ANZIAM J. 46(2005), 485–493

RAMASWAMI'S DUALITY AND PROBABILISTIC ALGORITHMS FOR DETERMINING THE RATE MATRIX FOR A STRUCTURED $GI/M/1$ MARKOV CHAIN

EMMA HUNT[1]

(Received 4 November, 2004)

Abstract

We show that Algorithm H* for the determination of the rate matrix of a block-$GI/M/1$ Markov chain is related by duality to Algorithm H for the determination of the fundamental matrix of a block-$M/G/1$ Markov chain. Duality is used to generate some efficient algorithms for finding the rate matrix in a quasi-birth-and-death process.

1. Introduction

In a companion article [8] we constructed a probabilistic algorithm for the determination of the rate matrix R of an irreducible block-$GI/M/1$ Markov chain and showed with benchmark numerical experiments that our procedure, Algorithm H*, compares favourably with existing methods in the literature. Algorithm H* assumes irreducibility of the Markov chain but makes no assumptions about ergodicity.

The notation was chosen to emphasise a duality, indicated by *, with Algorithm H for the determination of the fundamental matrix G for a block-$M/G/1$ Markov chain. The duality was not established in [8]. Algorithm H is derived in [7] and can be shown to reduce to a version of the cyclic reduction algorithm of Bini and Meini for the determination of G when further technical conditions are imposed. For an exposition of the cyclic reduction methodology the reader is referred to [2–4] and [11].

In this paper we take these ideas further. We manifest explicitly the duality between Algorithms H and H* and show how in the QBD case several other probabilistic algorithms can be constructed for the efficient calculation of R. We shall also find relations between Algorithm H*, the logarithmic reduction algorithm of Latouche and

[1]School of Mathematical Sciences, The University of Adelaide, Adelaide SA 5005, Australia; e-mail: emma.hunt@adelaide.edu.au.

© Australian Mathematical Society 2005, Serial-fee code 1446-1811/05

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

Ramaswami [10] and the cyclic reduction algorithms of Bini and Meini. To avoid repetition, we assume familiarity with the ideas and notation of [7] and [8].

2. The duality $\mathscr{A}_j \longleftrightarrow \mathscr{C}_j$

The duality * between processes of structured $GI/M/1$ and $M/G/1$ types was introduced by Ramaswami [13]. An alternative derivation based on time reversal was presented subsequently by Asmussen and Ramaswami [1]. Further developments are given in Bright [5].

The duality is between classes of block-$M/G/1$ chains and classes of block-$GI/M/1$ chains. The anomalous leading block row and column in the one-step transition matrices for these two paradigms do not enter into the duality and it is convenient to omit these and relabel the remaining rows and columns as $0, 1, \ldots$. So we replace the $M/G/1$ chain by a chain $\mathscr{A}_0$ on levels $\ell \geq 0$ given by the structured one-step transition matrix

$$P^{(0)} = \begin{bmatrix} A_1 & A_2 & A_3 & \cdots \\ A_0 & A_1 & A_2 & \cdots \\ 0 & A_0 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

and similarly replace the $GI/M/1$ chain by a chain $\mathscr{C}_0$ on levels $\ell \geq 0$ given by the structured one-step transition matrix

$$P^{*(0)} = \begin{bmatrix} C_1 & C_0 & 0 & 0 & \cdots \\ C_2 & C_1 & C_0 & 0 & \cdots \\ C_3 & C_2 & C_1 & C_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

It should be noted that these matrices are substochastic.

The blocks of both matrices are $k \times k$. The matrix $C = \sum_{j=0}^{\infty} C_j$ is taken as stochastic and irreducible and so has an invariant probability measure c, that is, $cC = c$. The entries of c are all positive since C is irreducible. Set $\Delta = \operatorname{diag}(c)$. Ramaswami's duality is given by $A_m = \Delta^{-1} C_m^T \Delta$ for all $m \geq 0$. Clearly the matrices A_m have nonnegative entries. For the notion of duality to be meaningful in the context of Markov chains, it is helpful for the matrix $A = \sum_{m=0}^{\infty} A_m$ to be stochastic. This is immediate. We have $A = \Delta^{-1} C^T \Delta$ or $C = \Delta^{-1} A^T \Delta$, so that $cC = c$ can be expressed as $c\Delta^{-1} A^T \Delta = c$ or $c\Delta^{-1} A^T = c\Delta^{-1}$. Since $c\Delta^{-1} = e^T$, where e is a suitable vector of units (here of length k), we thus have $Ae = e$ and so A is stochastic.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

The duality $A_n \longleftrightarrow C_n$ induces a correspondence between $P^{(0)}$ and $P^{*(0)}$. Denote by $(P^{(0)})_{n,m}$, $(P^{*(0)})_{n,m}$ respectively the (n, m) block entries in these two matrices $(n, m \geq 0)$. Then

$$\left(P^{(0)}\right)_{n,m} = \Delta^{-1}\left[\left(P^{*(0)}\right)_{m,n}\right]^T \Delta,$$

which we may express as $P^{(0)} \longleftrightarrow P^{*(0)}$ or

$$\mathscr{A}_0 \longleftrightarrow \mathscr{C}_0. \tag{2.1}$$

This provides a basis for an inductive proof of the following theorem, which extends the duality to one between the sequence $(\mathscr{A}_j)_{j\geq 0}$ of censored processes involved in the construction of Algorithm H and the sequence $(\mathscr{C}_j)_{j\geq 0}$ of censored processes used in the construction of Algorithm H*. This entails an extension of the duality to multistep transitions involving taboo levels.

THEOREM 2.1. *When A is irreducible, we have the duality*

$$\mathscr{A}_j \longleftrightarrow \mathscr{C}_j \quad \textit{for each } j \geq 0. \tag{2.2}$$

PROOF. Suppose (2.2) holds for some $j \geq 0$. We have

$$P^{(j)} = \begin{bmatrix} B_1^{(j)} & B_2^{(j)} & B_3^{(j)} & \cdots \\ A_0^{(j)} & A_1^{(j)} & A_2^{(j)} & \cdots \\ 0 & A_0^{(j)} & A_1^{(j)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad P^{*(j)} = \begin{bmatrix} D_1^{(j)} & C_0^{(j)} & 0 & 0 & \cdots \\ D_2^{(j)} & C_1^{(j)} & C_0^{(j)} & 0 & \cdots \\ D_3^{(j)} & C_2^{(j)} & C_1^{(j)} & C_0^{(j)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

so that $B_n^{(j)} = \Delta^{-1}\left(D_n^{(j)}\right)^T \Delta$, $n > 0$, and $A_n^{(j)} = \Delta^{-1}\left(C_n^{(j)}\right)^T \Delta$, $n \geq 0$.
Therefore

$$\begin{aligned} K_0^{(j+1)} &= \sum_{i=0}^{\infty}\left(A_1^{(j)}\right)^i = \sum_{i=0}^{\infty}\left\{\Delta^{-1}\left(C_1^{(j)}\right)^T \Delta\right\}^i = \Delta^{-1}\sum_{i=0}^{\infty}\left\{\left(C_1^{(j)}\right)^i\right\}^T \Delta \\ &= \Delta^{-1}\left(\left[I - C_1^{(j)}\right]^{-1}\right)^T \Delta = \Delta^{-1}\left(K_0^{(j+1)}\right)^T \Delta, \end{aligned}$$

so that $K_0^{(j+1)} \longleftrightarrow K_0^{(j+1)}$. This provides the basis for an (inner) induction that

$$K_m^{(j+1)} \longleftrightarrow K_m^{(j+1)} \quad \text{for } m \geq 0. \tag{2.3}$$

Suppose that (2.3) holds for $m = 0, 1, \ldots, n-1$ for some $n \geq 1$. Then from the equation

$$K_n^{(j+1)} = \sum_{m=0}^{n-1} K_m^{(j+1)} A_{2(n-m)+1}^{(j)} K_0^{(j+1)}$$

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

derived in [7, Section 3], we have

$$K_n^{(j+1)} = \sum_{m=0}^{n-1} \left\{\Delta^{-1} \left(K_m^{(j+1)}\right)^T \Delta\right\} \left\{\Delta^{-1} \left(C_{2(n-m)+1}^{(j)}\right)^T \Delta\right\} \left\{\Delta^{-1} \left(K_0^{(j+1)}\right)^T \Delta\right\}$$

$$= \Delta^{-1} \left(\sum_{m=0}^{n-1} K_0^{(j+1)} C_{2(n-m)+1}^{(j)} K_m^{(j+1)}\right)^T \Delta = \Delta^{-1} \left(K_n^{(j+1)}\right)^T \Delta$$

by [8, (3.7)], which gives the inductive step in the inner induction.

Using this result, we have similarly from the equation

$$L_n^{(j+1)} = \sum_{m=0}^{n} K_m^{(j+1)} A_{2(n-m)}^{(j)}$$

(derived in [7]) and [8, (3.4)] that $L_n^{(j+1)} \longleftrightarrow L_n^{(j+1)}$ for $n \geq 0$.

Finally we have from the relation

$$B_n^{(j+1)} = B_{2n-1}^{(j)} + \sum_{m=1}^{n} B_{2m}^{(j)} L_{n-m}^{(j+1)}$$

derived in [7] for $n \geq 1$ that

$$B_n^{(j+1)} = \Delta^{-1} \left(D_{2n-1}^{(j)}\right)^T \Delta + \sum_{m=1}^{n} \left\{\Delta^{-1} \left(D_{2m}^{(j)}\right)^T \Delta\right\} \left\{\Delta^{-1} \left(L_{n-m}^{(j+1)}\right)^T \Delta\right\}$$

$$= \Delta^{-1} \left[D_{2m}^{(j)} + \sum_{m=1}^{n} L_{n-m}^{(j+1)} D_{2m}^{(j)}\right]^T \Delta = \Delta^{-1} \left[D_n^{(j+1)}\right]^T \Delta$$

by [8, (3.3)], so that $B_n^{(j+1)} \longleftrightarrow D_n^{(j+1)}$ for $n \geq 1$.

A similar argument yields $A_n^{(j+1)} \longleftrightarrow C_n^{(j+1)}$ for $n \geq 0$, completing the outer induction.

COROLLARY 2.2. *It follows from Theorem* 2.1 *that*

$$\left[I - B_1^{(N)}\right]^{-1} A_0 \longleftrightarrow C_0 \left[I - D_1^{(N)}\right]^{-1}.$$

The successive approximants to G, R derived respectively by Algorithms H *and* H* *are of the forms of the left- and right-hand sides. Hence the N-th approximants T_N, T_N^* to G and R given by these algorithms satisfy*

$$T_N \longleftrightarrow T_N^*. \tag{2.4}$$

Relation (2.4) is a path-restricted version of the standard duality result $G \longleftrightarrow R$ between block-$M/G/1$ chains and block-$GI/M/1$ chains. It illustrates that when A is irreducible, duality can be applied directly to Algorithm H for G to produce Algorithm H* for R.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

3. Quasi-birth-and-death-chains

3.1. Algorithm H* In the case of a QBD, substantial simplifications occur in the equations prescribing Algorithm H*, as presented in [8, Section 4]. Since $C_n^{(j)} = 0$ for $n > 2$, we have $K_n^{(j)} = 0$ for $n > 0$ and so

$$L_n^{(j+1)} = \begin{cases} C_0^{(j)} K_0^{(j+1)}, & n = 0; \\ C_2^{(j)} K_0^{(j+1)}, & n = 1; \\ 0, & n > 1. \end{cases}$$

Also $D_n^{(j)} = 0$ for $n > 2$.

The relations linking $\mathscr{C}_{j+1}$ and $\mathscr{C}_j$ are thus

$$\begin{aligned} C_n^{(j+1)} &= C_n^{(j)} K_0^{(j+1)} C_n^{(j)} \quad (n = 0, 2), \\ C_1^{(j+1)} &= C_1^{(j)} + C_2^{(j)} K_0^{(j+1)} C_0^{(j)} + C_0^{(j)} K_0^{(j+1)} C_2^{(j)}, \\ D_1^{(j+1)} &= D_1^{(j)} + C_0^{(j)} K_0^{(j+1)} D_2^{(j)}, \\ D_2^{(j+1)} &= C_2^{(j)} K_0^{(j+1)} D_2^{(j)}. \end{aligned}$$

The initialisation is $C_n^{(0)} = C_n$ $(n = 0, 1, 2)$ and $D_n^{(0)} = C_n$ $(n = 1, 2)$.

3.2. Other QBD methods In the previous section we showed how Ramaswami's duality can be used to link Algorithm H for the determination of the fundamental matrix in a block-$M/G/1$ Markov chain and Algorithm H* for the determination of the rate matrix in a block-$GI/M/1$ Markov chain. Operationally, we could have used duality to induce Algorithm H* from Algorithm H.

We may also dualise the logarithmic reduction technique Algorithm LR for finding G for a QBD to obtain an Algorithm (LR)* that can be used for calculating R for a QBD. This was observed by Latouche and Ramaswami in their analysis [10].

There is a further possibility available for a QBD. We have the well-known relations

$$U = C_1 + C_0 G \quad \text{and} \quad R = C_0 (I - U)^{-1}$$

(see Hajek [6] and Latouche [9]). From these, R may be calculated via U once G has been determined. In fact, any Algorithm A for finding G in a QBD gives rise to an Algorithm AU for computing R.

This provides us with several methods for computing R for a QBD. Apart from the known Neuts method [12, page 13] and the Schur factorisation, both discussed in [8], we have the new Algorithms (LR)*, HU and H*.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

3.3. Relations between the algorithms We now consider together the probabilistic Algorithms HU, H*, LRU and (LR)* proposed for the determination of the rate matrix R. Of these only H* has general applicability beyond the QBD context. We examine the path contributions made to the estimates for R in these algorithms in the QBD case.

Recall that the matrix U is envisaged as referring to visits from level 0 to level 0, with level -1 taboo. For $\ell > 0$, the matrix $U(\ell)$ denotes the contribution to U arising from trajectories which do not reach level ℓ or higher. Similarly G refers to first visits to level -1 from level 0 and $G(\ell)$ is the contribution to G made by trajectories not attaining level ℓ (> 0) or higher. Finally, R refers to visits to level 0 from level -1 with -1 as a taboo level and $R(\ell)$ is the contribution to R made by trajectories not reaching level ℓ (> 0) or higher.

The estimate T_N of G made by iteration N of Algorithm H is then $G(2^{N+1})$ and is based on the determination of $U(2^{N+1})$. We have

$$U(2^{N+1} + 1) = A_1 + A_2 G(2^{N+1}). \tag{3.1}$$

Hence the estimate of U made in iteration N of Algorithm HU is $U(2^{N+1} + 1)$. Also

$$R(2^{N+1} + 1) = C_0 \left[I - U(2^{N+1} + 1) \right]^{-1},$$

so iteration N of Algorithm HU provides the estimate $R(2^{N+1} + 1)$ for R. The above argument shows incidentally that Algorithm HU is enhanced by the use of (3.1) to calculate a value for U rather than merely using the value of U already employed in estimating G.

In the same way, iteration N of Algorithm LR provides the estimate $G(2^{N+1} - 1)$ for G. Since

$$U(2^{N+1}) = A_1 + A_2 G(2^{N+1} - 1) \quad \text{and} \quad R(2^{N+1}) = C_0 \left[I - U(2^{N+1}) \right]^{-1},$$

the contribution to R from iteration N of Algorithm LRU is $R(2^{N+1})$.

Iteration N of Algorithm H* incorporates the contributions to R of all paths not involving level 2^{N+1} or higher, so that the estimate of R provided by that iteration is also $R(2^{N+1})$.

This shows that, in the QBD case, iteration N of Algorithms LRU and H* yields a common value to machine accuracy. Runs of the two algorithms with a number of examples confirmed this, so giving a useful check of our codes. Also the CPU times for Algorithm LRU and the simplified form of Algorithm H* for the QBD case were found to be the same, so that Algorithm LRU may be regarded as simply the QBD case of Algorithm H*.

Finally, iteration N of Algorithm (LR)* is readily seen to give for R the estimate $R(2^{N+1} - 1)$.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

Thus we have simple relationships between the estimates of R made by the various algorithms in the QBD case for a common iteration count N. Algorithm HU incorporates the contribution of trajectories involving one more level than does the general Algorithm H*, while Algorithm (LR)* involves one fewer. Algorithm LRU coincides with the general algorithm in the QBD case.

3.4. Numerics The preceding discussion indicates that one might expect results of very comparable but slightly decreasing accuracy as we move through use of Algorithm HU to Algorithms H* and (LR)* in turn to evaluate R for a QBD. Table 1 illustrates this for [8, Experiment 2].

TABLE 1. Results for [8, Experiment 2].

Case	Method	Iterations I	$\|R_I - C(R_I)\|_\infty$	CPU Time (sec.)
1	HU	12	9.5665e-13	0.010
	H*	12	9.5847e-13	0.010
	(LR)*	12	9.6029e-13	0.010
2	HU	12	8.8818e-16	0.010
	H*	12	8.8818e-16	0.010
	(LR)*	12	9.9920e-16	0.010
3	HU	10	4.2960e-12	0.010
	H*	10	4.3280e-12	0.010
	(LR)*	10	4.3605e-12	0.010
4	HU	10	4.6130e-14	0.010
	H*	10	4.6629e-14	0.010
	(LR)*	10	4.7073e-14	0.010

4. Error measures

In [11] Meini noted that, in the absence of an analysis of numerical stability, the common error measure $\|e - G_I e\|_\infty$ for an approximation G_I to a stochastic fundamental matrix G may not be appropriate for the invariant subspace method. She proposed instead the measure $\|G_I - A(G_I)\|_\infty$, which is also appropriate in the case of substochastic G. We now consider related isues for the error measure

$$\|R_I - C(R_I)\|$$

for an approximation R_I to R.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

We make use of the QBD given by

$$C_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1-rp \end{bmatrix}, \quad C_1 = \begin{bmatrix} 0 & p \\ rp & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1-p & 0 \\ 0 & 0 \end{bmatrix}$$

with $r \geq 1$ and $0 < p < 1/r$.

With these parameter choices, the QBD is irreducible. It is null recurrent for $r = 1$ and positive recurrent for $r > 1$, with rate matrix

$$R = \begin{bmatrix} 0 & 0 \\ (1-pr)/(1-p) & (1-pr)/(1-p) \end{bmatrix}.$$

We readily verify that, for a matrix

$$R_l = \begin{bmatrix} 0 & 0 \\ x & y \end{bmatrix} \quad \text{with} \quad 0 \leq x, y \leq \frac{1-pr}{1-p}, \tag{4.1}$$

we have

$$C(R_l) = \begin{bmatrix} 0 & 0 \\ rpy + (1-p)xy & 1 - pr + px \end{bmatrix}.$$

Take $r = 1$ and $p = 1/2$ and put $R_0 = \left[\begin{smallmatrix} 0 & 0 \\ 0.5 & 0.5 \end{smallmatrix}\right]$ and $R_1 = \left[\begin{smallmatrix} 0 & 0 \\ 0.6 & 0.9 \end{smallmatrix}\right]$. Then

$$\|R - R_1\|_\infty = 0.4 < 0.5 = \|R - R_0\|_\infty.$$

Also $C(R_0) = \left[\begin{smallmatrix} 0 & 0 \\ 0.375 & 0.75 \end{smallmatrix}\right]$, so that $\|R_0 - C(R_0)\|_\infty = 0.25$, and $C(R_1) = \left[\begin{smallmatrix} 0 & 0 \\ 0.57 & 0.95 \end{smallmatrix}\right]$, so that $\|R_1 - C(R_1)\|_\infty = 0.35$.

We thus have an example for which $\|R - R_1\|_\infty < \|R - R_0\|_\infty$, and in fact $0 \leq R_0 \leq R_1 \leq R$, but $\|R_0 - C(R_0)\|_\infty < \|R_1 - C(R_1)\|_\infty$.

Now let R_l be as in (4.1) with $x < (1-pr)/(1-r)$ and $y < 1$. Then

$$x(1-p)(1-y) < (1-pr)(1-y)$$

or

$$x - [rpy + xy(1-p)] < [1 - pr + px] - y,$$

so that $\Phi_1 < -\Phi_2$, where $\Phi_i := [R_l - C(R_l)]_{2,i}$ $(i = 1, 2)$. It follows at once that if $\Phi_2 > 0$, then $\Phi_1 < 0$.

References

[1] S. Asmussen and V. Ramaswami, "Probabilistic interpretations of some duality results for the matrix paradigms in queueing theory", *Comm. Statist. Stochastic Models* **6** (1990) 715–733.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press

[2] D. Bini and B. Meini, "On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems", in *Proc. 2nd Intern. Workshop on Numerical Solution of Markov Chains*, (Raleigh, North Carolina, 1995) 21–38.

[3] D. Bini and B. Meini, "Improved cyclic reduction for solving queueing problems", *Numer. Algorithms* **15** (1997) 57–74.

[4] D. A. Bini, G. Latouche and B. Meini, "Quadratically convergent algorithms for solving matrix polynomial equations", Technical Report 424, Université Libre Bruxelles, 2000.

[5] L. W. Bright, "Matrix-analytic methods in applied probability", Ph. D. Thesis, University of Adelaide, 1996.

[6] B. Hajek, "Birth-and-death processes on the integers with phases and general boundaries", *J. Appl. Probab.* **19** (1982) 488–499.

[7] E. Hunt, "A probabilistic algorithm for determining the fundamental matrix of a block $M/G/1$ Markov chain", *Math. Comput. Modelling* **38** (2003) 1203–1209.

[8] E. Hunt, "A probabilistic algorithm for finding the rate matrix of a block-$GI/M/1$ Markov chain", *ANZIAM J.* **45** (2004) 457–475.

[9] G. Latouche, "A note on two matrices occurring in the solution of quasi birth-and-death processes", *Stochastic Models* **3** (1987) 251–257.

[10] G. Latouche and V. Ramaswami, "A logarithmic reduction algorithm for Quasi-Birth-Death processes", *J. Appl. Probab.* **30** (1993) 650–674.

[11] B. Meini, "Solving QBD problems: the cyclic reduction algorithm versus the invariant subspace method", *Adv. Perf. Anal.* **1** (1998) 215–225.

[12] M. F. Neuts, *Matrix geometric solutions in stochastic models* (Johns Hopkins University Press, Baltimore, 1981).

[13] V. Ramaswami, "A duality theorem for the matrix paradigms in queueing theory", *Comm. Statist. Stochastic Models* **6** (1990) 151–161.

https://doi.org/10.1017/S1446181100009639 Published online by Cambridge University Press