Introduction to graphons and large dense graphs

Julien Weibel

Cermics, ENPC

november 4, 2022

Reminder: definition of a graph

A (non-directed) *graph* $G = (V, E)$ is composed of:

- a set of points V which are called *vertices*
- a set of *edges* $E \subset V \times V$ such that
  - E is symmetric, i.e. $(x, y) \in E$ if and only if $(y, x) \in E$
  - E has no self loop, i.e. $(x, x) \notin E$

Adjacency matrix of a graph

The set of edges E defines an adjacency relation on the vertex set V, which can be summarize as a symmetric matrix M, called the *adjacency matrix*, defined by:

$$\forall u, v \in V, \quad M(u, v) = \mathbb{1}(uv \in E).$$

Random graph 1: Erdös-Renyi graph

Let $n \in \mathbb{N}^*$, and $p \in [0, 1]$.

The *Erdös-Renyi graph* $\mathbb{G}(n, p)$ is a random graph with

- vertex set $V = \{v_1, \ldots, v_n\}$,
- for each couple of vertices $\{u, v\}$, the edge uv is included in E with probability p, independently of each other.

Random graph 2: Stochastic block model

Let $n \in \mathbb{N}^*$, $k \in \mathbb{N}^*$, and let

- $\pi = (\pi_1, \dots, \pi_k)$ be a probability distribution on $\{1, \dots, k\}$,
- and $p = (p(i,j))_{1 \leq i,j \leq k}$ be a $k \times k$ symmetric matrix with coefficients in $[0, 1]$.

The *stochastic block model* $\mathrm{SBM}(n, k, \pi, p)$ is a random graph with

- vertex set $V = \{v_1, \dots, v_n\}$, each vertex v gets a type variable X_v distributed as π and independent of each other,
- for each couple of vertices $\{u, v\}$, the edge uv is included in E with probability $p(X_u, X_v)$, independently of each other.

This model is also known as: planted partition model, or inhomogeneous random graph.

Definition of a graphon

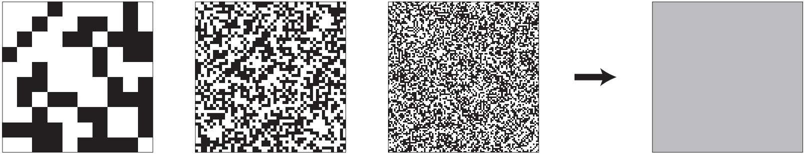
Let $(\Omega, \mathcal{F}, \mu)$ be a probability space.
A *graphon* is a measurable symmetric function $W : \Omega^2 \to [0, 1]$.
A *kernel* is a measurable symmetric function $W : \Omega^2 \to \mathbb{R}$.

We can think of W as weighted graph those vertex set is Ω, and with weight $W(x, y)$ on the edge xy.

Often, we take $\Omega = [0, 1]$ with the Borel σ-algebra and the Lebesgue measure.

Idea: a graphon is the limit of the adjacency matrix of graphs.

Example: convergence toward a graphon



Random graph 3: graphon model

Let $n \in \mathbb{N}^*$, $(\Omega, \mathcal{F}, \mu)$ be a probability space, and $W : \Omega \times \Omega \to [0, 1]$ be a graphon.

The *graphon model* $\mathbb{G}(n, \mu, W)$ is a random graph with

- vertex set $V = \{v_1, \ldots, v_n\}$, each vertex v gets a type variable $X_v \in \Omega$ distributed as μ and independent of each other,
- for each couple of vertices $\{u, v\}$, the edge uv is included in E with probability $W(X_u, X_v)$, independently of each other.

SBM as a graphon model

Consider the probability space $([0,1], \mathcal{B}, \text{Leb})$.

Partition $[0,1]$ into intervals $I_1, \ldots, I_k$ of length $\pi_1, \ldots, \pi_k$.

Define the graphon $W : [0,1]^2 \to [0,1]$ by $W(x,y) = p(i,j)$ for $x \in I_i, y \in I_j$.

Then, the random graphs $\text{SBM}(n, k, \pi, p)$ and $\mathbb{G}(n, \text{Leb}, W)$ have the same distribution.

SBM as a graphon model

Consider the probability space $([0,1], \mathcal{B}, \text{Leb})$.

Partition $[0,1]$ into intervals $I_1, \ldots, I_k$ of length $\pi_1, \ldots, \pi_k$.

Define the graphon $W : [0,1]^2 \to [0,1]$ by $W(x,y) = p(i,j)$ for $x \in I_i, y \in I_j$.

Then, the random graphs $\text{SBM}(n, k, \pi, p)$ and $\mathbb{G}(n, \text{Leb}, W)$ have the same distribution.

Remark: this also means that finite graphs can be represented as graphons.

Graphs as graphons

Let $G = (V, E)$ be a graph with n vertices.
Denote by M its adjacency matrix.

Consider the probability space $([0,1], \mathcal{B}, \mathrm{Leb})$.

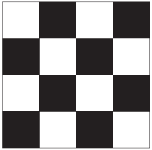
Partition $[0,1]$ into intervals $I_1, \dots, I_n$ each of length $1/n$.

Define the graphon $W_G : [0,1]^2 \to [0,1]$ by
$W_G(x, y) = M(v_i, v_j)$ for $x \in I_i, y \in I_j$.

Example: graph as a graphon

The adjacency matrix

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

is transformed into the step-function graphon 

Large dense graphs and sampling

- We want to study large dense graphs $G = (V, E)$
- Large: n is large
- Dense: if $n = |V|$ is the number of vertices,
  then the number of edges $|E|$ is $\Omega(n^2)$.

Large dense graphs and sampling

- We want to study large dense graphs $G = (V, E)$
- Large: n is large
- Dense: if $n = |V|$ is the number of vertices,
  then the number of edges $|E|$ is $\Omega(n^2)$.
- Problem: the cost to represent entirely those graphs is prohibitive
- Idea: sample a subgraph of smaller size
  $k << n$, $X_1, \cdots X_k$ iid sampled uniformly from V
  $\mathbb{G}(k, G) = G[X_1, \cdots X_k]$ the induced subgraph of G

Large dense graphs and sampling

- We want to study large dense graphs $G = (V, E)$
- Large: n is large
- Dense: if $n = |V|$ is the number of vertices,
  then the number of edges $|E|$ is $\Omega(n^2)$.
- Problem: the cost to represent entirely those graphs is prohibitive
- Idea: sample a subgraph of smaller size
  $k << n$, $X_1, \cdots X_k$ iid sampled uniformly from V
  $\mathbb{G}(k, G) = G[X_1, \cdots X_k]$ the induced subgraph of G
- Question : can we find a (continuous) limit to an (increasing) sequence of large dense graphs?

Sampling from graphons

For a graph: sample a subgraph of smaller size

$k << n, \qquad X_1, \cdots, X_k$ iid sampled uniformly from V

$\mathbb{G}(k, G) = G[X_1, \cdots X_k]$ the induced subgraph of G

For a graphon: remind the definition of $\mathbb{G}(k, \mathrm{Leb}, W)$

sample $X_1, \cdots, X_k \in \Omega$ iid according to μ

edge ij is included with probability $W(X_i, X_j)$,

independently of each other

Thus, the random graphs $\mathbb{G}(k, G)$ and $\mathbb{G}(k, \mathrm{Leb}, W_G)$ have the same distribution.

The cut norm for kernels

For a kernel W, we define its cut norm by:

$$\|W\|_\square = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)\,\mathrm{d}x\mathrm{d}y \right|$$

Remark the comparison with the L_1-norm:

$$\|W\|_\square \le \|W\|_1 = \int_{[0,1]^2} |W(x,y)|\,\mathrm{d}x\mathrm{d}y$$

The cut norm $\|\cdot\|_\square$ induces the cut distance $d_\square$ on kernels defined as $d_\square(U,W) = \|U - W\|_\square$.

The cut distance for kernels

Relabelling:

- for graphs: permutation of the vertices
- for graphons: measure-preserving bijections $\varphi : [0,1] \to [0,1]$

The cut distance for kernels

Relabelling:

- for graphs: permutation of the vertices
- for graphons: measure-preserving bijections $\varphi : [0,1] \to [0,1]$

We define

- $S_{[0,1]}$ the set of measure-preserving bijections $\varphi : [0,1] \to [0,1]$
- $\bar{S}_{[0,1]}$ the set of measure-preserving functions $\varphi : [0,1] \to [0,1]$

The cut distance between two kernels U and W is defined as:

$$\delta_\square(U, W) = \inf_{\varphi \in S_{[0,1]}} d_\square(U, W^\varphi) = \min_{\varphi, \psi \in \bar{S}_{[0,1]}} d_\square(U^\psi, W^\varphi).$$

Characterization of kernels by the cut distance

Corollaire 1

Let U and W be two kernels defined on the probability space $([0,1], \mathcal{B}, Leb)$. There is equivalence between:

1. $\delta_\square(U, W) = 0$
2. *there exist $\varphi, \psi \in \bar{S}_{[0,1]}$ such that $U^\varphi = W^\psi$ almost everywhere*
3. *for every $k \in \mathbb{N}^*$, the random graphs $\mathbb{G}(k, Leb, U)$ and $\mathbb{G}(k, Leb, W)$ have the same distribution.*

Characterization of convergence for the cut distance

Let W_n, $n \in \mathbb{N}$, and W be graphons defined on the probability space $([0,1], \mathcal{B}, \text{Leb})$.

There is equivalence between:

1. $\delta_\square(W_n, W) \to 0$
2. for every $k \in \mathbb{N}^*$, the random graphs $\mathbb{G}(k, \text{Leb}, W_n)$ converge in distribution to $\mathbb{G}(k, \text{Leb}, W)$ when $n \to \infty$.

Some properties of the cut distance

Théorème 2

The space of graphons with the cut distance $\delta_\square$ is compact.

Some properties of the cut distance

Théorème 2

The space of graphons with the cut distance $\delta_\square$ is compact.

Lemme 3

For every graphon W and $k \in \mathbb{N}^*$, there exist a step-function kernel U with k steps such that

$$\|W - U\|_\square \leq \frac{2}{\sqrt{\log k}}.$$

Estimation of the cut distance with sampling

Lemme 4 (Sampling lemma)

Let W be a graphon, and $k \in \mathbb{N}^*$.
With probability at least $1 - 2\exp(-k/(2\log k))$, we have:

$$\delta_\square(\mathbb{G}(k, Leb, W), W) \leq \frac{22}{\sqrt{\log k}} \ .$$

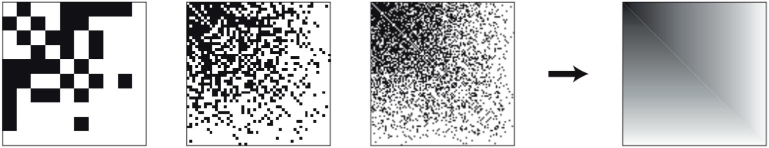
Graphons as limits for graphs

Let G_n, $n \in \mathbb{N}$, be graphs such that for every $k \in \mathbb{N}^*$, the random graphs $\mathbb{G}(k, G_n)$ converges in distribution to some random graph.

Then, there exist a graphon W such that $\delta_\square(W_{G_n}, W) \to 0$, i.e. for every $k \in \mathbb{N}^*$, the random graphs $\mathbb{G}(k, G_n)$ converges in distribution to $\mathbb{G}(k, \text{Leb}, W)$.

In this case, we say that the graphon W is the *limit* of the sequence of graphs G_n.

Example: convergence toward a graphon



Example: a diffusion process with a graphon

Without a graphon:

$$\mathrm{d}X(t,x) = \int_{[0,1]} X(t,y) - X(t,x) \; \mathrm{d}y$$

With a graphon or a kernel:

$$\mathrm{d}X(t,x) = \int_{[0,1]} W(x,y)(X(t,y) - X(t,x)) \; \mathrm{d}y$$

Example: an infection model with a graphon

$$\partial_t u(t,x) = (1 - u(t,x)) \int_\Omega u(t,y)\beta(x)W(x,y)\theta(y)\mu(\mathrm{d}y) - \gamma(x)u(t,x)$$

$$u(0,x) = u_0(x)$$

$$x \in \Omega,\ t \in]0, T]$$

Questions ?