On the Imbedding Problem for Three-state Time Homogeneous Markov Chains with Coinciding Negative Eigenvalues

CHEN Yong,　CHEN Jianmin

School of Mathematics and Computing Science, Hunan University of Science and Technology,

Xiangtan, Hunan, 411201, P.R.China. chenyong77@gmail.com

Abstract

For an indecomposable 3×3 stochastic matrix (i.e., 1-step transition probability matrix) with coinciding negative eigenvalues, a new necessary and sufficient condition of the imbedding problem for time homogeneous Markov chains is shown by means of an alternate parameterization of the transition rate matrix (i.e., intensity matrix, infinitesimal generator), which avoids calculating matrix logarithm or matrix square root. In addition, an implicit description of the imbedding problem for the 3×3 stochastic matrix in [13] is pointed out.

 keywords: imbedding problem; three-state time homogeneous Markov chains; negative eigenvalues

MSC numbers: 60J10; 60J27; 60J20

1

1   Introduction

The imbedding problem for finite Markov chains has a long history and was first posed by Elfving [5], which has applications to population movements in social science [21], credit ratings in mathematical finance [11], and statistical inference for Markov processes [1, 18]. For a review of the imbedding problem, the reader can refer to [2].

According to Kingman [11, 16], the imbedding problem is completely solved for the case of 2×2 matrices by D. G. Kendall, who proved that a 2×2 transition probability matrix is compatible with a continuous Markov process if and only if the sum of the two diagonal entries is larger than 1.

The explicit description of the imbedding problem for the 3×3 stochastic matrix with distinct eigenvalues or with coinciding positive eigenvalues is shown by Johansen in [13]. When the common eigenvalue is negative, P. Carette provides several necessary and sufficient conditions to characterize the imbedded stochastic matrix [2, Theorem 3.3, Theorem 3.6]. But in contrast to the above conclusions of Kingman or Johansen, it is not clear-cut.

Let I be the identity matrix. Let P, P^∞ be the 3×3 stochastic matrix and its limiting probability matrix respectively. Theorem 3.3 (resp. Theorem 3.6,) of [2] needs to calculate square root of the matrix $\mathsf{P}^\infty - \mathsf{I}$ (resp. P), and then to test whether each of the off-diagonal elements satisfies an inequality. But the matrix square root is many-valued, just like the matrix logarithm [21, p22].

In the present paper, a new necessary and sufficient condition is shown, which overcomes the difficulty of uncountably many versions of logarithm or square root (Theorem 2.6). We chiefly rely on an alternate parameterization of the transition rate matrix (Eq.(6)) for the proof. At the same time, for a fixed

P^∞, the exact lower bound of the eigenvalue of P that makes P embeddable is given (Remark 5)[1].

In the more general context of time-inhomogeneous Markov chains, the imbedding problem is dealt with by some authors [6, 7, 8, 9, 14]. But we only focus on the time-homogeneous Markov chains here.

2 The imbedding problem for 3-order transition Matrix with coinciding negative eigenvalues

The transition matrix P is called embeddable if there is a transition rate matrix Q for which $P = e^Q$.

Let $P = (p_{ij})$ be a 3×3 transition probability matrix. Suppose P is indecomposable, i.e., its state space does not contain two disjoint closed sets[4, P17]. Let the unique stationary probability distribution be $\mu' = (\mu_1, \mu_2, \mu_3)$ with $\mu_1 + \mu_2 + \mu_3 = 1$. Let $\vec{e} = (1, 1, 1)'$

Lemma 2.1. *Suppose* P *is indecomposable. If* Q *is a transition rate matrix such that* $P = e^{hQ}$, $h > 0$, *then* Q *has* $\mu'Q = 0$.

Proof. If a distribution ν has $\nu'Q = 0$ (i.e., the left eigenvector with eigenvalue 0), then $P = e^{hQ} = \sum_{n=0}^{\infty} \frac{1}{n!} h^n Q^n$ implies that $\nu'P = \nu'$. Since P is indecomposable, one has that $\nu = \mu$. □

Lemma 2.2. *Suppose the transition matrix* P *is embeddable with eigenvalues* $\{1, \lambda, \lambda\}$, $\lambda < 0$. *Then* P *is diagonalizable and satisfies*

$$P = P^\infty + \lambda(I - P^\infty), \tag{1}$$

---

[1] The existence of the lower bound is shown in [2, Theorem 3.7]

where $\mathsf{P}^\infty = \vec{e}\mu'$ *is the limiting probability matrix,* I *is the identity matrix.*[2]

Proof. If Q is a transition rate matrix such that $\mathsf{P} = e^{\mathsf{Q}}$, then Q has a pair of conjugate complex eigenvalues and is diagonalizable. Since $\mathsf{P} = e^{\mathsf{Q}}$, P is diagonalizable too.

The eigenvalues of P are $\{1,\,\lambda,\,\lambda\}$, thus the rank of the matrix $\lambda\mathsf{I} - \mathsf{P}$ is 1. Since $(\lambda\mathsf{I} - \mathsf{P})\vec{e} = (\lambda - 1)\vec{e}$, the three rows of $\lambda\mathsf{I} - \mathsf{P}$ are equal. Then

$$
\mathsf{P} = \begin{bmatrix} 1 - (x + y) & x & y \\ z & 1 - (y + z) & y \\ z & x & 1 - (z + x) \end{bmatrix}, \tag{2}
$$

and $1 - \lambda = x + y + z$. Note that $\mu' = (z,\,x,\,y)/(x + y + z)$, this ends the proof. $\square$

Corollary 2.3. *Suppose* P *satisfies the condition of Lemma 2.2, then the stationary probability distribution is positive and all elements of* P *are positive.*

Proof. Since $\mathsf{P} = \lambda\mathsf{I} + (1 - \lambda)\mathsf{P}^\infty$, one has that $\lambda + (1 - \lambda)\mu_i \geqslant 0$, $i = 1, 2, 3$. Then $\mu_i \geqslant \frac{-\lambda}{1 - \lambda} > 0$, and the off-diagonal elements satisfy $p_{ij} = (1 - \lambda)\mu_j > 0$, $i \neq j$. In addition, it was shown by Goodman that each of the diagonal elements of an embeddable matrix dominates the determinant, and that this determinant is positive: $p_{ii} \geqslant \det \mathsf{P} > 0$ [14, 10]. $\square$

The conclusions of Lemma 2.2 and Corollary 2.3 also appear in [2]. By Lemma 2.2, P is completely determined by its stationary distribution and the coinciding eigenvalues.

---

[2]This conclusion is asserted in [13].

Proposition 2.4. *Suppose that* P *satisfies Eq.(1).* P *can be imbedded if and only if there exists a transition rate matrix* Q *such that it has* $\mu'\mathsf{Q} = 0$ *and eigenvalues* θ *and* $\bar{\theta}$ *and it holds that* $e^{\theta h} = \lambda$ *for some* $h \in \mathbb{R}^+$.

The proof of Proposition 2.4 is presented in Section 2.1. Here θ is a complex eigenvalue $-p + \mathrm{i}q$ with $\frac{q}{p} = \frac{(2k+1)\pi}{-\log|\lambda|}$, $k \in \mathbb{Z}^+$.

Remark 1. By the Runnenberg condition in [21, 20], the complex eigenvalue $-p + \mathrm{i}q$ of the transition rate matrix Q satisfies that

$$\frac{|q|}{p} \leqslant \frac{1}{\sqrt{3}},$$

which the reader can also refer to [3] for detail. Then

$$|\lambda| \leqslant e^{-\sqrt{3}\pi} \doteq 0.0043,$$

and P is almost equal to its limiting probability matrix by Eq.(1).

For a transition rate matrix

$$\mathsf{Q} = \begin{bmatrix} -a_2 - a_3 & a_2 & a_3 \\ b_1 & -b_1 - b_3 & b_3 \\ c_1 & c_2 & -c_1 - c_2 \end{bmatrix}, \tag{3}$$

suppose that it has $\mu'\mathsf{Q} = 0$, that is to say,

$$\mu_1 a_2 - \mu_2 b_1 = \mu_2 b_3 - \mu_3 c_2 = \mu_3 c_1 - \mu_1 a_3. \tag{4}$$

Let

$$\nu = \frac{\mu_1 a_2 - \mu_2 b_1}{2}, \gamma = \frac{\mu_1 a_2 + \mu_2 b_1}{2}, \delta = \frac{\mu_2 b_3 + \mu_3 c_2}{2}, \kappa = \frac{\mu_3 c_1 + \mu_1 a_3}{2}. \tag{5}$$

We propose an alternate parameterization of the transition rate matrix as

$$\mathsf{Q} = \begin{bmatrix} -\frac{\kappa+\gamma}{\mu_1} & \frac{\gamma+\nu}{\mu_1} & \frac{\kappa-\nu}{\mu_1} \\ \frac{\gamma-\nu}{\mu_2} & -\frac{\gamma+\delta}{\mu_2} & \frac{\delta+\nu}{\mu_2} \\ \frac{\kappa+\nu}{\mu_3} & \frac{\delta-\nu}{\mu_3} & -\frac{\delta+\kappa}{\mu_3} \end{bmatrix}, \tag{6}$$

where

$$\kappa, \gamma, \delta \geqslant 0, \quad \kappa + \gamma, \gamma + \delta, \delta + \kappa > 0, \quad \text{and} \quad |\nu| \leqslant \kappa, \gamma, \delta. \tag{7}$$

The re-parameterization is one-to-one. Then the transition rate matrix Q with $\mu' \mathsf{Q} = 0$ must satisfy Eq.(6). [3]

Let the eigen-equation of Q be $\lambda(\lambda^2 + \alpha\lambda + \beta) = 0$. Then we have that

$$\alpha = \frac{\kappa + \gamma}{\mu_1} + \frac{\gamma + \delta}{\mu_2} + \frac{\delta + \kappa}{\mu_3}, \tag{8}$$

$$\beta = \frac{\kappa\gamma + \gamma\delta + \delta\kappa + \nu^2}{\mu_1\mu_2\mu_3}. \tag{9}$$

and the eigenvalue is

$$\theta = -\frac{\alpha}{2} + \mathrm{i}\sqrt{\beta - \alpha^2/4}. \tag{10}$$

Let the ratio between the imaginary (q) and real $(-p)$ parts of the nonzero eigenvalues be

$$
\begin{aligned}
H(\kappa, \gamma, \delta, \nu) &\triangleq \frac{|q|}{p} \\
&= \sqrt{\frac{4\beta - \alpha^2}{\alpha^2}} = \sqrt{4\frac{\beta}{\alpha^2} - 1} \\
&= \sqrt{\frac{4}{\mu_1\mu_2\mu_3} \frac{\kappa\gamma + \gamma\delta + \delta\kappa + \nu^2}{(\frac{\kappa+\gamma}{\mu_1} + \frac{\gamma+\delta}{\mu_2} + \frac{\delta+\kappa}{\mu_3})^2} - 1}.
\end{aligned}
\tag{11}
$$

Remark 2. Since there may be many different transition rate matrices such that $H(\kappa, \gamma, \delta, \nu) = \frac{(2k+1)\pi}{-\log|\lambda|}$, $k \in \mathbb{Z}^+$, the solution of $\mathsf{P} = e^{\mathsf{Q}}$ is not unique.

Remark 3. For the given $\mu' = (\mu_1, \mu_2, \mu_3)$, if $\nu = 0$ and $\kappa : \gamma : \delta = \frac{1}{\mu_2} : \frac{1}{\mu_3} : \frac{1}{\mu_1}$, then $H(\kappa, \gamma, \delta, \nu) = 0$. Note that when

$$\kappa, \gamma, \delta \geqslant 0, \quad \kappa + \gamma, \gamma + \delta, \delta + \kappa > 0, \quad \text{and} \quad |\nu| \leqslant \kappa, \gamma, \delta, \tag{12}$$

---

[3]This re-parameterization of the transition rate matrix appears in [3], and appears in [19, 12, 15, 17] implicitly, which is named the cycle decomposition by some authors.

$H(\kappa, \gamma, \delta, \nu)$ is a continuous real function of $(\kappa, \gamma, \delta, \nu)$.[4] Therefore, $H(\kappa, \gamma, \delta, \nu)$

takes the value $\frac{\pi}{-\log|\lambda|}$ if the maximum of $H(\kappa, \gamma, \delta, \nu)$ is greater than or equal

to $\frac{\pi}{-\log|\lambda|}$.

Therefore, an optimization problem is formulated, and the next Proposition solves it. We say that three positive numbers a, b, c satisfy the triangle inequality if $a + b > c$, $b + c > a$, $a + c > b$. Let

$$m = \min\{\mu_1, \mu_2, \mu_3\}. \tag{13}$$

Proposition 2.5. *Suppose that*

$$F(x_1, x_2, x_3, \mu_1, \mu_2, \mu_3) = \frac{x_1 x_2 + x_2 x_3 + x_3 x_1 + (\min\{x_1, x_2, x_3\})^2}{(\frac{x_1 + x_2}{\mu_1} + \frac{x_2 + x_3}{\mu_2} + \frac{x_3 + x_1}{\mu_3})^2}, \tag{14}$$

with $x_1, x_2, x_3 \geqslant 0$ and $x_1 + x_2$, $x_2 + x_3$, $x_3 + x_1 > 0$. Then the maximum of

$F(x_1, x_2, x_3, \mu_1, \mu_2, \mu_3)$ is

$$
\begin{cases}
\frac{1}{(\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3})^2}, & \text{when } \frac{1}{\mu_1}, \frac{1}{\mu_2}, \frac{1}{\mu_3} \text{ satisfy the triangle inequality,} \\
\frac{\mu_1 \mu_2 \mu_3}{4(1-m)}, & \text{otherwise.}
\end{cases} \tag{15}
$$

The maximum is attained at the points

$$
\begin{cases}
x_1 = x_2 = x_3, & \text{when } \frac{1}{\mu_1}, \frac{1}{\mu_2}, \frac{1}{\mu_3} \text{ satisfy the triangle inequality,} \\
x_1 = x_2 = \frac{\frac{1}{\mu_2} + \frac{1}{\mu_3}}{\frac{2}{\mu_1} - \frac{1}{\mu_2} - \frac{1}{\mu_3}} x_3, & \text{when } \frac{1}{\mu_1} \geqslant \frac{1}{\mu_2} + \frac{1}{\mu_3}, \\
x_2 = x_3 = \frac{\frac{1}{\mu_3} + \frac{1}{\mu_1}}{\frac{2}{\mu_2} - \frac{1}{\mu_3} - \frac{1}{\mu_1}} x_1, & \text{when } \frac{1}{\mu_2} \geqslant \frac{1}{\mu_3} + \frac{1}{\mu_1}, \\
x_3 = x_1 = \frac{\frac{1}{\mu_1} + \frac{1}{\mu_2}}{\frac{2}{\mu_3} - \frac{1}{\mu_1} - \frac{1}{\mu_2}} x_2, & \text{when } \frac{1}{\mu_3} \geqslant \frac{1}{\mu_1} + \frac{1}{\mu_2}.
\end{cases} \tag{16}
$$

Proof of Proposition 2.5 is presented in Section 2.1.

Theorem 2.6. *Suppose that* P *is a 3×3 stochastic matrix with eigenvalues* $\{1, \lambda, \lambda\}$, $\lambda < 0$, *and that* P *satisfies Eq.(1). Then* P *can be imbedded if and*

---

[4]It is needed that the term inside the $\sqrt{\cdot}$ of Eq.(11) is positive.

only if

$$\sqrt{\frac{4\mu_1\mu_2\mu_3}{(\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3)^2} - 1} \geqslant \frac{\pi}{-\log|\lambda|}, \quad \text{when} \ \frac{1}{\mu_1}, \frac{1}{\mu_2}, \frac{1}{\mu_3} \ \text{satisfy the triangle inequality,}$$

(17)

or

$$\sqrt{\frac{m}{1-m}} \geqslant \frac{\pi}{-\log|\lambda|}, \quad \text{otherwise.}$$

(18)

Proof of Theorem 2.6. Clearly, if the function $H(\kappa, \gamma, \delta, \nu)$ reaches its maximum then $\nu = \min\{\kappa, \gamma, \delta\}$. From Proposition 2.4 we have to find a transition rate matrix $Q(\kappa, \gamma, \delta, \nu)$ of the form just above (6), with eigenvalues $\theta(\kappa, \gamma, \delta, \nu)$ and $\bar{\theta}(\kappa, \gamma, \delta, \nu)$ for which $e^{\theta h} = \lambda$ for some $h > 0$. Using Proposition 2.5 we can find $(\kappa_0, \gamma_0, \delta_0, \nu_0)$, so that for the given μ', λ we have

$$H(\kappa_0, \gamma_0, \delta_0, \nu_0) = \frac{\pi}{-\log|\lambda|} = H_0.$$

(19)

The corresponding transition rate matrix $Q(\kappa_0, \gamma_0, \delta_0, \nu_0)$ has eigenvalue $\theta(\kappa_0, \gamma_0, \delta_0, \nu_0) = \theta_0$ and $\bar{\theta}_0$, where

$$\theta_0 = -\frac{\alpha_0}{2} + \mathrm{i}\sqrt{\beta_0 - \alpha_0^2/4} = \frac{\alpha_0}{2}[-1 + \mathrm{i}H_0] = \frac{\alpha_0}{2}\left[-1 + \mathrm{i}\frac{\pi}{-\log|\lambda|}\right].$$

(20)

We finally choose $h = -2\log|\lambda|/\alpha_0$ and find

$$\theta_0 h = -\frac{2\log|\lambda|}{\alpha_0}\frac{\alpha_0}{2}\left[-1 + \mathrm{i}\frac{\pi}{-\log|\lambda|}\right]$$

$$= \log|\lambda| + \mathrm{i}\pi$$

which satisfies $e^{\theta_0 h} = \lambda$. $\qquad\square$

Remark 4. If $\mu_1 + \mu_2 + \mu_3 = 1$, and $\frac{1}{\mu_1}, \frac{1}{\mu_2}, \frac{1}{\mu_3}$ satisfy the triangle inequality, then it can be shown easily that

$$4\mu_1\mu_2\mu_3 \geqslant (\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3)^2 \geqslant 3\mu_1\mu_2\mu_3,$$

(21)

i.e., the term inside the $\sqrt{\cdot}$ of Eq.(17) is positive.

Remark 5. For a fixed $\mathsf{P}^\infty = \vec{e}\mu'$, what are the possible values of $\lambda < 0$ that make the stochastic matrix P embeddable? In [2, Theorem 3.7], P. Carette puts the above question and shows that there exists $\Lambda < 0$ such that P is imbeddable if and only if $\Lambda \leqslant \lambda < 0$. As a consequence of Theorem 2.6, the exact value of Λ is that

$$\Lambda = \begin{cases} -\exp\left\{-\frac{\pi}{\sqrt{b}}\right\}, & \text{when } \frac{1}{\mu_1}, \frac{1}{\mu_2}, \frac{1}{\mu_3} \text{ satisfy the triangle inequality,} \\ -\exp\left\{-\sqrt{\frac{1-m}{m}}\,\pi\right\}, & \text{otherwise,} \end{cases} \tag{22}$$

where $b = \frac{4\mu_1\mu_2\mu_3}{(\mu_1\mu_2+\mu_1\mu_3+\mu_2\mu_3)^2} - 1$, $m = \min\{\mu_1, \mu_2, \mu_3\}$.

2.1 Proof of the propositions

Lemma 2.7. *If P satisfies Eq.(1), then its right eigenvectors with eigenvalue λ span the orthogonal complement of μ.*

Proof. If f is a right eigenvector of P with eigenvalue λ then

$$\lambda f = \mathsf{P}f = \vec{e}\mu' f + \lambda(\mathsf{I} - \vec{e}\mu')f = (1-\lambda)\vec{e}\mu' f + \lambda f, \tag{23}$$

so that $\mu' f = 0$. Hence the eigenvectors span the orthogonal complement of μ. $\qquad\square$

Proof of Proposition 2.4. The necessity. Since $\lambda < 0$ and $\mathsf{P} = e^{h\mathsf{Q}}$, Q has complex eigenvalues θ and $\bar{\theta}$ such that $e^{\theta h} = \lambda$. It follows from Lemma 2.1 that Q has $\mu'\mathsf{Q} = 0$.

The sufficiency. Since P satisfies Eq.(1), one obtains that

$$\mathsf{P} = \mathsf{F}\,\text{diag}\{1, \lambda, \lambda\}\,\mathsf{F}^{-1}, \tag{24}$$

where $\mathsf{F} = [\vec{e}, \varphi_1, \varphi_2]$, φ_1, φ_2 are any two linear independent vectors in the orthogonal complement of μ by Lemma 2.7.

Denote by $f + ig$ the eigenvector of Q with eigenvalue $\theta = p + iq$, $q \neq 0$. Clearly f, g are linear independent. $\mu'\mathsf{Q} = 0$ implies that f, g span the orthogonal complement of μ.

Let $\mathsf{P}_h = e^{h\mathsf{Q}}$. Hence $e^{h\mathsf{Q}} = \sum_{n=0}^{\infty} \frac{h^n}{n!}\mathsf{Q}^n$ implies that if $\mu'\mathsf{Q} = 0$ then $\mu'\mathsf{P}_h = \mu'$ and that

$$\mathsf{P}_h(f + ig) = e^{\theta h}(f + ig) = \lambda(f + ig). \tag{25}$$

Hence $\mathsf{P}_h f = \lambda f$, $\mathsf{P}_h g = \lambda g$. Any transition rate matrix Q has $\mathsf{Q}\vec{e} = 0$, so that $\mathsf{P}_h \vec{e} = \vec{e}$. Thus one obtains that

$$\mathsf{P}_h = [\vec{e},\, f,\, g]\,\mathrm{diag}\,\{1,\, \lambda,\, \lambda\}\,[\vec{e},\, f,\, g]^{-1}. \tag{26}$$

Note that in Eq.(24), one can choose that the matrix $\mathsf{F} = [\vec{e},\, f,\, g]$, which implies that $\mathsf{P}_h = \mathsf{P}$. $\qquad\square$

Lemma 2.8. *Suppose that* $f(x) = x + \frac{a}{x}$ *with* $0 < x \leqslant c$, *where* a, $c > 0$ *are two constants. Then the minimum of* $f(x)$ *is*

$$\begin{cases} 2\sqrt{a}, & \text{when } \sqrt{a} \leqslant c, \\ c + \frac{a}{c}, & \text{when } \sqrt{a} > c. \end{cases} \tag{27}$$

The minimum is attained at the point

$$\begin{cases} x = \sqrt{a}, & \text{when } \sqrt{a} \leqslant c, \\ x = c, & \text{when } \sqrt{a} > c. \end{cases} \tag{28}$$

Proof. It is trivial. $\qquad\square$

Denote that

$$\mathcal{D} = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1,\, x_2,\, x_3 \geqslant 0, x_1 + x_2,\, x_2 + x_3,\, x_3 + x_1 > 0\right\}, \tag{29}$$

$$\mathcal{E}_1 = \mathcal{D} \cap \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 \leqslant x_2 \leqslant x_3\right\}. \tag{30}$$

Since $F(x_1, x_2, x_3) = F(rx_1, rx_2, rx_3)$, $\forall r > 0$, the existence of the maximum of $F(x_1, x_2, x_3)$ on $\mathcal{D}$ is equal to the existence on the unit sphere which holds true since the unit sphere is compact.[5]

Lemma 2.9. *Restricted on $\mathcal{E}_1$, the maximum of $F(x_1, x_2, x_3)$ is*

$$\begin{cases} \frac{\mu_1 \mu_2 \mu_3}{4(1-\mu_1)}, & \text{when } \frac{1}{\mu_1} \geqslant \frac{1}{\mu_2} + \frac{1}{\mu_3}, \\ \frac{1}{(\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3})^2}, & \text{when } \frac{1}{\mu_1} < \frac{1}{\mu_2} + \frac{1}{\mu_3}. \end{cases} \tag{31}$$

The maximum is attained at the points

$$\begin{cases} x_1 = x_2 = \frac{\frac{1}{\mu_2} + \frac{1}{\mu_3}}{\frac{2}{\mu_1} - \frac{1}{\mu_2} - \frac{1}{\mu_3}} x_3, & \text{when } \frac{1}{\mu_1} \geqslant \frac{1}{\mu_2} + \frac{1}{\mu_3}, \\ x_1 = x_2 = x_3, & \text{when } \frac{1}{\mu_1} < \frac{1}{\mu_2} + \frac{1}{\mu_3}. \end{cases} \tag{32}$$

Proof. Restricted on $\mathcal{E}_1$, i.e., $x_1 \leqslant x_2 \leqslant x_3$, we obtain that

$$F(x_1, x_2, x_3) = \frac{x_1 x_2 + x_2 x_3 + x_3 x_1 + x_1^2}{\left(\frac{x_1+x_2}{\mu_1} + \frac{x_2+x_3}{\mu_2} + \frac{x_3+x_1}{\mu_3}\right)^2} = \frac{(x_1 + x_2)(x_3 + x_1)}{\left(\frac{x_1+x_2}{\mu_1} + \frac{x_2+x_3}{\mu_2} + \frac{x_3+x_1}{\mu_3}\right)^2}. \tag{33}$$

Let $r = x_1 + x_2$, $s = x_2 + x_3$, $t = x_3 + x_1$. Then $0 < r \leqslant t \leqslant s$, and

$$F(x_1, x_2, x_3) = \frac{rt}{\left(\frac{r}{\mu_1} + \frac{s}{\mu_2} + \frac{t}{\mu_3}\right)^2} \triangleq G(r, s, t) \leqslant G(r, t, t). \tag{34}$$

That is to say, $G(r, s, t)$ attains its maximum when $s = t$ since it is a decreasing function of s. Let $w = \sqrt{\frac{r}{t}}$. Then $0 < w \leqslant 1$ and we have that

$$\begin{aligned} G(r, t, t) &= \frac{rt}{\left[\frac{1}{\mu_1} r + \left(\frac{1}{\mu_2} + \frac{1}{\mu_3}\right) t\right]^2} \\ &= \left[\frac{\mu_1}{w + \mu_1 \left(\frac{1}{\mu_2} + \frac{1}{\mu_3}\right) \frac{1}{w}}\right]^2 \triangleq L(w). \end{aligned} \tag{35}$$

It follows from Lemma 2.8 that when $\mu_1\left(\frac{1}{\mu_2} + \frac{1}{\mu_3}\right) > 1$, the maximum of $L(w)$ is $\frac{1}{(\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3})^2}$, and when $\mu_1\left(\frac{1}{\mu_2} + \frac{1}{\mu_3}\right) \leqslant 1$, the maximum of $L(w)$ is $\frac{\mu_1}{4\left(\frac{1}{\mu_2} + \frac{1}{\mu_3}\right)}$. In addition, the direct computation yields the maximum points (32) from Lemma 2.8. $\qquad\square$

---

[5]The points $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ are not in $\mathcal{D}$, but the values of $F(x_1, x_2, x_3)$ are zero at these points, which do not alter the maximum of $F(x_1, x_2, x_3)$.

Proof of Proposition 2.5. In order to prove Proposition 2.5 we first assume that μ_i satisfies the triangular inequality. There is no loss of generality in assuming that $x_1 \leqslant x_2 \leqslant x_3$ because if this is not the case, the arguments can be permuted to satisfy the restriction. This will leave the value of F invariant if also μ_i are permuted correspondingly, but the triangular inequality condition is invariant to permutations of μ_i. Thus from Lemma 2.9 we get that F is bounded by the restricted maximum:

$$F(x_1, x_2, x_3) \leqslant \frac{1}{(\frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3})^2}. \tag{36}$$

If, however, μ_i does not satisfy the triangular inequality, then, without loss of generality we can assume that

$$\frac{1}{\mu_1} \geqslant \frac{1}{\mu_2} + \frac{1}{\mu_3}. \tag{37}$$

Now on the set $\mathcal{E}_1$ we can apply Lemma 2.9 and find that the function is bounded by the unrestricted maximum

$$F(x_1, x_2, x_3) \leqslant \frac{\mu_1 \mu_2 \mu_3}{4(1 - \mu_1)}. \tag{38}$$

Now when $\frac{1}{\mu_1} \geqslant \frac{1}{\mu_2} + \frac{1}{\mu_3}$, we have $\mu_1 \leqslant \frac{\mu_2 \mu_3}{\mu_2 + \mu_3} \leqslant \min(\mu_2, \mu_3)$, so that $\mu_1 = m = \min(\mu_1, \mu_2, \mu_3)$ and

$$F(x_1, x_2, x_3) \leqslant \frac{\mu_1 \mu_2 \mu_3}{4(1 - \mu_1)} \leqslant \frac{\mu_1 \mu_2 \mu_3}{4(1 - m)}. \tag{39}$$

$\square$

3 The imbedding problem for 3-order transition matrix with positive eigenvalues or complex eigenvalues

Proposition 3.1. *Suppose* P *be an indecomposable* 3×3 *transition probability matrix with eigenvalues* $\{1, \lambda_1, \lambda_2\}$. *Then it satisfies that*

$$\mathsf{P}^2 - (\lambda_1 + \lambda_2)\mathsf{P} + \lambda_1\lambda_2\mathsf{I} = (\lambda_1 - 1)(\lambda_2 - 1)\mathsf{P}^\infty. \tag{40}$$

It is exactly the Eq.(1.15) of [13]. Johansen S. points out that the condition of the imbedding problem can be given in terms of P and P^∞ (or μ) in [13]. The following two propositions are the direct corollary of Proposition 1.2, Proposition 1.4 of [13] and Eq.(40).

Proposition 3.2. *Let* P *be an indecomposable* 3×3 *transition probability matrix with positive eigenvalues* $\{1, \lambda_1, \lambda_2\}$. P *can be imbedded if and only if*

$$p_{ij} \geqslant \mu_j \frac{(\lambda_2 - 1)\log\lambda_1 - (\lambda_1 - 1)\log\lambda_2}{\log\lambda_2 - \log\lambda_1}, \quad i \neq j. \tag{41}$$

If $\lambda_1 = \lambda_2 = \lambda$, *then the right hand side of (41) is in the sense of limit, i.e., fix the value of* λ_1, *and let* $\lambda_2 \to \lambda_1$, *one has that*

$$p_{ij} \geqslant \mu_j(\lambda\log\lambda - \lambda + 1), \quad i \neq j. \tag{42}$$

Proposition 3.3. *Let* P *be an indecomposable* 3×3 *transition probability matrix with complex eigenvalues* $\{1, \lambda_1, \lambda_2\}$, $\lambda_1 = re^{i\theta}, \lambda_2 = re^{-i\theta}$, $\theta \in (0, \pi)$. P *can be imbedded if and only if*

$$p_{ij} \geqslant \mu_j(1 - r\cos\theta + \frac{\sin\theta}{\theta}r\log r), \quad i \neq j. \tag{43}$$

or

$$p_{ij} \leqslant \mu_j(1 - r\cos\theta + \frac{\sin\theta}{2\pi - \theta}r\log r), \quad i \neq j. \tag{44}$$

The two propositions appear implicitly in [13] and seem neater than Eq.(1.11-1.13,1.17) of Reference [13].

Remark 6. If P is reversible, then $\frac{p_{ij}}{\mu_j} = \frac{p_{ji}}{\mu_i}$. One should only test half number of the inequalities.

4   Conclusion

In the present paper, we solve the imbedding problem for 3-order transition matrix with coinciding negative eigenvalues by means of an alternate parameterization of the transition rate matrix, which is different from the traditional way to calculate the matrix logarithm or the matrix square root.

Acknowledgements

References

[1] Bladt M., Sφrensen M., Statistical inference for discretely observed Markov jump processes, J.R.S.S (B) 67 (3), 395-410. (2005)

[2] Carette P., Characterizations of Embeddable 3×3 Stochastic Matrices with a Negative Eigenvalue, New York J. Math. 1, 120-129. (1995)

[3] Chen Y., On the monotonicity of fluctuation spectra for three-state Markov processes, Fluctuation and Noise Letters, 7 (3) L181-192, (2007)

[4] Chung K. L., Markov chains with stationary Transition Probabiities, Second Edition, Springer-Verlag, Berlin, 1976

[5] Elfving G., Zur Theorie der Markoffschen Ketten, Acta Soc. Sci. Finn. A, 2, 1-17. (1937)

[6] Frydman H., Singer B., Total positivity and the embedding problem for Markov chains, Math. Proc. Camb. Phil. Soc. 86, 339-344. (1979)

[7] Frydman H., The embedding problem for Markov chains with 3 states, The Mathematical Proceedings of the Cambridge Philosophical Society, 87, 285-294. (1980)

[8] Frydman H., A structure of the bang-bang representation for 3×3 embeddable matrices, Z. Wahrscheinlichkeitstheorie Verw. Geb., 53, 305-316. (1980)

[9] Fuglede B., On the Imbedding Problem for Stochastic and Doubly Stochastic Matrices, Probab. Th. Rel. Fields 80, 241-260. (1988)

[10] Goodman G., An intrinsic time for non-stationary finite Markov chains. Z. Wahrscheinlichkeitstheorie Verw. Geb., 16, 165-180. (1970)

[11] Israel R., Rosenthal J., Wei J., Finding Generators for Markov Chains via Empirical Transition Matrices, with Applications to Credit Ratings, Math. Finance 11, 245-265. (2001)

[12] Jiang D.-Q., Qian M., Qian M.-P., Mathematical Theory of Nonequilibrium Steady States, Springer, Berlin, 2004

[13] Johansen S., Some Results on the Imbedding Problem for Finite Markov Chains, J. Lond. Math. Soc., 8, 345-351. (1974)

[14] Johansen S., Ramsey F. L., A bang-bang representation for 3x3 embeddable stochastic matrices. Z. Wahrscheinlichkeitstheor. Verw. Geb. 47, 107-118. (1979)

[15] Kalpazidou S., Cycle representations of Markov processes. Second edition. Springer, New York, 2006

[16] Kingman J. F. C., The imbedding problem for finite Markov chains, Z. Wahrscheinlichkeitstheorie Verw. Geb. 1, 14-24. (1962)

[17] McCausland W. J., Time reversibility of stationary regular finite-state Markov chains, Journal of Econometrics 136, 303-318. (2007)

[18] Metzner P., Dittmer E., Jahnke T., Schütte C, Generator estimation of Markov jump processes, Journal of Computational Physics, Volume 227, Issue 1, 10 November 2007, 353-375. (2007)

[19] Qian M.-P., Qian M., The decomplsition into a detailed blanced part and a circulation part of an irreversible stationary Markov chain, Scientia Sinica, Special Issue(II), 69-79.(1979)

[20] Runnenberg J. T., "On Elfving's Problem of Imbedding a Time-discrete Markov Chain in a Continuous Time One for Finitely Many States." Proceedings, Koninklijke Nederlandse Akademie van Wetenschappen, ser. A, Mathematical Sciences, 65(5), 536-541. (1962)

[21] Singer B., Spilerman S., The Representation of Social Processes by Markov Models, Am. J. Sociol. 82, 1-54. (1976)