Conversation 10: Where is Waldo?
More Applications of Markov Chains.

Winfried Just
Department of Mathematics, Ohio University

MATH3200: Applied Linear Algebra

Denny: Let me tell you a story about this guy Waldo. When he took MATH3200 he spent every evening working on the homework.

Frank: Hopeless nerd.

Denny: Yeah, but very sociable. He always operated as follows:

At 7p.m. he visited a randomly chosen student i among his six friends who took the same class and started working with that person.

After 10 minutes, he flipped a fair coin.

If the coin came up heads, he continued working with i for another 10 minutes before flipping the coin again.

If the coin came up tails, he moved to the room of a randomly chosen friend of i from the same class and repeated the procedure.

Cindy: This sounds so weird!!

Denny: Yeah, and the weirdest thing is that he always kept going like this until 1a.m. Every day.

Where should we go looking for Waldo at midnight?

Denny: The guy was so focused on his math problems that in order to avoid distractions he didn't even carry a cell phone. One day his Mom called me at midnight and asked me to get him on the phone right away. Some kind of emergency. I didn't know to which room I should go first looking for him.

Cindy: I wouldn't know either. With all this coin flipping and randomly moving around, his whereabouts would be rather unpredictable.

Bob: Neither would I have known. But why are you telling us this story Denny? Shouldn't we focus on our study of Markov chains?

Denny: I thought these Markov chains might have been rather useful for figuring out what to do.

Theo: You would have liked to know in which room Waldo would have been at midnight with the highest probability.

Yes, we can figure this out by building a Markov chain model.

Building a Markov chain for Waldo's studying habits

Cindy: Great idea! But how should we get started?

What do we need to specify first?

We first need to specify the set of states and the time steps.

Question C10.1: What set of states should we choose here?

Since we want to know the room where Waldo will be with the highest probability, and there are six possible rooms, we can make our states $i = 1, 2, \ldots, 6$ and let state i signify that Waldo is in room number i.

Question C10.2: What meaning should one time step have here?

Since Waldo may change locations every 10 minutes, one time step would last 10 minutes.

Question C10.3: If we conveniently set $t = 0$ to 7p.m., what time t in our Markov chain would correspond to midnight, when Waldo's Mom called?

Midnight would be after $t = (5)(6) = 30$ time steps.

Building the transition probability matrix $\mathbf{P}$ of Waldo

Question C10.4: What order should the transition probability matrix $\mathbf{P}$ have?

It should have order 6×6: $\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{bmatrix}$

Question C10.5: What should the value of each diagonal element p_{ii} be?

Since Waldo stays in the same room when the coin comes up heads, we have $p_{ii} = 0.5$ for all i:

$$\mathbf{P} = \begin{bmatrix} \mathbf{0.5} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{21} & \mathbf{0.5} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & \mathbf{0.5} & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & \mathbf{0.5} & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & \mathbf{0.5} & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & \mathbf{0.5} \end{bmatrix}$$

How about the off-diagonal elements in $\mathbf{P}$?

Denny: Finding p_{ij} for $i \neq j$ would be the tricky part, though.

Frank: Easy. Since there are six rooms and Waldo moves to a randomly chosen one, they should all be $1/6$.

Cindy: But wouldn't $i \neq j$ mean that the room j is different, so that we have only five possibilities? I think p_{ij} should be $1/5$ when $i \neq j$.

So we would get: $$\mathbf{P} = \begin{bmatrix} 0.5 & 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 0.5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 0.5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 0.5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0.5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0.5 \end{bmatrix}$$

Bob: Good point, Cindy! But this can't be quite right. The probabilities in each row add up to 1.5, but the matrix $\mathbf{P}$ should be stochastic, which means that the probabilities should add up to 1.

Cindy fixes her matrix

Cindy: Oh, I see! The probabilities of moving from room i to room j come into play *only when* the coin comes up tails, which happens with probability 0.5. So we need to multiply 1/5 with 0.5 here, which gives $p_{ij} = 1/10 = 0.1$ whenever $i \neq j$.

Then we get: $$\mathbf{P} = \begin{bmatrix} 0.5 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.5 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.5 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.5 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.5 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.5 \end{bmatrix}$$

Bob: Now the probabilities in each row add up to 1 and the matrix **P** is stochastic. It seems to me that we have built a plausible Markov chain model for the story as Denny has told it.

Denny: So how do we use this model for answering my question?

How about Denny's question?

Bob: We would need to calculate the probability distribution $\vec{\mathbf{x}}$ after 30 time steps, which is given by $\vec{\mathbf{x}}(30) = \vec{\mathbf{x}}(0)\mathbf{P}^{30}$.

Denny: Why 30?

Bob: Recall that $t = 30$ corresponds to midnight, and that $\mathbf{P}^{30}$ represents the matrix of transition probabilities after 30 time steps.

Denny: And what is $\vec{\mathbf{x}}(0)$?

Bob: The initial probability distribution at 7p.m., which is our $t = 0$. You said that Waldo then goes to a randomly chosen one among the six rooms. Did you mean that he would pick each of these rooms with equal probability at that time?

Denny: Yeah.

Bob: Then $\vec{\mathbf{x}}(0) = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$.

What do we find?

Denny: But I wouldn't like to calculate
$\vec{\mathbf{x}}(30) = \vec{\mathbf{x}}(0)\mathbf{P}^{30} = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \mathbf{P}^{30}$.

Theo: We can use MATLAB. I did this quickly on the side and got
$\vec{\mathbf{x}}(30) = [0.1667 \ \ 0.1667 \ \ 0.1667 \ \ 0.1667 \ \ 0.1667 \ \ 0.1667]$.

Question C10.6: So where should Denny go looking for Waldo?

This would mean that Waldo is equally likely to be in any of the rooms, so it doesn't matter where Denny starts searching for him.

Denny: But wait: Waldo told me afterwards that I should have gone to room 4 first, because that is the one he visits most often. He would have known. Your Markov model must be wrong!!!

Bob and Theo: But all the calculations were correct ...

Alice: We had ignored one important detail of Denny's story. Let's look at the story again.

Let's take another look at Denny's story

Denny: Let me tell you a story about this guy Waldo. When he took MATH3200 he spent every evening working on the homework.

He always operated as follows:

At 7p.m. he visited a randomly chosen student i among his six friends who took the same class and started working with that person.

After 10 minutes, he flipped a fair coin.

If the coin came up heads, he continued working with i for another 10 minutes before flipping the coin again.

If the coin came up tails, he moved to the room of a **randomly chosen friend of i** from the same class and repeated the procedure.

Alice: In our model, we had implicitly assumed that each two of Waldo's friends from this class are also friends of each other.

Denny: And that's not at all the case, I can tell you!

Always clarify the assumptions

Cindy: But you didn't tell us right away!

Denny: You could have asked.

Bob: Come on, Denny! Your story was too vague about this.

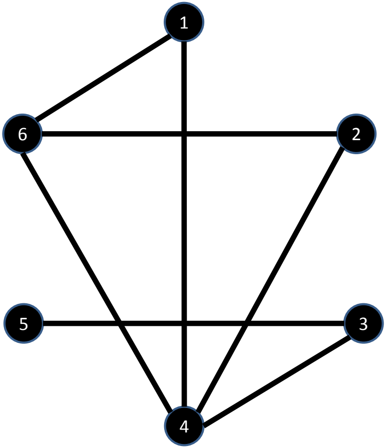
Alice: In fact, we *should* have asked. When building a model, we always need to ask the domain expert, in our case Denny, for clarification of all the important details that may have been left too vague at first.

Cindy: So let me ask you now, Denny:

Who is friends with whom among those six friends of Waldo?

Denny: You have already seen this. The friendships are exactly as in the graph that was shown in Lecture 3.

Recall our graph of friendships



Starting new transition probability matrix P

Cindy: So since 2, 3, and 5 are not friends of 1, Waldo will not move to their rooms from room 1 and I should have put $p_{12} = p_{13} = p_{15} = 0$, right?

$$\mathbf{P} = \begin{bmatrix} 0.5 & \mathbf{0} & \mathbf{0} & p_{14} & \mathbf{0} & p_{16} \\ p_{21} & 0.5 & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & 0.5 & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & 0.5 & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & 0.5 & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & 0.5 \end{bmatrix}$$

Bob: Right! Since 4 and 6 are friends of 1, Waldo will be equally likely to move to room 4 or to room 6 when the coin comes up tails, so that $p_{14} = p_{16} = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$:

$$\mathbf{P} = \begin{bmatrix} 0.5 & \mathbf{0} & \mathbf{0} & \mathbf{1/4} & \mathbf{0} & \mathbf{1/4} \\ p_{21} & 0.5 & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{31} & p_{32} & 0.5 & p_{34} & p_{35} & p_{36} \\ p_{41} & p_{42} & p_{43} & 0.5 & p_{45} & p_{46} \\ p_{51} & p_{52} & p_{53} & p_{54} & 0.5 & p_{56} \\ p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & 0.5 \end{bmatrix}$$

Let's find more elements of **P**

Question C10.7: What is p_{32}?

Since 2 is not a friend of 3 we have $p_{32} = 0$.

Question C10.8: What is p_{35}?

Since 5 is one of the two friends of 3, we have $p_{35} = \frac{1}{4}$.

Question C10.9: What is p_{53}?

Since 3 is the only friends of 5, Waldo will move to room 3 whenever he is in room 5 and the coin comes up tails, so we have $p_{35} = \frac{1}{2}$.

Where is Waldo at midnight?

Continuing like this, we can get the entire transition probability matrix **P**:

$$\mathbf{P} = \begin{bmatrix} 1/2 & 0 & 0 & 1/4 & 0 & 1/4 \\ 0 & 1/2 & 0 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1/2 & 1/4 & 1/4 & 0 \\ 1/8 & 1/8 & 1/8 & 1/2 & 0 & 1/8 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 1/6 & 1/6 & 0 & 1/6 & 0 & 1/2 \end{bmatrix} = [p_{ij}]_{6\times 6}$$

Now we let MATLAB calculate $\vec{\mathbf{x}}(30) = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \mathbf{P}^{30}$.

We get: $\vec{\mathbf{x}}(30) = [0.1427 \;\; 0.1427 \;\; 0.1432 \;\; 0.2856 \;\; 0.0717 \;\; 0.2140]$.

Question C10.10: Where should Denny go looking for Waldo?

Denny: To room 4, as it has the highest probability, exactly as Waldo said!

Question C10.11: How would you intuitively explain this finding based on the friendship network?

If Waldo weren't such great company ...

Frank: Your story is just BS, Denny.

Denny: Why don't you believe me?

Frank: For your story to be halfway believable, you would have to assume that Waldo is incredibly likable.

Alice: So what would happen if our six students were not all that fond of him?

Frank: In this case it would rather be the owner of the room i who flips the coin and decides whether to kick out Waldo.

Theo: From the mathematical perspective, this would be exactly the same though if then Waldo gets sent to the room of a randomly chosen friend of i.

Denny: Send Waldo to good a friend? Who would do that?

Frank: No, Waldo would be sent to the room of a randomly chosen student who is *not* i's friend. The Markov chain model for this must be rather different, I'd say!

Waldo late at night

Bob: I think we will be asked to explore the Markov chain for this alternative version of the story in Module 14, so let's talk about something else right now.

Denny: Believe it or not, this guy Waldo did even more weird stuff.

Cindy: Can you tell us? I mean, if it's not too gross.

Denny: At 1a.m., Waldo always went back to his own room and surfed the web.

Cindy: What's weird about that? Unless you mean ...

Denny: Not at all. He was only into math and computers.

Cindy: Then please tell us what Waldo did.

Waldo surfs the web

Denny: First he opened his home page and followed a randomly chosen link.

Then at each page that he visited:

- If the page had no link to another page, he teleported to a randomly chosen URL.
- If the page had links, he rolled a fair die.
  - If 6 came up, he teleported to a randomly chosen URL.
  - If any other number came up, he followed a randomly chosen link from the current page.

Bob: Do you mean: He chose the links to follow with equal probability, regardless of whether they seemed interesting or were pointing to the same page?

Denny: Yes, this was one odd thing about him. He didn't even seem to care whether the web page contained any math or, for that matter, any of the stuff that I cannot mention here in front of Cindy.

Frank: "And Alice," you forgot to say.

Alice: Thank you, Frank!

Denny: Yeah, that's whom I meant. There are are some rumors that Waldo made checklists of words that he found on the pages, but I don't know what these would be good for. In any case, he never actually read any of these pages or looked at any pictures. He just kept clicking on random links and teleporting.

Waldo surfs the web

Frank: You are giving us more BS here. How could he "teleport" to a randomly chosen URL?

Denny: Yeah, that's another odd aspect of it. I tried to ask him many times, but he never would tell me.

Alice: He must have had some sort of directory of the entire WWW and might have considered it proprietory information.

Denny: Maybe. And the absolute killer, or killer app so to speak, is that this surfing pattern, together with what he learned in MATH 3200, made Waldo rich and famous!

Question C10.12: Who is Waldo?

Believe it or not, we will derive a serious answer to this question in Module 15.

Take-home message

- When building a Markov chain, always start by specifying the set of the states and the meaning of one time step.
- Next find the matrix of transition probabilities.
- Then translate the given real-world question into a property of the Markov chain.
  - Here we wanted to know which state has the highest probability in the distribution $\vec{\mathbf{x}}(t) = \vec{\mathbf{x}}(0)\mathbf{P}^t$, where $t = 30$ was the number of time steps until midnight and the initial distribution $\vec{\mathbf{x}}(0)$ was determined by the information that Waldo is equally likely to visit any of the rooms first.
- Make sure to ask the domain expert wants you to answer the question for clarification of all relevant details.

Take-home message, continued

- Check the predictions of your model against available real-world data. If the model's predictions are way off, double-check the model's asumptions and revise the model if something is wrong or missing.
  - Here the initial version of our model did not match the info that room 4 is the one most often visited by Waldo.
- Seemingly very different real-world situations, like Waldo's studying and surfing patterns, may have the same underlying mathematical structure.