DATA NETWORKS

Dimitri Bertsekas

Massachusetts Institute of Technology

Robert Gallager

Massachusetts Institute of Technology

Prentice-Hall International, Inc.

A Division of Simon & Schuster
Englewood Cliffs, NJ 07632

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-196981-1 025

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*
Prentice-Hall, *Englewood Cliffs, New Jersey*

APPENDIX A: *Review of Markov Chain Theory*

The purpose of this appendix is to provide a brief summary of the results we need from discrete- and continuous-time Markov chain theory. We refer the reader to books on stochastic processes for detailed accounts.

3A.1 Discrete-Time Markov Chains

Consider a discrete-time stochastic process $\{X_n | n = 0, 1, 2, \ldots\}$ that takes values from the set of nonnegative integers, so the states that the process can be in are $i = 0, 1 \ldots$. The process is said to be a *Markov chain* if whenever it is in state i, there is a fixed probability P_{ij} that it will next be in state j regardless of the process history prior to arriving at i. That is, for all $n > 0$, $i_{n-1}, \ldots, i_0, i, j$

$$\begin{aligned} P_{ij} &= P\{X_{n+1} = j | X_n = i,\ X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} \\ &= P\{X_{n+1} = j | X_n = i\} \end{aligned}$$

We refer to P_{ij} as the *transition probabilities*. They must satisfy

$$P_{ij} \geq 0, \ \sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, \ldots$$

The corresponding transition probability matrix is denoted

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ \cdot & \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \cdot & \cdot & \cdot & \cdot \quad \cdot \quad \cdot \end{bmatrix}$$

Consider the n-step transition probabilities

$$P_{ij}^n = P\{X_{n+m} = j | X_m = i\}, \quad n \geq 0, i, j \geq 0 .$$

The *Chapman-Kolmogorov equations* provide a method for calculating P_{ij}^n. They are given by

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m, \quad n, m \geq 0, i, j \geq 0$$

From these equations, we see that P_{ij}^n are the elements of the matrix P^n (the transition probability matrix P raised to the n^{th} power).

We say that two states i and j *communicate* if for some n and n', we have $P_{ij}^n > 0$, $P_{ji}^{n'} > 0$. If all states communicate, we say that the Markov chain is *irreducible.* We say that the Markov chain is *aperiodic* if for each state i there is no integer $d \geq 2$ such that $P_{ii}^n = 0$ except when n is a multiple of d. A probability distribution $\{p_j | j \geq 0\}$ is said to be a *stationary distribution* for the Markov chain if

$$p_j = \sum_{i=0}^{\infty} p_i P_{ij}, \quad j \geq 0. \tag{3A.1}$$

We will restrict attention to irreducible and aperiodic Markov chains, since this is the only type we will encounter. For such a chain, denote

$$p_j = \lim_{n \to \infty} P_{jj}^n, \quad j \geq 0$$

It can be shown that the limit above exists and when $p_j > 0$, then $1/p_j$ equals the *mean recurrence time of j, i.e.*, the expected number of transitions between two successive visits to state j. If $p_j = 0$, the mean recurrence time is infinite. Another interpretation is that p_j represents the proportion of time the process visits j on the average. The following result will be of primary interest:

Theorem. In an irreducible, aperiodic Markov chain, there are two possibilities:

1. $p_j = 0$ for all $j \geq 0$ in which case the chain has no stationary distribution.
2. $p_j > 0$ for all $j \geq 0$ in which case $\{p_j | j \geq 0\}$ is the unique stationary distribution of the chain.

A typical example of case 1 above is an $M/M/1$ queueing system where the arrival rate λ exceeds the service rate μ.

In case 2, there arises the issue of characterizing the stationary distribution $\{p_j | j \geq 0\}$. For queueing systems, the following technique is often useful. Multiplying the equation $P_{jj} + \sum_{i=0, i \neq j}^{\infty} P_{ji} = 1$ by p_j and using Eq. (3A.1), we have

$$p_j \sum_{\substack{i=0 \\ i \neq j}}^{\infty} P_{ji} = \sum_{\substack{i=0 \\ i \neq j}}^{\infty} p_i P_{ij} \tag{3A.2}$$

These equations are known as the *global balance equations.* They state that, at equilibrium, the probability of a transition out of j (left side of Eq. (3A.2)) equals the probability of a transition into j (right side of Eq. (3A.2)).

The global balance equations can be generalized to apply to an entire set of states. Consider a subset of states S. By adding Eq. (3A.2) over all $j \epsilon S$, we obtain

$$\sum_{j \epsilon S} p_j \sum_{i \notin S} P_{ji} = \sum_{i \notin S} p_i \sum_{j \epsilon S} P_{ij} \tag{3A.3}$$

which means that *the probability of a transition out of the set of states S equals the probability of a transition into S.*

An intuitive explanation of these equations is based on the fact that when the Markov chain is irreducible, the state (with probability one) will return to the set S infinitely many times. Therefore, for each transition out of S there must be (with probability one) a reverse transition into S at some later time. As a result, the proportion of transitions out of S (over all transitions) equals the proportion of transitions into S. This is precisely the meaning of the global balance equations (3A.3).

3A.2 Detailed Balance Equations

As an application of the global balance equations, consider a Markov chain typical of queueing systems and, more generally, birth-death systems where two successive states can only differ by unity as in Fig. 3A.1. We assume that $P_{i,i+1} > 0$ and $P_{i+1,i} > 0$ for all i. This is a necessary and sufficient condition for the chain to be irreducible. Consider the sets of states

$$S = \{0, 1, \ldots, n\}$$

Application of Eq. (3A.3) yields

$$p_n P_{n,n+1} = p_{n+1} P_{n+1,n}\,, \quad n = 0, 1, \ldots \tag{3A.4}$$

i.e., in steady state, the probability of a transition from n to $n+1$ equals the probability of a transition from $n+1$ to n. These equations can be very useful in computing the stationary distribution $\{p_j | j \geq 0\}$ (see sections 3.3 and 3.4).

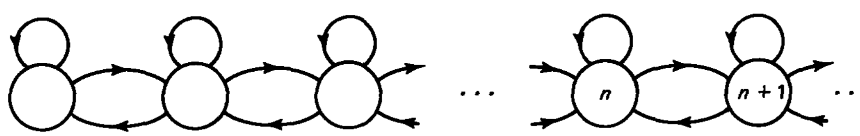

Figure 3A.1 Transition probability diagram for a birth-death process.

Equation (3A.4) is a special case of the equations

$$p_j P_{ji} = p_i P_{ij}\,, \quad i, j \geq 0 \tag{3A.5}$$

known as the ***detailed balance equations.*** These equations need not hold in any given Markov chain. However, in many important special cases, they do hold and greatly simplify the calculation of the stationary distribution. A common method of verifying the validity of the detailed balance equations for a given irreducible, aperiodic Markov chain is to hypothesize their validity and try to solve them for the steady-state probabilities p_j, $j \geq 0$. There are two possibilities; either the system (3A.5) together with $\sum_j p_j = 1$ is inconsistent or else a distribution $\{p_j | j \geq 0\}$ satisfying Eq. (3A.5) will be found. In the latter case, this distribution will clearly also satisfy the global balance equations (3A.2). These equations are equivalent to the condition

$$p_j = \sum_{i=0}^{\infty} p_i P_{ij}, \quad j = 0, 1, \ldots$$

so, by the theorem given earlier, $\{p_j | j \geq 0\}$ is the unique stationary distribution.

3A.3 Partial Balance Equations

Some Markov chains have the property that their stationary distribution $\{p_j | j \geq 0\}$ satisfies a set of equations which is intermediate between the global and the detailed balance equations. For every node j, consider a partition $S_j^1, \ldots, S_j^k$ of the complementary set of nodes $\{i | i \geq 0, i \neq j\}$ and the equations

$$p_j \sum_{i \epsilon S_j^m} P_{ji} = \sum_{i \epsilon S_j^m} p_i P_{ij}, \quad m = 1, 2, \ldots, k \tag{3A.6}$$

Equations of the form above are known as a set of *partial balance equations.* If a distribution $\{p_j | j \geq 0\}$ solves a set of partial balance equations, then it will also solve the global balance equations so it will be the unique stationary distribution of the chain. A technique that often proves useful is to guess the right set of partial balance equations satisfied by the stationary distribution and then proceed to solve them.

3A.4 Continuous-Time Markov Chains

A continuous-time Markov chain is a process $\{X(t) | t \geq 0\}$ taking values from the set of states $i = 0, 1, \ldots$ that has the property that each time it enters state i:

1. The time it spends in state i is exponentially distributed with parameter ν_i. We may view ν_i as the average rate (in transitions/sec) at which the process makes a transition when at state i.
2. When the process leaves state i, it will enter state j with probability P_{ij}, where $\sum_j P_{ij} = 1$.

   We will be interested in chains for which:

   1. The number of transitions in any finite length of time is finite with probability one (such chains are called *regular*).
   2. The discrete-time Markov chain with transition probabilities P_{ij} (called the *imbedded chain*) is irreducible.

Under the preceding conditions, it can be shown that the limit

$$p_j = \lim_{t \to \infty} P\{X(t) = j | X(0) = i\} \tag{3A.7}$$

exists and is independent of the initial state i. Furthermore if the imbedded chain has a stationary distribution $\{\pi_j | j \geq 0\}$, the steady-state probabilities p_j of the continuous chain are all positive and satisfy

$$p_j = \frac{\pi_j/\nu_j}{\sum_{i=0}^{\infty} \pi_i/\nu_i}, \quad j = 0, 1, \ldots \tag{3A.8}$$

The interpretation here is that π_j represents the proportion of visits to state j, while p_j represents the proportion of time spent in state j in a typical system run.

For every i and j, denote

$$q_{ij} = \nu_i P_{ij} \tag{3A.9}$$

Since ν_i is the rate at which the process leaves i and P_{ij} is the probability that it then goes to j, it follows that q_{ij} is the rate at which the process makes a transition to j when at state i. Consequently, q_{ij} is called the *transition rate* from i to j.

Since we will often analyze continuous-time Markov chains in terms of their time-discretized versions, we describe the general method for doing this.

Consider any $\delta > 0$, and the discrete-time Markov chain $\{X_n | n \geq 0\}$, where

$$X_n = X(n\delta), \qquad n = 0, 1, \ldots$$

The stationary distribution of $\{X_n\}$ is clearly $\{p_j | j \geq 0\}$, the stationary distribution of the continuous chain (cf. Eq. (3A.7)). The transition probabilities of $\{X_n | n \geq 0\}$ are

$$\overline{P_{ij}} = \delta q_{ij} + o(\delta), \quad i \neq j$$
$$\overline{P_{ii}} = 1 - \delta \sum_{j \neq i} q_{ij} + o(\delta)$$

Using these expressions in the global balance equations for the discrete chain (cf. Eq. (3A.2)) and taking the limit as $\delta \to 0$, we obtain

$$p_j \sum_{\substack{i=0 \\ i \neq j}}^{\infty} q_{ji} = \sum_{\substack{i=0 \\ i \neq j}}^{\infty} p_i q_{ij}, \quad j = 0, 1, \ldots \tag{3A.10}$$

These are the global balance equations for the continuous chain. Similarly, the detailed balance equations take the form

$$p_j q_{ji} = p_i q_{ij}, \quad i, j = 0, 1, \ldots, \tag{3A.11}$$

One can also write a set of partial balance equations and attempt to solve them for the distribution $\{p_j | j \geq 0\}$. If a solution is found, it provides the stationary distribution of the continuous chain.