Vincent's CS70 Discussion 13A Notes

April 20, 2016

1 Markov Chains!

1.1 Some definitions

Before we define what a Markov chain is, we first need to define stochastic matrices.

Definition 1. *An $n \times n$ right stochastic matrix P is a square matrix that satisfies the following:*

1. $P(i,j) \geq 0 \quad \forall i,j \in \{1,2,\ldots,n\}$
2. $\forall i \in \{1,2,\ldots,n\}, \quad \sum_{j=1}^{n} P(i,j) = 1.$

In other words, a right stochastic matrix is a square matrix with non-negative entries and row-sum equal to 1 for each row. We are now able to state the definition of a Markov chain.

Definition 2. *A (finite) Markov chain M is a 4-tuple $(\mathscr{X}, \pi_0, P, \{X_n\}_{n=0}^{\infty})$ consisting of*

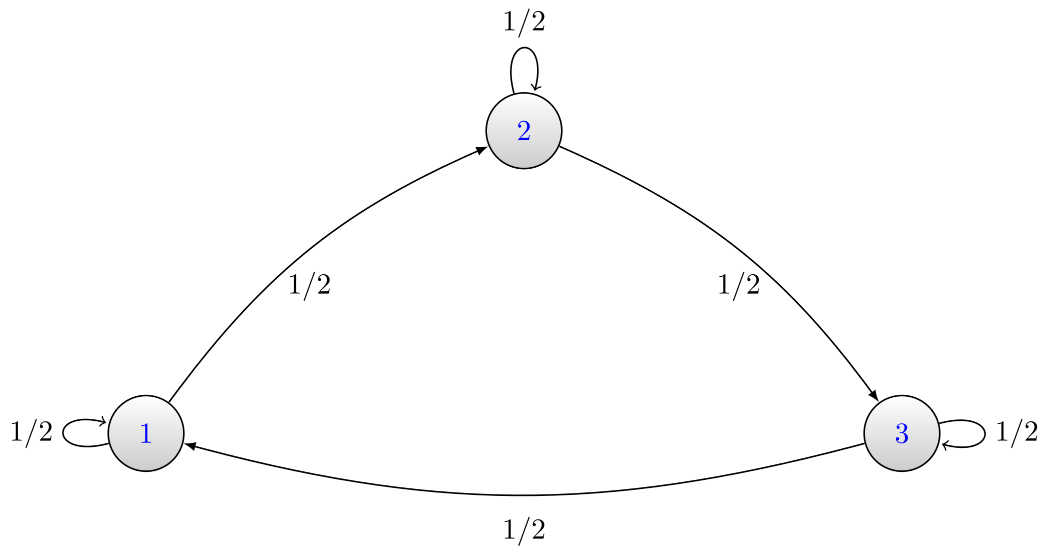
1. *A finite set of states* $\mathscr{X} = \{1,2,\ldots,K\}$.
2. *A probability distribution* $\pi_0 : \mathscr{X} \to [0,1]$ *such that* $\sum_{i \in \mathscr{X}} \pi_0(i) = 1$.
3. *A* $K \times K$ *right stochastic matrix* P.
4. *A sequence of random variables* $\{X_n\}_{n=0}^{\infty}$ *satisfying*
   - $\forall i \in \mathscr{X}, \quad \mathbb{P}(X_0 = i) = \pi_0(i)$
   - $\forall i,j \in \mathscr{X}, \quad \mathbb{P}(X_{n+1} = j | X_0, X_1, \ldots, X_n = i) = P(i,j)$.

There's quite a bit going on in this definition, so let's break it down to understand what's going on. Intuitively, a (finite) Markov chain represents a probabilistic walk through a given state space $\mathscr{X}$. At time $t = 0$, the probability that we find ourselves starting in state $i \in \mathscr{X}$ is given by $\pi_0(i)$. Our stochastic matrix P encodes the probability of transitioning from state i to state j for $i,j \in \mathscr{X}$. In other words, given that we are at state i at time $t = n$, the probability that we move to state j at time $t = n+1$ is $P(i,j)$.

It is important to note that we "forget" about where we were at times $t = 0,1,\ldots,n-1$: the probabilities of our next step only depend on the current state. Finally, the sequence $\{X_n\}_{n=0}^{\infty}$

represents the possible locations we may be at any point in time. For example, the random variable X_0 is our position at the start, and X_i denotes our position at time i. Of course, we have that $range(X_i) \subseteq \mathscr{X}$ for $i \in \mathbb{N}$.

1.2 A simple example



Consider the Markov chain represented by the figure above, and suppose that our initial distribution is given by $\pi_0(1) = 1$, $\pi_0(2) = \pi_0(3) = 0$. We have that $\mathscr{X} = \{1, 2, 3\}$, and finally

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

Let's compute a few things using our example above. By our assumptions, we know that at time $t = 0$, we are guaranteed to start at state 1. Suppose we want to know the amount of time we expect to take to move from state 1 to state 3. Let's begin by denoting the expected amount of time to move from state i to state j by $\beta(i, j)$. Our goal is to compute $\beta(1, 3)$.

A bit of thought gives us that

$$\beta(1, 3) = 1 + \frac{1}{2}(\beta(1, 3) + \beta(2, 3))$$

Why? We have yet to reach our goal, so we are guaranteed to have to move at least once. Furthermore, with probability $\frac{1}{2}$ we move to state 2 and from there expect to take $\beta(2, 3)$ steps to move to state 3, or we stay in state 1 (also with probability $\frac{1}{2}$) and from there expect to take $\beta(1, 3)$ steps to get to state 3. Similarly, we see that

$$\beta(2, 3) = 1 + \frac{1}{2}(\beta(2, 3) + \beta(3, 3))$$

and

$$\beta(3,3) = 0,$$

which we recognize as a system of three linear equations in three variables. Solving this system, we see that

$$\beta(1,3) = 4$$
$$\beta(2,3) = 2$$
$$\beta(3,3) = 0,$$

so we expect 4 units of time to pass before moving from state 1 to state 3.

We conclude by noting that an alternate way of viewing this problem would be to simply view it as calculating the expectation of $X = X_1 + X_2$, where $X_1, X_2 \sim Geom(\frac{1}{2})$. This is because the only way to get from state 1 to state 3 is to first move to state 2, and also there is no way of going "backwards". Thus, the expected number of steps to move from state 1 to 2 is 2, which is also the expected number of steps to move from state 2 to 3.