph H with vertex weights $\alpha_i = 1$ and edge weights $\beta_{ij} = G_{ij}$.

Given a graph G on $[n]$ and vertex sets $X,Y\subseteq[n]$, write $c_G(X,Y) := \sum_{i\in X}\sum_{j\in Y} G_{ij}$ for the number of edges across the cut X,Y. Then the **edge density** in G between X and Y is defined to be

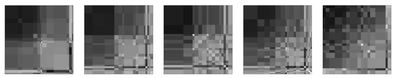
$$e_G(X,Y) := \frac{c_G(X,Y)}{|X|\,|Y|}; \tag{4}$$

when X and Y are disjoint, this quantity is the fraction of possible edges between X and Y that G contains.

Now suppose G is a graph on $[n]$, and $\mathcal{P} = \{P_0,\dots,P_{k-1}\}$ is a partition of the vertices of G into k classes. The **quotient graph** $G/\mathcal{P}$ is defined to be the weighted graph on $[k]$ with respective vertex weights $|P_i|/n$ and edge weights $\frac{e_G(P_i,P_j)}{|P_i|\,|P_j|}$. In our step-function estimator, we will routinely pass from a sampled graph G and a partition $\mathcal{P}$ of its vertex set to the graphon $W_{G/\mathcal{P}}$ formed from the quotient $G/\mathcal{P}$.

Iterative Step-Function Estimation

Given a partition $\mathcal{P}$ of a finite graph G (e.g., from k-means, HAC, random assignment, or trivially clustering all vertices in a single class), one may form the step-function $W_{G/\mathcal{P}}$. In Figure 1, we display the result of estimating a graphon according to several particular choices of clustering algorithms.



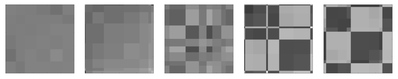
ISFE k-means HAC Affinity Spectral

Fig. 1: Graphons obtained from various clusterings. ISFE was applied to the trivial partition.

Iterative Step-Function Estimation (ISFE): Given a finite graph G, consider the following graphon estimator procedure:

1. Partition the vertices of G according to some clustering algorithm.
2. Improve this partition by iteratively running the following algorithm for $T\geq 0$ iterations.
   (a) **Input**: graph G, initial partition $\mathcal{P}^{(\text{old})}$, minimum number of classes ℓ, decay d
   (b) **Output**: new partition $\mathcal{P}^{(\text{new})}$
   (c) Initialize $Q = \{Q_0\}, Q_0 = \{0\}, c_0 = 0, \epsilon = 1, B = 1$.
   (d) While number of classes $B < \ell$:
      i. For vertices $i = 0,\dots,n-1$:
         A. Compute edge densities vector $e_i := \left\{e_G(P_j,\{i\})\right\}_j$ for all $j = 0,\dots,|\mathcal{P}^{(\text{old})}|-1$.
         B. Compute min distance between vector e_i and some vector e_{c_j}, where $j = 0,\dots,|\mathcal{P}^{(\text{old})}|-1$:
            $$j^* := \arg\min_j d_j(e_i,e_{c_j}) = \arg\min_j \sum |e(\{i\},\mathcal{P}_r^{(\text{old})}) - e(\{c_j\},\mathcal{P}_r^{(\text{old})})|$$
         C. If min distance $d_{j^*}(e_i,e_{c_j}) < \epsilon$: Add vertex i to existing class Q_{j^*}.
         D. Else: Create new class Q_B with $Q_B = \{i\}$ and centroid $c_B = i$; Add class Q_B to partition Q; $B \leftarrow B+1$
      ii. Decay epsilon $\epsilon \leftarrow \epsilon\cdot d$
   (e) $\mathcal{P}^{(\text{new})} \leftarrow Q$
3. Report the step-function graphon $W_{G/\mathcal{P}}$, where $\mathcal{P}$ is the final partition produced, with its classes sorted according to their average edge densities. For graphon value estimation, $\widehat{M}_{ij}$ is the edge density between the classes containing i and j.

In Figure 2, we show 4 iterations of ISFE on a sample from a stochastic block model.



$T=0$ $T=1$ $T=2$ $T=3$ $T=4$

Fig. 2: Iterations of ISFE on a 200 vertex sample from a SBM graphon with $p=0.5, q=(0.7,0.3)$, beginning with a random partition into 6 classes.

Examples

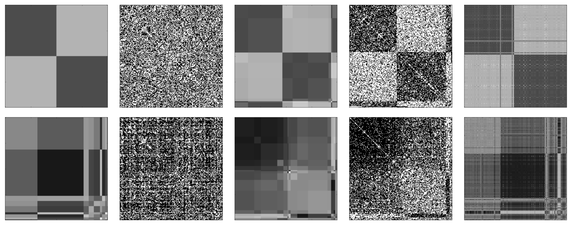


Fig. 3: Examples of graphon estimation using ISFE. Columns: (1) the original graphon; (2) a 200 vertex random sample from the original graphon; (3) ISFE-estimated graphon for $\ell = 8$ (SBM), 15 (IRM), $T=4$; (4) the random sample reordered according to the ISFE estimate; (5) ISFE estimate rearranged by increasing U_i. ISFE was applied to the trivial partition, i.e., all vertices were initially in a single class.

Stochastic block model graphon. In the stochastic block model, we assume there are k classes. We define the SBM graphon for $k=2$ as follows: given parameters $p\in[0,1], q = (q_0,q_1)\in[0,1]^2$, we partition $[0,1]$ into two pieces P_0,P_1 of length p and $1-p$, where $p\in[0,1]$. The value of the graphon is constant on $P_i\times P_i$ with value q_0 and constant on $P_i\times P_{1-i}$ with value q_1, for $i=0,1$. We show the result of ISFE on a graph sampled from an SBM graphon in the top row of Figure 3 for $p=0.5, q=(0.7,0.3)$.

Infinite relational model graphon. The infinite relational model (IRM) is a non-parametric extension of the SBM, where the (infinite) partition is generated by a Chinese restaurant process with concentration parameter α. For each class of the partition, the graphon is constant with value sampled from a beta distribution with parameters a,b. We show the result of ISFE on a graph sampled from an IRM graphon with $\alpha=3, a=3, b=2.9$ in the bottom row of Figure 3.

Results

Synthetic Data

We examine synthetic data sampled from several graphons: (1) a gradient given by the function $W(x,y) = \frac{(1-x)+(1-y)}{2}$; (2) an SBM graphon with $p=0.5, q=(0.7,0.3)$; (3) an SBM graphon with $p=0.3, q=(0.7,0.3)$; and (4) an IRM graphon with $\alpha=3, a=3, b=2.9$. In Figure 4, we display the results of ISFE, SAS (Airoldi and Chan, 2014) and USVT (Chatterjee, 2014) on a 200 vertex sample from each of the graphons. The original graphon is displayed in column 1. We evaluate all estimators using the MSE given in Equation 3, where the estimator has rearranged by increasing U_i.

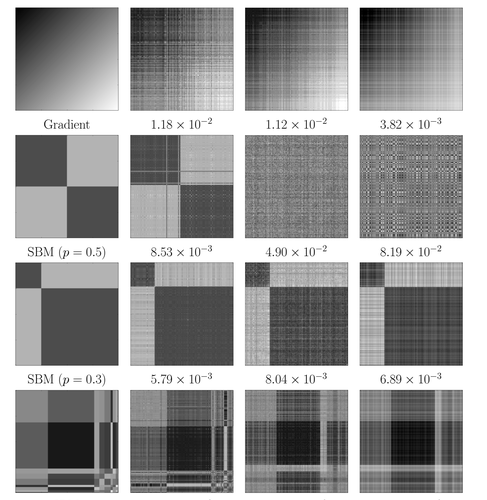


Fig. 4: Comparison of graphon estimation methods, ISFE (column 2), SAS (column 3), USVT (column 4), along with MSE.

Real Data

We demonstrate ISFE on three real datasets (Figure 5). For each dataset, we take a subset of nodes consisting of the top K highest-degree vertices and the edges between them.

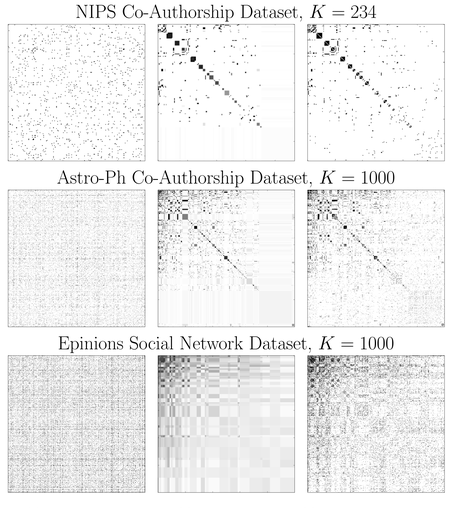


Fig. 5: Columns: (1) The K-vertex subgraph, (2) result of ISFE, and (3) sample rearranged according to ISFE.