Discrete-Time Markov Chains

Transient behavior

Evolution of the probability vector

Notations:

- $\pi_i(n) := \mathbf{P}(X_n = i)$
- $\pi(n) := [\pi_1(n) \quad \pi_2(n) \quad \pi_3(n) \quad \ldots]$
  (probability vector or distribution at period n)

Example: if we know we are in State 1 at time 0, then

$$\pi(0) = [1 \quad 0 \quad 0 \quad \ldots]$$

How to express $\pi_j(n+1)$ knowing $\pi(n)$, for any state j?

Now that we have defined our Markov chain, let's analyse how it evolves over time. Since the behavior of our process is random, we are interested in the evolution of the probabilities of being in each state i. We'll denote that by $\pi_i(n)$, that is, the probability to be in State i at the nth time period, which is a bit more compact than the original way of writing it. If I am interested in all states, then I can build a row vector $\pi(n)$ containing all the values, once I have chosen an ordering of the states. That vector will evolve with time: for example if I know that at period 0 I am in State 1, then I can write $\pi(0) = [10...0]$, while at the next period certain transitions may bring me to another state.

Let's say that I know the probability vector at period n (we can also say, the probability distribution at period n). Now I want to compute the probability of being in State j at the next period.

Evolution of the probability vector

For any j,

$$
\begin{aligned}
\underbrace{\mathbf{P}(\{X_{n+1}=j\})}_{\pi_j(n+1)} &= \quad \mathbf{P}(\ \{X_{n+1}=j\}\cap\{X_n=1\}\ ) \\
&\quad \uplus \mathbf{P}(\ \{X_{n+1}=j\}\cap\{X_n=2\}\ ) \\
&\quad \uplus \mathbf{P}(\ \{X_{n+1}=j\}\cap\{X_n=3\}\ ) \\
&\quad \uplus \quad \ldots \\
&= \sum_{\text{all states } i} \underbrace{\mathbf{P}(\ \{X_{n+1}=j\}\cap\{X_n=i\}\ )}_{\mathbf{P}(X_n=i)\ \mathbf{P}(X_{n+1}=j|X_n=i)}
\end{aligned}
$$

$$
\pi_j(n+1) = \sum_{\text{all states } i} \pi_i(n) P_{ij}
$$

But I was certainly somewhere at period n, so I can write the event $\{X_{n+1}=j\}$ as the union of all the events $\{X_{n+1}=j\}$ and $\{X_n=i\}$ for all the possible states i. Those events are disjoint, so the probability of their union is just the sum of their probabilities.
Now let's use the conditional probabilities: each term is the product "probability that X_n equals i" times "probability that X_{n+1} equals j knowing X_n equals i". The former term is what we denoted by $\pi_i(n)$, and the latter term is the probability transition P_{ij} , giving us a compact expression for $\pi_j(n+1)$.

So the jth component of the state probability vector $\pi(n+1)$ is the product of the state probability vector $\pi(n)$ with the jth column of the transition matrix.

Evolution of the probability vector: matrix form

Recursive equation

$$\pi(n+1) = \pi(n)P$$

Iterating, we have

$$\begin{aligned}
\pi(n) &= \pi(n-1) \times P \\
&= \pi(n-2) \times P \times P \\
&= \dots \\
&= \pi(0) \times P^n
\end{aligned}$$

Putting all these components together, to obtain the state probability vector at period $n+1$ we just need to take the state probability vector at period n and multiply it on the right by the transition matrix P.

This is a recursive expression, so we can write $\pi(n)$ as a function of $\pi(n?1)$, and go back up to time period 0 so that the state probability vector at period n is simply the product of the probability vector at period 0 and the transition matrix P to the nth power.

Evolution of the probability vector: example

Our weather forecast example: $\begin{cases} \text{State 1: clear} \\ \text{State 2: cloudy} \\ \text{State 3: rainy} \end{cases}$ $P = \begin{bmatrix} 0.7 & 0.3 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$

Start with clear weather: $\pi(0) = [1 \quad 0 \quad 0]$

Iterate:

$$\begin{aligned} \pi(1) &= \pi(0)P = [0.7 \quad 0.3 \quad 0] \\ \pi(2) &= \pi(1)P = [0.58 \quad 0.36 \quad 0.06] \\ \pi(3) &= \pi(2)P = [0.520 \quad 0.378 \quad 0.102] \\ &\dots \end{aligned}$$

Let's go back to our simple weather forecast model, and assume that the weather is clear at period 0 so that the state probability vector is [100]. Then we can compute the state probability vector at time 1, at time 2, and so on.