Random Walks on Graphs

Larry Fenn

DATE

Introduction

Goal: Finding the "center" of a graph (ex. the most popular person, or most influential group).
Example idea: Degree centrality.
Approach:

- Use abstract properties of the adjacency matrix.
- Figure out how to compute these properties.
- Figure out how to proceed even without the adjacency matrix!

Premises

Graph G is:

- Simple: Undirected, unweighted, with no graph loops or multiple edges between any two vertices.
- Connected: There is a path from any vertex to any other in the graph.
- Aperiodic: There is no integer $k > 1$ such that for any cycle in the graph k divides its length.
- Finite.

"Influence" centrality (ex. PageRank)

Degree centrality is too simple.
A vertex could be central without having the highest degree (you may only have two friends, but if your two friends are Barack Obama and Vladimir Putin...)
The centrality c_i of vertex i should take into consideration the centrality of its neighbors.
For some fixed K:

$$c_i = \frac{1}{K} \sum_{\text{all neighbors}} c_j$$

Eigenvalues & Eigenvectors

Given an $n \times n$ matrix $\boldsymbol{A}$:

- Scalar value λ and vector ψ are called *eigenvalues* and *eigenvectors* respectively of $\boldsymbol{A}$ if

$$\boldsymbol{A}\psi = \lambda\psi$$

- There will be from 1 to n distinct eigenvalue/eigenvector pairs.
- They are properties of the matrix (equivalently, the linear map the matrix represents).

The *spectral radius* of $\boldsymbol{A}$ is $\rho(\boldsymbol{A}) = \max_i |\lambda_i|$.

Requirements

The Perron-Frobenius theorem is a linear algebra theorem about eigenvalues and eigenvectors.
The version of the Perron-Frobenius theorem we will use has the following hypothesis for a matrix $\boldsymbol{A}$:

- $\boldsymbol{A}$ is a non-negative $n \times n$ matrix.
- $\boldsymbol{A}$ must be irreducible.
- $\boldsymbol{A}$ must be an aperiodic matrix.

In fact, the full theorem has far weaker hypothesis.

Graph-theoretic hypothesis

If we are working with a graph adjacency matrix $\boldsymbol{A}$, then the requirements mean:

- Non-negative: $\boldsymbol{A}$ represents a simple graph (since all entries of $\boldsymbol{A}$ are either 0 or 1).
- Irreducible: $\boldsymbol{A}$ represents a connected, undirected graph.
- Aperiodic: $\boldsymbol{A}$ represents a graph G where the greatest common divisor of all cycle lengths for cycles in G is 1.

Here, too, the graph conditions are stronger than what is needed to satisfy the hypothesis.

Statement of theorem

Perron-Frobenius theorem, special case
If the $n \times n$ matrix $\boldsymbol{A}$ is a non-negative, irreducible, and aperiodic, then the following hold:

- $\rho(\boldsymbol{A})$ is a positive number, and it must be an eigenvalue of $\boldsymbol{A}$.
- $\rho(\boldsymbol{A})$ is simple: it is distinct from the other eigenvalues.
- The eigenvector $\boldsymbol{\psi}$ associated with $\rho(\boldsymbol{A})$ has all positive components.
- The *only* eigenvector of $\boldsymbol{A}$ with all positive components is $\boldsymbol{\psi}$.

Proof is nontrivial.

Significance of ψ

ψ is the "influence" centrality measure we wanted earlier.

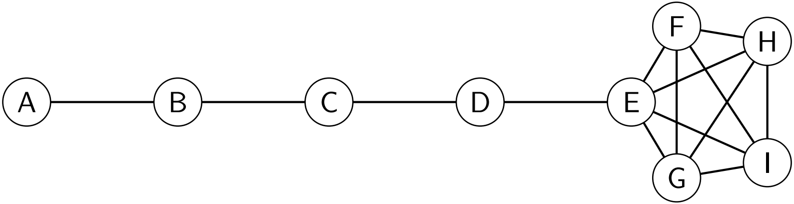
$$\boldsymbol{A}\psi = \rho(\boldsymbol{A})\psi$$

$$\frac{1}{\rho(\boldsymbol{A})}\boldsymbol{A}\psi = \psi$$

Take one component:

$$\frac{1}{\rho(\boldsymbol{A})}\sum_{j=1}^{n}\boldsymbol{A}_{ij}\psi_j = \psi_i$$

Example: “Lollipop”



$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \quad \psi = \begin{bmatrix} .002 \\ .008 \\ .032 \\ .122 \\ .462 \\ .439 \\ .439 \\ .439 \\ .439 \end{bmatrix}, \rho = 4.055$$

Requirements

The power iteration algorithm is an algorithm that finds an eigenvector and eigenvalue of a matrix $\boldsymbol{A}$.
In particular, it finds the eigenvector associated with the eigenvalue of largest absolute value.
The power iteration algorithm will converge (or have a subsequence that converges) to an answer subject to the following hypothesis:

- $\boldsymbol{A}$ has an eigenvalue *strictly greater* in absolute value than all of its other eigenvalues.
- The initial guess eigenvector $\vec{\boldsymbol{b}}_0$ satisfies $\vec{\boldsymbol{b}}_0 \cdot \psi \neq 0$.

The algorithm

The algorithm:

$$\vec{\boldsymbol{b}}_{k+1} = \frac{\boldsymbol{A}\vec{\boldsymbol{b}}_k}{\|\boldsymbol{A}\vec{\boldsymbol{b}}_k\|}$$

Proof sketch: Express $\vec{\boldsymbol{b}}_0$ over the eigenbasis for $\boldsymbol{A}$ ($\boldsymbol{A}$ is symmetric!), then apply $\boldsymbol{A}^k$ and use linearity & eigenvector properties.
Perron-Frobenius: $\rho(\boldsymbol{A})$ is both an eigenvalue of $\boldsymbol{A}$ and it is strictly greater than all of the other eigenvalues ($\rho(\boldsymbol{A})$ is simple).
ψ has all positive components, so take as an initial guess any positive vector.

Definition

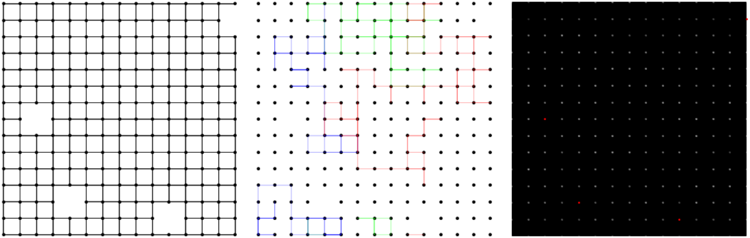
We can define a random walk process by assigning probabilties for travel from one vertex to another.

Basic random walk: uniformly select a neighbor.

Long-term behavior: diffuses to every part of the graph.

Example:

The following graphs are both horizontally and vertically periodic.



(a) Underlying graph (b) Three walks (c) Overall frequency

Definition

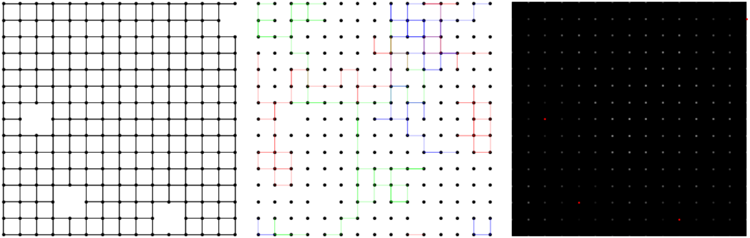
Rather than select uniformly among neighbors, we can select uniformly among *paths*.

Determine how many paths leave each neighboring vertex, and weigh the probability of travel to that vertex accordingly.

Long-term behavior: tends towards the more well-connected parts of the graph.

Example:

The following graphs are both horizontally and vertically periodic.



(a) Underlying graph (b) Three walks (c) Overall frequency

Derivation

Transition probability P_{ij} should be defined as:

$$P_{ij} = \lim_{k\to\infty} \frac{\boldsymbol{A}_{ij} \sum_{x=1}^{n} \boldsymbol{A}_{jx}^{k-1}}{\sum_{j'=1}^{n} \boldsymbol{A}_{ij'} \sum_{x=1}^{n} \boldsymbol{A}_{j'x}^{k-1}}$$

Intuitively: the denominator is all paths of increasing length (in the limit, infinite length) leaving i. The numerator is only paths leaving i that route through one of it's neighbors j.

Result

The transition probability P_{ij} ends up being:

$$P_{ij} = \frac{1}{\rho(\boldsymbol{A})} \frac{\psi_j}{\psi_i}$$

Proof sketch: Use matrix multiplication & the power iteration algorithm to compute the limit.

Motivation

Often, $\boldsymbol{A}$ is not explicitly known.

Example: a social network.

Thus, $\boldsymbol{\psi}$ and eigenvector centrality is not known.

Maximal entropy random walk tends towards the well-connected parts of a graph.

But the probabilities are defined based on $\boldsymbol{\psi}$.

Can we approximate centrality without global information?

Approximations of P_{ij}

Cut off the limit at fixed values of k to approximate P_{ij}; call the approximation P_{ij}^k

k represents something like search depth.

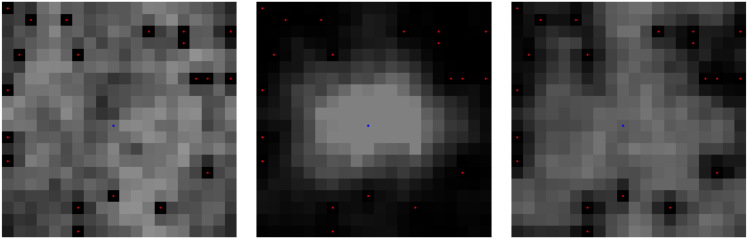
For reference:

$$P_{ij}^0 = \frac{\boldsymbol{A}_{ij}}{\deg v_i},\ \ P_{ij}^1 = \frac{\boldsymbol{A}_{ij} \deg v_j}{\displaystyle\sum_{j'=1}^{n} \boldsymbol{A}_{ij'} \deg v_{j'}},\ \ P_{ij}^2 = \frac{\boldsymbol{A}_{ij} \displaystyle\sum_{x=1}^{n} \boldsymbol{A}_{jx} \deg v_x}{\displaystyle\sum_{j'=1}^{n} \boldsymbol{A}_{ij'} \sum_{x=1}^{n} \boldsymbol{A}_{j'x} \deg v_x}$$

P_{ij}^0 is just the basic random walk.

In practice, P_{ij}^2 is “close enough”.

Comparison



(a) Basic (b) Maximal Entropy (c) Approximation

Conclusion

To find the “center” of a graph:

- If $\boldsymbol{A}$ is known, solve for ψ, using the power iteration algorithm if necessary.
- If $\boldsymbol{A}$ is not known, construct an approximation to the maximal entropy random walk and “follow” it.

Applications:

- Disease prediction.
- Infrastructure planning.
- Popularity contests.

Questions