MARKOV PROCESSES

In the Linear Algebra book by Lay, Markov chains are introduced in Sections 1.10 (Difference Equations) and 4.9. In this handout, we indicate more completely the properties of the eigenvalues of a stochastic matrix.

Markov processes concern fixed probabilities of making transitions between a finite number of states. We start by defining a probability transition matrix or stochastic matrix.

A *probability vector* $\mathbf{p}$ is vector with each component $p_i \geq 0$ and the sum of the components equal to one, $\sum_i p_i = 1$.

Definition 1. A $n \times n$ matrix $\mathbf{M}$ with real entries m_{ij} is called a *stochastic matrix* or *probability transition matrix* provided that each column of $\mathbf{M}$ is a probability vector. An entry m_{ij} is the probability of going from the j^{th} state to the i^{th} state and satisfies $0 \leq m_{ij} \leq 1$. The total probability of going from the j^{th} state to some other state is 1, $\sum_i m_{ij} = 1$ for all j. so each column sums to one.

Remark 2. *Warning*: In most applications of Markov processes (and Markov chains) m_{ij} is usually the probability of going from the i^{th} state to the j^{th} state so the $\mathbf{M}$ is replaced by its transpose $\mathbf{N} = \mathbf{M}^T$, which has row sum zero. Then, when we defined the action on a state vector (below), we have to multiply a row vector vector on the left rather than the usual column vector on the right, $\mathbf{x}^T\mathbf{N}$.

Note that if the i^{th} row has all zero entries, then there is no way to get to the i^{th} state. Therefore, we might as well drop it from consideration. Therefore, we always assume that each row has some nonzero entry.

If every column has only one nonzero entry, then all the entries are 0s and 1s: each state goes to a unique next state. If it also has the property that each row has some nonzero entry, then each row has one 1 and there is a unique state that comes to each i^{th} state. Such a matrix is called a *permutation matrix*; it is not "stochastic" so we assume some column has more than one nonzero entry.

Assume that $\mathbf{M}$ is an $n \times n$ stochastic matrix. Assume that some material is spread out among the n states, with $x_j^{(q)} \geq 0$ the amount of material at the j^{th} state at time q and $\mathbf{x}^{(q)} = (x_1^{(0)}, \dots, x_n^{(q)})^T$ be the column vector of the amount of material at time q in all the states. The material from the j^{th} state at time 0 that is returned to the i^{th} state at time 1 is given by $m_{ij}x_j^{(0)}$. The total amount at the i^{th} state at time 1 is the sum of the material from all the states, or

$$x_i^{(1)} = \sum_j m_{ij} x_j^{(0)}.$$

By the preceding equation for each component,

$$\mathbf{x}^{(1)} = \mathbf{M}\mathbf{x}^{(0)},$$

and, more generally,

$$\mathbf{x}^{(q)} = \mathbf{M}\mathbf{x}^{(q\text{-}1)},$$

for the transition from the material at the state at time q -1 to time q. By induction, $\mathbf{x}^{(q)} = \mathbf{M}^q\mathbf{x}^{(0)}$.

Notice that the total amount of material at time q is the same as at time q -1:

$$\begin{aligned} \sum_i x_i^{(q)} &= \sum_i \left(\sum_j m_{ij} x_j^{(q\text{-}1)} \right) \\ &= \sum_j \left(\sum_i m_{ij} \right) x_j^{(q\text{-}1)} \\ &= \sum_j x_j^{(q\text{-}1)}. \end{aligned}$$

(We use the fact that the column sums are 1.) By induction, the total amount X remains the same for all time periods, $\sum_j x_j^{(q)} = \sum_j x_j^{(q\text{-}1)} = \cdots = \sum_i x_i^{(0)}$. Call this total amount $X = \sum_i x_i^{(0)}$ Then, $p_j^{(q)} = {x_j^{(q)}}/X$ is the proportion of the material at the j^{th} state at time q. Letting

$$\mathbf{p}^{(q)} = (p_1^{(q)}, \dots, p_n^{(q)})^T = \frac{1}{X}\mathbf{x}^{(q)}$$

be the vector of these proportions, we get a probability vector. Then,

$$\begin{aligned}\mathbf{M}\mathbf{p}^{(q\text{-}1)} &= \mathbf{M}\frac{\mathbf{x}^{(q\text{-}1)}}{X}\\ &= \frac{1}{X}\mathbf{M}\mathbf{x}^{(q\text{-}1)}\\ &= \frac{1}{X}\mathbf{x}^{(q)}\\ &= \mathbf{p}^{(q)};\end{aligned}$$

thus, the probability vectors also transform through multiplication by the matrix $\mathbf{M}$.

Definition 3. For a stochastic matrix $\mathbf{M}$, the transformation $\mathbf{p}^{(q)} = \mathbf{M}\mathbf{p}^{(q\text{-}1)}$ on probability vectors is called a (finite) *Markov process*.

For a given initial probability vector $\mathbf{p}^{(0)}$, the sequence of resulting iterates $\mathbf{p}^{(q)} = \mathbf{M}^q\mathbf{p}^{(0)}$ is called a *Markov chain*.

Definition 4. A stochastic matrix $\mathbf{M}$ is called *regular* provided that there is a $q_0 > 0$ such that $\mathbf{M}^{q_0}$ has all positive entries, i.e., it is possible to make a transition from any state to any other state with exactly q_0 transitions. It then follows that $\mathbf{M}^q$ has all positive entries for $q \geq q_0$. A regular stochastic matrix automatically has the properties that each row has some nonzero entry and some column has more than one nonzero entry.

Definition 5. A stochastic matrix is called *irreducible* provided that it is possible to get from each state to any other state by making a finite number of transitions. More precisely, for any pair of states (i_0, j_0), there are $q \geq 1$ that can depend on (i_0, j_0) and indices j_k with $1 \leq j_k \leq n$ for $k = 1, \dots, q$, such that $j_q = i_0$ and $m_{j_k j_{k\text{-}1}} > 0$ for $k = 1, \dots, q$. This means that for each pair of states (i_0, j_0), there is a $q \geq 1$ such that the (i_0, j_0)-entry of $\mathbf{M}^q$ is nonzero.

Since a stochastic matrix $\mathbf{M}$ always has column sum one, $(1, \dots, 1)\mathbf{M} = (\sum_i m_{i1}, \dots, \sum_i m_{in}) = (1, \dots, 1)$. Taking the transpose, $\mathbf{M}^T(1, \dots, 1)^T = (1, \dots, 1)^T$, and $\mathbf{M}^T$ always has $(1, \dots, 1)^T$ as an eigenvector corresponding to the eigenvalue 1. Since $\mathbf{M}^T$ and $\mathbf{M}$ have the same eigenvalues, $\mathbf{M}$ always has 1 as an eigenvalue for some eigenvector $\mathbf{p}^*$,

$$\mathbf{p}^* = \mathbf{M}\mathbf{p}^*.$$

The following theorem gives conditions on the eigenvalues and eigenvectors for any stochastic matrix and for a regular stochastic matrix.

Theorem 6 (Perron–Frobenius). *(**a**) Assume that* $\mathbf{M}$ *is a stochastic matrix with each row containing a nonzero entry. Then,*

(i) $\mathbf{M}$ *has a probability vector* $\mathbf{p}^*$ *with all* $p_i > 0$ *as an eigenvector for the eigenvalue* 1, *and*

(ii) all the eigenvalues λ *of* $\mathbf{M}$ *satisfy* $|\lambda| \leq 1$.

*(**b**) Assume that* $\mathbf{M}$ *is a regular stochastic matrix.*

(i) The matrix $\mathbf{M}$ *has* 1 *as an eigenvalue of multiplicity one (i.e.,* 1 *is a simple root of the characteristic equation). An eigenvector* $\mathbf{p}^*$ *for the eigenvalue 1 can be chosen as a probability vector with all* $p_i > 0$.

(ii) All the other eigenvalues λ_i *have* $|\lambda_i| < 1$. *If* $\mathbf{v}^k$ *is a eigenvector for* λ_k, *then* $\sum_i v_i^k = 0$.

(iii) If $\mathbf{p}$ *is any probability vector with all* $p_i > 0$ *and* $\sum_i p_i = 1$, *then* $\mathbf{p} = \mathbf{p}^* + \sum_{j=2}^n y_k\mathbf{v}^k$ *for some choice of the* y_k. *Also,* $\mathbf{M}^q\mathbf{p}$ *converges to* $\mathbf{p}^*$ *as* q *goes to infinity.*

See [3] for a proof using iteration, [2] for a proof using norms of matrices, and Chapter XIII in Volume II of [1] for more general results.

We now give some examples.

Example 7. Let

$$\mathbf{M} = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{pmatrix}.$$

A direct calculations shows that $\mathbf{M}^2$ has all positive entries, so $\mathbf{M}$ is regular. This matrix has eigenvalues 1, 0.5, and 0.2. (We do not give the characteristic polynomial, but do derive an eigenvector for each of these eigenvalues.)

For $\lambda = 1$,

$$\mathbf{M} - \mathbf{I} = \begin{pmatrix} -0.5 & 0.2 & 0.3 \\ 0.3 & -0.2 & 0.3 \\ 0.2 & 0 & -0.6 \end{pmatrix}$$
$$\sim \begin{pmatrix} 1 & -0.4 & -0.6 \\ 0 & -0.08 & 0.48 \\ 0 & 0.08 & -0.48 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus, $v_1 = 3v_3$ and $v_2 = 6v_3$. Since we want $1 = v_1 + v_2 + v_3 = (3 + 6 + 1)v_3 = 10v_3$, $v_3 = 0.1$, and $\mathbf{p}^* = \mathbf{v}^1 = (0.3, 0.6, 0.1)^T$.

For $\lambda_2 = 0.5$,

$$\mathbf{M} - 0.5\,\mathbf{I} = \begin{pmatrix} 0 & 0.2 & 0.3 \\ 0.3 & 0.3 & 0.3 \\ 0.2 & 0 & -0.1 \end{pmatrix}$$
$$\sim \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 3 \\ 0 & 1 & 1.5 \end{pmatrix}$$
$$\sim \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus, $2v_1 = v_3$, $2v_2 = -3v_3$, and $\mathbf{v}^2 = (1, -3, 2)^T$ is an eigenvector. Notice that $v_1 + v_2 + v_3 = 1 - 3 + 2 = 0$, as Theorem 6 says must be the case.

For $\lambda_3 = 0.2$,

$$\mathbf{M} - 0.2\,\mathbf{I} = \begin{pmatrix} 0.3 & 0.2 & 0.3 \\ 0.3 & 0.6 & 0.3 \\ 0.2 & 0 & 0.2 \end{pmatrix}$$
$$\sim \begin{pmatrix} 1 & 0 & 1 \\ 3 & 6 & 3 \\ 3 & 2 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 6 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$
$$\sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus, $v_1 = -v_3$, $v_2 = 0$, and $\mathbf{v}^3 = (1, 0, -1)^T$ is an eigenvector. Again, $v_1 + v_2 + v_3 = 1 + 0 - 1 = 0$.

If the original probability vector is given by $\mathbf{p}^{(0)} = (0.45, 0.45, 0.1)^T$, then

$$\mathbf{p}^{(0)} = (0.3, 0.6, 0.1)^T + \frac{1}{20}(1, -3, 2)^T + \frac{1}{10}(1, 0, -1)^T$$

and

$$\mathbf{M}^q \mathbf{p}^{(0)} = (0.3, 0.6, 0.1)^T + \frac{1}{20}\left(\frac{1}{2}\right)^q (1, \text{-}3, 2)^T + \frac{1}{10}\left(\frac{1}{5}\right)^q (1, 0, \text{-}1)^T,$$

which converges to the probability vector $\mathbf{p}^* = (0.3, 0.6, 0.1)^T$ as q goes to infinity. This convergence of the iterates holds for any initial probability vector $\mathbf{p}^{(0)}$. See Theorem 6. ■

Example 8 (Complex Eigenvalues). The following stochastic matrix

$$\mathbf{M} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.1 & 0.6 & 0.3 \\ 0.3 & 0.1 & 0.6 \end{pmatrix}.$$

illustrates the fact that an regular stochastic matrix can have complex eigenvalues. The eigenvalues are $\lambda = 1$ and $0.4 \pm i\, 0.1\sqrt{3}$. Notice that $|0.4 \pm i\, 0.1\sqrt{3}| = \sqrt{0.16 + 0.03} = \sqrt{0.19} < 1$. The eigenvectors are $(1/3, 1/3, 1/3)^T$ and $(-1, -1, 2)^T \pm i\,(\sqrt{3}, -\sqrt{3}, 0)^T$. ■

Example 9 (Not irreducible). An example of a stochastic matrix that is not irreducible is given by

$$\mathbf{M} = \begin{pmatrix} 0.8 & 0.3 & 0 & 0 \\ 0.2 & 0.7 & 0 & 0 \\ 0 & 0 & 0.6 & 0.3 \\ 0 & 0 & 0.4 & 0.7 \end{pmatrix},$$

which has eigenvalues $\lambda = 1, 1, 0.5$, and 0.3. Notice that it is possible to go between states 1 and 2, and it is possible to go between and states 3 and 4, but it is not possible to go from the states 1 and 2 to the states 3 and 4. Since it is not irreducible, it is also not regular. The eigenvectors for $\lambda = 1$ are $\mathbf{v}^1 = (0.6, 0.4, 0, 0)^T$, $\mathbf{v}^2 = (0, 0, 3/7, 4/7)^T$, and any linear combinations of $\mathbf{v}^1$ and $\mathbf{v}^2$. In particular, the average of $\mathbf{v}^1$ and $\mathbf{v}^2$ is an eigenvector that satisfies the conditions of the Perron-Frobenius Theorem, $\mathbf{p}^* = (0.3, 0.2, 3/14, 4/14)^T$. ■

Example 10 (Not Regular). An example of a stochastic matrix that is irreducible, but not regular, is given by

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 0.8 & 0.3 \\ 0 & 0 & 0.2 & 0.7 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

which has eigenvalues $\lambda = 1$, -1, and $\pm\sqrt{0.5}$. Here, it is possible to get from any state to any other state, but starting at state one, the odd iterates are always at either states 3 or 4 and the even iterates are always at either states 1 or 2. Thus, there is no one power for which all the transition probabilities are positive. Therefore, $\mathbf{M}$ is not regular. Also, this matrix has another eigenvalue -1 with absolute value equal to one. The probability eigenvector for the eigenvalue 1 is $(0.3, 0.2, 0.3, 0.2)^T$. ■

References

[1] Gantmacher, F.R., *The Theory of Matrices*, Chelsea Publ. Co., New York, 1959.
[2] Olver, P. and C. Shakiban, *Applied Linear Algebra*, Pearson Prentice Hall, Upper Saddle River, New Jersey, 2006.
[3] Robinson, C., *Introduction to Dynamical Systems: Continuous and Discrete*, Pearson Prentice Hall, Upper Saddle River, New Jersey, 2004.