

## PROBLEM SET 2

MARK DITSWORTH

### 1. PROBLEM 1

Given graph  $G = (V, E, W)$  consider a random walk on  $V$  with transition probabilities  $M_{ij} = PX(t+1) = j|X(t) = i = \frac{w_{ij}}{\deg(i)}$ .

Partition the vertex set as  $V = V_+ \cup V_- \cup V_*$ . Suppose that every node in  $V_*$  is connected to at least one node in either  $V_+$  or  $V_-$ . Given a node  $i \in V$  let  $g(i)$  be the probability that a random walker starting at  $i$  reaches a node in  $V_+$  before reaching one in  $V_-$ . If  $i \in V_+$ , then  $g(i) = 1$  and if  $i \in V_-$ , then  $g(i) = 0$ . Find  $g(i)$  for  $i \in V_*$ .

There are two scenarios in which the random walker will reach a node in  $V_+$  before one in  $V_-$ : Moving from  $V_*$  to  $V_+$  immediately, or moving from  $V_*$  to  $V_*$  repeatedly until moving to  $V_+$ . We essentially view leaving  $V_*$  as an absorbing barrier, and want the probability that we end in  $V_+$ , which is the sum of all probabilities that will end in that scenario:

$$\sum_{j \in V_+} M_{ij} + \sum_{j \in V_*} \sum_{k \in V_+} M_{ij} M_{jk} + \sum_{j \in V_*} \sum_{k \in V_*} \sum_{l \in V_+} M_{ij} M_{jk} M_{kl} + \dots$$

Define  $\nu_* \in \mathbb{R}^n$ ,  $n = |V|$  where the  $j$ th element is 1 if  $j \in V_*$ , and 0 otherwise. Similarly define  $\nu_+ \in \mathbb{R}^n$  for the partition  $V_+$ . Let  $\Psi \in \mathbb{R}^n$ ,  $n = |V|$  have elements everywhere equal 0 except for the  $i$ th node that is the starting node. The probability that a random walker starting at node  $i$  is absorbed in the  $V_+$  partition at time-step  $k$  is expressed as

$$\Psi^T [M \text{diag}(V_*)]^{(k-1)} M V_+$$

The  $[M \text{diag}(V_*)]^{(k-1)}$  matrix expresses the probability of starting at node  $i$ , and ending at node  $j$  after  $k-1$  steps, discounting the intermediate nodes that would absorb the random walker earlier (nodes belonging to  $V_+$  or  $V_-$ ). For convenience, we will denote this matrix as  $M_*$ , as in only accounting for transitions within partition  $V_*$ . The resulting matrix is dotted with  $M$  for the last transition time step, and then dotted with  $V_+$  to attain the vector of probabilities end in partition  $V_+$ .  $\Psi^T$  dotted with this vector selects the probability stemming from starting at node  $i$ . Thus, summing the probabilities across all  $k$  yields

$$\Psi^T \left[ \sum_{k=0}^{\infty} (M_*)^k \right] M V_+$$

With each element of  $M_*$  less than 1, and the summation across each row bounded above by 1, it is clear that  $\lim_{n \rightarrow \infty} (M_*)^n = \mathbf{0}$ . Thus, the infinite sum converges to

$(I - M_*)^{-1}$ . In conclusion, the probability that a random walker starting at node  $i \in V_*$  will reach a node in  $V_+$  before a node in  $V_-$  is calculated by

$$g(i) = \Psi^T (I - M_*)^{-1} M V_+.$$

Note that for graphs with large  $|V|$ , the inverse operation can be quite expensive. Thus  $g(i)$  should be approximated either via pseudo-inverse operations or monte carlo simulations.

## 2. PROBLEM 2