Inter-Section

A Note to the Reader

Until this point in the book, results have been derived or proven, or at least the nature of the proof has been indicated. However, proofs of the techniques of Sections 34, 35 and 36 as well as the core concept of Chapter 37 are beyond this book and have been omitted. Do not let this deter you. This material is still the most exciting material so far. Learning to use these techniques is rewarding and leads to substantial understanding. The fact is most people who use these methods have probably never seen the proofs. Similarly, you can read and verify a proof and still not understand the theorem. Do not let yourself be deterred by math snobs. You can always get to the proofs later. The idea that it is always best for semantics to precede syntax is quite arguable and probably is a lousy generalization. I will give sources for the proofs for those who are ready now, and for the rest when they are ready. To be able to read these sources you should have had a course in linear algebra and you should have taken the usual calculus sequence. The methods of Sections 34 and 35 are justified in:

Finite Markov Chains. John G. Kemeny and J. Laurie Snell. Springer-Verlag, New York, N.Y. 1976.

The core concept of Section 37 (that every two person zero-sum game has a value) is demonstrated in:

Linear Programming. Vašek Chvátal. W. H. Freeman. New York, N.Y. 1983.

An Introduction to Linear Programming and Game Theory, 2nd. ed. Paul R. Thie. Wiley. New York, N.Y. 1988.

34

Ergodic Chains

Earlier we characterized ergodic chains as having both of the following two properties:

- Each node can reach any other node (the chain is one communicating class).
- The chain is not periodic.

An ergodic chain can be characterized in another way. The long term state probabilities are independent of the initial state vector.[1] These long term probabilities are also known as the *steady state* probabilities. Consider the Markov chain in **Figure 1**. This chain has the quality that after roughly 40 transitions, there is a 9/68 probability of being in state A; there is a 5/68 probability of being in state B; there is a 9/68 probability of being in state C; and there is a 45/68 probability of being in state D. Note, it does not matter in what state you started. With an ergodic chain, if we wait for a sufficient number of transitions where that number of transitions depends on the chain, then the probability of being in any state is independent of where we started.[2]

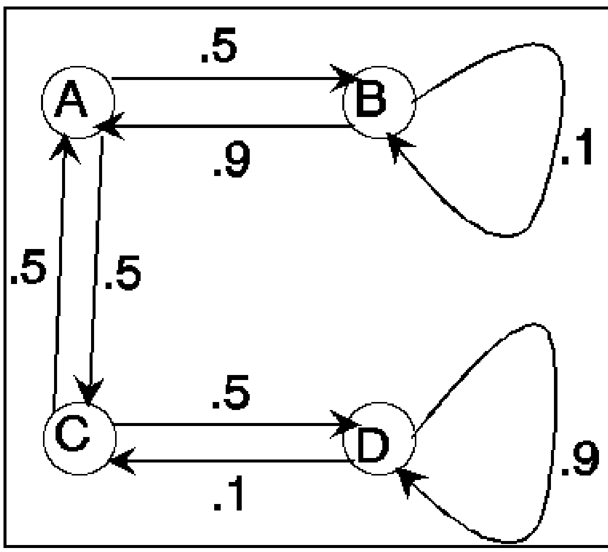

Figure 1 An Ergodic Chain

[1]I remind you that we are only considering finite chains.

[2]This is not **precisely** true. What state we start in does effect the probability of being in any state no matter how far ahead we look. However, this effect becomes extremely small, and in any practical sense is insignificant after a sufficient number of transitions.

$$M := \begin{bmatrix} 0 & .5 & .5 & 0 \\ .9 & .1 & 0 & 0 \\ .5 & 0 & 0 & .5 \\ 0 & 0 & .1 & .9 \end{bmatrix}$$

$$M^{40} = \begin{bmatrix} 0.133 & 0.074 & 0.132 & 0.662 \\ 0.132 & 0.074 & 0.132 & 0.662 \\ 0.132 & 0.074 & 0.132 & 0.662 \\ 0.132 & 0.074 & 0.132 & 0.662 \end{bmatrix}$$

Figure 2 Transition Matrix to the 40'th Power

One way of finding the long term probabilities of an Ergodic chain is to take the transition matrix M and an arbitrary initial vector v, say v = (1,0,...,0). Compute the state vector for after n transitions, that is $v \cdot M^n$. When you reach a level n, where $v \cdot M^n \doteq v \cdot M^{n+1}$, that is the state vector after n transitions is nearly identical to the state vector after n + 1 transitions, then that vector gives the long term probabilities. Another way to find the long term probabilities of an ergodic chain is to consider high powers of the transition matrix M until the matrix *stabilizes*, that is, it does not change (within, say, ten decimal places). When that happens, each row of the matrix M^n will be identical with the steady state vector (the vector giving long term probabilities). In the case of **Figure 1** the matrix begins to stabilize at about n = 40; that is it stabilizes at 40 transitions as is shown in **Figure 2**. This shows the steady state vector for the problem of **Figure 1** to be v = (.132, .074, .132, .662). Note that these long term probabilities are consistent with the fractions given above.

Solution of Ergodic Chains by Linear Equations

The steady state vector can be solved for using linear algebra at the high school level. I am going to use this technique in the examples in this chapter. I want to show you how it is done. However, in practice, you should use a computer program to do this. Anything with more than about four states is usually too cumbersome to do by hand.[1] It so happens that the steady state vector, v, is the only probability vector that will satisfy the equation $v \cdot M = v$. Such a vector

[1]This raises the question why am I showing how to do these problems by hand when I suggest using computer programs? The reason is that I feel that seeing how it is done by hand can add to the student's understanding of the problem and its solution.

is known as an *eigenvector*.[1] There can be many eigenvectors that satisfy $v \cdot M = v$, but only one will be a probability vector. In the case of our problem, we want to solve:

$$(a,b,c,d)\begin{pmatrix} 0 & .5 & .5 & 0 \\ .9 & .1 & 0 & 0 \\ .5 & 0 & 0 & .5 \\ 0 & 0 & .1 & .9 \end{pmatrix} = (a,b,c,d)$$

In the above matrix equation, a, b, c, d, represent the long term probabilities of being in states A, B, C, and D respectively. Multiplying out the left side, we get a system of four equations in four unknowns:

$$\begin{array}{lclcl} .9b & + & .5c & = & a \qquad (1) \\ .5a & + & .1b & = & b \qquad (2) \\ .5a & + & .1d & = & c \qquad (3) \\ .5c & + & .9d & = & d \qquad (4) \end{array}$$

In a systems of equations for an ergodic chain such as this, it is **always** true that one equation is redundant. We can and we should eliminate one equation. Often it is the case that one equation will be more onerous that the others, and that equation will be eliminated. Such is not the case here, and we will arbitrarily eliminate Equation 4. That leaves us with three equations in four unknowns. However, the long term probabilities must satisfy $a + b + c + d = 1$. Such is **always** the case (and this is what makes (a, b, c, d) a probability vector). In every such problem, you should eliminate one of the original equations and add the equation that has the

[1]Such terminology is of no use to us in this book. However, it is useful to make yourself sound knowledgeable when discussing these topics. Essentially, *eigenvector* is *cool* terminology. If you were to refer to *that vector that when multiplied times the transition matrix gives itself*, that would be gauche and people might stop associating with you.

variables adding to 1. Doing this here and multiplying Equations 1, 2, and 3 by 10 and moving all the variables to the left hand side, we get the system of equations:

$$\begin{array}{llll} -10a + 9b + 5c & = & 0 & (1) \\ 5a - 9b & = & 0 & (2) \\ 5a - 10c + d & = & 0 & (3) \\ a + b + c + d & = & 1 & (4) \end{array}$$

The usual technique for solving a system such as this is to solve for each variable in terms of a single variable, then substitute in Equation 4. We will solve for each variable in terms of variable a. Equation 2 gives us $b = \frac{5}{9}a$. Substituting this in Equation 1, we get **c = a**.

Substituting this in Equation 3, we get **d = 5a**. Substituting for b, c, and d in Equation 4, we get $a + \frac{5}{9}b + a + 5a = 1$. From this we get the values given earlier:

$$a = \frac{9}{68}; \quad b = \frac{5}{68}; \quad c = \frac{9}{68}; \quad d = \frac{45}{68} .$$

An Ergodic Example: Market Share[1]

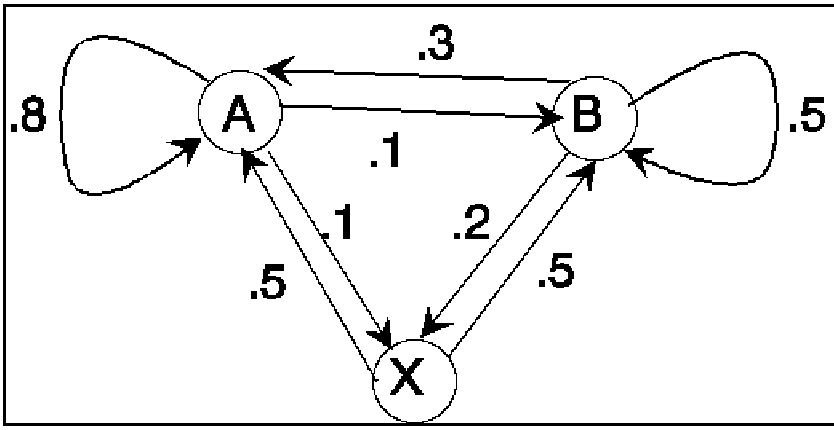

Figure 3 The Product Loyalty Graph for Slamclaw Customers

Consider the market positions of three brands of slamclaws: Brand A, Brand B, and Brand X. Customers who use Brand A return to it 80% of the time; 10% go to Brand B; and the remaining 10% go to Brand X. Customers who use Brand B return to it 50% of the time; 30% go to brand A; and 20% go to Brand X. Customers who use Brand X go to Brand A 50% of the time and to Brand B 50% of the time. The market shares of the three brands are given in **Figure 3**. We would like to analyze the long term prospects in the slamclaw business, and in particular we want to examine the potentials of the three *players*.

To solve for the long term probabilities of the Markov chain of **Figure 3**, we need to solve the matrix equation:

$$(a,b,x)\begin{pmatrix} .8 & .1 & .1 \\ .3 & .5 & .2 \\ .5 & .5 & 0 \end{pmatrix} = (a,b,x)$$

[1]The market share example and the finite queue example (in a less general form) both occur in a long-out-of-print book by Guillermo Owen titled *Finite Mathematics* (I think).

In this equation the variables a, b, and x represent the long term (or steady state) probabilities for Brands A, B, and X. The corresponding linear equations are:

$$\begin{aligned} .8a + .3b + .5x &= a \qquad (1) \\ .1a + .5b + .5x &= b \qquad (2) \\ .1a + .2b \qquad &= x \qquad (3) \end{aligned}$$

Again, one of these equations is always redundant. In this case Equation 3 is slightly simpler than the other equations, so that is the equation that we definitely will **not** remove. Rather arbitrarily we will remove Equation 2. We will multiply each equation by 10 to get rid of the decimal points, and we will move all the variables to the left hand side of the equal signs. Lastly, we must add the probability equation **a + b + x = 1** so that again we have three equations in three unknowns. This gives us:

$$\begin{aligned} -2a + 3b + 5x &= 0 \qquad (1) \\ a + 2b - 10x &= 0 \qquad (2) \\ a + b + x &= 1 \qquad (3) \end{aligned}$$

Adding 2 times Equation 1 to Equation 2 (and simplifying) we get $b = \frac{3}{8}a$. Substituting for b in Equation 2 (and simplifying) we get $x = \frac{7}{40}a$. Lastly, substituting for both b and x in Equation 3 we can solve for a. From that we can solve directly for b, and x. The values are:

$$a = \frac{20}{31}; \quad b = \frac{15}{62}; \quad x = \frac{7}{62}$$

As expected Brand A totally dominates. Surprisingly, Brand B's long term market share is closer to Brand X's than Brand A's.

☐ **Exercise 1** Analyze the market share example by studying the transition matrix. Look at high powers of the matrix. (Keep squaring the matrix until you reach at least the 16th power (use software).) Use different initial state vectors and derive the state vectors after many transitions. Either way, you should get results consistent with the analysis above.

☐ **Exercise 2** Find the long term probabilities for the graph in **Figure 4**.

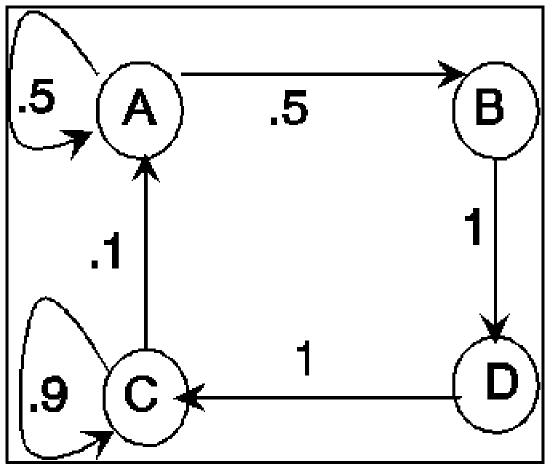

Figure 4 A Four Node Ergodic Chain

☐ **Exercise 3** Find the long term probabilities for the graph in **Figure 4**. (Even though this is a five node chain, its equations are particularly easy to solve by hand.) Note that without the loop from a to a, that it would be periodic with period 2.

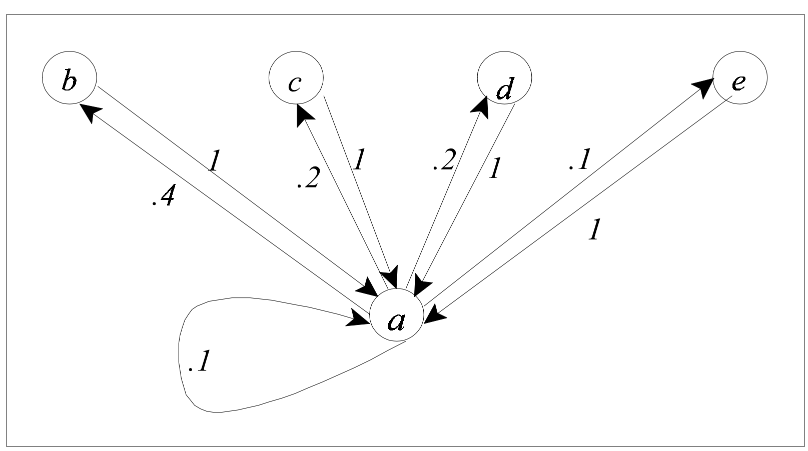

Figure 5 A Five Node Markov Chain

The Finite Queue

By a queue, we mean a line. In formal mathematics and engineering they are called queues, and their study is called queueing theory.[1] This example is nice and I think it is important. After we have finished with the work on the example, I will tell you why the example is more realistic (and practical) than it seems.

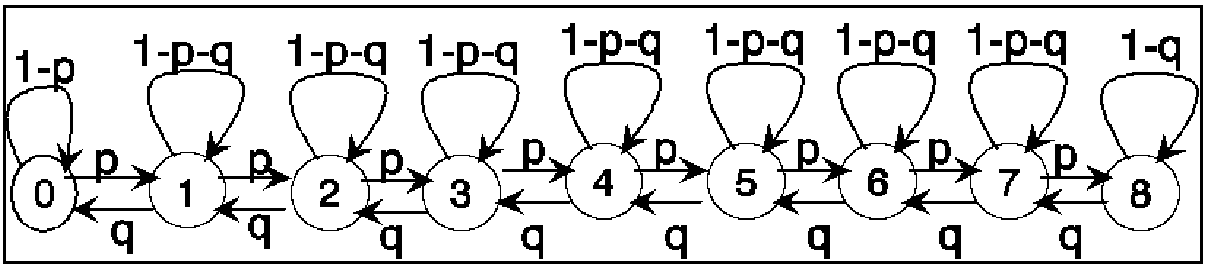

Figure 6 A Finite Queue

The reason that we have a finite queue in **Figure 6** is that whenever there are 8 people in the shop, the doors are locked. The state of the queue is the number of people in the queue. State 0 thus corresponds to an empty queue. We assume that in an interval of time, that there is a probability of p that a customer arrives. Customers always arrive singly.[2] In the same interval of time there is a probability of q that a customer is served. It never happens that there is a service and an arrival in the same interval of time. Whenever there is an arrival, since that means a new customer in the shop, the state increases by 1. Whenever, there is a service, the customer served leaves the shop, and the state decreases by 1. In states 1 through 7 the probability that the state does not change in an interval of time is the probability, $1-p-q$, that there is no service nor an arrival during that interval. In state 0 there can be no service since there is no one in the shop to be served. Hence the probability of staying in state 0, during an interval of time, is the probability, $1-p$ that there is no arrival. In state 8 the doors are locked and there can be no arrival, therefore the probability $1-q$, of staying in state 8 during an interval

[1]In Britain lines are actually called queues, and one must presume they study *line theory*.

[2]This is a realistic assumption with many criminal enterprises.

of time is the probability that there is no service. Letting x_i represent the long term probability of being in state i, the matrix form for solution of the long term probabilities of each state is:

$$(x_0,x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8)\begin{pmatrix} 1-p & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ q & 1-p-q & p & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & q & 1-p-q & p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & q & 1-p-q & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q & 1-p-q & p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q & 1-p-q & p & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & q & 1-p-q & p & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & q & 1-p-q & p \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 1-q \end{pmatrix}$$

$$= \quad (x_0,x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8)$$

This is equivalent to the nine equations:

$$\begin{aligned} (1-p)x_0 + qx_1 &= x_0 && (1) \\ px_0 + (1-p-q)x_1 + qx_2 &= x_1 && (2) \\ px_1 + (1-p-q)x_2 + qx_3 &= x_2 && (3) \\ px_2 + (1-p-q)x_3 + qx_4 &= x_3 && (4) \\ px_3 + (1-p-q)x_4 + qx_5 &= x_4 && (5) \\ px_4 + (1-p-q)x_5 + qx_6 &= x_5 && (6) \\ px_5 + (1-p-q)x_6 + qx_7 &= x_6 && (7) \\ px_6 + (1-p-q)x_7 + qx_8 &= x_7 && (8) \\ px_7 + (1-q)x_8 &= x_8 && (9) \end{aligned}$$

As always one equation is redundant and is to be replaced by:

$$x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1 \qquad (10)$$

Ordinarily one would replace any equation other than Equations 1 and 9, since they are simpler than the other equations. However, there is a pattern to the equations that we will utilize and it is in fact Equation 9 that we will replace.

Solving Equation 1 for x_1 in terms of x_0, we get: $x_1 = \frac{p}{q}x_0$. Replacing x_1 in Equation 2 and solving for x_2 we get: $x_2 = \left(\frac{p}{q}\right)^2 x_0$. Using the same strategy through Equation 7 we get in general: $x_i = \left(\frac{p}{q}\right)^i x_0; \quad i = 1, 2, \ldots, 8$. Substituting into Equation 10, we get the finite geometric series:

$$x_0 + x_0\left(\frac{p}{q}\right)^1 + x_0\left(\frac{p}{q}\right)^2 + x_0\left(\frac{p}{q}\right)^3 + x_0\left(\frac{p}{q}\right)^4 + x_0\left(\frac{p}{q}\right)^5 + x_0\left(\frac{p}{q}\right)^6 + x_0\left(\frac{p}{q}\right)^7 + x_0\left(\frac{p}{q}\right)^8 = 1$$

Factoring out the x_0, we get:

$$x_0\left(1 + \left(\frac{p}{q}\right)^1 + \left(\frac{p}{q}\right)^2 + \left(\frac{p}{q}\right)^3 + \left(\frac{p}{q}\right)^4 + \left(\frac{p}{q}\right)^5 + \left(\frac{p}{q}\right)^6 + \left(\frac{p}{q}\right)^7 + \left(\frac{p}{q}\right)^8\right) = 1$$

Using the finite geometric series formula we get: $x_0 = \dfrac{1 - \frac{p}{q}}{1 - \left(\frac{p}{q}\right)^9}$, from which it follows that:

$$x_i = \frac{\left(\frac{p}{q}\right)^i - \left(\frac{p}{q}\right)^{i+1}}{1 - \left(\frac{p}{q}\right)^9}; \quad i = 0, 1, \ldots, 8 \ .$$

☐ **Exercise 4** The last two formulas do not work in the case where p = q. What is x_i when p = q?

☐ **Exercise 5** The queue example has treated a shop where the doors close when there are 8 people in the shop. Give the general formula for when there are N people in the shop.

The queue example seems to have several unreasonable assumptions. In the period of time concerned there can be exactly one arrival, or one service, or neither of these. There cannot be two arrivals, or two services, or an arrival and a service. Thus the problem seems unrealistic. However, these assumptions become far more realistic when the period of time is very short, such as 1 second. In queueing theory the analysis is often done much as we have done it here, with the time period then taken to an infinitesimal limit. Thus the model **can** be practical. Also, the methods used here are similar in other ways to many of the methods used in queueing theory. Lastly, by closing the doors of the shop, we made the queue space finite. However, these methods can be extended. We can consider a potentially infinite queue and an infinite Markov chain. For such a queue to make sense the arrival probability must be less than the service probability (otherwise the queue actually goes towards infinity). With such an assumption our denominator $(p/q)^n$ goes to 0 and the formula derived here still works.

1.

$$M := \begin{pmatrix} .8 & .1 & .1 \\ .3 & .5 & .2 \\ .5 & .5 & 0 \end{pmatrix}$$

$$M^{64} = \begin{pmatrix} 0.645 & 0.242 & 0.113 \\ 0.645 & 0.242 & 0.113 \\ 0.645 & 0.242 & 0.113 \end{pmatrix}$$

Long term probabilities are approximately (.645, .242, .113).

2. The long-term probabilities are: a = 1/7; b = 1/14; c = 5/7; d = 1/14

3. a = 10/19; b = 4/19; c = 2/19; d = 2/19; e = 1/19

4. If p = q each state is equally likely. We get $x_i = 1/9$. (We basically get the finite geometric series when x = 1. $1 + x + .. x^n = n+1$.

5. If $p \neq q$ let $\rho = p/q$. Let N be the number of the people we allow into the shop. Then $x^i = (\rho^i - \rho^{i+1})/(1 - \rho^{N+1})$ for i = 0, .. , N.

   If p = q then $x^i = 1/(N+1)$ for i = 0, .. , N.