Markov Chains

As an application of matrix multiplication unrelated to systems of equations, let's consider Markov chains. But first, a quick reminder of some probability basics.

Suppose the probability of event E_1 is p_1 and the probability of event E_2 is p_2. For example, if a die is rolled, event E_1 could be rolling a 3; the probability of that would be $p_1 = 1/6$.

1. If E_1 and E_2 are independent events (so whether or not one event occurs does not depend on whether or not the other event occurs), then the probability that both E_1 *and* E_2 occur is $p_1 p_2$.

   Example. A die is rolled, so the probability of rolling a 3 is $1/6$. A coin is flipped, so the probability of getting a heads is $1/2$. These two events are independent, so the probability of rolling a 3 and getting a heads is $\frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$. ☑

2. If E_1 and E_2 are mutually exclusive (so they can't both occur), then the probability that either E_1 *or* E_2 occurs is $p_1 + p_2$.

   Example. A die is rolled, so the probability of rolling a 3 is $p_1 = 1/6$ and the probability of rolling an even number is $3/6$. These two events are mutually exclusive, so the probability of rolling a 3 or an even number is $\frac{1}{6} + \frac{3}{6} = \frac{4}{6}$. ☑

Before we talk about Markov chains in general, let's start with an example. Suppose each morning for breakfast, you drink either milk (M), water (W) or soda (S). Suppose the probability that you have milk is 0.2, the probability that you have water is 0.3 and the probability that you have soda is 0.5. Since these cover all possibilities, these probabilities need to add up to 1. We could store these values as a ***probability vector***:

$$\begin{array}{ccc} M & W & S \\ \end{array}$$
$$\begin{pmatrix} 0.2 & 0.3 & 0.5 \end{pmatrix}$$

This tells us at a glance the probability of drinking milk, water and soda. It's likely, however, that what we choose to drink depends on what we had to drink the previous breakfast. We might get the above probability vector if we had milk the previous breakfast:

$$\begin{array}{c c} & \begin{array}{ccc} M & W & S \end{array} \\ M & \begin{pmatrix} 0.2 & 0.3 & 0.5 \end{pmatrix} \end{array}.$$

If we drank water for the previous breakfast, we might get the probability vector

$$\begin{array}{c c} & \begin{array}{ccc} M & W & S \end{array} \\ W & \begin{pmatrix} 0.4 & 0.3 & 0.3 \end{pmatrix} \end{array},$$

and if we had soda, we might get the probability vector

$$\begin{array}{c c} & \begin{array}{ccc} M & W & S \end{array} \\ S & \begin{pmatrix} 0.1 & 0.7 & 0.2 \end{pmatrix} \end{array}.$$

The complete situation can be described by stacking the probability vectors to get a matrix:

$$\begin{array}{c c} & \begin{array}{ccc} M & W & S \end{array} \\ \begin{array}{c} M \\ W \\ S \end{array} & \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.4 & 0.3 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{pmatrix} \end{array}.$$

Note that the rows and columns are labeled in the same order. This is an example of a Markov chain, and the matrix is called a stochastic matrix. Since each row of a stochastic matrix is a probability vector, the elements of each row must add up to 1.

More generally, we have the following definition.

Definition. A ***Markov Chain*** or ***Markov Process***is a system with the following properties:

1. There are a finite number of possible states: $\{E_1, E_2, \ldots, E_n\}$.
2. Changes occur in discrete steps.

3. When in state E_i, the probability p_{ij} of going to state E_j depends only on E_i and E_j. $\square$

Each state in a Markov chain will have a probability vector

$$\begin{array}{cccccc} & E_1 & E_2 & E_3 & \dots & E_n \\ E_i & (\, p_{i1} & p_{i2} & p_{i3} & \dots & p_{in}\,) \end{array}$$

It's convenient to display these probability vectors as a stochastic matrix:

$$\begin{array}{cc} & \begin{array}{ccccc} E_1 & E_2 & E_3 & \dots & E_n \end{array} \\ \begin{array}{c} E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_n \end{array} & \begin{pmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \dots & p_{2n} \\ p_{31} & p_{32} & p_{33} & \dots & p_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \dots & p_{nn} \end{pmatrix} \end{array}$$

This matrix is called the ***transition matrix*** for the Markov chain; it stores all the probabilities involved in going from one state to another.

The system needs to start off in some state. (Going back to the breakfast drink situation, you need to have something to drink for the first breakfast.) Suppose the probability it starts off in state E_1 is q_1, the probability it starts off in state E_2 is q_2, etc.; so we start off with the probability vector

$$\vec{q} = \begin{pmatrix} q_1 & q_2 & q_3 & \dots & q_n \end{pmatrix}.$$

Now, the probability of starting off in state E_1 is q_1 and the probability of going from state E_1 and staying in state E_1 is p_{11}; putting this together, the probability of starting in state E_1 and then staying in state E_1 is $q_1 p_{11}$. Similarly, we get the following:

1. The probability of starting in state E_1 and staying in state E_1 is $q_1 p_{11}$.
2. The probability of starting in state E_2 and going to state E_1 is $q_2 p_{21}$.
3. The probability of starting in state E_3 and going to state E_1 is $q_3 p_{31}$.
4. etc.

Altogether, the probability of being in state E_1 after the first step is

$$q_1 p_{11} + q_2 p_{21} + \cdots + q_n p_{n1} = \begin{pmatrix} q_1 & q_2 & q_3 & \dots & q_n \end{pmatrix} \begin{pmatrix} p_{11} \\ p_{21} \\ p_{31} \\ \vdots \\ p_{n1} \end{pmatrix}.$$

Example. Suppose for your first breakfast, there is 0.5 probability you will have milk, 0.25 probability you will have water, and 0.25 probability that you will have soda, so the initial probability vector is

$$\begin{array}{rc} & \begin{array}{ccc} M & W & S \end{array} \\ \vec{q} = & \begin{pmatrix} 0.5 & 0.25 & 0.25 \end{pmatrix}. \end{array}$$

Recalling the stochastic matrix for this system, the probability of having milk for your second breakfast is

$$0.5 \cdot 0.2 + 0.25 \cdot 0.4 + 0.25 \cdot 0.1 = 0.225.$$

☑

Still assuming that the initial probability vector is $\vec{q}$, the probability vector for the state after 1 step is

$$\left(\begin{pmatrix} q_1 & q_2 & \dots & q_n \end{pmatrix} \begin{pmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{pmatrix} \quad \begin{pmatrix} q_1 & q_2 & \dots & q_n \end{pmatrix} \begin{pmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{n2} \end{pmatrix} \quad \dots \quad \begin{pmatrix} q_1 & q_2 & \dots & q_n \end{pmatrix} \begin{pmatrix} p_{1n} \\ p_{2n} \\ \vdots \\ p_{nn} \end{pmatrix} \right)$$

which can be rewritten

$$\begin{pmatrix} q_1 & q_2 & \dots & q_n \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} = \vec{q}P$$

where P is the transition matrix.

Proposition. *Suppose that a Markov process with transition matrix P has an initial probability vector $\vec{q}$. Then the probability vector for the state after one step is $\vec{q}P$.* $\square$

Example. The probability vector for the first breakfast is $\vec{q} = \begin{pmatrix} .5 & .25 & .25 \end{pmatrix}$. Since the transition matrix is

$$P = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.4 & 0.3 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{pmatrix},$$

the probability vector for the second breakfast is

$$\vec{q}P = \begin{array}{c} \begin{array}{ccc} M & W & S \end{array} \\ \begin{pmatrix} .225 & .4 & .375 \end{pmatrix} \end{array}.$$

So, for example, the probability that you will have water for your second breakfast is 0.4. ☑

Continuing, since the probability vector after one step is $\vec{q}P$, the probability vector after two steps will be $(\vec{q}P)P = \vec{q}P^2$, etc. Each step introduces another power of P.

Corollary. *If a Markov process has transition matrix P, then the transition matrix corresponding to k steps is P^k.* $\square$

Example. Using our example with transition matrix

$$P = \begin{array}{c c} & \begin{array}{ccc} M & W & S \end{array} \\ \begin{array}{c} M \\ W \\ S \end{array} & \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.4 & 0.3 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{pmatrix} \end{array}$$

the transition matrix for two steps is

$$P^2 = \begin{array}{c c} & \begin{array}{ccc} M & W & S \end{array} \\ \begin{array}{c} M \\ W \\ S \end{array} & \begin{pmatrix} 0.21 & 0.50 & 0.29 \\ 0.23 & 0.42 & 0.35 \\ 0.32 & 0.38 & 0.3 \end{pmatrix} \end{array}$$

So, for example, the probability of drinking soda for breakfast two days after having water for breakfast is 0.35. ☑