

Computational Discrete Time Markov Chain with Correlated Transition Probabilities

Peerayuth Charnsethikul

Operations Research and Management Science Units, Industrial Engineering Department
Kasetsart University, Bangkok 10903, Thailand

Abstract: This study presents a computational procedure for analyzing statistics of steady state probabilities in a discrete time Markov chain with correlations among their transition probabilities. The proposed model simply uses the first order Taylor's series expansion and statistical expected value properties to obtain the resulting linear matrix equations system. Computationally, the bottleneck is $O(n^4)$ but can be improved by distributed and parallel processing. A preliminary computational experience is reported.

Key words: Markov chain, steady state analysis, correlated transition probabilities, computational methods.

INTRODUCTION

Markov chain has been around for almost one and a half century. A.A. Markov initiation (from the appendix of the book by Howard^[1]) has been continuously received attentions and extended theoretically and practically by various branches of science and technology. Basically, the discrete time Markov chain is probably the most popular in term of field foundation and linkage. Mathematically, the problem statement is to determine the vector π such that

$$\pi^T = \pi^T P \quad (1)$$

$$e^T \pi = 1, e^T = [1 \ 1 \ 1 \ \dots]_{n \times 1} \quad (2)$$

where $\pi = [\pi_i]$ is an $n \times 1$ vector representing steady state probability of each system state i , $i=1,2,\dots,n$ and $P = [p_{ij}]$ is an $n \times n$ matrix representing the transition probabilities among states with the property that $\sum_{j=1}^n p_{ij} = 1, \forall i$.

To solve the equations system (1) and (2), a given deterministic P must be provided. Practically, each element in P can be determined by either the applied problem assumptions or statistical approximations. In both cases, some parameters must be assumed or estimated from a finite sample and most of the time, these parameters are not deterministic and independent. For an example in marketing, the transition probability that customers of product brand i become customers of product brand j , estimated from a market survey report may consist of its mean, variance and co-variances with other transitions. Therefore, the decision problem becomes determining an uncertain π according to an uncertain P .

This study was aimed to propose the use of first order Taylor's series approximations to obtain an

approximated corresponding linear equations system for determining means, variances and co-variances statistics of π and analyze its computational complexity with a preliminary computation experiments. In the following sections, related literatures will be reviewed and then, the details of model development with related computational procedures will be proposed, tested and concluded.

Literature reviews: Solving equations (1) and (2) based on the assumptions of irreducible Markov chain has been the main focus for many researchers in various fields. Trends and directions have changed according to available computing technology. An introduction level can be found in Kao^[2] and Kulkarni^[3]. Another original state-space approach founded by Howard^[1] can be considered as a classic. In terms of numerical stability, the approach of Sheskin^[4] and Grassman *et al.*^[5] are basic foundations for numerical solutions with further extensions on other special structural transition matrix types by Grassman and Heyman^[6] and on their passage time statistics by Grassman^[7]. A more complete detail on state of the art of Markov chain computations was investigated and systematically organized by Stewart^[8]. With current technology, MATLAB programming language (2000) is an appropriate alternative for solving problems with small and intermediate state sizes (see some available implemented programs in Kao^[2]). However, the number of states in real life models is usually very large (much more than a million) and the transition probability matrix can be either sparse or dense. This leads to the use of parallel processing to increase computing efficiency. Some of these works in recent years are Knottenbelt and Harrison^[9], Benzi and Tuma^[10] and Dingle *et al.*^[11].

Statistical linear modeling with multiple random elements by matrix approaches is summarized and

illustrated in Moser^[12]. Moreover, a good introduction on the use of expected value on linear transformation can be found in Strark and Woods^[13]. Charnsethikul^[14] used such foundations to develop approaches for solving n by n linear equations with correlations among right hand side values and equation coefficients. One of the aims in this was to simplify the approach and apply to the case of discrete time Markov chain analysis for obtaining an approximated linear model with variables concerning with means and variances of steady state probabilities and co-variances among transition and steady state probabilities. In the next section, the resulting model will be proposed and directed to the involved computing schemes for further algorithmic analysis and developments.

Model development and a computational procedure:

A statistical model for determining statistics involving with π and P from equations (1) and (2) is developed based on the given input of $E[P] = [E[p_{ij}]]$ as the expected value of the matrix P, $K[p_{kl}, P]$ as the co-variance matrix between p_{kl} and all other elements in P and $K[p_k p_l^T]$ as the co-variance matrix between elements in k^{th} and l^{th} column of P for $k, l = 1, 2, \dots, n$.

First, expand the right hand side of equation (1) using the truncated first order Taylor series approximations around the mean values and take the expected value on both sides as shown below.

$$\begin{aligned} \pi^T &\approx E[\pi^T]E[P] + (\pi^T - E[\pi^T])E[P] + E[\pi^T](P - E[P]) \\ E[\pi^T] &\approx E[\pi^T]E[P] \end{aligned} \tag{3}$$

The higher order term truncation is due to the fact that both $E[p_{ij}]$ and $E[\pi^T]$ values are less than one and each uncertain value is close to its expected value. Now, take the expected value on both sides of equation (2) and obtain the following.

$$e^T E[\pi] = 1 \tag{4}$$

Next, multiply both sides of equation (1) by p_{kl} , $k, l = 1, 2, \dots, n$ and take the expected value resulting the equations system as follows.

$$\begin{aligned} p_{kl} \pi^T &= p_{kl} \pi^T P \\ K[p_{kl}, \pi^T] &= K[p_{kl}, \pi^T]E[P] + E[\pi^T] K[p_{kl}, P] \\ K[p_{kl}, \pi^T][I - E[P]] &= E[\pi^T] K[p_{kl}, P] \end{aligned} \tag{5}$$

For equation (2), multiply both sides by p_{kl} and take the expected value lead to a redundant relation. Therefore, at this moment, $E[\pi]$ can be solved using equations (3) and (4). The right hand side of equation (5) can be obtained for each k and l and $K[p_{kl}, \pi^T]$ as the co-variances between p_{kl} and all elements in π , can be computed by solving equation (5) n^2 times according to all possible choice of k and l. Finally, rearrange equation (1) and obtain the multiplicative form as

follows and then, take the expected value getting the following matrix relation.

$$\pi^T = \pi^T P \rightarrow \pi^T [P - I] = 0 \text{ or } [P^T - I] \pi = 0$$

Therefore,

$$[P^T - I] \pi \pi^T [P - I] = 0$$

If $W^T = [P^T - I]$ then $W = [P - I]$. Next, apply the first order approximation to obtain the following matrix equation for solving variances/co-variances among the steady state probabilities as follows.

$$\begin{aligned} E[W^T]K[\pi\pi^T]E[W] + [E[\pi^T]K[w_i w_j^T]E[\pi]] \\ + [E[w_i^T]K[\pi w_j^T]E[\pi]] + [E[\pi^T]K[w_i \pi^T]E[w_j]] = 0 \tag{6} \\ E[W^T]K[\pi\pi^T]E[W] = [q_{ij}] + [r_{ij}] + [s_{ij}] \\ K[\pi\pi^T] = E[W^T]^{-1} [[q_{ij}] + [r_{ij}] + [s_{ij}]] E[W]^{-1} \end{aligned} \tag{7}$$

where

$$\begin{aligned} E[W] &= E[P] - I, \\ K[w_i w_j^T] &= K[p_i p_j^T], \\ K[\pi w_j^T] &= K[\pi p_j^T], \\ q_{ij} &= -E[\pi^T]K[w_i w_j^T]E[\pi], \\ r_{ij} &= -E[w_i^T]K[\pi w_j^T]E[\pi], \\ s_{ij} &= -E[\pi^T]K[w_i \pi^T]E[w_j], \quad i, j = 1, 2, \dots, n, \\ K[w_i w_j^T] \text{ and } K[p_i p_j^T] &\text{ represents the co-variance matrix between elements of columns } i \text{ and } j \text{ of matrices } W \text{ and } P \text{ respectively,} \end{aligned}$$

$E[w_j]$ represents the mean vector of elements of column j of matrix W and

$K[\pi w_j^T]$ and $K[\pi p_j^T]$ represents the co-variance matrix between elements of column j of matrices W and P respectively and elements of vector π .

Therefore, in summary, a numerical algorithm for obtaining statistics of steady state probabilities in case of correlated transition probabilities can be stated step by step as follows.

Step 1: Find $E[\pi^T]$ by solving equations (3) and (4).

Step 2: Substitute $E[\pi^T]$ found from step 1 to equation (5) and solve for $K[p_{kl}, \pi^T]$ for all possible k and l.

Step 3: Substitute $E[\pi^T]$ and $K[p_{kl}, \pi^T]$ for all possible k and l found from steps 1 and 2, respectively to equation (7) and finally, solve for $K[\pi\pi^T]$.

In step 1, the complexity of $O(n^3)$ is required for general dense P since the computation is involved with solving n by n simultaneous linear equations while step 2 requires at most solving n by n equations n^2 times with the complexity of $O(n^4)$ since each equation has the same coefficient matrix $[I - P]$. Therefore, only one n by n system is needed to be solved with $O(n^3)$ while

others require $O(n^2)$. Finally in the last step, solving the corresponding matrix equation requires solving the inverse matrix of $E[W]$ and two matrix multiplications with size n by n , leading to the complexity of $O(n^3) + O(n^3) = O(n^3)$. Therefore, the bottleneck complexity of the whole procedure can be identified at step 2 with the complexity of $O(n^4)$. However, for each k and l in step 2, solving for each $K[p_{kl}, \pi^T]$ is independent and can be computed in parallel for efficiency improvements.

RESULTS AND DISCUSSIONS

The proposed computational procedure in the previous section was coded as a MATLAB program utilizing its numerical linear algebra objects on solving corresponding equations and matrix operations. The input test data, $E[P]$ and $K[p_{kl}, P]$ for all k and l were randomly generated on various n . The corresponding co-variance matrix was generated such that the diagonal dominant property holds and also maintains the semi-positive definite assumption. The observed response is the computational time in each procedure step. The maximum n is up to 200 due to a clear illustrated bottleneck on step 2 of the procedure and excessive computation time limit of 2 hours. All generated problem were tested on a microcomputer with Pentium IV, 2.4 Ghz and 1Gbytes RAM. The results are as shown in Table 1.

Table 1: Average time consumed in case of dense matrices

N	Computational Time (sec.)		
	Step 1	Step 2	Step 3
10	<0.000	0.062	<0.000
25	<0.000	2.516	<0.000
50	<0.000	33.813	0.031
75	<0.000	190.172	0.056
100	<0.000	528.360	0.150
150	0.005	2,161.423	0.562
200	0.022	6,152.513	1.786

From the Table 1, the results are consistent with our initial assumption that the computation bottleneck is in step 2. Moreover, the time response on step 2 can be fitted using direct regression analysis and the appropriate relationship with n also hold with the theoretical results of $O(n^4)$. For a larger scale problem, the computations on step 2 can be independently distributed among parallel processors by partitioning equally the sub-problem set of all possible k and l under the assumption that computations in steps 1, 2 (for each k and l) and 3 can be efficiently solved in a single processor. However, in a much larger scale problem, sparse matrix technology with more complicated parallel designs and programming will be required.

CONCLUSION

This study uses the first order Taylor’s series approximation to create a linear model for determining

statistics of steady state probabilities in a discrete time irreducible Markov Chain. The proposed procedure to solve the corresponding model can be proven theoretically and experimentally within $O(n^4)$. Moreover, distributed computing can be directly applied to the procedure for efficiency improvements. For an extensive study, the effects of neglecting higher order terms in Taylor’s series should be furthered investigated especially in case of non-Gaussian assumptions with higher order correlations among transition probabilities.

REFERENCES

- Howard, R.A., 1971. Dynamic Probabilistic Systems. Vol. I, New York: John Wiley & Sons.
- Kao, E.P.C., 1997. An Introduction to Stochastic Processes, Duxbury Press, ITP.
- Kulkarni, V.G., 1995. Modeling and Analysis of Stochastic Systems, Chapman & Hall.
- Sheskin, T.J., 1985. A Markov chain partitioning algorithm for computing steady state probabilities. Oper. Res., 33: 228-235.
- Grassman, W.K., M.I. Taksar and D.P. Heyman, 1985. Regenerative analysis and steady state distributions for Markov chain. Oper. Res., 33: 1107-1116.
- Grassman, W.K. and D.P. Heyman, 1990. Equilibrium distribution of block-structured Markov chains with repeated rows. J. Appl. Prob., 27: 557-576.
- Grassman, W.K., 1987. Means and variances of time averages in Markovian environments. Eur. J. Oper. Res., 31: 132-139.
- Stewart, W.J., 1994. Introduction to the Numerical Solution of Markov Chains. Princeton University Press.
- Knottenbelt, W.J. and P.G. Harrison, 1999. Distributed disk-based solution techniques for large Markov models. Proc. 3rd Intl. Conf. on the Numerical Solution of Markov Chains, Zaragoza, Spain, pp: 58-75.
- Benzi, M. and M. Tuma, 2002. A parallel solver for large-scale Markov chains. Appl. Numer. Math., 41: 135-153.
- Dingle, N.J, P.G. Harrison and W.J. Knottenbelt, 2004. Uniformization and hypergraph partitioning for the distributed computation of response time densities in very large Markov models. J. Parallel and Distributed Computing, 64: 908-920.
- Moser, B.K., 1996. Linear Models: A Mean Approach. Academic Press.
- Stark, H. and J.W. Woods, 1986. Probability, Random Processes and Estimation Theory for Engineers. Prentice-Hall.
- Charnsethikul, P., 2004. Solving $(n \times n)$ linear equations with parameters covariances. Intl. J. Computational Methods, 1: 471-489.