Random Walks on Graphs and Directed Graphs: Course Notes

Eric Tressler

28 September 2005

1 Review and Motivation

Recall that for a graph $G = (V, E)$ we define the *transition probability matrix* $P(u, v)$ by

$$P(u, v) = \begin{cases} \frac{1}{d_u} & \text{if } u \sim v \\ 0 & \text{otherwise} \end{cases},$$

where d_u denotes the degree of a vertex u and $u \sim v$ is notation for $\{u, v\} \in E$. Recall further that we may represent a distribution on G by a row vector f whose ith entry gives the "amount" residing at the ith vertex v_i. A typical initial distribution might be $\overrightarrow{\mathbb{1}} = (1, 1, \ldots, 1)$, or the row vector e_i^* with all entries zero except for a one in the ith entry. Some papers like to use the uniform distribution $\overrightarrow{\mathbb{1}}$; however, we will see that this distribution is not stationary if the graph in question is not regular.

The notion of a transition probability matrix allows us to examine distributions on G after some number of steps in our walk: if f is the initial distribution, then fP is the distribution after one step, and fP^t is the distribution after t steps. In a sense, we are considering all possible random walks simultaneously; for instance, if $f = e_i^*$, the jth entry of fP^t is the likelihood that if we start a walk at vertex v_i and continue for t random steps that we will find ourselves at vertex v_j.

However, before this tool becomes useful to us, we must consider the mathematical behavior of iteratively multiplying by P on the right. Does this sequence converge? Does it converge to a stationary distribution? Is such a distribution necessarily unique? Finally, what is the rate of convergence of this process? (We say that a distribution π is stationary or stable if $\pi P = \pi$, in which case $\pi P^t = \pi$ for all t.)

2 Examples

Example 1. If G is a k-regular graph, then $P = \frac{1}{k}A$, where A is the adjacency matrix of G. Let $\overrightarrow{\mathbb{1}}$ be the initial distribution, and observe that $\overrightarrow{\mathbb{1}}P = \overrightarrow{\mathbb{1}}\left(\frac{1}{k}A\right) = \overrightarrow{\mathbb{1}}$, so that $\overrightarrow{\mathbb{1}}$ is a stable distribution for any regular graph.

Example 2. Let G be the disjoint union of two copies of K_m, and let distributions π_1 and π_2 be defined by

$$\pi_1 = (\underbrace{1, \ldots, 1}_{m}, \underbrace{0, \ldots, 0}_{m})$$

and

$$\pi_2 = (\underbrace{0, \ldots, 0}_{m}, \underbrace{1, \ldots, 1}_{m}).$$

Then with the obvious ordering on the vertices of G, we have that $\pi_1 P = \pi_1$ and $\pi_2 P = \pi_2$. Thus we see that stable distributions are not necessarily unique.

Example 3. Let G be the star graph on $k + 1$ vertices (that is, G is a tree with k vertices of degree one, each adjacent to a single vertex of degree k).

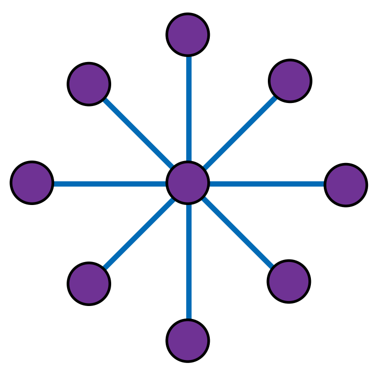

Figure 1: The star graph on nine vertices.

Let distribution π be defined by

$$\pi = (\underbrace{1, \ldots, 1}_{k}, k).$$

Then it is easy to see that π is a stable distribution for G.

3 Finding Stable Distributions

These examples give us some rudimentary insight into how stable distributions relate to a graph; before we continue, however, we must define the notion of volume for a graph. If $G = (V, E)$ is an undirected graph and $S \subseteq V$ is arbitrary, then define the *volume* of S by

$$\begin{aligned} \operatorname{vol} S &= \sum_{v \in S} d_v \\ &= \sum_{v \in S} \operatorname{vol} v. \end{aligned}$$

Of course, $\operatorname{vol} G$ is defined to be $\operatorname{vol} V$. If G is directed, then let d_u^+ and d_u^- denote the outdegree and indegree of the vertex u, respectively. Here we do not define the volume of an arbitrary subset of V; we define only

$$\operatorname{vol} G = \sum_{u \in V} d_u^+ = \sum_{v \in V} d_v^-.$$

Note that volume is a measure on G, and that it is not unique in this respect. This measure comes naturally from the theory, but of course any measure is useful as long as it produces good results.

Now for an undirected graph $G = (V, E)$ define the distribution π by

$$\begin{aligned} \pi &= \left(\frac{d_u}{\operatorname{vol} G} \right)_{u \in V} \\ &= \left(\frac{d_1}{\operatorname{vol} G}, \dots, \frac{d_n}{\operatorname{vol} G} \right). \end{aligned}$$

Then we have that

$$\begin{aligned} (\pi P)(v) &= \sum_{u \sim v} \pi(u) P(u, v) \\ &= \sum_{u \sim v} \frac{d_u}{\operatorname{vol} G} \cdot \frac{1}{d_u} \\ &= \frac{d_v}{\operatorname{vol} G} \\ &= \pi(v). \end{aligned}$$

Since this holds for all $v \in V$, it follows that π is a stable distribution.

3.1 The Directed Case

We are not so lucky in the directed case. Define π^+ and π^- by

$$\pi^+ = \left(\frac{d_u^+}{\operatorname{vol} G} \right)_{u \in V}$$

and

$$\pi^- = \left(\frac{d_u^-}{\operatorname{vol} G} \right)_{u \in V}.$$

Then $\pi^+ P = \pi^-$, so that there is no hope of obtaining a stable distribution as we have done above for undirected graphs. In fact, the equation $fP = f$ has no closed form solution for f in general.

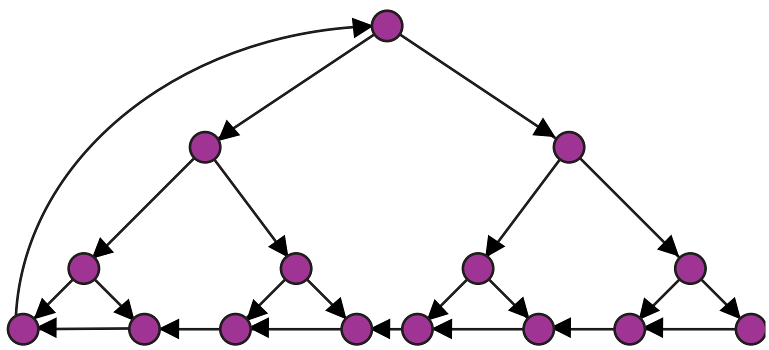

Figure 2: The graph considered in Exercise 1.

Exercise 1. Try to find a stable distribution for the graph G in Figure 2. This example demonstrates the difficulty of trying to find a stable distribution. Observe that if one starts at the top vertex v_0 and begins to walk randomly on this graph, every path visits v_0 infinitely many times. It is easy to see that the number of steps before the first return to v_0 is, with equal likelihood, any number in the interval $[4, 11]$. This gives some indication of the difficulty of finding a stable distribution in G.

4 The Issue of Convergence

Now we will attempt to answer our questions from earlier in the lecture for undirected graphs, so consider again the process of taking an initial distribution f and multiplying it on the right by P^t for some t. This is a particularly easy operation if P happens to be a diagonal matrix; happily, we will achieve something similar for arbitrary P. Note that we may write

$$\begin{aligned} P = D^{-1}A &= D^{-\frac{1}{2}}(D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}} \\ &= D^{-\frac{1}{2}}MD^{\frac{1}{2}}, \end{aligned}$$

where D is the diagonal matrix of degrees, A is the adjacency matrix, and $M = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ is a symmetric matrix. Now we need the following theorem.

Theorem 1 (Eigenvalue Decomposition Theorem)**.** *If $M \in M_n(\mathbb{R})$ is symmetric, then $M = U\Delta U^*$, where U is unitary and Δ is diagonal with entries the eigenvalues of M.*

Proof. A complete proof can be found in [1]. □

Given the result above, more manipulation will be useful. So, to continue, let ρ_i be the eigenvalues of M (necessarily real since M is real and symmetric) and let $f_0^*, \ldots, f_{n-1}^*$ be the columns of U if we write $M = U\Delta U^*$. Then we have that

$$\begin{aligned} f_i M &= e_i^* \operatorname{diag}(\rho_1, \ldots, \rho_{n-1}) \begin{pmatrix} f_0 \\ \vdots \\ f_{n-1} \end{pmatrix} \\ &= \rho_i e_i^* \begin{pmatrix} f_0 \\ \vdots \\ f_{n-1} \end{pmatrix} \\ &= \rho_i f_i. \end{aligned}$$

Therefore, we may write

$$M = \sum_{i=0}^{n-1} \rho_i f_i^* f_i = \sum_{i=0}^{n-1} \rho_i \mathbb{P}_i,$$

where $\mathbb{P}_i$ is projection onto the ith eigenvector. Now if

$$f = \sum_{i=0}^{n-1} a_i f_i = \sum_{i=0}^{n-1} f f_i^* f_i$$

is arbitrary, then

$$fM = \sum_{i=0}^{n-1} \rho_i (f f_i^*) f_i.$$

Let

$$f_0 = \left(\sqrt{\frac{d_v}{\operatorname{vol} G}} \right)_{v \in V} = \frac{\vec{\mathbb{1}} D^{\frac{1}{2}}}{\sqrt{\operatorname{vol} G}}$$

and observe that

$$\begin{aligned} f_0 M &= \frac{\vec{\mathbb{1}} D^{\frac{1}{2}}}{\sqrt{\operatorname{vol} G}} \cdot \left(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \right) \\ &= \frac{\vec{\mathbb{1}} D^{\frac{1}{2}}}{\sqrt{\operatorname{vol} G}} = f_0, \end{aligned}$$

so that $\rho_0 = 1$.

Finally we can apply this to answer our earlier questions (the norm below is the

2-norm):

$$
\begin{aligned}
\|fP^t - \pi\| &= \left\| fD^{-\frac{1}{2}} M^t D^{\frac{1}{2}} - \frac{\vec{\mathbb{1}} D}{\operatorname{vol} G} \right\| \\
&= \left\| fD^{-\frac{1}{2}} \sum_{i=0}^{n-1} (\rho_i^t f_i^* f_i) D^{\frac{1}{2}} - \frac{\vec{\mathbb{1}} D}{\operatorname{vol} G} \right\| \\
&= \left\| fD^{-\frac{1}{2}} \sum_{i=1}^{n-1} (\rho_i^t f_i^* f_i) D^{\frac{1}{2}} + fD^{-\frac{1}{2}} \rho_0 f_0^* f_0 D^{\frac{1}{2}} - \frac{\vec{\mathbb{1}} D}{\operatorname{vol} G} \right\| \\
&= \left\| fD^{-\frac{1}{2}} \sum_{i=1}^{n-1} (\rho_i^t f_i^* f_i) D^{\frac{1}{2}} \right\| \\
&\leq \max_{i>0} |\rho_i|^t \cdot \left(\frac{\max_x \sqrt{d_x}}{\max_y \sqrt{d_y}} \right).
\end{aligned}
$$

If $|\rho_i| < 1$ for $i \neq 0$, then this inequality gives the existence and uniqueness of a stationary distribution, as well as telling us that our iterative process converges, and a bound on the rate of its convergence. This condition ($|\rho_i| < 1$ for $i \neq 0$) does not hold in general, and so it seems natural now to consider for which classes of graphs this condition is satisfied.

References

[1] Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1999.