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
Exchangable 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. Our general framework for estimating a graphon uses step-function estimators 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 an attempt 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
$$W(x, y) : [0, 1]^2 \rightarrow [0, 1]$$
(1)

Exchangable random graphs can be sampled from graphons as follows:
$$U_i \sim \text{Uniform}[0,1]$$
$$G_{ij} \mid U_i, U_j \sim \text{Bernoulli}(W(U_i, U_j)) \quad \text{for } i < j$$
(2)

Every exchangable graph arises from a mixture of such sampling processes (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 $G$ in (W) using $(U_i)_{i \in [n]}$ as in Equation (2). Graphon value estimation consists of estimating a $\hat{M} = (\hat{M}_{ij})_{i,j}$ for the matrix $M = (M_{ij})_{i,j}$. 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
$$\text{MSE}(\hat{M}) = \mathbf{E} \left[ \sum_{i,j} (M_{ij} - \hat{M}_{ij})^2 \right]$$
(3)

Step-Function Approximation
A graphon $U$ is called a step-function when there is a partition $S = \{S_0, \ldots, S_N\}$ 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 $a_i$ and edge weights $b_{ij}$ for $i, j \in [n]$, the step-function graphon $W_{G,H}$ associated with $H$ is $W_{G,H}(x,y) = \beta_{ij}$ for $x \in S_i$ and $y \in S_j$, where the steps $eta_{ij}$ are a partition of $[0, 1]$ into consecutive intervals of size $\frac{\text{area}(S_i)}{\text{area}(S_j)}$, for $i, j \in [N]$ for an infinite graphon $G$ used the graph $H$ with vertex weights $a_{ij} = 1$ and edge weights $b_{ij} = G_{ij}$.

Given a graph on $[n]$ and vertex sets $X, Y \subseteq [n]$, write $e_{G}(X,Y) = \sum_{x,y} G_{xy}$ for the number of edges across the cut $X, Y$. Then the edge density in $G$ between $X$ and $Y$ is defined to be
$$e_{G}(X,Y) \mid \text{dist}(X,Y)$$
(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 $\mathcal{P} = \{P_0, \ldots, P_{k-1}\}$ 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 $\sum_{x \in P_i} a_x$ and edge weights $\sum_{i,j} b_{ij}$. 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 (ISFE)
Given a finite graph $G$, consider the following graphon estimation 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 $G$ and a partition $\mathcal{P}_{(t)}$, minimum number of classes $K$, decay $\alpha$.
   (b) Output: new partition $\mathcal{P}_{(t+1)}$.
      (i) Initialize $Q = (0,1)$. $Q = (\alpha, 1, \beta)$.  
      (ii) While number of classes $K < T$:
        (A) Compute edge density vector $e = (e_{ij})_i$, for all $i,j \in [k]$. 
        (B) Compute min-distance between vector $e$ and some vector $e^*$, where $e^* = (e_{ij}^*)_i, j$. 
        (C) Compute min-distance $d(e, e^*)$ = $\sum_{i,j} |e_{ij} - e^*_{ij}|^2$.
        (D) Add vertex $v$ to existing class $Q_r$.
        (E) Break. Create new class $Q_r + 1$ with $G_{ij} = 0$, and centroid $\bar{w} = \text{Acc}(G_{ij})$ partition $Q_{r+1}$.
      4. Decay $\alpha$ until $K < T$.

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

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

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.

Examples

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 4-means HAC Affinity Spectral
Fig 1. Graphons of data from various datasets. ISFE was applied to the trivial partition.

Iterative Step-Function Estimation
Fig 3. Comparison of graphon estimation methods. (HAC, column 2); SAS (column 3); USMT (column 4); along with ISFE.

Results

Fig 4. Examples of graphon estimation using ISFE. Column 1: the original graphons (2) 200 vertex samples from the original graphons (3) ISFE estimated graphons for $T = 4$ (4) ISFE + annealing (5) ISFE annealing restricted to increasing $Q$. 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 $P_1 \times P_1$ with value $q_0$ and constant on $P_0 \times P_0$ with value $q_1$. 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 $\alpha$. For each class of the partition, the partition 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 $\alpha = 3, \beta = 2.9$ in the bottom row of Figure 3.