〈Research Paper〉

A Matrix Method for the Analysis of Two -Dimensional Markovian Queues*

KIM SUNG SHICK

Abstract

This paper offers an alternative to the common probability generating function approach to the solution of steady state equations when a Markovian queue has a multivariate state space. Identifying states and substates and grouping them into vectors appropriately, we formulate a two-dimensional Markovian queue as a Markov chain. Solving the resulting matrix equations the transition point steady state probabilities (SSPs) are obtained. These are then converted into arbitrary time SSPs. The procedure uses only probabilistic arguments and thus avoids a large and cumbersome state space which often poses difficulties in the solution of steady state equations. For the purpose of numerical illustration of the approach we solve a Markovian queue with one server and two classes of customers.

I. INTRODUCTION

In the analysis of Markovian queues, if the states of the system are represented by a single variable, steady state probabilities (SSPs) are obtained usually directly from steady state equations without much difficulty. But for multivariate state queues this is possible only in a few simple cases. In most multivariate queues steady state equations are too complex to be handled without advanced mathematical techniques. The customery approach is to introduce probability generating functions (PGF's) and to derive a single function involving a smaller set of unknowns or to reduce the original set of equations into a smaller set of equations in PGF's. In the former case solutions are derived by appealing to the analytical properties of the function and extending solutions to the entire set of unknowns. In the latter case the PGF equations are solved using matrix methods as well as analytical properties of PGF's. Avi-Itzhak [1], Bhat and Fischer [2] and Taylor and Templeton [13] are good examples of the latter approach. However, there are models for which these methods are not applicable. In such cases either formulating the PGF equations is not practically possible, or even with the successfully formulated equations, the approach leads to highly unstable numerical computations.

In this paper, for a two-dimensional Markovian queue, we first reduce the number of steady state equations into a smaller set and apply a probabilistic matrix method for its solution. General properties of such

* This work is the first phase of the research program which is sopported by Ministry of Education.

matrix solutions in block - partitioned stochastic matrices have been given by Neuts [8] (Also see [5], [6], [7], [9] for their related papers) and his results form the theoretical basis for the method developed in this paper.

The formulation of the problem discussed in this paper is quite general and the method as developed here has the following advantages over those mentioned earlier

1) Instead of dealing with individual state - substate combinations, the number of variables is reduced by grouping all substates that correspond to a state or by considering the states that have a common set of substates as a group.
2) The transition probabilities between and within the groups of states are represented by matrices and submatrices. To obtain SSPs of the system we solve matrix equations which are much simpler to handle than complex equations involving PGF's.
3) The solution technique employs well - known probabilistic arguments rather than generating functions and roots of functional equations.

In queueing systems, some of the factors that determine the state space are the number of customers in the system, customer types and the number of servers. For this study we shall assume that the state space consists of states and substates where states correspond to the number of customers in the system and the substates are associated with factors that depend on the specific states. Thus if we let J_m^n be the m^{th} substate when the system has n customers, the state space of the Markovian queue may be written as $\{(n, J_m^n), 0 \leq n \leq N, 1 \leq m \leq k(n)\}$, where $N(>1)$ is the system capacity and $k(n)$ is the number of all possible substates when there are n customers in the system.

The Markovian queues studied in this paper have countably infinite statespace (i. e. $N=\infty$) with changes of state in unit steps and the following general characteristics.

1) $k(n)$ is finite for all n and there exists an integer U such that, for all $n=U, U+1, \cdots\cdots$, $J_m^n = J_m^U$ and $k(n)=k(U)$, i. e., if n is greater than U the substates associated with n are identical to the case of U.
2) Let $P^+_{mm'}(n)$ and $P^-_{mm''}(n)$ be the one step transition probabilities from a state (n, J_m^n) to $(n+1, J_{m'}^{n+1})$ and $(n-1, J_{m''}^{n-1})$, respectively ; then $P^+_{mm'}(n)=P^+_{mm'}(U)$ and $P^-_{mm''}(n)=P^-_{mm''}(U+1)$ for all $n= U, U+1, \cdots\cdots$.
3) No substates are associated with n if $n=0$. Thus the state of an empty system is written as 0.

II. Transition Probability Matrix and Equilibrium Condition

In the queueing systems under consideration. let X_i and Y_i denote the number of customers in the system and the substate formed by the customers, respectively, after the i^{th} transition (service completion or new arrival). Clearly the process (X_i, Y_i) is a Markov chain with state space $\{(n, J_m^n), n \geq 0, 1 \leq m \leq k(n)\}$. The transition probability matrix T of the Markov chain is of the general form

$$T = \begin{bmatrix} 0 & A_0 & & & & & \\ B_1 & 0 & A_1 & & & & \\ \cdot & \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & & & & \\ & B_{U-1} & 0 & A_{U-1} & & & \\ & & B_U & 0 & W & & \\ & & & M & 0 & W & \\ & & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \end{bmatrix} \qquad (1)$$

where $A_n : k(n) \times k(n+1)$ matrix

B_n : $k(n)\times k(n-1)$ matrix, and M and W are $k(U)\times k(U)$ matrices.

It is easily seen that, by the nature of the system, T is an irreducible aperiodic stochastic matrix. We denote the SSP of the state (n, J_m^n) as $\pi(n, J_m^n)$ and $\pi(n)=\sum \pi(n, J_m^n)$. Also the vector $\Pi(n)$ is defined as $\Pi(n)=[\pi(n, J_1^n), \cdots\cdots, \pi(n, J_k^n(n))]$

Consider the existence of an invariant probability vector Π satisfying

$$\Pi \geq 0$$

$$\Pi T=\Pi \text{ and } \Pi e=1 \qquad (2)$$

where e is the transpose of the vector $(1, 1, \cdots\cdots, 1)$. The second equation in (2) may be written as

$$\Pi(0)=\Pi(1)B_1$$

$$\Pi(n)=\begin{cases}\Pi(n-1)A_{n-1}+\Pi(n+1)B_{n+1} & 1\leq n\leq U\\ \Pi(n-1)A_{n-1}+\Pi(n+1)M, & n=U\\ \Pi(n-1)W+\Pi(n+1)M, & n>U\end{cases} \qquad (3)$$

$$\text{and } \sum_{n=0}^{\infty}\Pi(n)e=1.$$

Let the matrix $D=M+W$; then D is an irreducible stochastic matirx and there exists a non-negative $k(U)\times k(U)$ matrix R, which satisfies

$$R=W+R^2M,$$

$$sp(R)\leq 1 \text{ and } dR\leq d, \qquad (4)$$

where $sp(R)$ denotes the spectral radius of R and d is the invariant probability vector of the matrix D. Furthermore, since every row of the matrix M should have at least one positive element the matrix R is irreducible. (For further details see Neuts [8]). For $n>U$, we try a solution of the form $\pi(n)=\pi(U)X^{n-U}$. Substituting this in the equation for $n>U$ in (3) and removing common factors yield

$$X^2M+W=X \qquad (5)$$

showing that X satisfies (4).

Thus for $n>U$, R satisfies

$$\Pi(n)=\Pi(U)R^{n-U}$$

The last equation in (3) now may be rewritten as

$$\sum_{n=0}^{U-1}\Pi(n)e+\sum_{n=U}^{\infty}\Pi(U)R^{n-U}e=1 \qquad (6)$$

For the existence of $\Pi(n)>0$, $n=0, 1, 2 \cdots$ which satisfy equation(6), the second term in the LHS of the equation must converge. It can be easily seen that the convergency is assured if $sp(R)<1$.

Theorem 1. A Markovian queueing system with transition probability matrix (1) reaches equilibrium if and only if

$$dMe > dWe$$

Proof. Let $\beta=2Me$, then by Theorem 2 of [8], the irreducibility of the matrix D makes $R>0$ and $sp(R)<1$ if and only if $dB>1$. Also Theorem 3 in [8] assures that all $\pi(n, J_m^n)$ are strictly positive. Hence the equilibrium condition can be written as $2dMe>1$. Since this condition is equivalent to $2dMe>dDe=1$, we can write

$$dMe>dDe-dMe=dWe.$$

We can consider this theorem as the equilibrium condition for the discrete two dimensional random walk where one of coordinates has two reflecting barriers while the other has one. The random walk is in equilibrium if and only if the total probability of upward jump (dWe) is smaller than that of downward jump (dWe)

At this point we define the traffic intensity of the system as the following :

Definition 1. The traffic intensity p of the Markovian queue characterized by the transition probability matrix (1) is defined as

$$p=\frac{dWe}{dMe}$$

Clearly, $p<1$ if and only if $dMe>dWe$.

Finally, since $sp(R)<1$, the matrix $(I-R)^{-1}$ exists. Also by the irreducibility of R, the matrix $(I-R)^{-1}$ is strictly positive. Hence we can write

$$\sum_{n=U}^{\infty}\Pi(n)=\Pi(U)(I-R)^{-1} \qquad (7)$$

Ⅲ. Steady-State Probabilities

If we substitute $\Pi(U+i)$ by $\Pi(U)R^i$, equation (3) becomes

$$\Pi(0) = \Pi(1) B_1$$

$$\Pi(n) = \Pi(n-1) A_{n-1} + \Pi(n+1) B_{n+1} \quad n<U$$

$$\Pi(U) = \Pi(U-1) A_{U-1} + \Pi(U) RM, \quad n=U \qquad (8)$$

$$\Pi(n) = \Pi(U) R^{n-U}, \quad n \geq U+1$$

Writing the first U+1 equations of (8) in matrix form, we have

$$[\Pi(0), \Pi(1), \cdots, \Pi(U)] \begin{bmatrix} 0 & A_0 & & & \\ B_1 & 0 & A_1 & & \\ \cdot & & \cdot & \cdot & \\ & \cdot & & \cdot & A_{U-1} \\ & & B_{U-1} & 0 & RM \\ & & & B_U & \end{bmatrix} = [\Pi(0), \Pi(1), \cdots, \Pi(U)] \qquad (9)$$

In the above system of equations, since all states communicate with each other, the matrix is irreducible and $A_n e + B_n e = e$, $n=0, \cdots, U-1$. Thus if $(B_U + RM)e = e$ then all row sums of the matrix are is and therefore the matrix is an irreducible stochastic matrix. To see $(B_U + RM)e = e$, we first multiply both sides of (4) by e to get

$$Re = We + R^2 Me.$$

Substituting We by e−Me, we have

$$Re = (e - Me) + R^2 Me,$$

and this leads to

$$RMe = e - Me$$

thus finally we get

$$B_U e + RMe = B_U e + e - Me$$
$$= B_U e + We = e$$

Hence the matrix in (9) is indeed an irreducible stochastic matrix. Since $\sum_{n=U}^{\infty} \Pi(U)(I-R)^{-1}$, when we apply the normalizing condition

$$\sum_{n=0}^{U-1} \pi(n)e + \Pi(U)(I-R)^{-1} e = 1 \qquad (10)$$

to equation (9), the resulting solution $\Pi(0)$, $\Pi(1), \cdots, \Pi(n)$ is unique and give precisely the first U+1 SSP vectors of the Markov chain (X_t, X_t).

These are the state probabilities at transition epochs. However, in many Markovian queues, the mean sojourn time at each state can be different from others. In general, for a given state (n, J_m^n), the arbitrary time SSP $q(n, J_m^n)$ is not identical to $\pi(n, J_m^n)$. Since $q(n, J_m^n)$'s give more general information, we give below the complete procedure of obtaining these probabilities rather than the π's. We do this by first establishing relationships between $\Pi(n)$'s and the arbitrary time SSP vector $Q(n) = q(n, 1), \cdots q(n, J_k^n(n))$.

Since each row of the matrices M and W has at least one positive element and M+W is the stochastic matrix D, all row sums of the matrix M are positive and strictly less than 1. This fact gives $sp(M) < 1$ and hence assures the existance of $(I-RM)^{-1}$ under equilibrium condition, $sp(R) < 1$. Hence we are able to write

$$\Pi(U) = \Pi(U-1) A_{U-1} + \Pi(U) RM$$

i. e., $$\Pi(U) = \Pi(U-1) A_{U-1} (I-RM)^{-1} \qquad (11)$$

Also since R and M are irreducible, $(I-RM)^{-1}$ is positive. Now assuming that the spectral radius of the square matrix

$$A_{U-1}(I-RM)^{-1} B_U < 1,$$

we may write

$$\Pi(U-1) = \Pi(U-2) A_{U-2} [I - A_{U-1}(I-RM)^{-1} B_U]^{-1}.$$

Hence, in the reverse order, the first U probability vectors may be written as

$$\Pi(U) = \Pi(U-1) V_U$$

$$\Pi(n) = \Pi(n-1) V_n, \quad 1 \leq n \leq U-1 \qquad (12)$$

where $V_U = A_{U-1}(I-RM)^{-1}$ and $V_n = A_{n-1}(I - V_{n+1} B_{n+1})^{-1}$ $n = 1, 2, \cdots U-1$. This relationship is valid only under the assumption that the spectral radii of the matrices $V_{n+1} B_{n+1}$ $(n=0, 1, \cdots$

U-2) are all less than one and $(I-V_{n+1}B_{n+1})^{-1}$'s are positive. This assumption, however, does not hold for general matrix problems. For example, if $sp(I-RM)^{-1} \gg 1$ and the row sum of the matrices A_{U-1} and B_U are close to 1, then the assumptions are violated. However, in our problem, this is not the case. In fact, for the existence of SSPs the assumptions must hold. Otherwise we may get negative values as SSPs. Now let $V(1)=V_1$ and $V(n)=V_1V_2\cdots V_n$, then from (12) we can represent all the SSP vectors in terms of $\Pi(0)$ as

$$\Pi(n)=\Pi(0)\ V(n) \quad ,1 \leq n \leq U$$
$$\Pi(n)=\Pi(0)\ V(n)R^{n-U} \quad , n \geq U.$$

Since V_1 is lxk(1) matrix and V_i is a k(i-1)xk(i) (i=1, ……, U) matrix, V(n) is a lxk(n) vector

Now we are in a position to calculate the arbitrary time point SSP vectors. Define Y_n (n=0,1, ……, U) as the k(n)xk(n) diagonal matrix whose entry (j, j) is the mean sojourn time at state (n, J_j^n), j=1, 2 … k(n). It is easily seen that, if we let λ be the customers' overall arrival rate, Y_0 is the single value $1/\lambda$. From Theorem 5.16 of page 104 of Ross [11], the relationship between $Q(i)$ and $\Pi(i)$ is given by

$$Q(i)=\frac{\Pi(i)\ Y_i}{\sum_{n=0}^{\infty} \Pi(n)\ Y_n e}, \quad i=0,1,2,\cdots \quad (14)$$

Now we let $a=[\sum_{n=0}^{\infty} \Pi(n)\ Y_n e]^{-1}$ and multiply both sides of equation (13) by a and Y_n. This yields

$$a\Pi(n)\ Y_n = a\Pi(0)\ V(n)\ Y_n \quad ,n < U$$
$$a\Pi(n)\ Y_U = a\Pi(0)\ V(U)R^{n-U}\ Y_U \quad , n \geq U \quad (15)$$

The second equation is possible sine $Y_n=Y_u$ for n=U+1, ……. Noting that $Y_0=\lambda$ and substituting for $\Pi(n)$, n=0, 1, 2, … from (14) we get

$$Q(n)=\lambda Q(0)\ V(n)\ Y_n \quad , 1 \leq n < U$$
$$Q(n)=\lambda Q(0)\ V(U)\ R^{n-U} Y_U$$

To find Q(0) we need the nanalizin: co id-tion

$$\sum_{n=0}^{\infty} Q(n)e=1$$

Since the condition is equivalent to

$$Q(0)\ [1+\lambda \sum_{n=0}^{U-1} V(n)\ Y_n e+\lambda \sum_{n=U}^{\infty} V(U)R^{n-U}\ Y_U e]=1$$

we get

$$Q(0)=[1+\lambda \sum_{n=0}^{\infty} V(n)\ Y_n e+\lambda V(U)\ (I-R)^{-1}\ Y_U e]^{-1}$$

Other probability vectors are obained using equation (16). We note that the probability q(n) of n customers in the system, is simply Q(n)e for n=1, 2, …

In the course of the numerical implementation of the procedure developed above to find V_n (n=1, 2, …U) we need inverses of (I-RM) and $(I-V_{n+1}B_{n+1})$ n=0, 1, 2, … U-1. Since the inverse of each of these matrices can be represented as an infinite series of real positive matrices we do not have to deal with complex values in numerical computations. Thus this approach is simple and requires only straightforward numerical calculations which can be easily implemented on a computer.

Ⅳ. Example

For the purpose of illustration of the procedure developed we adopt the M/M/1 system with two types of customers as an example. Readers interested in the application of the method to more complex systems may refer to [3, 4]. Even though the system is relatively simple, if we use the usual method of probability generating functions as in [2] and [12], the system provides SSPs only after complicated mathematical maneuvers.

Consider the M/M/1 system where two types of customers (type 1, 2) arrive in Poisson processes with mean rates λ_1 (type1) and λ_2 (type 2). Let X_i be the probability that an incomming customer is type i ; then the overall arrival rate λ is $\sum \lambda_i X_i$. Also we assume the service times to have exponential distributions with mean rates $\frac{1}{\mu_1}$ and $\frac{1}{\mu_2}$ for customer types 1 and 2 respectively We difine the states of the system as (n, m), where n is the number of customers in the system and m, the type of customer in service. Thus the state space is {(n,m), n=0, 1, 2,…; m=1,2}. When the states of the system are defined this way, for each $n \geqq 1$, the number of substates, k (n) is 2. Moreover, the transition probabilities $P^+_{mm}(1)$ and $P^-_{mm}(2)$ defined earlier are given by

$$P^+_{mm}(1) = \{\lambda/(\lambda_0+\mu_m) \quad , m=m' \quad , \text{otherwise},$$

$$P^-_{mm}(2) = (\mu_m/(\lambda+\mu_m)) \cdot X_m$$

and are the same for n=2, 3, …….
Thus we have U=1,

$$W= \begin{bmatrix} \lambda/(\lambda+\mu_1) & 0 \\ 0 & \lambda(\lambda+\mu_2) \end{bmatrix}$$

and

$$M= \begin{bmatrix} X_1\mu_1/(\lambda+\mu_1) & X_2\mu_1/(\lambda+\mu_1) \\ X_1\mu_2/(\lambda+\mu_2) & X_2\mu_2/(\lambda+\mu_2) \end{bmatrix}$$

Also, with little effort, we find

$$A_0 = [X_1, X_2]$$

$$B_1 = [\mu_1/(\lambda+\mu_1), \ \mu_2/(\lambda+\mu_2)]$$

At this stage, to proceed further, we need numerical values for the parameters. Let $\lambda_1=1$, $\lambda_2=1$, $X_1=0.3$, $X_2=0.7$, $\mu_1=1.5$, and $\mu_2=2$. Then $\lambda=1$,

$$M= \begin{bmatrix} 0.9/5 & 2.1/5 \\ 0.6/3 & 1.4/3 \end{bmatrix}$$

$$W= \begin{bmatrix} 1/2.5 & 0 \\ 0 & 1/3 \end{bmatrix}$$

Thus

$$D=M+W= \begin{bmatrix} 2.9/5 & 2.1/5 \\ 0.6/3 & 2.4/3 \end{bmatrix}$$

and the invariant probability vector d is found as d=[0.323, 0.677].
Therefore, dMe = 0.8257, dWe = .3397 and the traffic intensity p = 0.4114, thus the system is in equilibrium. To compute the matrix R we use the successive substitutions discussed by Neuts [8], which is given by the sequence

$$R_0=W, \ R_{n+1}=W+R_n^2 M, \text{ for } n \geqq 0.$$

This is a highly efficient method. Other methods are also found in [6]. The sequence $\{R_n\}$ is monotone increasing and converges to R. With the method we find

$$R = \begin{bmatrix} .580 & .077 \\ .125 & .387 \end{bmatrix}$$

Consequently

$$(I-R)^{-1} = \begin{bmatrix} 2.413 & .3107 \\ 0.504 & 1.6947 \end{bmatrix}$$

and

$$(I-RM)^{-1} = \begin{bmatrix} 1.185 & .4319 \\ .154 & 1.360 \end{bmatrix}$$

Thus we have

$$V_1 = A_0(I-RM)^{-1} = [0.4633, \ 1.0816] = V$$

Utilizing this and the mean sojourn time matrices

$$Y_0 = 1$$

$$Y_1 = \begin{bmatrix} 1/2.5 & 0 \\ 0 & 1/3 \end{bmatrix}$$

we get

$$Q(0) = [1+\lambda V(1) \ (I-R)^{-1} = 0.4302,$$

$$Q(1) = [0.797, \ .155]$$

$$Q(n) = [0.797, .155] \begin{bmatrix} .580 & .077 \\ .125 & .387 \end{bmatrix}^{n-1}$$

$$\begin{bmatrix} 1/2.5 & 0 \\ 0 & 1/3 \end{bmatrix}, n = 2, 5, \cdots\cdots.$$

Furthermore if we let

$$[y_1, y_2] = \sum_{n=1}^{\infty} Q(n)\ I$$

then y_1 and y_2 are the unconditional probabilities that the server is occupied by the type 1 and the type 2 customers, respectively. We obtain that

$$[y_1, y_2] = \sum_{n=1}^{\infty} Q(n)\ I$$
$$= \lambda Q(0)\ V(1)\ (I-R)^{-1} Y_2$$
$$= (.290,\ 0.283).$$

Hence the conditional probabilities that the server is busy by the type 1 and the type 2 customers when the system is busy are 0.506 and 0.494, respectively.

References

1. Avi-Itzhak, B. and Mitrang, I. L., "A Many-Server Queue with Service Interruption," Opns. Res. 16, 628-638, 1968.
2. Bhat, U. N. and Fisher, M. J., "Multichannel Queueing Systems with Heterogeneous Class of Arrivals," Naval Res. Logist. Quart., 23, 271-282, 1976.
3. Kim, S. S., "M/M/S Queueing System where Customers Demand Multiple Server Use," Ph. D. Thesis, Southern Methodist University, 1979.
4. Kim, S. S., "Multiserver Markovian Queue Where Customers Demand Multiple Server Use," forthcoming.
5. Neuts, M. F., "The Markov Renewal Branching Process," Proc. Conf. on Math. Methods in the Theory of Queues, Kalamazoo, MI, Springer Verlag, NY., 1-21, 1974.
6. Neuts, M. F., "Moment Formulas for the Markov Renewal Branching Process," Adv. Appl. Prob. 8, 690-711, 1976.
7. Neuts, M. F., "The M/G/1 Queue with Several Types of Customers and Change-Over Times," Adv. Appl. Prob., 9, 604-644, 1977.
8. Neuts, M. F., "Markov Chainswith Application in Queueing Theory, which have a Matrix-Geometric Invariant Probability Vector," Adv. Appl. Prob., 10, 185-121, 1978.
9. Neuts, M. F., "The Probilistic Significance of the Rate Matrix in Matrix-Geometric Invariant Vectors," J. Appl. Prob., 17, 291-296, 1980.
10. Purdue, P., "A Queue with Poisson Input and Semi-Markov Service Times : Busy Period Analysis," J. Appl. Prob., 12, 353-357, 1975.
11. Ross, S. M., Applied Probability Models with Optimization Applications, Holden Day, 1970.
12. Slater, N. B. & Kotiah, T. C. T., "The Steady State of a Multi-Server Mixed Queue," Adv. Appl. Prob., 5, 614-631, 1973.
13. Taylor, I. D. S. and Templeton, J. G. C., "Waiting Time in a Multi-Server Cutoff-Priority Queue, and its Application to an Urban Ambulance Service," Opns. Res. 28, 1168-1188, 1980.