

PROFESSOR: So some of the standard questions that we've examined already about random graphs are the probability of getting from one place to another, or the expected time to get from one place to another. But a different kind of question that comes up in a fundamental way is the probability of being someplace. So let's examine that.

Here is the graph with states blue, orange, and green that we've seen before. And suppose that I start at state B. And I ask, what's the probability of being at each of these states after one step?

So to start with, I'm interested in p_B , p_O , and p_G -- which is the probability of being at state B, the probability of being at state O, and the probability of being at state G. The sum of the probabilities is going to be 1. And initially when I tell you that I'm at state B, it means the probability of being at B is 1, and the other two is 0.

And I'm interested in the way that these probabilities update after one step. If p'_B is the probability of being in state B after one step, and p'_O is the probability of being in the orange state one step later-- and likewise for green-- what are these probabilities? Well it's easy to see just reading off this graph that the only place you're at is B. So the only way to get probability of being somewhere is by following an edge out of B.

So the probability of being at one step at the orange vertex is $1/4$. And it's likewise $1/4$ for being at the green state. And it's $1/2$ for staying at the blue state. So what we can say is that the updated probabilities of being at these different states is $1/2$, $1/4$, and $1/4$, as we've just reasoned.

OK. Let's keep going. Given that the probability that I'm at the states blue, orange, and green are given by this vector of probabilities, what's the distribution after two steps? So let p''_B be the probability of being at state B after two steps, starting from B. Well, the way we can figure that out is by using conditional probabilities.

Let's look at the example of calculating the probability of being in the orange state two steps after you've started at the blue state. So here was the probabilities of being at the different states after one step. How do I get to the orange state?

Well I can get to the orange state from the blue state. And so the probability of being in the

orange state after two steps is the probability of being at the blue state after one step times the probability that I take this edge to the orange state. That is, it's the probability of going from B to O-- given that I'm at B-- times the probability of being in B after one step. This then is the probability of being in O after two steps.

And likewise, another component of the probability of being at O is that if you're at O, and what's the probability of going from O to O? And that is this $1/3$ times the probability of being at O at all, which is $1/4$.

And the final case, using again the law of total probability, breaking it up into cases, the third way that I can get to the orange state on step two is by being at the green state on step one following the green to O edge-- of which there isn't any, so that's going to be probability 0-- times the probability of being at the green state, which is $1/4$, but it won't matter.

So let's just fill in these amounts looking at the first term. The probability of going from B to O when you're at B is simply the probability of this edge. It's $1/4$. And likewise, the probability of going from O to O given that you're at O is the probability of this edge-- namely $1/3$. So we can fill that term in. And finally the probability of going from G to O is 0, given that you're at G, because there isn't any vertex there. And then you fill in those probabilities and do the arithmetic. You come out with $5/24$ probability of being in the orange state after two steps.

Well the same calculation, you can figure out what's the probability of being at the blue state or the green state after two steps. And there's the answer. There's a $50/50$ chance of being at the blue state after two steps, $5/24$ as we saw at the orange state, and the rest of it is $7/24$ is the probability of being at the green state.

OK. So what's going on in general? And we can explain how to do these calculations by using a little bit of linear algebra. So let's define the edge probability matrix of a random walk graph is just the adjacency matrix of the graph, except that instead of using 0's and 1's to indicate whether an edge is not present or present, I'll use in the i, j position of the matrix-- the probability of the edge that goes from i to j -- if there is an edge-- and 0 if there isn't any edge.

Let's look at an example. So here is the way we'd fill it in abstractly for our three-state graph. It'll be a 3 by 3 matrix with the probabilities of the successive edges in the corresponding position. So this is the position in the B, B coordinate is the probability of the edge from B to B. The O, B coordinate, if you think of the columns as labeled blue, orange, green; and the rows as labeled blue, orange, green. And this is the orange, blue coordinate. And it's the probability

of the edge from O to B.

Let's fill in the first row, which was-- this is just read directly off the graph. It was the edges out of B that went from B to B, from B to O, and from B to green-- G. And it had those weights. And if I fill in the rest of it, I get the edge probability matrix for our simple three-state graph. And there it is.

So this last one shows the fact that there is a certain edge from green to blue. The only place you can go from green is to blue, and you can't go to either orange or green in one step.

OK. So why are we bringing up the matrix? Well if you looked at the way we updated the state to go from the one-step distribution to the two-step distribution, it was really a matrix multiply. And in general, to do an update, you're just going to do a vector matrix multiplication.

If you have the probabilities of being in the successive states B, O, and G, and you do a vector matrix multiplication using the probability matrix of the graph, you get the updated vector of distributions. And that's easy to check just from the definitions, and from the definition of vector times matrix, which I assume you're familiar with.

So now we can ask what's the distribution after t steps, starting from some particular given distribution-- say, starting at state B, or starting at any possible distribution of probabilities to the different states. And the way that we can figure that out-- so I'm interested in other words is the probability of being in O after t steps G after t steps in B after t steps, say, starting from state B. And what happens also as t approaches infinity? And these are sort of two basic questions that we're going to be asking.

So first of all, how do you calculate starting at a given distribution p_B, p_O, p_G where you're going to be after t steps? Well, you're just continually updating, which means multiplying by M t times. So the distribution after t steps is gotten by taking the initial distribution times the t -th power of M .

Now this is actually already useful computationally, because it means that since you can compute a matrix power by successive squarings, you actually only need about \log of t matrix multiplications in order to be able to figure out what's the distribution of probabilities after t steps of the graph.

Then the crucial concept that we want to examine-- and we'll make a lot of use of in the next

video when we talk about a page rank-- is the idea of a stationary distribution. So a stationary distribution means that once you're in the stationary distribution, it's stable. You're going to stay in that distribution. You're not going to be in any particular state, but you'll have a vector of probabilities of being in the different states. And one step later, that vector's not going to change. So what it means is that the next-step distribution is the same as the current distribution.

What's the stationary distribution here? Well, the way we're going to have to calculate that is here's how you update. This is the result of the vector matrix multiplication. But let's just spell it out in terms of the conditional probabilities. After one step, if the original distribution is p_B , p_O , p_G , then the new probability of being in state B, the only way you can get there is by following the edge from B to B with probability $1/2$. And that's times the probability of being at B.

And the other way you can get to B is by being at the green state. And then one step later you're certain to be at B. So that adds a contribution of 1 times p_G . And likewise for p_O -- the updated probability of being at the orange state and the green state.

And what we want is that these updated probabilities are the same as the ones that I'm starting with. That's the definition of stability. You update the vector p_B , p_O , p_G , and you get the same vector. That's what makes it stable.

And of course, a side constraint. Since you can always solve a system of equations like this by letting all the p 's be 0 , which is degenerate, we add the constraint that the sum of the probabilities of being in the states has to be 1 . Well if we solve that simple 3 by 3 system of equations, then it turns out that the stable distribution is there's an $8/15$ chance of being in state B, a $3/15$ chance of being in state orange, and a $4/15$ chance of being in state green.

And you should check that yourself by asking what's the probability of being in p_B after one step given these probabilities? And I'm not going to talk you through that. But just to verify and imprint the idea of stability, that's one that's worth stopping the video for a moment to check and do a little arithmetic with a pencil and paper.

OK. So in general, what we're going to do is we're trying to find the stationary distribution vector-- call it \bar{s} , for vector. And we get this by solving the vector matrix equation-- that the distribution vector times the edge probability matrix is equal to that same distribution vector. We want to solve this system of equations. If there are n states, then this is an n by n system

of equations, with an additional constraint that we want the norm of the stable vector to be 1, because that's to avoid the degenerate 0 solution.

Well there are some problems with stationary distributions that we want to think about. First of all, what happens in this example where you have just two states, and the probability of being in the first state at 1 and the second state is 0? Well if you update that state, what happens is you just go to the second state with probability 1.

And you can keep doing that. And there may be a stable distribution here, but this particular pattern doesn't converge to it. As you go through time, at every other step you're at state 1, and every other step you're at state 0.

But you never get to a stable distribution where step after step you are at equal probability of being at these two places. I'm assuming here that this is a certain edge, and that's a certain edge. It has to be. There's only one edge out. So a stable distribution would be $1/2, 1/2$. But this thing doesn't converge to it.

OK. Here's a slightly more complicated example, where again assume that all the edges are equally likely. There's exactly two edges out of each of these vertices so that each edge has weight $1/2$. And the problem with this graph is that when you ask what's the stable distribution and, well, if you look at it, if you assume that the probability of being in the middle is 0, and the two places that you get stuck at have probability p and $1 - p$, then that's stable, because once you're at this state with probability p you're following the one certain edge that goes back around to this vertex. And therefore there's probability of p of being there one step later, and likewise probability q of one step later. So the split between p and q is a stable distribution for this thing, with probability 0 and 0 there.

And of course p and q can be any real numbers between 0 and 1. So there's actually an uncountable number of stable distributions for this graph. Problem here is it's not strongly connected. And that turns out to be a sufficient condition that it's got a single stable distribution whenever it's strongly connected.

So in general we can ask the question, is there always a stationary distribution for any random graph? well, if the graph is finite, yes, there's guaranteed to be a stationary distribution. But is it unique? Well sometimes, sometimes not. If the graph is strongly connected, it will be unique. But we've seen examples in the previous slide where it's not unique. In fact, it could be uncountably many.

And another crucial question is, does a random walk approach the stable distribution no matter how you start? And that first example was one where you went between the first state and the second state and oscillated. And it never converged on the stable distribution of $1/2$ and $1/2$.

In general, it's nice when you can say that no matter how you start, after a while things stabilize, and you wind up at the unique stable distribution. So sometimes it'll be the case that every initial distribution will eventually converge on the stable one or the stationary one. Sometimes not.

And then another crucial question will be, how quickly does this convergence happen? If we start off at some arbitrary probability distribution, or some particular state, how long does it take before by and large the probabilities that were in the different states has become pretty stationary? And the rate at which that happens again varies depending on the graph.