Math for CS 2015/2019 solutions to "In-Class Problems Week 14, Wed. (Session 35)"

https://github.com/spamegg1

April 19, 2023

Contents

1 Problem 1

1.1 (a)

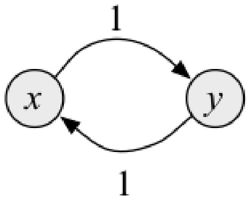


Figure 1

Find a stationary distribution for the random walk graph in Figure 1.

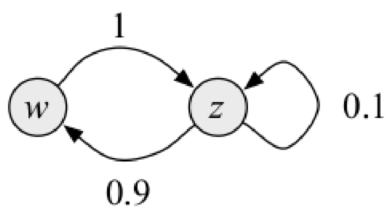
Proof. $d(x) = d(y) = 1/2$ □

1.2 (b)

Explain why a long random walk starting at node x in Figure 1 will not converge to a stationary distribution. Characterize which starting distributions will converge to the stationary one.

Proof. It won't converge to a stationary distribution, because you just alternate between nodes x and y. □

1.3 (c)



Find a stationary distribution for the random walk graph in this figure.

Proof. $d(w) = 9/19, d(z) = 10/19$.

You can derive this by setting

$d(w) = (9/10)d(z)$,

$d(z) = d(w) + (1/10)d(z)$, and

$d(w) + d(z) = 1$.

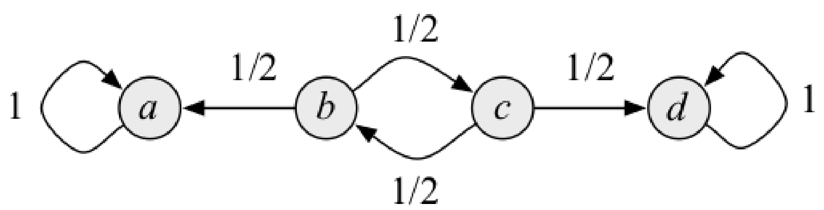
There is a unique solution. □

1.4 (d)

If you start at node w above and take a (long) random walk, does the distribution over nodes ever get close to the stationary distribution? You needn't prove anything here, just write out a few steps and see what's happening.

Proof. Yes, it does. □

1.5 (e)



Explain why the random walk graph in this figure has an uncountable number of stationary distributions.

Proof. For any real number $0 < p < 1$ there is a stationary distribution: $d(b) = d(c) = 0$, $d(a) = p$, $d(d) = 1 - p$. $\square$

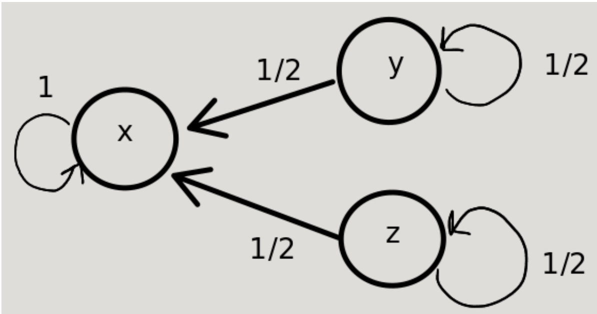
1.6 (f)

If you start at node b in the last figure and take a long random walk, the probability you are at node d will be close to what fraction? Explain.

Proof. 1/3. $\square$

1.7 (g)

Give an example of a random walk graph that is not strongly connected but has a unique stationary distribution. Hint: There is a trivial example.



Proof. Consider this graph. It's not strongly connected (there is no directed path between y and z).

To solve for the stationary distribution, we have the equations:

$$\begin{aligned} x &= x \cdot 1 \\ y &= (1/2) \cdot y \\ z &= (1/2) \cdot z \end{aligned}$$

The last two equations force $y = z = 0$, which forces $x = 1$. This is a unique solution, so it's a unique stationary distribution. $\square$

2 Problem 2

Prove that for finite random walk graphs, the uniform distribution is stationary iff the probabilities of the edges coming into each vertex always sum to 1, namely

$$\sum_{u \in \text{into}(v)} p(u, v) = 1$$

where $\text{into}(w) ::= \{v \mid \langle v \to w \rangle \text{ is an edge}\}$.

Proof. 1. Assume G is a finite random walk graph with vertex set $V = \{v_1, \ldots, v_n\}$.

2. Assume the uniform distribution $\Pr[\text{at } v_i] = \frac{1}{n}$ on G is stationary.

Want to prove $\sum_{u \in \text{into}(v_i)} p(u, v_i) = 1$ for all $1 \leq i \leq n$.

3. Since the distribution is stationary, $\Pr[\text{at } v_i] = \Pr[\text{go to } v_i \text{ at next step}]$.

So by (2) we have $\frac{1}{n} = \Pr[\text{go to } v_i \text{ at next step}]$.

4. Notice that by definition, $\Pr[\text{go to } v_i \text{ at next step}] = \sum_{u \in \text{into}(v_i)} (p(u, v_i) \cdot Pr[\text{at } u])$.

So we have $\frac{1}{n} = \sum_{u \in \text{into}(v_i)} (p(u, v_i) \cdot Pr[\text{at } u])$ by (3).

5. Again, since the distribution is stationary, $\Pr[\text{at } u] = \frac{1}{n}$ for all $u \in \text{into}(v_i)$ and all $1 \leq i \leq n$.

6. Combining 4,5 we get $\frac{1}{n} = \sum_{u \in \text{into}(v_i)} (p(u, v_i) \cdot Pr[\text{at } u]) = \frac{1}{n} \cdot \sum_{u \in \text{into}(v_i)} p(u, v_i)$.

7. Cancelling $1/n$ we get $\sum_{u \in \text{into}(v_i)} p(u, v_i) = 1$.

8. Conversely, assume $\sum_{u \in \text{into}(v_i)} p(u, v_i) = 1$ for all $1 \leq i \leq n$.

Want to prove that the uniform distribution $\Pr[\text{at } v_i] = \frac{1}{n}$ on G is stationary.

The proof is very similar to the above steps. □

3 Problem 3

A Google-graph is a random-walk graph such that every edge leaving any given vertex has the same probability. That is, the probability of each edge $\langle v \to w \rangle$ is $1/\text{outdeg}(v)$.

A digraph is symmetric if, whenever $\langle v \to w \rangle$ is an edge, so is $\langle w \to v \rangle$. Given any finite, symmetric Google-graph, let $d(v) ::= \text{outdeg}(v)/e$ where e is the total number of edges in the graph.

3.1 (a)

If d was used for webpage ranking, how could you hack this to give your page a high rank? ...and explain informally why this wouldn't work for "real" page rank using digraphs?

Proof. ??? □

3.2 (b)

Show that d is a stationary distribution.

Proof. 1. Assume there are e edges in total in the graph G. We need to show for evert vertex v: $\Pr[\text{at } v] = \Pr[\text{go to } v \text{ at next step}]$. Assume v is a vertex in G.

2. By the definition of our distribution we have $\Pr[\text{at } x] = d(x) = \text{outdeg}(x)/e$ for all vertices x. In particular $\Pr[\text{at } v] = d(v) = \text{outdeg}(v)/e$.

3. Since G is symmetric, $\text{outdeg}(x) = \text{indeg}(x)$ for all vertices x. So in particular, $\Pr[\text{at } v] = \text{indeg}(v)/e$.

4. Also remember that by definition of a Google-graph we have $p(u, v) = 1/\text{outdeg}(u)$ for all other vertices u.

5. Then by (2), (3), (4), $\Pr[\text{go to } v \text{ at next step}]$ is equal to:

$$\sum_{u \in \text{into}(v)} Pr[\text{at } u] \cdot p(u, v) = \sum_{u \in \text{into}(v)} \frac{\text{outdeg}(u)}{e} \cdot \frac{1}{\text{outdeg}(u)} = \frac{|\text{into}(v)|}{e} = \frac{\text{indeg}(v)}{e}$$

6. By (3) and (5) we see that $\Pr[\text{at } v] = \text{indeg}(v)/e = \Pr[\text{go to } v \text{ at next step}]$. So d is a stationary distribution. □