

Markov Processes

1. INTRODUCTION

Before we give the definition of a Markov process, we will look at an example:

Example 1: Suppose that the bus ridership in a city is studied. After examining several years of data, it was found that 30% of the people who regularly ride on buses in a given year do not regularly ride the bus in the next year. Also it was found that 20% of the people who do not regularly ride the bus in that year, begin to ride the bus regularly the next year. If 5000 people ride the bus and 10,000 do not ride the bus in a given year, what is the distribution of riders/non-riders in the next year? In 2 years? In n years?

First we will determine how many people will ride the bus next year. Of the people who currently ride the bus, 70% of them will continue to do so. Of the people who don't ride the bus, 20% of them will begin to ride the bus. Thus:

$$5000(0.7) + 10,000(0.2) = \text{The number of people who ride bus next year.} = b_1$$

By the same argument as above, we see that:

$$5000(0.3) + 10,000(0.8) = \text{The number of people who don't ride the bus next year.} = b_2$$

This system of equations is equivalent to the matrix equation: $M\mathbf{x} = \mathbf{b}$ where

$$M = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}, \mathbf{x} = \begin{pmatrix} 5000 \\ 10,000 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

Note $\mathbf{b} = \begin{pmatrix} 5500 \\ 9500 \end{pmatrix}$. For computing the result after 2 years, we just use the same matrix M , however we use \mathbf{b} in place of \mathbf{x} . Thus the distribution after 2 years is $M\mathbf{b} = M^2\mathbf{x}$. In fact, after n years, the distribution is given by $M^n\mathbf{x}$.

The forgoing example is an example of a Markov process. Now for some formal definitions:

Definition 1. A **stochastic process** is a sequence of events in which the outcome at any stage depends on some probability.

Definition 2. A **Markov process** is a stochastic process with the following properties:

- (a.) The number of possible outcomes or states is finite.
- (b.) The outcome at any stage depends only on the outcome of the previous stage.
- (c.) The probabilities are constant over time.

If \mathbf{x}_0 is a vector which represents the initial state of a system, then there is a matrix M such that the state of the system after one iteration is given by the vector $M\mathbf{x}_0$. Thus we get a chain of state vectors: $\mathbf{x}_0, M\mathbf{x}_0, M^2\mathbf{x}_0, \dots$ where the state of the system after n iterations is given by $M^n\mathbf{x}_0$. Such a chain is called a **Markov chain** and the matrix M is called a **transition matrix**.

The state vectors can be of one of two types: an absolute vector or a probability vector. An **absolute vector** is a vector whose entries give the actual number of objects in a give state, as in the first example. A **probability vector** is a vector where the entries give the percentage (or probability) of objects in a given state. We will take all of our state vectors to be probability vectors from now on. Note that the entries of a probability vector add up to 1.

Theorem 3. Let M be the transition matrix of a Markov process such that M^k has only positive entries for some k . Then there exists a unique probability vector \mathbf{x}_s such that $M\mathbf{x}_s = \mathbf{x}_s$. Moreover $\lim_{k \rightarrow \infty} M^k \mathbf{x}_0 = \mathbf{x}_s$ for any initial state probability vector \mathbf{x}_0 .

The vector \mathbf{x}_s is called a the **steady-state vector**.

2. THE TRANSITION MATRIX AND ITS STEADY-STATE VECTOR

The transition matrix of an n -state Markov process is an $n \times n$ matrix M where the i, j entry of M represents the probability that an object is state j transitions into state i , that is if $M = (\mathbf{m}_{ij})$ and the states are S_1, S_2, \dots, S_n then \mathbf{m}_{ij} is the probability that an object in state S_j transitions to state S_i .

What remains is to determine the steady-state vector. Notice that we have the chain of equivalences:

$$M\mathbf{x}_s = \mathbf{x}_s \Leftrightarrow M\mathbf{x}_s - \mathbf{x}_s = \mathbf{0} \Leftrightarrow M\mathbf{x}_s - I\mathbf{x}_s = \mathbf{0} \Leftrightarrow (M - I)\mathbf{x}_s = \mathbf{0} \Leftrightarrow \mathbf{x}_s \in N(M - I)$$

Thus \mathbf{x}_s is a vector in the nullspace of $M - I$. If M^k has all positive entries for some k , then $\dim(N(M - I))=1$

and any vector in $N(M - I)$ is just a scalar multiple of \mathbf{x}_s . In particular if $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is any non-zero vector in

$N(M - I)$, then $\mathbf{x}_s = \frac{1}{c}\mathbf{x}$ where $c = x_1 + \dots + x_n$.

Example:

A certain protein molecule can have three configurations which we denote as C_1, C_2 and C_3 . Every second the protein molecule can make a transition from one configuration to another configuration with the following probabilities:

$$\begin{array}{ll} C_1 \rightarrow C_2, P = 0.2 & C_1 \rightarrow C_3, P = 0.5 \\ C_2 \rightarrow C_1, P = 0.3 & C_2 \rightarrow C_3, P = 0.2 \\ C_3 \rightarrow C_1, P = 0.4 & C_3 \rightarrow C_2, P = 0.2 \end{array}$$

Find the transition matrix M and steady-state vector \mathbf{x}_s for this Markov process.

Recall that $M = (\mathbf{m}_{ij})$ where \mathbf{m}_{ij} is the probability of configuration C_j making the transition to C_i . Therefore

$$M = \begin{pmatrix} 0.3 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.4 \end{pmatrix} \text{ and } M - I = \begin{pmatrix} -0.7 & 0.3 & 0.4 \\ 0.2 & -0.5 & 0.2 \\ 0.5 & 0.2 & -0.6 \end{pmatrix}$$

Now we compute a basis for $N(M - I)$ by putting $M - I$ into reduced echelon form:

$$U = \begin{pmatrix} 1 & 0 & -0.8966 \\ 0 & 1 & -0.7586 \\ 0 & 0 & 0 \end{pmatrix} \text{ and we see that } \mathbf{x} = \begin{pmatrix} 0.8966 \\ 0.7586 \\ 1 \end{pmatrix} \text{ is the basis vector for } N(M - I)$$

Consequently, $c = 2.6552$ and $\mathbf{x}_s = \begin{pmatrix} 0.3377 \\ 0.2857 \\ 0.3766 \end{pmatrix}$ is the steady-state vector of this process.

Note also that the nullspace of $M - I$ can be found using MATLAB and the function `null()`:

$$\mathbf{x} = \text{null}(M - \text{eye}(3))$$

$$\mathbf{x}_s = \mathbf{x} / \text{sum}(\mathbf{x})$$