Stochastic modelling of evolving networks using graph limits

Anton Klimovsky

U of Stuttgart

20220623 @ Paris

Graphons

For $n \in \mathbb{N}$, let $\mathcal{G}_n$ denote the set of all $2^{\binom{n}{2}}$ undirected labelled simple graphs with n vertices.

$$G \in \mathcal{G}_n \xleftrightarrow{\text{representation}} \text{symmetric adjacency matrix } \quad A^G(i,j) := \begin{cases} 0, & i \not\sim j, \\ 1, & i \sim j. \end{cases}$$

Q: How to describe $n \to \infty$ limit & the limiting objects?

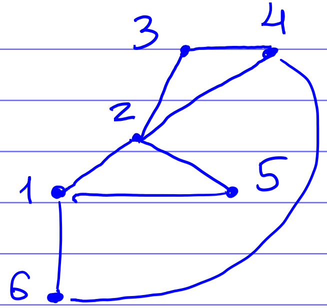
A natural way to embed A^G into a set of bivariate symmetric functions called graphons:

$$W := \{ h : [0,1]^2 \to [0,1] \mid h(x,y) = h(y,x), \ \forall x, y \in [0,1] \}$$

and represent $G \rightsquigarrow h^G \in W$ as follows:

$$h^G(x,y) := \begin{cases} 0, & \lceil nx \rceil \not\sim \lceil ny \rceil, \\ 1, & \lceil nx \rceil \sim \lceil ny \rceil. \end{cases}$$

Example

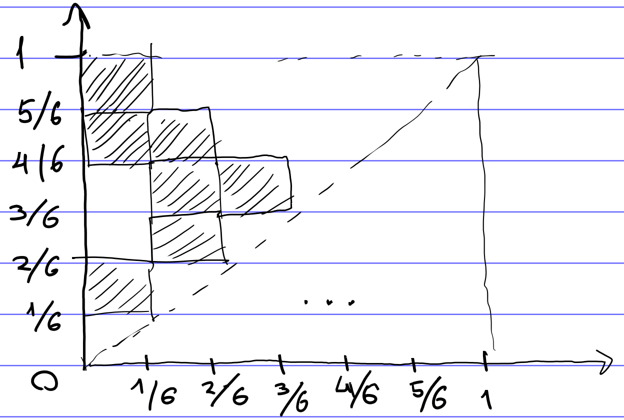


$G_6 \in \mathcal{G}_6$

$n = 6$

representation

$\longleftrightarrow$



$$h := \begin{cases} 1, & x, y \in \blacksquare \\ 0, & \text{otherwise} \end{cases}$$

Cut distance on $\mathcal{W}$

$$d_\square(h_1, h_2) := \sup_{S,T\subset[0,1]} \left| \int_{S\times T} dx\,dy\,[h_1(x,y) - h_2(x,y)] \right|$$

Equivalence relation on $\mathcal{W}$: Let Σ be the set of measure-preserving bijections (isomorphisms) $\sigma : [0,1] \to [0,1]$.

Define an equivalence relation on $\mathcal{W}$ as

$$h_1 \equiv h_2 \iff \exists\, \sigma \in \Sigma : \; h_1^\sigma \equiv h_2,$$

where $h_1^\sigma(x,y) = h_1(\sigma(x), \sigma(y))$.

Quotient space: $(\tilde{\mathcal{W}}, \delta_\square)$, where

$$\delta_\square(\tilde{h}_1, \tilde{h}_2) = \inf_{\sigma_1,\sigma_2\in\Sigma} d_\square(h_1^{\sigma_1}, h_2^{\sigma_2})$$

is called the **cut metric**.

$\boxed{\text{NB!}}$ The equivalence classes correspond to unlabelled graphs

Subgraph counts

Let $F \subset G$ be two finite graphs with vertex sets $V(F)$, $V(G)$.
Let $\hom(F, G)$ be the # of graph homomorphisms from $F \to G$
The homomorphism density is defined as

$\leftarrow$ density of F in G

$$t(F, G) := \frac{\hom(F, G)}{\underbrace{|V(G)|^{|V(F)|}}_{\text{max \# of homs possible}}} \in [0, 1]$$

def: A sequence of labelled simple graphs $(G_n)_{n \in \mathbb{N}}$ converges, when

$$\underbrace{(t(F, G_n))_{n \in \mathbb{N}}}_{\text{seq. of reals}} \text{ converges } \forall \text{ finite simple graphs } F.$$

Q: Converges to what?

A limiting object: Graphon

Consider a finite simple graph F with $|V(F)|=k$ and edge set $E(F)$.
Let $h \in W$. Define the density of F in h as

$$t(F,h) := \int_{[0,1]^k} dx_1 \cdot \ldots \cdot dx_k \prod_{(i,j)\in E(F)} h(x_i, x_j).$$

Then,

$$t(F,h^G) = \int_{[0,1]^k} dx_1 \cdot \ldots \cdot dx_k \prod_{(i,j)\in E(F)} h^G(x_i,x_j)$$

$$= \frac{\hom(F,G)}{|V(G)|^{|V(F)|}} = t(F,G).$$

That is, $(G_n)_{n\in\mathbb{N}}$ converges to $h\in W$, when

$$\lim_{n\to\infty} t(F,G_n) = \lim_{n\to\infty} t(F,h^{G_n}) = t(F,h), \quad \forall F \text{ finite, simple}$$

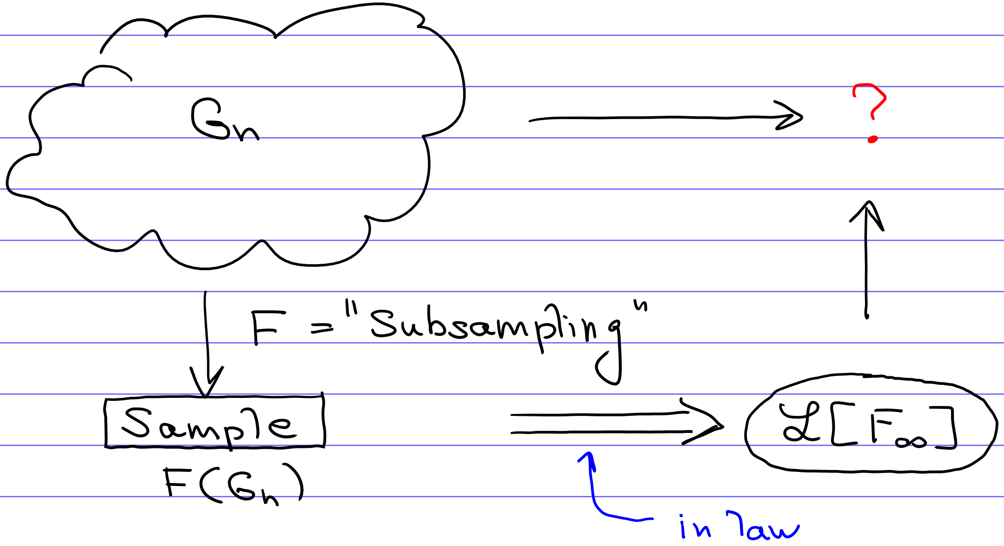
graphon

$$\Rightarrow \text{"} h = \lim_{n\to\infty} h^{G_n} \text{"}.$$

$\boxed{\text{NB!}}$ Works well for dense graphs; i.e., $\deg(i) = \Theta(n) \ldots$

Another viewpoint: Sampling convergence

def (Sampling convergence, informal) G_n converges as $n \to \infty$,
iff all $\underbrace{\text{"random substructures"}}_{\text{e.g. subsamples}}$ $F(G_n)$ converge in distribution
as $n \to \infty$.



Example: Uniformly sampled subgraphs

$$G_n = (\underbrace{[n]}_{\text{vertex set}}, \underbrace{E_n}_{\text{edge set}}), \quad n \in \mathbb{N}. \quad \text{Fix } k \in \mathbb{N}.$$

Sample: $X_1, \dots, X_k \underset{\text{i.i.d.}}{\sim} \text{UNIFORM}([n])$

Subgraph: $F_k(G_n) = ([k], E_k(G_n))$, where

$$E_k(G_n) := \{ \{i,j\} \subset [k] : \{X_i, X_j\} \in E_n \}$$

def: (Sampling convergence for graphs) G_n sampling converges, iff $F_k(G_n)$ converges in distribution $\forall k \in \mathbb{N}$

$\boxed{\text{NB!}}$ $F_k(G_n)$ is exchangeable, i.e., its distribution is invariant under relabelling of the vertex set.

Basic properties

Proposition 1: (Borgs et al. 2008) The following are equivalent:

(i) $(G_n)_{n\in\mathbb{N}}$ converges.

(ii) $(\tilde{h}^{G_n})_{n\in\mathbb{N}}$ is Cauchy in the $\delta_\square$-metric.

(iii) $(t(F,h^{G_n}))_{n\in\mathbb{N}}$ converges, $\forall F$ finite, simple.

(iv) $\exists h\in W$: $\lim_{n\to\infty} t(F,h^{G_n}) = t(F,h)$, $\forall F$ finite, simple.

(v) $(G_n)_{n\in\mathbb{N}}$ sampling converges.

Proposition 2: (Lovász & Szegedy 2006)

$(\tilde{W}, \delta_\square)$ is compact.

Dynamic graphs & graphons

Sequences of dense dynamic random graphs

$$\{G_n(t)\}_{t\geq 0}, \qquad n\in\mathbb{N},$$

and their $n\to\infty$ limits, e.g., dynamic random graphons.

$$\{h_t\}_{t\geq 0}$$

$\boxed{\text{NB!}}$ $\quad G_n(t) \rightsquigarrow h^{G_n(t)} \in \tilde{W}$.

Key challenges:

- Construct examples with rich enough limiting dynamics.
- Statistical inference, estimation, uncertainty quantification...
- Sparse networks.

Motivation: Evolution of complex networks,
Stochastic processes on evolving networks.

Networks evolve over time:

- Social networks: Friendships come & go, interaction intensity varies (cf. the pandemic's)
- Economic networks: Business relations evolve over time
- Biological networks: Environmental changes, evolution, population genetics, ...

...

(Some) Previous work

- "Easy": construct random dynamics of random graphs in $\mathcal{G}_n$ for fixed $n \in \mathbb{N}$.

- Challenge: construct rich limiting dynamics for dense graphs as $n \to \infty$.

- Crane 2016: constructed limiting dynamics based on exchangeable incidence matrices and Aldous-Hoover theory of exchangeable arrays.

- Černý & K. 2020: extended ↑ to weighted graphs

Markovian dynamics of exchangeable arrays

Černy & K. 2020

Assumptions:

- Network = weighted graph
- Vertex-exchangeable array
- (globally) Markovian infinite array (discrete & cont. time)

Proposition • Subsamples of Markovian exch. arrays are (in general) not Markov (even in discrete time)

• However, any subsample of a continuous-time Feller infinite exch. array is Markov.

Proposition: (Structure of jumps) Given there is a jump, either

(i) [Macroscopic event] A positive fraction of edge weights in the array jumps.
(ii) [Mesoccopic event] A positive proportion of edge weights in a raw/col jumps.
(iii) [Microscopic event] A single edge weight jumps.

Graphon dynamics from population genetics

Athreya, den Hollander, Roellin 2020

constructed a class of diffusions in the space of densities.

Idea: Introduce types of nodes and use population genetic dynamics to let them evolve.
Connect individuals by edges with prob. depending on their types.

Example (Moran model) • Consider n individuals carrying one of two types, say, $\{0,1\}$.

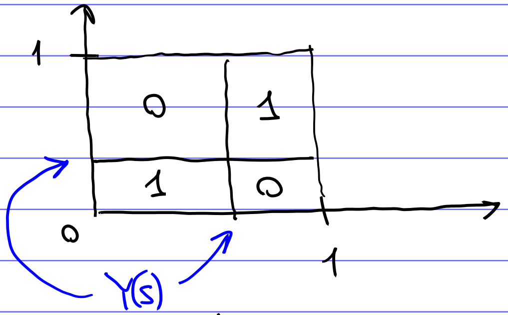
- At rate 1, an individual randomly draws an individual from the population (including itself) and adopts its type.
- Let $X_n(s) :=$ # individuals of type 0 at time s.
- $Y_n(s) := \frac{1}{n} X_n(ns) \underset{n\to\infty}{\Longrightarrow} Y(s)$, where $Y(s) = Y(0) + \int_0^s \sqrt{Y(u)(1-Y(u))}\, dW(u)$

Wright-Fisher diffusion

A Wright–Fisher graphon

- $G_n(s) := ([n], E_n(s))$, where
- $\{i,j\} \in E_n(s)$ iff i,j of the same type.
- $\forall$ connected graph F_k on k vertices:

$$t(F_k, G_n) = \frac{X_n(s)^k + (n - X_n(s))^k}{n} \underset{n\to\infty}{\Longrightarrow} Y(s)^k + (1 - Y(s))^k$$



- This is just the simplest population dynamics...
- Using this setup, one can construct a whole class of limiting dynamics driven by various population genetics models.

Challenges:

This is just the beginning!

- Write down the generator of the limiting dynamics.
- Random dynamic sparse models.
- Stochastic processen on evolving graphs.
- Statistical inference, estimation, "uncertainty quantification", ...

(Stochastic processes on evolving networks)

a network of 10 researches

Supported by DFG