| STA4513: Statistical Models of Networks | Fall 2014 |
|---|---|
| Lecture 5 — 8 September, 2014 | |
| *Prof. Daniel M. Roy* | *Scribe: Namdar Homayounfar* |

# 1 Overview

In this lecture, we introduce the concept of a graphing and review its associated measures. The development here closely follows that of *Large Networks and Graph Limits* by L. Lovász [Lov12].

# 2 Graphing

**Definition 1.** Let $\Omega$ be a Borel space (e.g., the unit interval) with Borel sets $\mathcal{S}$. A graph $G$ of bounded degree $D$ with a node set $\Omega$ and edge set $E \subseteq \Omega \times \Omega$ is **measure-preserving** if the edge set $E \in \mathcal{S}$ and there exists a probability measure $\lambda$ on $\Omega$ such that, for $A, B \in \mathcal{S}$

$$\int_A deg_B(x)\lambda(dx) = \int_B deg_A(x)\lambda(dx). \tag{1}$$

A **graphing** is a quadruple $(\Omega, \mathcal{S}, G, \lambda)$ where $G$ is measure-preserving with respect to $\lambda$.

**Notes:**

- For $B \subseteq \Omega$, define $deg_B(x) = \#\{y \mid (x, y) \in E \text{ and } y \in B\}$, which counts how many of edges out of $x$ land in the vertex set $B$
- (1) can be written as $\mathrm{E}[1(U \in A)deg_B(U)] = \mathrm{E}[1(U \in B)deg_A(U)]$, where $U \sim \lambda$.
- $\Omega$ need not be finite or even countable. Indeed, the main case of interest is $\Omega = [0, 1]$.

# 3 Important Measures Associated with a Graphing $G$

## 3.1 A measure over vertices

To better understand graphings, we have to define some measures associated with them. One such measure is the **volume measure** defined on $\Omega$:

$$Vol(A) = \int_A deg(x)\lambda(dx)$$

where $deg(x)$ is the number of edges from $x$ to $A$.
We can write

$$Vol(A) = \mathrm{E}\Big(deg(U); A\Big) \text{ where } U \sim \lambda$$

The average degree of a node $U \sim \lambda$ is given by

$$d_0 := Vol(\Omega) = \int_\Omega deg(x)\lambda(dx)$$

However, $Vol$ is not a probability measure since $Vol(\Omega)$ is not necessarily equal to 1. But

$$\lambda^*(A) = \frac{Vol(A)}{Vol(\Omega)}$$

is a probability measure since $Vol(\Omega) \leq D < \infty$. The probability measure $\lambda^*$ is also known as the **stationary distribution** of $G$.

$\lambda$ vs. $\lambda^*$

- How can we sample from $\lambda^*$ using $\lambda$ ?

  1. Sample $U \sim \lambda$
  2. Accept with probability $\frac{deg(U)}{D}$

  Why does this work? Notice that

$$\begin{aligned} \text{P}(U \in A \text{ and accept } U) &\propto \int \frac{deg(x)}{D} 1_A(x)\lambda(dx) \\ &\propto \int_A deg(x)\lambda(dx) \end{aligned}$$

- The other way around, how can we sample from $\lambda$ using $\lambda^*$?

  1. Sample $U \sim \lambda^*$
  2. Accept with probability $\frac{1}{deg(U)}$

  This works provided that no vertex is isolated (has degree zero).

### 3.2 A measure over edges

Now, in order to better understand $\lambda^*$, define a finite measure over edges $\Omega \times \Omega$ by

$$\eta(A \times B) = \int_A deg_B(x)\lambda(dx)$$

This defines a measure on all Borel sets $C \subset \Omega \times \Omega$ by Caratheodory's theorem. Then $\frac{\eta}{d_0}$ where $d_0 = Vol(\Omega)$ is a probability measure.

Next we might consider how to sample from $\eta/d_0$:

1. Sample $U \sim \lambda^*$
2. Sample $V \sim neighbours(U)$

Notice that

$$\begin{aligned}
\mathrm{P}((U,V)\in A\times B) &\propto \int_A \frac{deg_B(x)}{deg(x)}\lambda^*(dx)\\
&= \int_A \frac{deg_B(x)}{deg(x)}\frac{deg(x)}{deg(\Omega)}\lambda(dx)\\
&= \frac{\eta}{d_0}
\end{aligned}$$

Now, suppose $(U,V)\sim\frac{\eta}{\lambda_0}$. We have

$$\mathrm{P}(U\in A)=\lambda^*(A)=\frac{\eta(A\times\Omega)}{Vol(\Omega)}$$

and, using the measure-preserving property of $G$,

$$\mathrm{P}(V\in A)=\frac{\eta(\Omega\times A)}{Vol(\Omega)}=\frac{\eta(A\times\Omega)}{Vol(\Omega)}=\lambda^*(A).$$

The measure $\lambda^*$ is called stationary because if one begins a simple random walk from a random vertex distributed according to $\lambda^*$, then the distribution of the particle after its first (and then every) step is also $\lambda^*$.

## 4 Turning a Finite Graph to a Graphing on [0,1]

Let $F$ be a finite graph and assume $V(F)=[n]$. Define $G_F$ as follows:

1. Split $[0,1]$ into $n$ chunks $J_i=\left[\frac{(i-1)}{n},\frac{i}{n}\right)$
2. For every edge $ij\in E(F)$ with $i<j$, connect every point $x\in J_i$ to $x+(j-i)/n\in J_j$

**Theorem 2.** *$G_F$ is measure-preserving.*

**Proof Sketch:** Let $\lambda$ be the lebesque measure. Then

$$\begin{aligned}
\int_A deg_B(x)dx &= \sum_i\sum_j\int_{A\cap J_i} deg_{B\cap J_j}(x)dx\\
&= \sum_j\sum_i\int_{B\cap J_i} deg_{A\cap J_i}(x)dx \quad \text{by definition of } G_F\\
&= \int_B deg_A(x)dx
\end{aligned}$$

Example) Figure below represents a pentagon turned into a graphing:

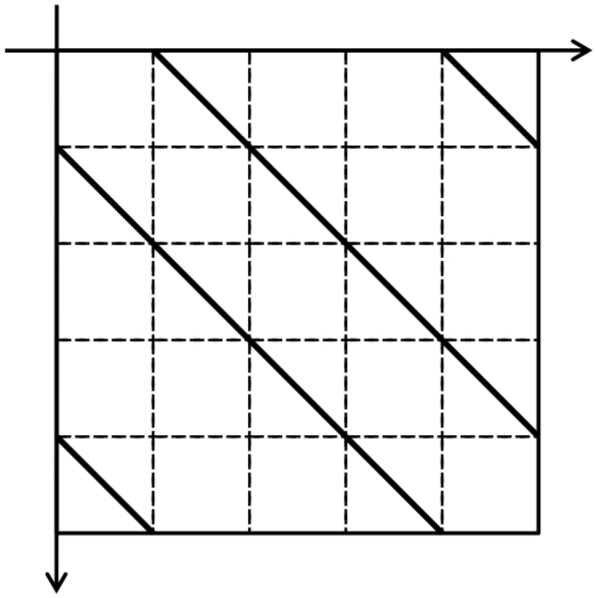

Figure 1: Figure taken from [Lov12].

# 5 Random Rooted Graphs

- Let $G^\bullet$ be the set of connected countable graphs $H$ of bounded degree $D$ with a distinguished node called the root. If $H \in G^\bullet$ denote its root by $root(H)$.
- $deg(H) = deg(root(H))$
- Let $\mathcal{G}^\bullet$ be the set of finite rooted graphs in $G^\bullet$
- $B_r \in \mathcal{G}^\bullet$ be the set of all finite graphs such that the nodes are within distance $r$ from the root; These are the so-called "r-balls".
- Let $B_{H,r} \in B_r$ denote the r-neighbourhood of $H$'s root.
- Let $G_x$ denote the connected component rooted at $G$.

$\mathbb{H}$: the graph of graphs

- $V(\mathbb{H}) = G^\bullet$
- Let $(H, v) \in G^\bullet$. For every edge $e = vv' \in E(\mathbb{H})$, connect $(H, v)$ by an edge to the rooted graph $(H, v') \in G^\bullet$
- All degrees in $\mathbb{H}$ are bounded by $D$.
- Let $\sigma$ be a probability measure on $G^\bullet$. Define

$$\sigma^*(A) = \frac{\int_A deg(H)\sigma(dH)}{\int_{G^\bullet} deg(H)\sigma(dH)}$$

(This is well defined since both integrands are less than $D$.)

**Sampling Process**

1. Sample $H \in \sigma^*$
2. Sample an edge $E$ uniformly among the edges incident with $root(H)$

$(H, E)$ is a graph with an oriented edge. (Let $\vec{G}$ be the space $G^\bullet$ with an oriented edge leaving the root.) Let $\vec{\sigma} = \mathrm{P}(H, E)$ be the distribution on $\vec{G}$.

**Definition 3.** We say that $\sigma$ is **involution invariant** (or $H$ is **unimodular**) when the map $\vec{G} \to \vec{G}$ obtained by reversing the orientation of the root edge is measure preserving for $\sigma^*$.

What does it all mean?

Let $G$ be a graphing. Let $u$ be a random vertex and let $G_u$ be the connected component in $G^\bullet$. $G_u$ is a "random rooted component of $G$". If $U \sim \lambda^*$ then $(G, U) \sim \sigma^*$.

**Theorem 4** ([Lov12, Thm 18.37])**.** *Every involution invariant probability distribution on $G^\bullet$ can be represented by a graphing.*