PROFESSOR: Hi, everyone. Welcome back. So today, I'd like to tackle a problem in Markov matrices.

Specifically, we're going to start with this problem which almost has a physics origin. If we have a particle that jumps between positions A and B with the following probabilities-- I'll just state it-- if it starts at A and jumps to B with probability 0.4 or starts at A and stays at A with probability 0.6, or if it starts at B then it goes to A with probability 0.2 or stays at B with probability 0.8, we'd like to know the evolution of the probability of this particle over a long period of time.

So specifically the problem we're interested today is: if we have a particle and we know that it starts at position A, what is the probability that it is at position A and the probability that it's at position B after one step, after n steps, and then finally after an infinite number of steps? So I'll let you think about this problem for a moment and I'll be back.

Hi everyone. Welcome back. So the main difficulty with this problem is that it's phrased as a physics problem. And we need to convert it into some mathematical language to get a handle on it.

So specifically, what we'd like to do is to convert this into a matrix formalism. So what we can do is we can write this little graph down and describe everything in this graph using a matrix. So I'm going to call this matrix A, and I'm going to associate the first row of A with particle position A and particle position B. And I'll associate the first and second columns with particles positions A and B. And then what I'm going to do is I'm going to fill in this matrix with the probability distributions.

So, specifically what's going to go in this top left corner? Well, the number 0.6, which represents the probability that I stay at position A, given that I start at position A. What's going to go here in the bottom left-hand corner? Well, we're going to put 0.4, because this is the probability that I wind up at B, given that I start at A. And then lastly, we'll fill in these other two columns or the second column with 0.8 and 0.2.

So I'll just state briefly this is what's called a Markov matrix. And it's called Markov, because first off, every element is positive or 0. And secondly, the sum of the elements in each column is 1. So if we note 0.4 plus 0.6 is 1, 0.8 plus 0.2 is 1. And these matrices come up all the time when we're talking about probabilities and the evolution of probability distributions.

OK. So now, once we've encoded this graph using this matrix A, we now want to tackle this problem. So I'm going to introduce the vector p, and I'm going to use a subscript 0 is to denote the probability that the particle is at time 0. So we're told that the particle starts at position A. So at time 0, I'm going to use the vector [1, 0].

Again, I'm going to match the top component of this vector with the top component of this matrix and the first column of this matrix. And then likewise, the second component of this vector with the second row and second column of this matrix. And we're interested in: how does this probability evolve as the particle takes many steps?

So for one step, what's the probability of the particle going to be? Well, this is the beauty of introducing matrix notation. I'm going to denote p_1 to be the probability of the particle after one step. And we see that we can write this as the matrix A multiplied by p_0.

So the answer is 0.6 and 0.4. And I achieve this just by multiplying this matrix by this vector. OK? So this concludes part one.

Now part two is a little trickier. So part two is n steps. And to tackle this problem, we need to use a little more machinery.

So first off, I'm going to note that p_1 is A times p_0. Likewise, p_2-- so this is the position of the-- the probability distribution of the particle after two steps. This is A times p_0, which is A squared times p_0. And we note that there's a general trend. After n steps-- so P_n-- the general trend is, it's going to be this matrix A raised to the n-th power, multiply the vector P0.

So how do we take the n-th power of a matrix? Well, this is where we use eigenvectors and eigenvalues. So recall, that we can take any matrix A that's diagonalizable and write it as U D U inverse, where D is a diagonal matrix and this matrix U is a matrix whose columns correspond to the eigenvectors of A.

So for this problem, I'm just going to state what the eigenvalues and eigenvectors are. And I'll let you work them out. So because it's a Markov matrix, we always have an eigenvalue which is 1. And in this case, we have an eigenvector u which is 1 and 2.

In addition, the second eigenvalue is 0.4. And the eigenvector corresponding to this one is [1, -1]. And I'll just call these u_1 and u_2, like that.

OK, we can now write this big matrix U as 1, 2; 1, -1. D is going to be-- now I have to match

things up. If I'm going to put the first eigenvector in the first column, we have to stick 1 in the first column as well and then 0.4 like this. And then lastly, we also have U inverse which I can just work out to be minus 1/3, one over the determinant, times -1, -1; -2, and 1, which simplifies to this.

OK, so now if we take A and raise it to the power of n, we have this nice identity that all the U and U inverses collapse in the middle. And we're left with U, D to the n, U inverse, p_0. Now raising the a diagonal matrix to the power of n is a relatively simple thing to do. We just take the eigenvalues and raise them to the power of n.

So when we compute this product, there's a question of what order do we do things? Now these are 2 by 2 matrices, so in theory we could just multiply out, 2 by 2 matrix, 2 by 2 matrix, 2 by 2 matrix, and then on a vector which is a 2 by 1 matrix.

But if you're in a test and you're cramped for time, you want to do as little computations as possible. So what you want to do is you want to start on the right-hand side and then work backwards. So if we do this, we end up obtaining 1, 2, this is going to be to the power of n, 1/3, [1, 2].

OK, so for this last part, I'm just going to write down the final answer. And I'll let you work out the multiplication of matrices. So we have for p_n: 1/3, 2 times 0.4 to the n plus 1, -2 0.4 to the n plus 2. And this is the final vector for p of n.

So this finishes up Part 2. And then lastly, for Part 3, what happens when n goes to infinity? Well, we have the answer for any n. So we can just take the limit as n goes to infinity. Now, specifically as n goes to infinity, 0.4 raised to some very large power vanishes. So these two terms drop off. And at the end of the day, we're left with p_∞ is 1/3 [1, 2]. OK?

So just to recap, we started off with a particle starting at A, and then after a very long time, the particle winds up with a probability distribution which is 1/3, 1 and 2. And this is quite characteristic of Markov matrix chains. Specifically, we note that 1/3 * [1, 2] is a multiple of the eigenvector corresponding to eigenvalue 1. So even though the particle started at position A, after a long period of time, it tended to forget where it started and approached, diffused into this uniform distribution.

OK. I'd like to finish up here. And I'll see you next time.