CMPSCI 240: Reasoning about Uncertainty

Lecture 15: Steady-State Theorem

Andrew McGregor

University of Massachusetts

Last Compiled: March 23, 2017

Outline

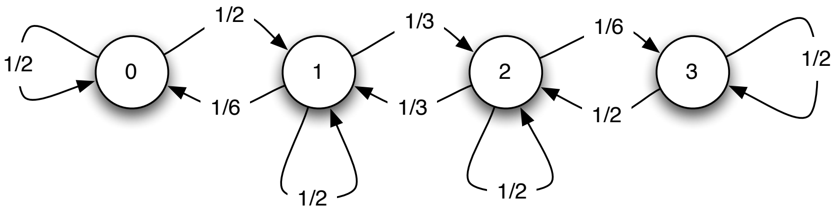
Analyzing the Queue at Earth Foods Cafe

- Consider a queue at Earth Foods Cafe
- Every minute, someone joins the queue...
  - With probability 1 if the queue has length 0
  - With probability $2/3$ if the queue has length 1
  - With probability $1/3$ if the queue has length 2
  - With probability 0 if the queue has length 3.
- Every minute, the server serves a customer with probability $1/2$.

Suppose 1 person in line at noon. How many people in line at 12:10pm?

States with Transition Probabilities

- Weight p_{ij} on arrow from state i to state j indicates the probability of transitioning to state j given we're in state i.



- Can work out things like "what's the probability we're in state 2 after two steps if we're currently in state 3."

Markov Chain

A Markov Chain consists:

- A set of states: $\{s_1, \ldots, s_k\}$
- A matrix P of transition probabilities $\{p_{ij} : 1 \leq i, j \leq k\}$
- An initial state s_i.

A Markov Chain defines a series of random variables $X_0, X_1, X_2, \ldots$ where

- $X_0 = i$
- For all $t = 1, 2, 3, \ldots$

$$P\left(X_t = j | X_{t-1} = i\right) = p_{ij}$$

- Write the distribution of each X_t as

$$v_t = \left(P\left(X_t = 1\right), P\left(X_t = 2\right), \ldots, P\left(X_t = k\right)\right)$$

Note: The value of X_t only depends on the value of X_{t-1}, e.g.,

$$P\left(X_t = j | X_{t-1} = i, X_{t-2} = h\right) = P\left(X_t = j | X_{t-1} = i\right) = p_{ij}$$

Analyzing Markov Chains via Matrices

- Define transition probability matrix:

$$A = \begin{pmatrix} p_{0,0} & p_{0,1} & p_{0,2} & p_{0,3} \\ p_{1,0} & p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,0} & p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,0} & p_{3,1} & p_{3,2} & p_{3,3} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/6 & 1/2 & 1/3 & 0 \\ 0 & 1/3 & 1/2 & 1/6 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

- Markov Chain Theorem Given v_{t-1}, we can compute

$$v_t = v_{t-1}A$$

and so $v_t = v_{t-1}A = v_{t-2}AA = v_{t-3}AAA = \ldots = v_0A^t$.

- Proof: By the law of total probability

$$v_t[j] = P(X_t = j) = \sum_i P(X_t = j | X_{t-1} = i) P(X_{t-1} = i) = \sum_i p_{i,j} v_{t-1}[i]$$

and so $v_t = v_{t-1}A$ as claimed.

Simulation of the queue if there is initially one person

$$
\begin{aligned}
v_0 &= \langle 0.000, 1.000, 0.000, 0.000 \rangle \\
v_1 &= \langle 0.167, 0.500, 0.333, 0.000 \rangle \\
v_2 &= \langle 0.167, 0.444, 0.333, 0.056 \rangle \\
v_3 &= \langle 0.158, 0.416, 0.342, 0.084 \rangle \\
v_4 &= \langle 0.148, 0.401, 0.352, 0.099 \rangle \\
v_5 &= \langle 0.142, 0.391, 0.359, 0.109 \rangle \\
v_6 &= \langle 0.136, 0.386, 0.364, 0.114 \rangle \\
v_7 &= \langle 0.133, 0.382, 0.368, 0.118 \rangle \\
v_8 &= \langle 0.130, 0.380, 0.370, 0.120 \rangle \\
\vdots & \vdots \vdots \\
v_\infty &= \langle 0.125, 0.375, 0.375, 0.125 \rangle
\end{aligned}
$$

Simulation of the queue if there is initially three people

$$
\begin{aligned}
v_0 &= \langle 0.000, 0.000, 0.000, 1.000 \rangle \\
v_1 &= \langle 0.000, 0.000, 0.500, 0.500 \rangle \\
v_2 &= \langle 0.000, 0.167, 0.500, 0.333 \rangle \\
v_3 &= \langle 0.028, 0.250, 0.472, 0.251 \rangle \\
v_4 &= \langle 0.056, 0.296, 0.404, 0.204 \rangle \\
v_5 &= \langle 0.078, 0.324, 0.423, 0.177 \rangle \\
v_6 &= \langle 0.093, 0.341, 0.407, 0.159 \rangle \\
v_7 &= \langle 0.104, 0.353, 0.397, 0.148 \rangle \\
v_8 &= \langle 0.111, 0.360, 0.389, 0.140 \rangle \\
\vdots & \quad \vdots \quad \vdots \\
v_\infty &= \langle 0.125, 0.375, 0.375, 0.125 \rangle
\end{aligned}
$$

Outline

Question

Do all Markov chains have the property that eventually the distribution settles to the "same steady" state regardless of the initial state?

Definition

If $v = vA$, we say v is a steady state distribution for the Markov Chain.

For the queueing example, if $v = (a, b, c, d)$ then

$$(a,b,c,d) = (a,b,c,d)\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (\frac{a}{2}+\frac{b}{6}, \frac{a}{2}+\frac{b}{2}+\frac{c}{3}, \frac{b}{3}+\frac{c}{2}+\frac{d}{2}, \frac{c}{6}+\frac{d}{2})$$

Solving $a = \frac{a}{2} + \frac{b}{6}, b = \frac{a}{2} + \frac{b}{2} + \frac{c}{3}, c = \frac{b}{3} + \frac{c}{2} + \frac{d}{2}$ and $d = \frac{c}{6} + \frac{d}{2}$ gives

$$v = (0.125, 0.375, 0.375, 0.125)$$

Question

Theorem

"Most" Markov Chains have a unique steady state distribution regardless of initial state that is approached by successive iterations from any starting distributions.

Irreducible Markov Chains

Example

Consider the Markov Chain with transition matrix:

$$A = \begin{pmatrix} 0 & 0.9 & 0.05 & 0.05 \\ 0.2 & 0.8 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This Markov chain doesn't converge to a unique steady state.

Definition

A Markov chain is irreducible if in the transition graph there exists a path from every state to every other state, i.e., you can't get stuck in a small group of nodes.

Irreducible Markov Chains

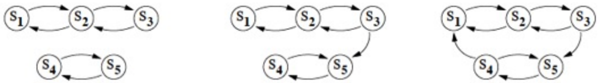


Figure 2.1: Examples of three Markov chains the one to the right is irreducible while the other two are not.

Periodic Markov Chains

Example

Consider the Markov Chain with transition matrix:

$$\begin{pmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.75 & 0 & 0.25 & 0 \\ 0 & 0.75 & 0 & 0.25 \\ 0.75 & 0 & 0.25 & 0 \end{pmatrix}$$

This Markov chain doesn't converge at all!

Definition

An irreducible Markov chain with transition matrix A is called *periodic* if there is some $t \in \{2, 3, \ldots\}$ such that there exists a state s which can be visited only at time $\{t, 2t, 3t, 4t, \ldots\}$ steps, that is with a period of t. If A is not periodic $(t = 1)$ we call the chain *aperiodic*.

Periodic/Aperiodic Markov Chains

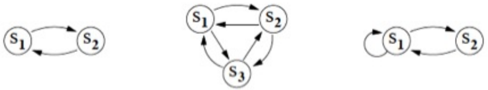


Figure 2.2: Examples of three Markov chains of which the left one has period 2 while the other two both are aperiodic.

Steady State Theorem

Definition

Consider Markov chain with k states and let $v^* = (v^*[1], \ldots, v^*[k])$ be a probability distribution on the states of C. We say that the Markov chain *approaches* v^* if for any arbitrarily small $\epsilon > 0$, there exists a sufficiently large t such that for any i and any starting distribution, then

$$|P(X_t = i) - v^*[i]| \leq \epsilon$$

Theorem

If a Markov Chain is aperiodic and irreducible then there exists a distribution v^* such that v_t approaches v^*.

Steady State Distribution: 2 state case

- Consider a Markov chain C with 2 states and transition matrix

$$A = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

  for some $0 \leq a, b \leq 1$
- Since C is irreducible: $a, b > 0$
- Since C is aperiodic: $a + b < 2$
- Let $v^* = (c, 1-c)$ be a steady state distribution, i.e., $v^* = v^* A$
- Solving $v^* = v^* A$ gives:

$$v^* = \left(\frac{b}{a+b}, \frac{a}{a+b} \right)$$

Steady State Distribution: 2 state case (continued)

- We say v_t converges to v^* if for any $\epsilon > 0$, there exists t^* such that for all $t \geq t^*$ corresponding entries of v_t and v^* differ by at most ϵ.
- Suppose the start distribution is

$$v = (c + \gamma, 1 - c - \gamma)$$

  i.e., entries are $|\gamma|$ away from the corresponding entry in v^*
- After one more step the distribution is

$$vA = (c + \gamma(1 - a - b), 1 - c - \gamma(1 - a - b))$$

  i.e., entries are now only $|\gamma(1 - a - b)| < |\gamma|$ away.
- Hence if we pick t^* such that $|1 - a - b|^{t^*} < \epsilon$ then after t^* or more steps entries will differ by at most

$$|\gamma(1 - a - b)^{t^*}| \leq \epsilon|\gamma| \leq \epsilon$$