

Let $G = (V, E)$ be an undirected connected graph. By *random walk* on such a graph we simply mean the random walk on the directed graph obtained by replacing each edge of G with two oppositely directed edges. So this walk is on a strongly connected directed graph. So we already know that $\mathbf{1}$ is a right eigenvector, and the eigenvalue 1 has multiplicity 1, and the other values lie between -1 (inclusive) and 1 (exclusive).

It turns out that we can immediately write down the left eigenvector of eigenvalue 1. This is simply the vector d of degrees, i.e. d_i is the degree of vertex i . To see this, note that as before the walk matrix is $W = D^{-1}A$ where D is a diagonal matrix with the ii entry equal to d_i , and A the adjacency matrix. Now note that $dD^{-1} = \mathbf{1}$. Thus $dW = dD^{-1}A = \mathbf{1}A = d$. By scaling we can make the entries add up to one, giving the vector π for the stationary distribution for the walk, i.e. $\pi_i = d_i / \sum_j d_j$.

It is not necessary, however, that if we start with an arbitrary distribution we will converge to π . Indeed, by taking the graph to be a single edge, and starting at one of the nodes, the distribution will be $(1, 0)$ and $(0, 1)$ alternately. The unique stationary distribution $(0.5, 0.5)$ will not be reached. The key problem here is that the walk is *periodic*, at even steps it is at one node and at odd steps at the other.

1 Lazy Random Walk

In this, at each step we toss a fair coin. If coin = head, we do nothing (hence lazy). If coin = tail, we pick a neighbour at random and move there. Let x^0 be the initial probability distribution, and x^t the distribution at step t . Then the probability of being at node j in step $t + 1$ can be written as:

$$x_j^{t+1} = 0.5x_j^t + 0.5 \sum_i x_i^t a_{ij} / d_i$$

So we define the lazy walk matrix Z :

$$Z = 0.5 \cdot (I + D^{-1}A)$$

whereupon we have $x^{t+1} = x^t Z = x^0 Z^t$. We first observe that Z is also stochastic and has the same eigenvectors as $W = D^{-1}A$. Suppose x is

an (row) eigenvector of W of eigenvalue α . Then $xZ = 0.5x(I + W) = 0.5x + 0.5xW = x0.5(1 + \alpha)$. Hence x is an eigenvector of eigenvalue $\beta = 0.5(1 + \alpha)$. Thus it follows that there is a one to one correspondence between the eigenvalues of W and Z . Since W is known to have eigenvalue 1 with multiplicity 1, the corresponding eigenvalue for Z is $0.5(1 + 1) = 1$ and clearly this has multiplicity 1. Note further that the eigenvalues of W lie in the range $[0,1]$. Thus the eigenvalues of Z must lie in the range $[0.5(1 + (-1)), 0.5(1 + 1)] = [0, 1]$.

Theorem 1 *Consider the lazy random walk on any connected graph G , with d_i denoting the degree of vertex i . Let π to be the (row) vector $d/\sum_i d_i$, where $d = (d_1, \dots, d_n)$. Then π is the unique stationary distribution for the (ordinary) walk with matrix W and the lazy walk with matrix Z . Further, starting from any initial distribution x^0 , let $x^{t+1} = x^t Z$ be the distribution for the lazy random walk at step $t + 1$. Then x^t converges to π . Specifically*

$$\|\pi - x^t\| \leq \sqrt{(\max_j d_j / \min_j d_j) \beta_2^t}$$

where β_i is the i th largest eigenvalue of Z .

Note that for the ordinary walk we may not have convergence to the stationary distribution; the walk may cycle, e.g. if the walk is on C_2 , starting at $[1,0]$ the probability distribution will alternate between $[1,0]$ and $[0,1]$, even though the stationary distribution $[0.5,0.5]$ is unique.

Proof: Uniqueness is already proved.

We prove the theorem only when the graph is k -regular. In this case $W = D^{-1}A = A/k$, where A is the adjacency matrix. Since A is symmetric, so are W and $Z = (I + W)/2$. Also $\pi = \mathbf{1}/n$.

We note the following important facts about real symmetric matrices. An $n \times n$ real symmetric matrix M has n eigenvalues (including multiplicity), and a corresponding set of eigenvectors which are orthogonal. In other words, $M = Q^T \Lambda Q$ where Q is an orthogonal matrix (rows are unit vectors and mutually orthogonal), Λ is a diagonal matrix with eigenvalues of M on the diagonal. Further, the rows of Q are the eigenvectors of M .

So we may write Z as $Q^T \Lambda Q$. Let q_i denote the i th row of Q , i.e. the i th eigenvector. Assume that the eigenvectors are arranged in non-increasing order of the eigenvalues. Thus Λ has $(\beta_1, \beta_2, \dots, \beta_n)$ along the diagonal. We already know $d/\sum d_i = \mathbf{1}/n$ is an eigenvector of eigenvalue 1. But since we require q_i to be unit vectors, we normalize and get $q_1 = \mathbf{1}/\sqrt{n}$.

The key idea of the proof is to consider everything in the basis Q . In this basis, multiplying by Q is equivalent to scaling the components along q_i by β_i . Thus the component along q_1 will remain the same, while the components along the other directions will decrease, giving us the final distribution to be in the direction q_1 .

Algebraically, we have $x^t = x^0 Z^t$. But $Z^t = (Q^T \Lambda Q)^t = Q^T \Lambda^t Q$ noting that $Q^T Q = I$. Thus

$$x^t = x^0 Q^T \Lambda^t Q = \sum_i (x^0 \cdot q_i^T) \beta_i^t q_i$$

Note that $x^0 \cdot q_i^T$ is simply the component of the vector x^0 in the direction q_i which we then scale by β_i^t as explained. Noting that $\beta_1 = 1$ we get:

$$x^t = (x^0 \cdot q_1^T) q_1 + \sum_{i=2}^n (x^0 \cdot q_i^T) \beta_i^t q_i$$

The first term is $(x^0 \cdot \mathbf{1}^T / \sqrt{n}) \mathbf{1} / \sqrt{n} = \mathbf{1} / n = \pi$. Thus the second term in the above expression gives the vector difference between x^t and the stationary distribution π . Since q_i are orthogonal, we can find the Euclidean length of the difference vector by the Pythagorean formula.

$$\|x^t - \pi\| = \left\| \sqrt{\sum_{i=2}^n ((x^0 \cdot q_i^T) \beta_i^t)^2} \right\| \leq \beta_2^t \left\| \sqrt{\sum_{i=2}^n ((x^0 \cdot q_i^T))^2} \right\| \leq \beta_2^t \|x^0\|$$

The last inequality follows because the length of x^0 is the same sum but starting from $i = 1$. Finally, $\|x^0\| \leq 1$ since it is a probability vector. ■