October 19, 2005
Kevin Costello

Random Walks on Directed Graphs

1 Comments on a Method of Santosh Ventala

The Eigencluster Algorithm given by Santosh Ventala is essentially a recursive algorithm; it uses spectral decomposition to compute an approximate best cut into two parts, then recursively applies the algorithm to each piece of the cut. This may be inefficient, as you are forced to keep computing eigenfunctions after every division. There are possibly simpler approaches

The algorithm is essentially based on finding a good cut in the graph. Finding "good cuts" (an ambiguous term) is the mother of all algorithms. This is the basis for the divide and conquer technique used in clustering, sparse approximations, and so on.

2 Random Walks on Directed Graphs – An Example

A good example to start with is the directed cycle C_n. This is the Cayley graph of the integers mod n, with an edge directed from i to $i+1$, so that the adjacency matrix is a circulent matrix with

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Since for $\theta = e^{\frac{2\pi ij}{n}}$ we have $A \begin{pmatrix} 1 \\ \theta \\ \theta^2 \\ \vdots \\ \theta^{n-1} \end{pmatrix} = \theta \begin{pmatrix} 1 \\ \theta \\ \theta^2 \\ \vdots \\ \theta^{n-1} \end{pmatrix}$, all eigenvalues of A are n^{th} roots of unity and all have norm 1. In this case, there is no single dominating eigenvalue, which is not surprising since this not-so-random "random walk" doesn't converge to any stationary distribution but oscillates the initial distribution with a period of n.

The cycle is an example of a graph which corresponds to a group. Groups are useful because we can give a short description to a very large structure. Starting with elements of our group as vertices and placing edges via generators we can produce Cayley graphs. Such graphs usually have high symmetry.

Another way to give a short description to a large graph is to assume it can be approximated somehow by a "random-like" structure (as in [1]). This leads to random graphs which we discussed in the last lecture.

Our goal is to understand complicated things with complicated structure (e.g., the internet graph or the telephone graph). There is little hope of doing this directly, so we use combinations of groups and random graphs to give simple models we can work with. Our goal, as always, is to make the graph easier to work with, but powerful enough to still give meaningful results.

3 Markov Chains and Transition Probabilities for a Directed Walk

The general rule for convergence of Markov Chains is that they are "ergodic" (meaning the random walk converges) if and only if they are both "irreducible" and "aperiodic". Our goal here is to translate the irreducibility and aperiodicity requirements to properties of a graph that will determine the convergence of the random walk on it.

Recall that the transition matrix for a random walk on an (unweighted) directed graph is:

$$P(u,v) = \begin{cases} \frac{1}{d_u} & \text{if } u \to v \text{ is an edge in } G; \\ 0 & \text{otherwise;} \end{cases}$$

where here d_u denotes the out-degree of vertex u (we will use d_u^- if we want to denote the in-degree of u). The transition for a weighted version is similar, only we replace d_u by $\sum_{u \to v} w_{uv}$ and the transition probabilities by

$$P(u,v) = \begin{cases} \frac{w_{uv}}{d_u} & \text{if } u \to v \text{ is an edge in } G; \\ 0 & \text{otherwise;} \end{cases}$$

We still have $P = D^{-1}A$, as in the undirected case. When G was undirected, however, A was a symmetric matrix, so the corresponding Laplacian $D^{-1/2}AD^{-1/2}$ was also symmetric, guaranteeing that its spectrum had nice properties (a complete orthonormal basis of real eigenvectors, for example). In the directed case A is no longer symmetric, so we lose many of these properties. Some people try to symmetrize P by considering the spectrum of, for example, $P + P^*$ or $PP*$. Certainly this will get back to symmetric matrices, but the problem then is to take the eigenvalue that is calculated and relate it back to the graph. Without a graphical meaning to these numbers the approach won't be useful.

4 The Perron-Frobenius Theorem

Although the transition matrix of a random walk on a directed graph may not be symmetric, it does have the useful property that all of its entries are non-negative. The Perron-Frobenius theorem gives useful information about the eigenvalues of such matrices in many cases.

Theorem 1. *Let M be a matrix with nonnegative entries. Assume furthermore that M is **irreducible**, meaning that it is impossible to permute the rows and columns of M to place it in the form*

$$M' = \begin{pmatrix} A & 0 \\ B & C \end{pmatrix}$$

with the upper right block having dimension $k \times (n-k)$ for some k. Then there is a real $\rho_0 > 0$ such that the following hold:

1. ρ_0 is an eigenvalue of M, and all other eigenvalues satisfy $|\rho| \leq \rho_0$.

2. The eigenvector corresponding to the eigenvalue ρ_0 has nonnegative entries.

3. If there are $k-1$ other eigenvalues with $|\rho_i| = \rho_0$, then they are of the form $\rho_0 \theta^j$, where $\theta = e^{\frac{2\pi i}{k}}$.

Exercise: Prove the Perron-Frobenius theorem. Hint: this is not trivial (a proof can be found in [2].)

5 Translating Perron-Frobenius into Graph Theoretic Terms

In order to apply Perron-Frobenius to a random walk on a graph, we need to make sure that the matrix we apply it to is irreducible. With this in mind, we have the following:

Lemma 1. *A graph G is **strongly connected** if any of the following equivalent conditions hold:*

1. For every u and v in $V(G)$ there exist paths in G from u to v and from v to u

2. For any partition of $V(G)$ into two disjoint nonempty sets A and B there is an edge from A to B and an edge from B to A.

3. The adjacency matrix A of G is irreducible.

4. The transition probability matrix P of G is irreducible.

Condition 1 is usually taken as the definition of strongly connected.

Exercise: Show that the four definitions above are equivalent. Some of these are obvious, but the most interesting one is showing that $2 \Rightarrow 1$.

The other condition for convergence in a Markov Chain was aperiodicity. For a graph we will again describe it by a series of equivalent definitions:

Lemma 2. *A strongly connected graph G is* ***periodic*** *if any of the following equivalent conditions hold:*

1. The transition matrix P has an eigenvalue $\rho \neq 1$ such that $|\rho| = 1$.

2. There is a $k > 1$ such that all eigenvalues with norm 1 are of the form $e^{\frac{2\pi ij}{k}}$.

3. There is an edge-preserving map from $V(G)$ to the vertices of C_k with $k > 1$ (a map such that $u \to v$ in G implies $f(u) \to f(v)$ in C_k).

4. The GCD of all cycle lengths in G is $k > 1$.

Note that 1 will always be an eigenvalue of P since by construction we have for any fixed u that $\sum_v P(u, v) = 1$. In other words, $P\mathbf{1} = \mathbf{1}$, so the all 1 vector is a right eigenvector of P with eigenvalue 1.

$1 \Leftrightarrow 2$ follows from the Perron-Frobenius theorem. We have $3 \Rightarrow 4$ since cycles in G get mapped to cycles in C_k (so the GCD is certainly at least k). To show $4 \Rightarrow 3$ we simply pick the destinations for each vertex in G inductively by first sending an arbitrary vertex x to 0. If u has already been placed and $u \to v$, we set $f(v) = f(u) + 1$, condition 4 guarantees that these placings are well-defined.

Exercise: Complete the proof of equivalence for Lemma 2. In particular, show that $3 \Leftrightarrow 1$.

As a general rule it is good to have many equivalent definitions. This allows us to look at problems from many different angles. This also gives us some flexibility in that we can use the easiest definition that might apply in a particular case. This is the motivation behind quasirandom graphs, which enable us to pick whichever of many equivalent properties is easiest to check or work with.

We can now restate the Perron-Frobenius and convergence theorems in terms of these two properties.

Theorem 2. *A strongly connected directed graph G has a maximum eigenvalue ρ_0, and that eigenvalue has a corresponding non-negative eigenvector. Furthermore, $|\rho_i| < \rho_0$ for all $i > 1$ iff the GCD of the cycle lengths of G is 1.*

Theorem 3. *The random walk on a directed graph G converges to a unique stationary distribution if G is strongly connected and is not periodic.*

Our cycle graph was strongly connected, but it was also periodic. This is what led to the problems in the spectrum and the oscillation of the random walk.

References

[1] F. Chung, L. Lu, and V. Vu, "The Spectra of Random Graphs With Given Expected Degree" , *Proceedings of National Academy of Sciences* (2003), 100:6313-6318.

[2] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1999.