Conversation 9: More on Markov chains
Making forecasts with weather.com light

Winfried Just
Department of Mathematics, Ohio University

MATH3200: Applied Linear Algebra

Review: Markov chains and probabilities

- Real-world situations that involve not entirely predictable change can be modeled as *stochastic processes.*
- Here we introduced *Markov chains,* which are one kind of stochastic processes.
- In Markov chains, time is assumed to proceed in discrete *steps,* and the real-world situations of interest are categorized into nonoverlapping *states.*
- When constructing a Markov chain model of a real-world situation, we need to start by specifying the meaning of each state and the meaning of one time step.
- *Probabilities* are numbers between 0 and 1. The probability of an event can be thought of as the proportion of times this event would occur if we conduct a large number of observations of very similar situations.

Review: Time steps and states in our Markov chain for weather.com light

- Time advances in discrete steps $t = 0, 1, 2, \dots$ that last one day each.
- All weather patterns on a given day are unambiguously partitioned into two nonoverlapping states that we named as follows:
  1. State 1: sunny day.
  2. State 2: rainy day.

Review: Transition probabilities and transition probability matrices

- We let p_{ij} denote the *transition probability* from state i to state j. This means p_{ij} is the probability that **if** the process is in state i at time t, **then** it will be in state j at time $t+1$.

- These probabilities can be arranged into the *transition probability matrix* $\mathbf{P} = [p_{ij}]$ that has order $n \times n$, where n is the number of states.

- In our example for weather.com light the number of states is $n = 2$, and transition probability matrix takes

  the form $\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$ For example, $\mathbf{P} = \begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{bmatrix}$

- The transition probability matrix **P** must be a *stochastic matrix,* which means that all its elements p_{12} are probabilities and each row adds up to 1.

Probability distributions

Bob: So how would we use these transition probability matrices to make a forecast?

Theo: First recall what I told you about probability distributions. The vector $\vec{x}(t) = [x_1(t), \ldots, x_n(t)]$ would be the *probability distribution* for the states on a given day. The individual probabilities $x_i(t)$ of the states must add up to 1.

Cindy: So, in our example of weather.com light, we have $n = 2$ and $\vec{x}(t) = [x_1(t), x_2(t)]$, right?

Theo: Right! The distribution $\vec{x}(t)$ is a measure of our uncertainty about the weather on day t.

Denny: Are you saying, Theo, that if we would give it a fifty-fifty chance that it will rain on Friday, the 13^{th}, then $\vec{x}(13) = [x_1(13), x_2(13)] = [0.5, 0.5]$?

Theo: We can think about the concept in terms of this example.

Next-day forecasts

Theo: Now let's think about making a next-day forecast as figuring out the probability distribution $\vec{x}(t+1)$ for day $t+1$, which would be Saturday, the 14^{th} in Denny's example.

Denny: That would also be a fifty-fifty chance, obviously, since I already don't know what would happen on Friday, the 13^{th}. So $\vec{x}(14) = [0.5, 0.5]$ in Theo's notation.

Frank: Not obvious to me!

Theo: We will see. The general formula for obtaining the probability distribution for the next day is $\vec{x}(t+1) = \vec{x}(t)\mathbf{P}$. Thus when $\vec{x}(13) = [0.5, 0.5]$ and if we use the matrix $\mathbf{P}$ from our numerical example, we get:

$$\vec{x}(14) = \vec{x}(13)\mathbf{P} = [0.5, 0.5]\begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{bmatrix} = [0.35, 0.65] \neq [0.5, 0.5]$$

Denny: If you say so. But why? Why isn't $\vec{x}(14) = [0.5, 0.5]$?

Forecasts for tomorrow

Question C9.1: What is the explanation for this result?

Cindy: In our numerical example for **P**, transitioning into state 2 is more likely than transitioning into state 1, so the chance of rain the next day will no longer be fifty-fifty.

Frank: But wait! Theo threw this formula $\vec{x}(t+1) = \vec{x}(t)\mathbf{P}$ for the probability distribution on the next day at us.
Where does it come from?

Theo: This formula is true in every Markov chain.

Frank: But why?

Alice: Good question Frank! Think about our example weather.com light.

We could arrive at a sunny day $t+1$ in two different ways:

If both days t and $t+1$ are sunny days, and if day t is a rainy day and then we transition into a sunny day at time $t+1$.

Forecasts for tomorrow, continued

Alice: Look at Theo's formula $\vec{x}(t+1) = \vec{x}(t)\mathbf{P}$ for two states:

$$[x_1(t+1), x_2(t+1)] = [x_1(t), x_2(t)] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

$$[x_1(t+1), x_2(t+1)] = [x_1(t)p_{11} + x_2(t)p_{21}, x_1(t)p_{12} + x_2(t)p_{22}]$$

The scenario of both days t and $t+1$ being sunny has a probability that is the product $x_1(t)p_{11}$ of the probability $x_1(t)$ that day t is sunny with the probability p_{11} of then transitioning into another sunny day.

The other scenario has a probability that is the product $x_2(t)p_{21}$ of the probability $x_2(t)$ that day t is rainy with the probability p_{21} of then transitioning into a sunny day.

Since exactly one of these scenarios must occur, we need to add these two probabilities to obtain $x_1(t+1)$, exactly as in Theo's formula.

Forecasts for tomorrow, continued

Frank: Oh, I see! And the probability $x_2(t+1)$ of $t+1$ being rainy can similarly be broken up into two scenarios, two rainy days in a row, and a sunny day followed by a rainy day!

Bob: But I don't see how we could use Theo's formula $\vec{x}(t+1) = \vec{x}(t)\mathbf{P}$ for predicting tomorrow's weather.

Frank: Why not? Then $\vec{x}(t)$ would be the probability distribution for today, $\vec{x}(t+1)$ would be the probability distribution for tomorrow, and that's the vector of probabilities $x_1(t+1)$ that tomorrow will be a sunny day and $x_2(t+1)$ that it would be a rainy day. So that's your forecast!

Bob: I got that. But we already know that I got drenched today. So there is no uncertainty about today's state. How can we then have a probability distribution $\vec{x}(t)$ for today's weather?

Question C9.2: What would you tell Bob here?

Forecasts for tomorrow, continued

Alice: Since you are **certain** that today we are in state 2, rainy day, we have $x_2(t) = 1$ and $x_1(t) = 0$, which means $\vec{x}(t) = [0, 1]$. This is still a probability distribution.

Denny: Now we get from Theo's formula

$$\vec{x}(t+1) = \vec{x}(t)\mathbf{P} = [0, 1] \begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{bmatrix} = [0.3, 0.7]$$

Cindy: The distribution is the same as the second row of $\mathbf{P}$!

Theo: And that is exactly as it should be, since the elements p_{21}, p_{22} of that row were defined as the probabilities of observing states 1 or 2 the next day **if** the current state is state 2.

Bob: So, I wouldn't even need to multiply $\vec{x}(t)$ with $\mathbf{P}$ and could read the forecast right off the relevant row of $\mathbf{P}$?

Theo: Right. If today were a sunny day, the probability distribution for tomorrow would be given by the first row of $\mathbf{P}$.

Denny: Cool! Less work this way.

How about a two-day forecast?

Cindy: But how about long-range forecasts? We know it was rainy today, so what are the chances of rain on Saturday?

Theo: Since today is Monday, you will need the matrix of transition probabilities from the state at time t to the state at time $t + 5$.

Cindy: Can I call it $\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$ so that I don't get mixed up by using the same symbols as in $\mathbf{P}$?

Alice: Very good, Cindy! But I would recommend that you try to work it out for transitioning from the state today, at time t, to the state the day after tomorrow, at time $t + 2$. A two-day forecast. This is easier.

Cindy: I'm more interested in Saturday, though. But OK. Let's keep it simple.

Working out a two-day forecast

Cindy: So let

- q_{11} be the probability that if day t is a sunny day, then day $t+2$ will also be sunny.
- q_{12} be the probability that if day t is a sunny day, then day $t+2$ will be rainy.
- q_{21} be the probability that if day t is a rainy day, then day $t+2$ will be sunny.
- q_{12} be the probability that if day t is a rainy day, then day $t+2$ will also be rainy.

How can we express the matrix $\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$

in terms of the next-day transition probabilities $p_{11}, p_{12}, p_{21}, p_{22}$? I think I would need some of Theo's probability theory here.

Bob: But this wasn't a prerequisite for this course!

Theo: I will help you with the facts about probabilities that you need, Cindy. First of all, if the state at time t is i, then the probability that the state at time $t+1$ is j and the state at time $t+2$ is k is given by $p_{ij}p_{jk}$, because the process would first transition from state i into state j and then from state j into state k.

Cindy: So, in particular, for q_{11}, the probability that if day t is a sunny day, then day $t+2$ will also be sunny, we would have $i = j = k = 1$ in your notation, and $q_{11} = p_{11}p_{11} = p_{11}^2$, right?

Question C9.3: Did Cindy give us the correct formula for q_{11}?

Alice: You correctly expressed the probability that a sunny day is followed by two sunny days in a row.

Denny: And that's q_{11}, so Cindy is right!

Bob: But isn't there also another scenario?

Working out a two-day forecast, continued

Cindy: Oh I see! The weather could switch from sunny today to rainy tomorrow, and then back to sunny! This would occur with probability $p_{12}p_{21}$ by Theo's formula.

Theo: Exactly! And since one of these two scenarios must occur if the process is to transition from state 1 into state 1 within two time steps, and since they are *mutually exclusive,* which means they cannot both occur, you must add their probabilities.

Cindy: So it will be $q_{11} = p_{11}p_{11} + p_{12}p_{21}$.

Bob: Very good, Cindy! Now express q_{12}.

Cindy: q_{12} is the probability that if day t is a sunny day, then day $t + 2$ will be rainy. This could happen if day $t + 1$ is sunny, with probability $p_{11}p_{12}$, or if day $t + 1$ is rainy, with probability $p_{12}p_{22}$. So, I obtain $q_{12} = p_{11}p_{12} + p_{12}p_{22}$.

Alice: You got the hang of it, Cindy!

Working out a two-day forecast, completed

Cindy: Now, please, someone else do q_{21} and q_{22}!

Question C9.4: Find expressions for q_{21} and q_{22}.

The transition probability matrix from day t to day $t+2$ is

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \begin{bmatrix} p_{11}p_{11} + p_{12}p_{21} & p_{11}p_{12} + p_{12}p_{22} \\ p_{21}p_{11} + p_{22}p_{21} & p_{21}p_{12} + p_{22}p_{22} \end{bmatrix}$$

Cindy: Done!! But I really wouldn't want to work out a five-day forecast. You were sooooo right about this, Alice!

Alice: Let's take a look at our result. Have we seen this before?

Bob: Yes! This is simply the product of $\mathbf{P}$ with itself!

$$\mathbf{Q} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \mathbf{PP} = \mathbf{P}^2$$

Frank: Yeah, but we could have figured this out more easily, with fewer calculations.

Can we get away with fewer calculations?

Denny: Sounds good to me, but how?

Frank: Look Denny: $t + 2$ is the day after $t + 1$.

Denny: Yeah. So what?

Frank: So, $\vec{x}(t + 2) = \vec{x}(t + 1)\mathbf{P}$.

Denny: But how would this help? We don't know $\vec{x}(t + 1)$.

Frank: That's $\vec{x}(t + 1) = \vec{x}(t)\mathbf{P}$. So now we substitute and get:
$\vec{x}(t + 2) = \vec{x}(t + 1)\mathbf{P} = (\vec{x}(t)\mathbf{P})\mathbf{P} = \vec{x}(t)\mathbf{P}\mathbf{P} = \vec{x}(t)\mathbf{P}^2$.

Denny: Wow! I thought you don't like math, Frank?

Frank: I don't. But I like doing more calculations than necessary even less.

Theo: That's an excellent argument, Frank! We can use it to prove by mathematical induction that for all positive integers k we have $\vec{x}(t + k) = \vec{x}(t)\mathbf{P}^k$.

Working out a five-day forecast, anyone?

Denny: What does this mean in plain English?

Theo: I will be happy to explain the details of the proof.

Bob: Maybe not now. I think we can read about this proof in Module 13.

Cindy: This formula looks very useful though! When we let $k = 5$, we get a five-day forecast, right?

Theo: Right!

Cindy: So then I can calculate the forecast for Saturday as

$$\vec{x}(t+5) = \vec{x}(t)\mathbf{P}^5 = [x_1(t), x_2(t)] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}^5 = [0, 1] \begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{bmatrix}^5$$

Frank: This would be too tedious to calculate!

Bob: In Module 13 we will do this by using MATLAB. So let's wait until that module comes online!

Take-home message

Consider a Markov chain with n states and transition probability matrix $\mathbf{P}$.

- The *probability distribution* at time t is a row vector $\vec{x}(t) = [x_1(t), \dots, x_n(t)]$ that gives the probabilities $x_i(t)$ that the system is in state i at time t.
- The probability distribution for the next step is given by $\vec{x}(t+1) = \vec{x}(t)\mathbf{P}$.
- More generally, for any $k \geq 1$ and state $\vec{x}(t)$, the distribution $\vec{x}(t+k)$ after k time steps is given by $\vec{x}(t+k) = \vec{x}(t)\mathbf{P}^k$, where $\mathbf{P}^k$ is the *k-step transition probability matrix.*