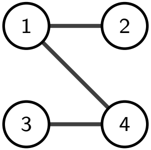
How Distribution & Continuity Assumptions Effect Graphon Approximation

Intern: Taj Patel, Supervisor: Amiee Rice

September 4, 2020

What is a Network/Graph?

- A Graph consists of a set of points (referred to as Nodes) and a set of edges which tells us which nodes should be connected.



Adjacency Matrix

If we have a graph that has n nodes, then we can assign it an nxn matrix. Assign each node a value from { 1,2,....n }. Construct a matrix A st for $i, j \in \{1, 2, ...n\}$

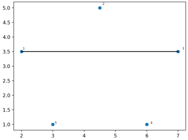
$$A_{i,j} = \begin{cases} 1 & \text{Nodes i and j are connected} \\ 0 & else \end{cases}$$

Erdos Renyi Model

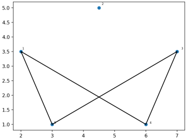
- The Erdos Renyi model essentially constructs a random graph given a set of vertices. It does this by joining edges based off simulations of bernouilli random variables. ie. $G(n, p)$ model states (n vertices given) and there is an edge between any 2 vertices (u, v) iff $Bernouilli(p) = 1$.
- The Erdos Renyi model gives us a way of generating sequences of graphs. $e.g$ can define $R_n := G(n, 0.5)$

Erdos Renyi Model Cont...

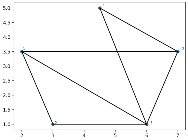
Figure: $G(5, p)$, edge formation given by $Bernoulli(p)$ random variables



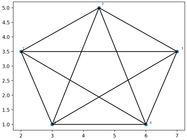
((a)) $G(5, 0.1)$



((b)) $G(5, 0.3)$



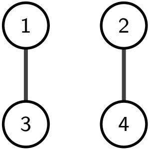
((c)) $G(5, 0.6)$

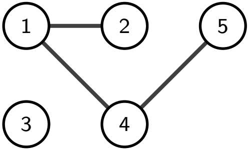


((d)) $G(5, 1)$

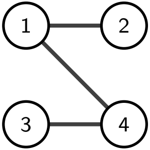
Convergence

- Just like sequences of numbers, we can think of a sequence of graphs also converging.
- How do we define the distance between 2 graphs?

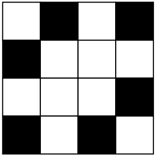




Pixel Pictures



$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$



Cut Metric

- So from a graph, we can construct a function $g : [0,1]^2 \longrightarrow [0,1]$.
- Hence if we have a sequence of graphs, we also have a sequence of maps.

Thus we define the distance between 2 graphs (G, W) by:

$$\delta(G, W) := inf_{\sigma \in \mathcal{M}} \left(\int_0^1 \int_0^1 |g(\sigma(x), \sigma(y)) - w(x, y)|^2 dxdy \right)$$

Where $\mathcal{M}$ denotes the set of all measuring preserving maps from [0,1] to [0,1], g and w are the maps corresponding to the pixel pictures of the graphs G and W.

Graphons

Definition(Graphon): $f : [0,1]^2 \longrightarrow [0,1]$, be a symmetric, Lebesgue-Measurable function.

$G(n, f) Model$

This model starts by being given n nodes, in some space. Label the nodes from 1...m.

1. For each $i \in 1, ...n$, pick $u_i \sim U[0, 1]$ to correspond to node i.
2. To determine edge formation between nodes i and j, Simulate a $Bernouilli(f(u_i, u_j))$ random variable.
3. Similarly to the Erdos Renyi model, if this random variable is 1, then form an edge between the 2 nodes, else don't.
4. Repeat 2 and 3 for all pairs of nodes, to then generate a graph.

Key Result

Proposition: Let f be a graphon, and for every n $\geq$1, let R_n be an instance of $G(n, f)$. Then with probability 1, $(R_n)_{n\geq 1}$ converges to f in the cut distance. [1]

Little Summary

- A Graph is just a set points, where some are connected and these connections are called edges.

Little Summary

- A Graph is just a set points, where some are connected and these connections are called edges.
- A Graphon is a symmetric Lebesgue measurable function from $[0,1]^2$ to $[0,1]$ and is often considered as the limit of graphs.

Little Summary

- A Graph is just a set points, where some are connected and these connections are called edges.
- A Graphon is a symmetric Lebesgue measurable function from $[0,1]^2$ to $[0,1]$ and is often considered as the limit of graphs.
- If f is a graphon, then the sequence of graphs generated by $G(n,f)$ converge in the cut metric to f.

Little Summary

- A Graph is just a set points, where some are connected and these connections are called edges.
- A Graphon is a symmetric Lebesgue measurable function from $[0,1]^2$ to $[0,1]$ and is often considered as the limit of graphs.
- If f is a graphon, then the sequence of graphs generated by $G(n,f)$ converge in the cut metric to f.
- Suppose $x, y \in [0,1]$, then we can think of f(x,y) as the probability that two nodes that have been assigned values x and y are connected.

Approximation of a Graphon

- Let $(R_n)_{n\geq 1}$ be a convergent sequence of graphs, with corresponding adjacency matrices $(A^{(n)})_{n\geq 1}$.
- For graph R_n, construct an estimator $\hat{f}_n$ using the following:

Approximation of a Graphon

- Let $(R_n)_{n\geq 1}$ be a convergent sequence of graphs, with corresponding adjacency matrices $(A^{(n)})_{n\geq 1}$.
- For graph R_n, construct an estimator $\hat{f}_n$ using the following:

1. Assign each node a value **uniformly** from $[0, 1]$.

Approximation of a Graphon

- Let $(R_n)_{n\geq 1}$ be a convergent sequence of graphs, with corresponding adjacency matrices $(A^{(n)})_{n\geq 1}$.
- For graph R_n, construct an estimator $\hat{f}_n$ using the following:

1. Assign each node a value **uniformly** from $[0, 1]$.
2. Divide $[0, 1]^2$ into smaller rectangles.

Approximation of a Graphon

- Let $(R_n)_{n\geq 1}$ be a convergent sequence of graphs, with corresponding adjacency matrices $(A^{(n)})_{n\geq 1}$.
- For graph R_n, construct an estimator $\hat{f}_n$ using the following:

1. Assign each node a value **uniformly** from $[0, 1]$.
2. Divide $[0, 1]^2$ into smaller rectangles.
3. Plot the cartesian product of the uniform values minus the points that lie on y=x. Each coorindate corresponds to a potential connection in our graph.

Approximation of a Graphon

- Let $(R_n)_{n\geq 1}$ be a convergent sequence of graphs, with corresponding adjacency matrices $(A^{(n)})_{n\geq 1}$.
- For graph R_n, construct an estimator $\hat{f}_n$ using the following:

1. Assign each node a value **uniformly** from $[0,1]$.
2. Divide $[0,1]^2$ into smaller rectangles.
3. Plot the cartesian product of the uniform values minus the points that lie on y=x. Each coorindate corresponds to a potential connection in our graph.
4. The value of our estimator is a scalar multiple of the proportion of points connected within each rectangle.

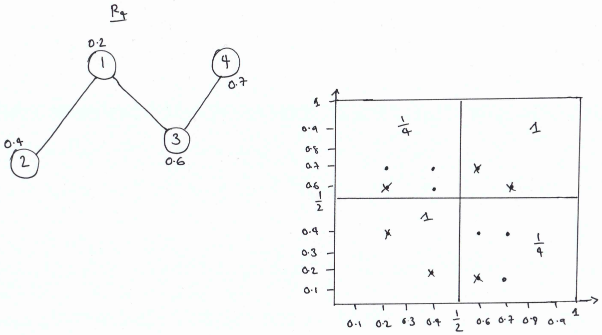
Approximation of a Graphon

- Let $(R_n)_{n\geq 1}$ be a convergent sequence of graphs, with corresponding adjacency matrices $(A^{(n)})_{n\geq 1}$.
- For graph R_n, construct an estimator $\hat{f}_n$ using the following:

1. Assign each node a value **uniformly** from $[0, 1]$.
2. Divide $[0, 1]^2$ into smaller rectangles.
3. Plot the cartesian product of the uniform values minus the points that lie on y=x. Each coorindate corresponds to a potential connection in our graph.
4. The value of our estimator is a scalar multiple of the proportion of points connected within each rectangle.

Remark: Indeed $\hat{f}_n$ converges to the limit, wrt the cut metric.[2]

Quick Example:



Problems:

Cut Metric (For my case):

$$inf_{\sigma \in \mathcal{M}} \left(\int_0^1 \int_0^1 |f(\sigma(x), \sigma(y)) - \hat{f}_n(x,y)|^2 dxdy \right)$$

Where $\mathcal{M}$ denotes the set of measure preserving maps from $[0,1]$ to $[0,1]$, and $\hat{f}_n$ denotes our estimator.

- The cut norm/distance is rather hard to work with, as it's unclear when the infimum is attained (although it is known that the infimum is attained).
- Hence I made the mistake of omitting the infimum and tried to proceed by approximating error with the L2 norm.

Initial Idea + What I did:

- I initially wanted to investigate rates of convergence and how these are effected by varying regularity assumptions of the underlying graphon.

Initial Idea + What I did:

- I initially wanted to investigate rates of convergence and how these are effected by varying regularity assumptions of the underlying graphon.
- However this didn't work out, the assumptions I required to investigate rates didn't hold. (could've been due to the lack of me using the cut-norm).

Initial Idea + What I did:

- I initially wanted to investigate rates of convergence and how these are effected by varying regularity assumptions of the underlying graphon.
- However this didn't work out, the assumptions I required to investigate rates didn't hold. (could've been due to the lack of me using the cut-norm).
- In our graphon estimation, the construction relies on us assigning values uniformly to the nodes, but why is this the case? Hence I thought I'd investigate this, vary the distribution to see how the L2 Error is effected. Furthermore I'd also like to see if continuity assumptions of the underlying graphon effect the L2 errors.

How I did this?

Distributions I'll be using to assign values:

- $Uniform(0, 1)$
- $Normal(0.5, \left(\frac{1}{6}\right)^2)$
- Deterministic values: If the graph we want to construct an estimator for has n nodes. Assign nodes the values $\{ \frac{1}{n}, \frac{2}{n}, \dots 1 \}$
- $Exponential(10)$

How I did this Cont...

1. Let f be a graphon. (Shall vary the continuity assumptions of this). Use the $G(n, f)$ model to construct a sequence of graphs that converge to f.

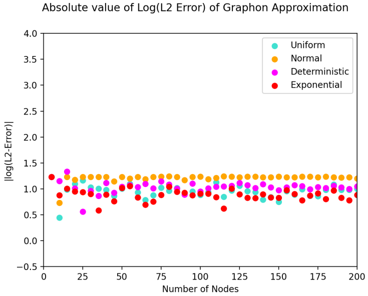
How I did this Cont...

1. Let f be a graphon. (Shall vary the continuity assumptions of this). Use the $G(n, f)$ model to construct a sequence of graphs that converge to f.
2. For every 5^{th} graph in this sequence, use the graphon approximation method to construct estimators for each distribution.

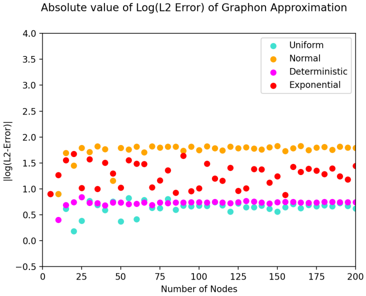
How I did this Cont...

1. Let f be a graphon. (Shall vary the continuity assumptions of this). Use the $G(n, f)$ model to construct a sequence of graphs that converge to f.
2. For every 5^{th} graph in this sequence, use the graphon approximation method to construct estimators for each distribution.
3. Plot the L2 errors / $|log(L2Error)|$ against the number of nodes.

Plots:

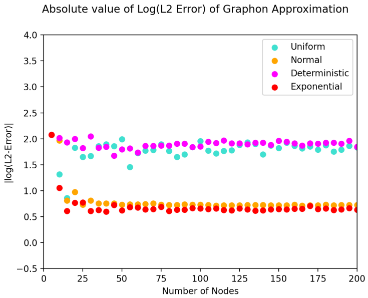


(a) f $= \dfrac{x+y}{2}$

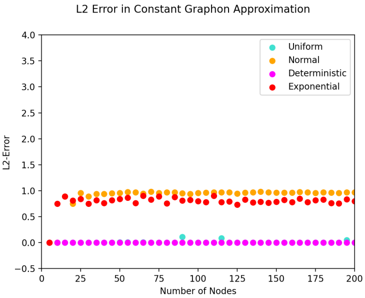


(b) Step function

Plots



(c) Step function



(d) $f = 1$

Conclusion:

- Doesn't seem continuity assumptions effect the L2 Error, hence may not effect convergence rates.
- If we have some prior knowledge about the limit/graphon, then a normal/ exponential distribution may be beneficial.
- In general, normal and any other distributions that generate values that are mainly centered around a value, tend to have larger L2 errors.
- Rather surprisingly, in the case of a constant graphon, we don't seem to have convergence for normal/exponential. This could be explained by clustering of the points.

Criticisms Improvements

*C*riticisms:

- The biggest one, being me not using the cut-norm.
- I've also tried to comment on trends, by only looking at the first 200 terms.

*F*urther Work:

- I'd like to somehow incorporate the cut norm.
- Improve the efficiency of my code to get information about the size of these errors.
- Use the graphon:

$$f(x,y) = \begin{cases} 1 & x, y \in \mathbb{Q} \\ 0 & \text{else} \end{cases}$$

References

- Daniel Glasscock. "What is a Graphon?" In: *Notices of the AMS* 62.1 (2015).
- Patrick J Wolfe and Sofia C Olhede. "Nonparametric graphon estimation". In: *arXiv preprint arXiv:1309.5936* (2013).