Journal of Mathematical Modeling
Vol. 9, No. 4, 2021, pp. 573-583. Research Article

On the spectral properties and convergence of the bonus-malus Markov chain model

Kenichi Hirose*

10-17 Moto-machi, Ono City, Fukui 912-0081, Japan

Email(s): kenichihirose@gmail.com

Abstract. In this paper, we study the bonus-malus model denoted by $BM_k(n)$. It is an irreducible and aperiodic finite Markov chain but it is not reversible in general. We show that if an irreducible, aperiodic finite Markov chain has a transition matrix whose secondary part is represented by a nonnegative, irreducible and periodic matrix, then we can estimate an explicit upper bound of the coefficient of the leading-order term of the convergence bound. We then show that the $BM_k(n)$ model has the above-mentioned periodicity property. We also determine the characteristic polynomial of its transition matrix. By combining these results with a previously studied one, we obtain essentially complete knowledge on the convergence of the $BM_k(n)$ model in terms of its stationary distribution, the order of convergence, and an upper bound of the coefficient of the convergence bound.

Keywords: Bonus-malus system, Markov chains, convergence to stationary distribution, the Perron-Frobenius theorem.
AMS Subject Classification 2010: 60J10, 60J20, 15B48, 15B51.

1 Introduction

For an irreducible, aperiodic finite Markov chain, the spectral properties of its transition matrix are closely related with its asymptotic behavior. Indeed, such a matrix has the leading simple eigenvalue 1 and an associated positive eigenvector which, after normalizing, gives a unique stationary distribution. In addition, the second largest eigenvalues, more precisely their common absolute value and the maximal multiplicity in the minimal polynomial, determine the order of convergence to the stationary distribution. However, the convergence bound includes an indeterminacy due to a positive constant.

In this paper, we show that if the transition matrix of such a finite Markov chain has a certain property, which we call $(P(N))$, then we can estimate an explicit upper bound of the

*Corresponding author.
Received: 22 February 2021 / Revised: 9 March 2021 / Accepted: 15 March 2021
DOI: 10.22124/jmm.2021.18991.1625

 http://jmm.guilan.ac.ir

leading-order coefficient of the convergence bound.

Next we introduce the bonus-malus finite Markov chain model $BM_k(n)$, which is named after a system of insurance premiums (see [2] for background). We then show that the transition matrix of this model has the above-mentioned property. Furthermore, we determine the characteristic polynomial of the transition matrix. Thus, for the bonus-malus Markov chain model, we gain more detailed knowledge on its spectrum and the convergence to its stationary distribution.

2 Convergence theorem for irreducible, aperiodic finite Markov chains with periodic secondary parts

For an integer $n \geq 2$, let F denote an $n \times n$ real matrix which is nonnegative, irreducible and aperiodic. The Perron-Frobenius theorem implies the following facts:

- The matrix F has a positive eigenvalue $\lambda_* > 0$ of algebraic multiplicity 1 such that λ_* equals the spectral radius of F, and an associated positive eigenvector $x \in \mathbb{R}^n$, $x > 0$.
- The transpose of F, denoted by F^T, also has λ_* as an eigenvalue and an associated positive eigenvector $u \in \mathbb{R}^n$, $u > 0$.
- We normalize $x = (x_i)$ and $u = (u_i)$ such that $\|x\|_1 = \sum_{i=1}^n x_i = 1$ and $u^T x = \sum_{i=1}^n u_i x_i = 1$.
- Suppose furthermore that F is a column stochastic matrix, then we have $\lambda_* = 1$ and $u = (1, 1, \ldots, 1)^T$.

Proposition 1. *Keeping the above notation, let $W_1(F)$ be the eigenspace of F associated to λ_*; equivalently, $W_1(F)$ is the one-dimensional subspace in $\mathbb{R}^n$ spanned by x, and let $W_2(F)$ be the hyperplane in $\mathbb{R}^n$ orthogonal to u; namely, $W_2(F) = \left\{\xi \in \mathbb{R}^n \;\middle|\; u^T\xi = \sum u_i\xi_i = 0\right\}$. Then, identifying F as a linear transformation on $\mathbb{R}^n$, both $W_1(F)$ and $W_2(F)$ are invariant under the action of F, and we have a direct sum decomposition $\mathbb{R}^n = W_1(F) \bigoplus W_2(F)$; equivalently, for each $i = 1, 2$ the restriction of F to $W_i(F)$, denoted by F_i, is a linear transformation of $W_i(F)$ and $F = F_1 \bigoplus F_2$. Clearly $F_1 = \lambda_* \mathrm{id}_{W_1(F)}$. We call F_2 the secondary part of F. If additionally F is column stochastic, then $\lambda_* = 1$ and $W_2(F) = \{\xi \in \mathbb{R}^n \mid \sum \xi_i = 0\}$ as noted above.*

Proof. Consider the rank-one $n \times n$ matrix $P = xu^T$. Then P is idempotent and P commutes with F. Indeed, P is idempotent because $P^2 = x(u^T x)u^T = xu^T = P$. Next $FP = Fxu^T = \lambda_* xu^T$. Since $F^T u = \lambda_* u$, we have $u^T F = \lambda_* u^T$. Hence, $PF = xu^T F = \lambda_* xu^T$. Thus $FP = PF$. Therefore we have an F-invariant direct sum decomposition $\mathbb{R}^n = \mathrm{Im}\,(P) \bigoplus \mathrm{Ker}\,(P)$.

Clearly $\mathrm{Im}\,(P) = W_1(F)$. Hence, if we show that $\mathrm{Ker}\,(P) = W_2(F)$ then we are done. Let $\xi \in \mathrm{Ker}\,(P)$. Then $P\xi = x(u^T\xi) = 0$. Since $x > 0$, we have $u^T\xi = 0$. Hence $\xi \in W_2(F)$. The converse is also true. Thus $\mathrm{Ker}\,(P) = W_2(F)$. This completes the proof. ☐

Now consider the following property for nonnegative, irreducible and aperiodic matrices. Let F and F_2 be as in Proposition 1. Let $n \geq 3$ and let N denote an integer such that $2 \leq N \leq n-1$.

We say that F has property $(P(N))$ if the secondary part F_2 is represented by an $(n-1)\times(n-1)$ matrix which is nonnegative, irreducible and periodic with period N. In the next section we present an example of a matrix which has this property.

Assume that F has property $(P(N))$. Let $\varphi(t)$ denote the characteristic polynomial of F, and $\psi(t)$ the characteristic polynomial of F_2. Since $F = F_1 \bigoplus F_2$, we have $\varphi(t) = (t-\lambda_*)\psi(t)$. The Perron-Frobenius theorem for nonnegative, irreducible and periodic matrices applied to F_2 implies the following facts (See [1, Chapter XIII, §5]):

- The polynomial $\psi(t)$ is of the form $\psi(t) = t^r\overline{\psi}(t^N)$ for some polynomial $\overline{\psi}(X) \in \mathbb{R}[X]$ and an integer $r \geq 0$.
- The polynomial $\overline{\psi}(X)$ has a unique maximal positive real root $\alpha > 0$ which is simple and $0 < \alpha^{1/N} < \lambda_*$.

Henceforth we consider the case where F is a transition matrix of an irreducible, aperiodic finite Markov chain. In particular, F is column stochastic and hence its leading eigenvalue $\lambda_* = 1$. Then, a well-known convergence theorem for irreducible, aperiodic finite Markov chains says that starting from any initial distribution $\xi \in \mathcal{D}_n$, the series $\{F^\nu\xi\}_{\nu=0}^{\infty}$ converges to a unique stationary distribution $0 < x \in \mathcal{D}_n$. Here $\mathcal{D}_n = \{\xi \in {\mathbb{R}_+}^n \mid \|\xi\|_1 = \sum \xi_i = 1\}$ is the set of probability distributions on $\{1,2,\ldots,n\}$, and ${\mathbb{R}_+}^n$ is the nonnegative orthant in $\mathbb{R}^n$. In the following we present a proof of the convergence theorem when F has property $(P(N))$. Furthermore, we show that in this case we can gain more knowledge on the leading-order coefficient of the convergence bound.

Recall that the L^1-operator norm of an $n\times n$ complex matrix M is defined by $\|M\|_1 = \max_{\|\xi\|_1=1}\|M\xi\|_1$ where the norm on the right-hand side is the L^1-vector norm on $\mathbb{C}^n$. It is known that $\|M\|_1 = \max_{1\leq j\leq n}\sum_{i=1}^n |M_{ij}|$. See [3, Section 5.6]. Therefore, if M is an $n\times n$ real column stochastic matrix, then $\|M\|_1 = 1$.

Theorem 1. *Let F be the transition matrix of an irreducible, aperiodic finite Markov chain. Suppose that F has property $(P(N))$. Let $P = xu^T$ where x is the normalized positive eigenvector of F associated to its leading eigenvalue 1 and $u = (1,1,\ldots,1)^T$. Then, there exists a positive constant $C > 0$, and for any $\varepsilon > 0$ there exists an integer $N' > 0$ such that*

$$\|F^\nu - P\|_1 \leq (C+\varepsilon)\rho^\nu \quad \text{for all } \nu \geq N', \tag{1}$$

where $0 < \rho < 1$ is the second largest positive eigenvalue of F (namely, $\rho = \alpha^{1/N}$ in the above notation). Explicitly, one can take

$$C = 2^{n-1}\sum_{l=0}^{N-1} |\varphi'(\lambda_l)|^{-1}, \tag{2}$$

where $\varphi(t)$ is the characteristic polynomial of F, $\lambda_l = \rho\omega^l, 0 \leq l \leq N-1$, $\omega = \exp(2\pi i/N)$.

Proof. Considering the spectral decomposition of F, seen as a linear transformation on $\mathbb{C}^n$, we have the following formula valid for $\nu \geq n$

$$F^\nu = \sum_{\lambda\in\sigma(F)}\sum_{j=0}^{s_\lambda-1}\binom{\nu}{j}\lambda^{\nu-j}(F-\lambda I_n)^j\pi_\lambda(F),$$

where $\sigma(F) \subset \mathbb{C}$ is the set of eigenvalues of F, and s_λ denotes the multiplicity of $\lambda \in \sigma(F)$ in the minimal polynomial (note that s_λ is the maximal size of Jordan blocks of F associated to λ) and $\pi_\lambda(F)$ denotes the projection matrix onto the generalized eigenspace associated to $\lambda \in \sigma(F)$. See [4, (6.1.39), (6.1.41)].

The set $\sigma(F)$ contains 1 and $\lambda_l, 0 \leq l \leq N-1$. Since these eigenvalues are simple, we have $s_1 = s_{\lambda_l} = 1, 0 \leq l \leq N-1$. Note also that $P = \pi_1(F)$. Therefore we have

$$F^\nu = P + \sum_{l=0}^{N-1} {\lambda_l}^\nu \pi_{\lambda_l}(F) + \sum_{|\lambda|<\rho} \sum_{j=0}^{s_\lambda - 1} \binom{\nu}{j} \lambda^{\nu-j} (F - \lambda I_n)^j \pi_\lambda(F).$$

Hence, we have

$$\|F^\nu - P\|_1 \leq C_1 \rho^\nu + C_2 \sum_{|\lambda|<\rho} \nu^{s_\lambda - 1} |\lambda|^\nu,$$

where $C_1 = \sum_{l=0}^{N-1} \|\pi_{\lambda_l}(F)\|_1$ and C_2 is a positive number which does not depend on ν. In the above inequality, the second term on the right-hand side is dominated by ρ^ν as ν tends to infinity. For, if $|\lambda| < \rho$ then $\nu^{s_\lambda-1}|\lambda|^\nu \rho^{-\nu} = \nu^{s_\lambda-1}(|\lambda|/\rho)^\nu \to 0$ (as $\nu \to \infty$).

Finally we show that C given by (2) is an upper bound of C_1 by using the properties of the Mahler measure of polynomials. It suffices to show that $\|\pi_{\lambda_l}(F)\|_1 \leq 2^{n-1}|\varphi'(\lambda_l)|^{-1}$ for each $0 \leq l \leq N-1$. We divide the proof in three steps.

Step 1: Each projection matrix $\pi_\lambda(F)$ can be expressed as a polynomial in F with complex coefficients via partial fraction decomposition of $1/\varphi(t)$. See [4, (6.1.38a), (6.1.38b)]. Since every λ_l is a simple eigenvalue, the polynomial expression of $\pi_{\lambda_l}(F)$ is $\pi_{\lambda_l}(F) = c_l \hat{\varphi}_l(F)$ where $c_l = \mathrm{Res}(1/\varphi(t), \lambda_l) = 1/\varphi'(\lambda_l)$ and $\hat{\varphi}_l(t) = \varphi(t)/(t - \lambda_l) \in \mathbb{C}[t]$. Note that $\hat{\varphi}_l$ is monic and of degree $n-1$.

Step 2: For a complex polynomial $f(t)$ of degree d, written by

$$f(t) = a_0 t^d + a_1 t^{d-1} + \cdots + a_d = a_0(t - \alpha_1) \cdots (t - \alpha_d),$$

with $a_0 \neq 0$, the length is defined by $L(f) = \sum_{i=0}^d |a_i|$ and the Mahler measure is defined by $M(f) = |a_0| \prod_{i=1}^d \max\{|\alpha_i|, 1\} = |a_0| \prod_{|\alpha_i| \geq 1} |\alpha_i|$. It is known that $L(f)$ and $M(f)$ satisfy the inequality $M(f) \leq L(f) \leq 2^d M(f)$. See [5, Sections 6-8]. The set of roots of $\hat{\varphi}_l(t)$ is $\sigma(F) \setminus \{\lambda_l\}$. Therefore, all the roots of $\hat{\varphi}_l(t)$ except for 1 lie inside the unit disk. Hence, by definition, we have $M(\hat{\varphi}_l) = 1$. The above inequality implies that $L(\hat{\varphi}_l) \leq 2^{n-1} M(\hat{\varphi}_l) = 2^{n-1}$.

Step 3: Since F is column stochastic, $\|F\|_1 = 1$. Therefore, for any $f(t) = \sum_{i=0}^d a_i t^{d-i} \in \mathbb{C}[t]$, we have

$$\|f(F)\|_1 = \left\| \sum_{i=0}^d a_i F^{d-i} \right\|_1 \leq \sum_{i=0}^d |a_i| \|F^{d-i}\|_1 \leq \sum_{i=0}^d |a_i| {\|F\|_1}^{d-i} = \sum_{i=0}^d |a_i| = L(f).$$

By combining the above results, we have

$$\|\pi_{\lambda_l}(F)\|_1 = \|c_l \hat{\varphi}_l(F)\|_1 \leq |c_l| L(\hat{\varphi}_l) \leq 2^{n-1} |\varphi'(\lambda_l)|^{-1}.$$

This completes the proof. □

We add a remark. Without assuming that F has property $(P(N))$, if all second largest eigenvalues of F are simple, then inequality (1) still holds where ρ should be taken as the common absolute value of the second largest eigenvalues of F. The estimate of an upper bound of the leading-order coefficient in the above proof is also applicable to this case. If F has second largest eigenvalues that are not simple, then in the right-hand side of inequality (1), ρ^{ν} should be replaced by $\nu^{s-1}\rho^{\nu}$ where s denotes the maximal multiplicity of the second largest eigenvalues in the minimal polynomial of F. In this case, we can estimate an upper bound of the leading-order coefficient as a sum of the length of the polynomials corresponding to $\lambda^{-(s-1)}(F-\lambda I_n)^{s-1}\pi_{\lambda}(F)$ over the second largest eigenvalues λ such that $s_{\lambda}=s$.

3 The bonus-malus Markov chain model

Following [2], we introduce a family of time-homogeneous finite Markov chains, called the bonus-malus model and denoted by $BM_k(n)$, where $n \geq 2$ is the number of states and k is an integer parameter such that $1 \leq k \leq n-1$.

The $BM_k(n)$ model can be seen as a random walk on a bounded integer line, denoted by $\{1, 2, \ldots, n\}$, with the following transition rules:

- From interior states $i \in \{2, \ldots, n-1\}$, the probability of moving one step down is p and moving k steps up is $q(=1-p)$. If $i+k>n$, then the terminal state is n.
- There are two boundary states $i=1$ and n. From $i=1$, the probability of staying in the same state is p and moving k steps up is q. From $i=n$, the probability of moving one step down is p and staying in the same state is q.

This model arose from the bonus-malus system for insurance premiums used in automobile insurance and workers' compensation insurance. Under this system, the insurance premium of an individual insured member in a given period is adjusted upwards or downwards according to the occurrence of accidents in the preceding period. For more background information of this model, see [2].

The transition matrix of the $BM_k(n)$ model is given by the following $n \times n$ matrix

$$A = \begin{array}{r} k\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \\ n-k\left\{\begin{array}{c} \\ \\ \\ \\ \end{array}\right. \end{array} \begin{bmatrix} p & p & & & & & & \\ & 0 & p & & & & & \\ & & \ddots & \ddots & & & & \\ & & & 0 & p & & & \\ q & 0 & & & \ddots & \ddots & & \\ & q & \ddots & & & \ddots & \ddots & \\ & & \ddots & 0 & 0 & \cdots & 0 & p \\ & & & q & q & \cdots & q & q \end{bmatrix},$$

where

$$A_{ij} = \begin{cases} p, & \text{if } i = \max\{j-1, 1\}, \\ q, & \text{if } i = \min\{j+k, n\}, \\ 0, & \text{otherwise}, \end{cases}$$

for $i, j \in \{1, 2, \ldots, n\}$. Observe that A is a nonnegative and column stochastic matrix. Further, as shown in [2], A is irreducible and aperiodic.

When $k = 1$ and $p = 1/2$, the $BM_k(n)$ model reduces to a simple symmetric random walk on a bounded integer line with self-loops (laziness) at both boundary states. In this special case A is symmetric and has a uniform stationary distribution; hence, it is reversible. However, in general $BM_k(n)$ Markov chains are not reversible, nor do they have uniform stationary distributions.

Note that in the special case of $k = n - 1$, the convergence to the stationary distribution of the model is fully analyzed (see [2, Section 3.3]). Therefore, in the remainder of this paper we focus on the cases in which $n \geq 3$ and $1 \leq k \leq n - 2$.

Theorem 2. *Let n and k be integers such that $n \geq 3$ and $1 \leq k \leq n - 2$. Then the transition matrix A of the $BM_k(n)$ model has property $(P(k+1))$. Explicitly, we have*

$$U^{-1}AU = \begin{bmatrix} 1 & \\ & \tilde{A}_2 \end{bmatrix}.$$

Here

$$\tilde{A}_2 = \begin{array}{r} k\left\{\vphantom{\begin{matrix}0\\0\\ \ddots\end{matrix}}\right. \\ n-k-1\left\{\vphantom{\begin{matrix}q\\ \ddots\\ q\end{matrix}}\right. \end{array} \begin{bmatrix} 0 & p & & & & \\ & 0 & p & & & \\ & & \ddots & \ddots & & \\ q & & & \ddots & \ddots & \\ & \ddots & & & \ddots & p \\ & & q & 0 & \cdots & 0 \end{bmatrix}, \tag{3}$$

where

$$(\tilde{A}_2)_{ij} = \begin{cases} p, & \text{if } i = j - 1, \\ q, & \text{if } i = j + k, \\ 0, & \text{otherwise}, \end{cases}$$

for $i, j \in \{1, 2, \ldots, n-1\}$ is nonnegative, irreducible and periodic with period $k+1$; and, U is the following $n \times n$ nonsingular matrix composed of n column vectors

$$U = \big[x, e_1 - e_2, e_2 - e_3, \ldots, e_{n-1} - e_n\big],$$

where $x \in \mathbb{R}^n$, $x > 0$ is a positive eigenvector of A associated to its eigenvalue 1, and $\{e_i\}_{i=1}^n$ is the standard basis of $\mathbb{R}^n$. We call $\tilde{A}_2$ the period matrix of the $BM_k(n)$ model.

Proof. We know that $Ax = x$. Note that $\mathcal{E} = \{e_i - e_{i+1}\}_{i=1}^{n-1}$ forms a basis of the subspace $W_2(A) = \{\xi \in \mathbb{R}^n \mid \sum \xi_i = 0\}$. Thus U is nonsingular.

To find $\tilde{A}_2$, we note

$$A(e_i - e_{i+1}) = \begin{cases} q(e_{k+1} - e_{k+2}), & \text{for } i = 1, \\ p(e_{i-1} - e_i) + q(e_{k+i} - e_{k+i+1}), & \text{for } 2 \leq i \leq n-k-1, \\ p(e_{i-1} - e_i), & \text{for } n-k \leq i \leq n-1. \end{cases}$$

This shows that the linear transformation A_2 on $W_2(A)$ is represented by the matrix $\tilde{A}_2$ of the form (3) with respect to the basis $\mathcal{E}$. The following example shows the calculation of $\tilde{A}_2$ for $n = 5,\ k = 2$.

$$
\begin{aligned}
A\underline{\mathcal{E}} &= \begin{bmatrix} p & p & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 \\ q & 0 & 0 & p & 0 \\ 0 & q & 0 & 0 & p \\ 0 & 0 & q & q & q \end{bmatrix} \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & -1 & 1 & \\ & & -1 & 1 \\ & & & -1 \end{bmatrix} = \begin{bmatrix} 0 & p & 0 & 0 \\ 0 & -p & p & 0 \\ q & 0 & -p & p \\ -q & q & 0 & -p \\ 0 & -q & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & -1 & 1 & \\ & & -1 & 1 \\ & & & -1 \end{bmatrix} \begin{bmatrix} 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ q & 0 & 0 & p \\ 0 & q & 0 & 0 \end{bmatrix} = \underline{\mathcal{E}}\tilde{A}_2,
\end{aligned}
$$

where $\underline{\mathcal{E}}$ denotes the 5×4 matrix listing horizontally the column vectors of $\mathcal{E}$.

The matrix $\tilde{A}_2$ is nonnegative although not stochastic. Observe that the directed graph associated with matrix $\tilde{A}_2$ is strongly connected and periodic with period $k+1$. Hence $\tilde{A}_2$ is irreducible and periodic with period $k+1$. Figure 1 depicts the directed graph of the period matrix for $n = 6, k = 3$, which has period 4.

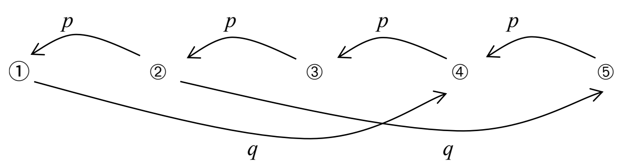

Figure 1: Directed graph associated with the period matrix of $BM_3(6)$.

This completes the proof. $\square$

4 The characteristic polynomial of the $BM_k(n)$ transition matrix

Let $\varphi_{n,k}(t)$ and $\psi_{n-1,k}(t)$ denote the characteristic polynomials of the transition matrix A and the period matrix $\tilde{A}_2$ of the $BM_k(n)$ model, respectively. We know that $\varphi_{n,k}(t) = (t-1)\psi_{n-1,k}(t)$. Next we determine $\psi_{n-1,k}(t)$.

Theorem 3. *Let n and k be integers such that $n \geq 3$ and $1 \leq k \leq n-2$. Let $n-1 = m(k+1)+r$ where $m \geq 0$ and $0 \leq r \leq k$. Then we have*

$$\psi_{n-1,k}(t) = t^r f_{n-1,k}(t),$$

where $f_{n-1,k}(t) = (p^k q)^m \overline{f_{n-1,k}}(t^{k+1}/(p^k q))$ for some $\overline{f_{n-1,k}}(X) \in \mathbb{Z}[X]$ of degree m, constructed in the following way. For integers $\mu \geq 0, 0 \leq \nu \leq k$, write $\overline{f_{(\mu,\nu)}}(X) = \overline{f_{\mu(k+1)+\nu,k}}(X)$ and define *them recursively by*

$$\begin{cases} \overline{f_{(0,\nu)}}(X) = 1, & \text{for } 0 \le \nu \le k, \\ \overline{f_{(\mu+1,\nu)}}(X) = X.\overline{f_{(\mu,k)}}(X) - \sum_{\alpha=0}^{\nu} \overline{f_{(\mu,\alpha)}}(X), & \text{for } \mu \ge 0, 0 \le \nu \le k. \end{cases} \tag{4}$$

The proof of Theorem 3 requires some preliminary steps. For each integer j such that $1 \le j \le n-1$, let $[\tilde{A}_2]_j$ denote the $j \times j$ matrix obtained from $\tilde{A}_2$ by deleting its first $(n-j-1)$ rows and columns, and put $Y_j = tI_j - [\tilde{A}_2]_j$ where I_j is the identity matrix of size j. Observe that for j such that $k+1 \le j \le n-1$, $[\tilde{A}_2]_j$ is the period matrix of $BM_k(j+1)$. Hence $\psi_{j,k}(t) = \det Y_j$ for $k+1 \le j \le n-1$. We extend this notation by setting $\psi_{j,k}(t) = \det Y_j$ for $1 \le j \le k$ as well and $\psi_{0,k}(t) = 1$. For simplicity, henceforth we will write $\psi_j = \psi_{j,k}(t)$ and $c = p^k q$.

Lemma 1. *With the above notation, $\psi_j \in \mathbb{R}[t], 0 \le j \le n-1$, satisfy the following equations*

$$\begin{cases} \psi_j = t^j, & \text{for } 0 \le j \le k, \\ \psi_j = t\psi_{j-1} - c\psi_{j-k-1}, & \text{for } k+1 \le j \le n-1. \end{cases} \tag{5}$$

Proof. For $0 \le j \le k$, we have

$$\psi_j = \det Y_j = \begin{vmatrix} t & -p & 0 & 0 \\ & t & \ddots & 0 \\ & & \ddots & -p \\ & & & t \end{vmatrix} = t^j.$$

For $k+1 \le j \le n-1$, we have

$$\psi_j = \det Y_j = \begin{vmatrix} t & -p & & & & \\ & t & -p & & & \\ & & \ddots & \ddots & & \\ -q & & & \ddots & \ddots & \\ & \ddots & & & \ddots & -p \\ & & -q & 0 & \cdots & t \end{vmatrix}.$$

Expand the above determinant along the first column, which has only two nonzero entries. The contribution from $(1,1)$ entry equals $t \det Y_{j-1} = t\psi_{j-1}$. The contribution from $(k+1,1)$ entry is

$$(-1)^k(-q) \left|\begin{array}{ccc|c} -p & & & \\ * & \ddots & & O_{k,j-k-1} \\ * & * & -p & \\ \hline * & \cdots & * & \\ \vdots & \vdots & \vdots & Y_{j-k-1} \\ * & \cdots & * & \end{array}\right|,$$

where the upper-left block is a $k \times k$ lower triangular matrix whose diagonal entries are all equal to $-p$, the upper-right block is a $k \times (j-k-1)$ matrix whose entries are all zero, and the lower-right block is identical to Y_{j-k-1}. Hence the contribution from $(k+1, 1)$ entry equals $(-1)^k(-q)(-p)^k \det Y_{j-k-1} = -c\psi_{j-k-1}$. Therefore we have $\psi_j = t\psi_{j-1} - c\psi_{j-k-1}$. This completes the proof. □

The difference equation $\psi_j = t\psi_{j-1} - c\psi_{j-k-1}$ is written in the matrix form

$$\begin{bmatrix} \psi_j \\ \vdots \\ \psi_{j-k} \end{bmatrix} = M \begin{bmatrix} \psi_{j-1} \\ \vdots \\ \psi_{j-k-1} \end{bmatrix},$$

where M is the following $(k+1) \times (k+1)$ matrix

$$M = \begin{bmatrix} t & & & & -c \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix}.$$

Lemma 2. *The $(k+1)$-th power of the above matrix M is given by*

$$M^{k+1} = \begin{bmatrix} t^{k+1} - c & -ct & -ct^2 & \cdots & -ct^k \\ t^k & -c & -ct & \cdots & -ct^{k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ t^2 & 0 & \vdots & \cdots & -ct \\ t & 0 & 0 & \cdots & -c \end{bmatrix}.$$

Proof. Observe first that

$$M^k = \begin{bmatrix} t^k & -c & -ct & \cdots & -ct^{k-1} \\ t^{k-1} & 0 & -c & \cdots & -ct^{k-2} \\ \vdots & \vdots & \vdots & & \vdots \\ t & 0 & 0 & \cdots & -c \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

We show the $k = 3$ case by direct calculation, which explains the proof in the general case.

$$M = \begin{bmatrix} t & 0 & 0 & -c \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad M^2 = \begin{bmatrix} t^2 & 0 & -c & -ct \\ t & 0 & 0 & -c \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad M^3 = \begin{bmatrix} t^3 & -c & -ct & -ct^2 \\ t^2 & 0 & -c & -ct \\ t & 0 & 0 & -c \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Next we claim that $M^{k+1} = tM^k - cI_{k+1}$. This follows from the Cayley-Hamilton theorem by noting that the characteristic polynomial of M is $\det(xI_{k+1} - M) = x^{k+1} - tx^k + c$. By combining these two results, we have the required expression of M^{k+1}. □

Proof of Theorem 3. Note first that from Eq. (4) it follows that $\overline{f_{(\mu,\nu)}}(X) \in \mathbb{Z}[X]$ and $\deg \overline{f_{(\mu,\nu)}}(X) = \mu$ for all integers $\mu \geq 0, 0 \leq \nu \leq k$. Write $\psi_{(\mu,\nu)} = \psi_{\mu(k+1)+\nu}$ and

$$y_\mu = \begin{bmatrix} \psi_{(\mu,k)} \\ \vdots \\ \psi_{(\mu,0)} \end{bmatrix} \in \mathbb{R}[t]^{k+1}.$$

It then suffices to show that

$$y_\mu = \begin{bmatrix} t^k f_{(\mu,k)}(t) \\ \vdots \\ t f_{(\mu,1)}(t) \\ f_{(\mu,0)}(t) \end{bmatrix} \quad \text{for } \mu \geq 0.$$

We induct on μ. For $\mu = 0$, indeed we have $y_0 = (\psi_k, \ldots, \psi_1, \psi_0)^T = \left(t^k, \ldots, t, 1\right)^T$. Assume that the above relation holds for μ. Then by Lemma 1, $y_{\mu+1}$ satisfies $y_{\mu+1} = M^{k+1} y_\mu$. For $0 \leq \nu \leq k$, let $(y_\mu)_\nu$ denote y_μ's $(\nu+1)$-th component from the bottom. Then by Lemma 2, we have

$$(y_{\mu+1})_\nu = t^{\nu+1}(y_\mu)_k - c(y_\mu)_\nu - ct(y_\mu)_{\nu-1} - \cdots - ct^\nu (y_\mu)_0.$$

By the induction hypothesis, $(y_\mu)_j = t^j f_{(\mu,j)}(t)$ for $0 \leq j \leq k$. Hence we have

$$\begin{aligned}(y_{\mu+1})_\nu &= t^{\nu+1} t^k f_{(\mu,k)}(t) - ct^\nu f_{(\mu,\nu)}(t) - ct^\nu f_{(\mu,\nu-1)}(t) - \cdots - ct^\nu f_{(\mu,0)}(t) \\ &= t^\nu \left[t^{k+1} f_{(\mu,k)}(t) - c\left(f_{(\mu,\nu)}(t) + f_{(\mu,\nu-1)}(t) + \cdots + f_{(\mu,0)}(t)\right)\right].\end{aligned}$$

By the construction of $\overline{f_{(\mu+1,\nu)}}(X)$ and noting that $f_{(\mu,\nu)}(t) = c^\mu \overline{f_{(\mu,\nu)}}(X)$ where $X = t^{k+1}/c$, we have

$$\begin{aligned}(y_{\mu+1})_\nu &= t^\nu c^{\mu+1} \left(X.\overline{f_{(\mu,k)}}(X) - \sum_{\alpha=0}^{\nu} \overline{f_{(\mu,\alpha)}}(X)\right) \\ &= t^\nu c^{\mu+1} \overline{f_{(\mu+1,\nu)}}(X) \\ &= t^\nu f_{(\mu+1,\nu)}(t).\end{aligned}$$

This completes the proof. □

Note that in the $k = n-1$ case, we have $\varphi_{n,n-1}(t) = t^{n-1}(t-1)$ (see [2]).

5 Convergence of the $BM_k(n)$ model

We have already known how to construct the stationary distribution of the $BM_k(n)$ model. The main result in [2, Theorem 1] states that a prenormalized eigenvector $x = (x_i)$ of A associated to eigenvalue 1 is given by $x_i = p^{n-i}\gamma_{i-1}$ where $\gamma_{i-1} \in \mathbb{R}$, $1 \leq i \leq n$, satisfy the following equations

$$\begin{cases} \gamma_0 = 1, & \\ \gamma_j = q, & \text{for } 1 \leq j \leq k, \\ \gamma_j = \gamma_{j-1} - c\gamma_{j-k-1}, & \text{for } k+1 \leq j \leq n-1. \end{cases} \tag{6}$$

By normalizing x we obtain the stationary distribution $\hat{x} \in \mathcal{D}_n$ of the $BM_k(n)$ model. Attention should be drawn to the similarity between Eq. (5) in Lemma 1 and Eq. (6) above.

By virtue of Theorem 2, we can apply Theorem 1 to A. In addition, the spectrum of A is completely known from Theorem 3. Let us denote the total variation distance between two distributions $y, z \in \mathcal{D}_n$ by $d(y, z) = (1/2)\|y - z\|_1$ with the L^1-vector norm. Then, by taking $\varepsilon = C$ in Theorem 1 and using other results established in this paper, for $n \geq 3$ and $1 \leq k \leq n-2$, the convergence of the $BM_k(n)$ model is described as follows.

Starting from any initial distribution $\xi \in \mathcal{D}_n$, the series $\{A^\nu \xi\}_{\nu=0}^{\infty}$ converges to the above stationary distribution $\hat{x} > 0$ with respect to this metric. In more detail, there exist a positive constant $C > 0$ and an integer $N > 0$ such that

$$d(A^\nu \xi, \hat{x}) \leq C\rho^\nu \quad \text{for all } \nu \geq N.$$

Here, the order of convergence $\rho = (\alpha p^k q)^{1/(k+1)}$ where α is the maximal positive root of $\overline{f_{n-1,k}}(X)$ defined by Eq. (4) in Theorem 3. This is because $\psi_{n-1,k}(t) = t^r f_{n-1,k}(t)$ and that $f_{n-1,k}(t) = c^m \overline{f_{n-1,k}}(X)$ where $X = t^{k+1}/c$ and $c = p^k q$. This affirms the conjectural statements on the convergence of the $BM_k(n)$ model in [2, Section 3.4]. Moreover, an upper bound of coefficient C is calculated by applying formula (2) in Theorem 1 to the characteristic polynomial $\varphi_{n,k}(t)$ that was determined in Theorem 3.

In practical applications, our results will be instrumental for the insurers adopting the bonus-malus insurance premium system to gain knowledge on the stationary distribution of insurance premiums and quantitative estimates of the speed of convergence to the stationarity based on the predetermined number of risk classes n, the predetermined adjustment parameter k, and the experience value of the probability of nonoccurrence of insurance claims p.

References

[1] F.R. Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea Publishing Company, New York, 1959.

[2] K. Hirose, *Stationary distributions of Markov chain models of the bonus-malus system*, Math. Scientist **42** (2017) 51–59.

[3] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[4] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, 1991.

[5] K. Mahler, *Lectures on Transcendental Numbers,* Lecture Notes in Mathematics 546, Springer-Verlag Berlin Heidelberg, 1976.