10 (1990), 2, 217—22

Acta Mathematica Scientia
数学物理学报

LARGE DEVIATIONS FOR MARKOV CHAINS*

Chen Mufa (陈木法) *Lu Yungang* (卢云刚)
Dept. of Math., Beijing Norm. Univ., Beijing, China

Abstract

Donsker–Varadhan's large deviation theory is applied to Markov chains with continuous parameter and countable state space. Since the lower estimate given by Donsker and Varadhan [5; part 4] is general enough in our context, we study only the upper estimate, and make the conditions for the estimate explicit.

§ 1. Introduction

Throughout this paper, we take $E=\{0, 1, 2, \cdots\}$ as our state space. Suppose that we are given a totally stable conservative Q-matrix $Q=(q_{ij}: i, j\in E)$ (i.e., $0\leqslant q_{ij}<\infty$, $i\neq j$, $\sum_{j\neq i} q_{ij}=q_i\equiv -q_{ii}<\infty$, $i\in E$), which determines uniquely a jump process (more precisely, its transition probability function) $P(t)=(P_{ij}(t): i, j\in E)$, $t\geqslant 0$. Such Q-matrix is called regular.

Various necessary and sufficient conditions on the Q-matrix for the above uniqueness assumption to hold have been obtained and can be found eleswhere (see Hou[6], Chen and Zheng [4], Yan and Chen [13]).

The problem we are interested is when the upper and lower estimates of large deviation hold for the $P(t)$ determined by the Q-matrix $Q=(q_{ij})$. The answer to the lower estimate given by Donsker and Varadhan [5; part 4] is very general. For the upper estimate, their hypotheses are described in terms of the infinitesimal generator of the Markov process. Unfortunately, the domain of an infinitesimal generator is usually hard to figure out exactly even for Markov chains, so the picture for the upper estimate is still not very clear yet.

Our discussions are divided in the following two parts:

Case (I) There are some absorbing states, i.e., $q_i=0$ for some $i\in E$;

Case (II) There are no absorbing states.

For the first case, we use the theory of the minimal nonnegative solutions and coupling technique; for the second, we use martingale approach. Either of the method is different to the traditional one. Our starting point is a result obtained by Stroock [9; Theorem (8.12)]. The main results of this paper are organized by Theorem 2.10, 2.11 and 3.1—3.6.

* Received Sep. 4, 1985. This work is supported in part by NSFC and the Foundation of Zhongshan University, Advanced Research Centre.

§ 2. Case (I)

Definition 2.1. We call

$$\begin{cases} x_i \geqslant \sum_k c_{ik} x_k, \\ x_i \geqslant d_i, \end{cases} \quad i \in E \tag{1}$$

a constrained system of homogenous nonnegative linear inequalities (abbr. a constrained system) if

$$0 \leqslant c_{ik} \leqslant +\infty, \quad 0 \leqslant d_i \leqslant +\infty, \quad i, k \in E \tag{2}$$

and

$$\sum_k c_{ik} d_k \geqslant d_i, \quad i \in E. \tag{3}$$

Definition 2.2. We call $\{x_i^*: i \in E\}$ (may be valued $+\infty$) the minimal (nonnegative) solution to the constrained system (1), if $\{x_i^*: i \in E\}$ is a solution to (1) and for any solution $\{x_i: i \in E\}$ of (1), we have

$$x_i^* \leqslant x_i, \quad i \in E.$$

By induction, it is elementary to prove the following three results (cf. Hou and Guo [7]):

Theorem 2.3. The minimal solution to the constrained system (1) exists uniquely. Indeed, it can be obtained in the following way: set

$$x_i^{(0)} = d_i, \quad i \in E,$$

$$x_i^{(n+1)} = \sum_k c_{ik} x_k^{(n)}, \quad i \in E, \ n \geqslant 0,$$

then

$$x_i^{(n)} \uparrow x_i^*, \quad i \in E.$$

Theorem 2.4. Let $\{\bar{x}_i: i \in E\}$ be a solution to

$$\begin{cases} y_i \geqslant \sum_k \bar{c}_{ik} y_k, \\ y_i \geqslant \bar{d}_i, \end{cases} \quad i \in E, \tag{4}$$

where $(\bar{d}, \bar{c})$ satisfies the conditions (2), (3) and

$$c_{ik} \leqslant \bar{c}_{ik}, \quad d_i \leqslant \bar{d}_i, \quad i, k \in E. \tag{5}$$

Then $\bar{x}_i \geqslant x_i^*$ for each $i \in E$.

Theorem 2.5. Let (c, d) and $(\bar{c}, \bar{d})$ satisfy (2), (3) and (5). Then each solution of (4) is a solution of (1). In particular, if (4) has a finite nonnegative (resp. non-decreasing) solution, then so does (1).

Now, we endow E with the discrete topology, then $P(t)$ is a Feller semigroup. Let $\{X(t): t \geqslant 0\}$ be the sample process defined on the usual space $(\Omega \equiv D([0, \infty), E), \mathscr{B}_\Omega, \boldsymbol{P})$ associated with the semigroup $P(t)$. Denote the successive jump times of $\{X(t)\}$ by $\tau_0 \equiv 0, \tau_1, \tau_2, \cdots$.

Theorem 2.6. Let Φ be a function on E satisfying

$$\Phi(i) < q_i, \quad \text{if} \quad q_i > 0; \ = 0, \quad \text{if} \quad q_i = 0 \tag{6}$$

and set

$$\varphi_i = \boldsymbol{E}_i\Big(\exp\Big[\int_0^\infty \Phi(X(s))ds\Big]\Big), \quad i \in E. \tag{7}$$

Then $\{\varphi_i: i \in E\}$ is the minimal solution to (1) with $d_i=1$, $c_{ii}=1$, $c_{ij}=0$ $(i \neq j)$ if $q_i=0$; and

$$d_i = \frac{q_i}{q_i - \Phi(i)},$$

$$c_{ij} = \begin{cases} 0, & i=j, \\ \dfrac{q_{ij}}{q_i - \Phi(i)}, & i \neq j, \end{cases}$$

if $q_i > 0$.

Proof. Clearly, (d, c) satisfies (2) and (3). Put

$$\varphi_i^{(n)} = \boldsymbol{E}_i\Big(\exp\Big[\int_0^{\tau_{n+1}} \Phi(X(s))ds\Big]\Big), \quad i \in E,\ n \geqslant 0.$$

By Theorem 2.3, it suffices to prove that

$$\varphi_i^{(0)} = d_i, \quad \varphi_i^{(n+1)} = \sum_k c_{ik}\varphi_k^{(n)}, \quad n \geqslant 0,\ i \in E. \tag{8}$$

If $q_i = 0$, then $\varphi_i^{(n)} \equiv 1$ $(n \geqslant 0)$ since $\Phi(i) = 0$, so (8) holds. We now assume $q_i > 0$. Then

$$\varphi_i^{(0)} = \boldsymbol{E}_i\Big(\exp\Big[\int_0^{\tau_1} \Phi(X(s))ds\Big]\Big) = \int_0^\infty q_i \exp[\Phi(i)t - q_i t]\,dt$$

$$= \frac{q_i}{q_i - \Phi(i)} = d_i,$$

from this, induction assumption and the strong Markov property, we have

$$\varphi_i^{(n+1)} = \boldsymbol{E}_i\Big(\boldsymbol{E}_i\Big(\exp\Big[\int_0^{\tau_{n+2}} \Phi(X(s))ds\Big] \Big| \mathscr{F}_{\tau_1}\Big)\Big)$$

$$= \boldsymbol{E}_i\Big(\exp\Big[\int_0^{\tau_1} \Phi(X(s))ds\Big] \boldsymbol{E}_{X(\tau_1)}\Big(\exp\Big[\int_0^{\tau_{n+1}} \Phi(X(s))ds\Big]\Big)\Big)$$

$$= \sum_{k\neq i} \frac{q_{ik}}{q_i} \varphi_k^{(n)} \boldsymbol{E}_i\Big(\exp\Big[\int_0^{\tau_1} \Phi(X(s))ds\Big]\Big)$$

$$= \sum_{k\neq i} \frac{q_{ik}}{q_i} \varphi_k^{(n)} \cdot \frac{q_i}{q_i - \Phi(i)} \varphi_k^{(n)}$$

$$= \sum_k c_{ik}\varphi_k^{(n)}.$$

Lemma 2.7. Let $Q = (q_{ij})$ be a birth-death matrix (i.e., $q_{ij} = 0$ if $|i-j| > 1$ and $q_i = \sum_{j\neq i} q_{ij}$) satisfying

$$q_0 = 0, \quad q_{i,i+1} \equiv b_i > 0, \quad q_{i,i-1} = \alpha b_i, \quad i \in E, \quad \lim_{i\to\infty} b_i = \infty,$$

where $\alpha > 1$. Then there exists a function Φ on E such that

$$\Phi(0) = 0,\ \Phi(i) < q_i,\ i \geqslant 1;\ \Phi(i) \to \infty \quad \text{as } i \to \infty,$$

and $\varphi_i (i \in E)$, defined by (7) with this Φ, is finite.

Proof. The regularity of the Q-matrix follows from $\alpha > 1$. For each $s \in (0, 1)$, define

$$\Phi_s(i) = \begin{cases} 0, & \text{if } i = 0, \\ s(1+\alpha)b_i, & \text{if } i \geqslant 1. \end{cases}$$

Then Φ_s satisfies (6) for each $s\in(0,1)$. By Theorem (2.6) and (2.5), it is enough to show that there exists an $s\in(0,1)$ such that the following constrained system:

$$\begin{cases} x_i=\dfrac{1}{(1-s)(1+\alpha)}[\alpha x_{i-1}+x_{i+1}] \\ x_0=1\equiv d_0, & i\geqslant 1, \\ x_i\geqslant d_i\equiv\dfrac{1}{1-s}, \end{cases} \tag{9}$$

has a finite nonnegative solution. To this end, set

$$\lambda(s)=\frac{1}{2}[(1-s)(1+\alpha)+\sqrt{(1-s)^2(1+\alpha)^2-4\alpha}].$$

Then, $\{\lambda(s)^i\colon i\in E\}$ is a solution to the difference equation:

$$(1-s)(1+\alpha)x_i=\alpha x_{i-1}+x_{i+1},\quad i\geqslant 1.$$

Since $\lambda(0)=\alpha>1$,

$$\left(\lambda(s)-\frac{1}{1-s}\right)\Big|_{s=0}=\alpha-1>0,$$

there exists an $s_0\in(0,1)$ such that

$$\lambda(s)>\frac{1}{1-s}>1$$

for all $s\in(0,s_0)$. Thus, $\{x_i=\lambda(s_0/2)^i\colon i\in E\}$ gives the required solution

Lemma 2.8. Let $Q=(q_{ij})$ be a single birth matrix (i.e., $q_{ij}=0$ if $j>i+1$, $q_i=\sum_{j\neq i}q_{ij}$, $i, j\in E$) satisfying

$$q_0=0,\quad q_{i,i+1}\equiv b_i>0,\quad \sum_{j<i}q_{ij}=\alpha b_i,\quad i\in E,\ \alpha>1,\quad \lim_{i\to\infty}b_i=\infty.$$

Then the assertions of Lemma 2.7 hold.

Proof. From Yan and Chen[12] or Chen[3], it follows that the Q-matrix is regular.

For the remainders, simply observe that the solution $\{\lambda(s)^i\colon i\in E\}$ constructed in the last lemma also satisfies

$$\lambda(s)^0=1=d_0,$$

$$\lambda(s)^i\geqslant d_i=\frac{q_i}{q_i-\Phi_s(i)}=\frac{1}{1-s},\quad i\geqslant 1$$

and

$$\begin{aligned}\lambda(s)^i&=\frac{1}{(1-s)(1+\alpha)}[\alpha\lambda(s)^{i-1}+\lambda(s)^{i+1}]\\ &=\sum_{k<i}\frac{q_{ik}}{(1-s)(1+\alpha)b_i}\lambda(s)^{i-1}+\frac{1}{(1-s)(1+\alpha)}\lambda(s)^{i+1}\\ &\geqslant\sum_{k<i}\frac{q_{ik}}{b_i(1-s)(1+\alpha)}\lambda(s)^k+\frac{1}{(1-s)(1+\alpha)}\lambda(s)^{i+1},\quad i\geqslant 1.\end{aligned}$$

Similarly, one may prove

Corollary 2.9. Let $Q=(q_{ij})$ be a single birth Q-matrix with some but finite number of absorbing states. Suppose that there exists an $\alpha>1$ such that

$$q_{i,i-1}=\alpha b_i\equiv\alpha q_{i,i+1},\quad i>i_0,\quad \lim_{i\to\infty}b_i=\infty,$$

where $i_0=\max\{i: q_i=0\}$, then the conclusions of Lemma 2.7 hold.

Theorem 2.10. Let $Q=(q_{ij})$ be a single birth Q-matrix (cf. Lemma 2.8). Suppose that there exists an $\alpha>1$ such that

$$\sum_{j<i} q_{ij}\geqslant \alpha q_{i,i+1},\quad i\in E;\quad \max\{i: q_i=0\}<\infty;\quad \sum_{j<i} q_{ij}\to\infty,\quad \text{as } i\to\infty.$$

Then there exists a function Φ on E such that

$$\Phi(i)\uparrow\infty \quad \text{as } i\uparrow\infty$$

and

$$\varphi_i\equiv \boldsymbol{E}_i\left(\exp\left[\int_0^\infty \Phi(X(s))ds\right]\right)<\infty,\quad i\in E.$$

Proof. Let $Q=(\bar{q}_{ij})$ be a birth-death matrix, defined by

$$\bar{q}_{i,i-1}=\sum_{j<i} q_{ij},\quad i\geqslant 1;$$

$$\bar{q}_{i,i+1}=\frac{1}{\alpha}\sum_{j<i} q_{ij},\quad i\in E.$$

By Corollary 2.9, there exists a function $\bar{\Phi}$ on E having the properties mentioned in the theorem (if necessary, using $\tilde{\Phi}(i)\equiv\inf_{k\geqslant i}\bar{\Phi}(k)$ $(i\in E)$ instead of $\bar{\Phi}$, we may assume that $\bar{\Phi}(i)\uparrow\infty$ as $i\uparrow\infty$). Now, the coupling for jump processes (cf. [3]) gives us

$$P^{(i_1,i_2)}[X(t)\leqslant \bar{X}(t)]=1,\quad i_1\leqslant i_2$$

where $\{X(t), t\geqslant 0\}$ and $\{\bar{X}(t), t\geqslant 0\}$ are determined by $Q=(q_{ij})$ and $Q=(\bar{q}_{ij})$ respectively, and $P^{(i_1,i_2)}$ is the probability measure on the common probability space. Indeed, we can take the coupling generator as follows:

$$\begin{aligned}\tilde{\Omega}f(i_1, i_2)=&\sum_{k>0}(q_{i_1,i_1-k}-\bar{q}_{i_2,i_2-k})^+[f(i_1-1, i_2)-f(i_1, i_2)]\\&+\sum_{k>0}(\bar{q}_{i_2,i_2-k}-q_{i_1,i_1-k})^+[f(i_1, i_2-k)-f(i_1, i_2)]\\&+\sum_{k>0}q_{i_1,i_1+1}\wedge \bar{q}_{i_2,i_2+1}[f(i_1-1, i_2-k)-f(i_1, i_2)]\\&+(q_{i_1,i_1+1}-\bar{q}_{i_2,i_2+1})^+[f(i_1+1, i_2)-f(i_1, i_2)]\\&+(\bar{q}_{i_2,i_2+1}-q_{i_1,i_1+1})^+[f(i_1, i_2+1)-f(i_1, i_2)]\\&+q_{i_1,i_1+1}\wedge \bar{q}_{i_2,i_2+1}[f(i_1+1, i_2+1)-f(i_1, i_2)]\end{aligned}$$

here we use the convention: $q_{ij}=0$ if $j<0$. Thus, if we take $\Phi=\bar{\Phi}$, then

$$\begin{aligned}\varphi_i=&\boldsymbol{E}_i\left(\exp\left[\int_0^\infty \Phi(X(s))ds\right]\right)=\tilde{\boldsymbol{E}}^{(i,i)}\left(\exp\left[\int_0^\infty \Phi(X(s))ds\right]\right)\\ \leqslant&\tilde{\boldsymbol{E}}^{(i,i)}\left(\exp\left[\int_0^\infty \Phi(\bar{X}(s))ds\right]\right)=\boldsymbol{E}_i\left(\exp\left[\int_0^\infty \Phi(\bar{X}(s))ds\right]\right)\\ =&\bar{\varphi}_i<\infty,\quad i\in E.\end{aligned}$$

This finishes our proof.

Let $\mathfrak{M}_1(E)$ be the space of all probability measures on E. We shall regard $\mathfrak{M}_1(E)$ as a Polish space endowed with the weak topology.

The following result is now a straightforward consequence of Theorem 2.10 and [9; Theorem (8.12)]:

Theorem 2.11. Everything is the same as in Theorem 2.10, then, for every closed set C in $\mathfrak{M}_1(E)$, we have

$$\overline{\lim_{t\to\infty}} \log Q_{t,i}(C) \leqslant -\inf_{\mu\in C} I(\mu), \tag{10}$$

where $Q_{t,i}=P_i\circ L_t^{-1}$,

$$L_t(\cdot,\ A)=\frac{1}{t}\int_0^t \chi_A(X(s))ds,$$

$$I(\mu)=-\inf_{u\in D^+(L)}\int\frac{Lu}{u}d\mu,$$

and L is the strong infinitesimal operator of the process, its domain is $D(L)$. $D^+(L)$ denotes the set of all uniformly positive functions in $D(L)$.

Example 2.12. Let us consider the birth-death matrix:

$$q_{i,i+1}=b_i=\lambda_1 i,\quad q_{i,i-1}=a_i=\lambda_2 i,\quad i\geqslant 0.$$

It is known that the Q-matrix is always regular, and is ergodic iff $\lambda_1\leqslant\lambda_2$. Theorem 2.11 works for the case $\lambda_1<\lambda_2$.

§ 3. Case (II)

In this section, we use martingale approach to study the upper estimate of large deviation for Markov chains. It is well known that

$$f(X(t))-\int_0^t (Lf)(X(s))ds,\quad f\in D(L) \tag{11}$$

and

$$f(X(t))\exp\left[-\int_0^t\left(\frac{Lf}{f}\right)(X(s))ds\right],\quad f\in D^+(L) \tag{12}$$

are all $(\Omega,\{\mathscr{F}_t\}_{t\geqslant 0},\boldsymbol{P}_i)$-martingales for quite general Markov process $\{X(t): t\geqslant 0\}$ with strong or weak infinitesimal generator L (cf., Stroock and Varadhan [10]). In our case, we are given Q-matrix only, the generator is hard to handle. Our goal is to introduce some martingales associated with the successive jump times $\tau_0\equiv 0$, τ_1, τ_2, $\cdots$ of the jump process $\{X(t): t\geqslant 0\}$.

Again, we assume that the given Q-matrix is regular. In the case (II), $q_i>0$ for each $i\in E$. Hence

$$\tau_0<\tau_1<\cdots<\tau_\infty\equiv\lim_{n\to\infty}\tau_n \tag{13}$$

and

$$\boldsymbol{P}[\tau_n<\infty]=1,\quad n\geqslant 0;\quad \boldsymbol{P}[\tau_\infty=\infty]=1 \tag{14}$$

(cf., Wang [12; p.245, Theorem 1]).

Set

$$D(Q)=\{f:\ Qf\in C_b(E)\},$$

$$C_b^+(E)=\{f:\ \exists\ m \text{ and } M \text{ such that } 0<m\leqslant f\leqslant M<\infty\}.$$

Theorem 3.1. For each $f\in D(Q)$, define

$$Y_n=f(X(\tau_n))-\int_0^{\tau_n}(Qf)(X(t))dt. \tag{15}$$

Then, for each $i\in E$, $\{Y_n\}_0^\infty$ is a $(\Omega, \{\mathscr{F}_{\tau_n}\}_{n\geq 0}, \boldsymbol{P}_i)$-martingale. In particular,

$$\boldsymbol{E}_i[Y_n]=f_i,\quad i\in E,\ n\geqslant 0. \tag{16}$$

Proof. Fix $f\in D(Q)$, $n\geqslant 1$ and $A\in\mathscr{F}_{\tau_n}$. Since

$$\begin{aligned}\infty>&\boldsymbol{E}_i[|f(X(\tau_{n+1}))-f(X(\tau_n))|,\ A]\\ &=\sum_j \boldsymbol{E}_i[I_{[X(\tau_n)=j]A}|f(X(\tau_{n+1}))-f(X(\tau_n))|]\\ &=\sum_j \boldsymbol{E}_i[I_{[X(\tau_n)=j]A}\boldsymbol{E}_{X(\tau_n)}(|f(X(\tau_1))-f(X(0))|)]\\ &=\sum_j \boldsymbol{E}_i\Big[I_{[X(\tau_n)=j]A}\Big|\sum_{k\neq j}\frac{q_{jk}}{q_j}f_k-f_j\Big|\Big]\\ &=\sum_j \boldsymbol{E}_i\Big[I_{[X(\tau_n)=j]A}\Big|\frac{(Qf)_j}{q_j}\Big|\Big]\end{aligned}$$

and so

$$\begin{aligned}&\boldsymbol{E}_i\Big[\int_{\tau_n}^{\tau_{n+1}}|(Qf)(X(t))|dt,\ A\Big]\\ &=\sum_j \boldsymbol{E}_i\Big[I_{[X(\tau_n)=j]A}\int_{\tau_n}^{\tau_{n+1}}|(Qf)(X(t))|dt\Big]\\ &=\sum_j \boldsymbol{E}_i\Big[I_{[X(\tau_n)=j]A}\boldsymbol{E}_{X(\tau_n)}\Big(\int_0^{\tau_1}|(Qf)(X(t))|dt\Big)\Big]\\ &=\sum_j \boldsymbol{E}_i[I_{[X(\tau_n)=j]A}|(Qf)_j|\boldsymbol{E}_j(\tau_1)]\\ &=\sum_j \boldsymbol{E}_i\Big[I_{[X(\tau_n)=j]A}\Big|\frac{(Qf)_j}{q_j}\Big|\Big]<\infty.\end{aligned}$$

These facts not only show that each Y_n is integrable, but also show that $\{Y_n\}_0^\infty$ is a $(\Omega, \{\mathscr{F}_{\tau_n}\}_{n\geq 0}, \boldsymbol{P}_i)$-martingale.

Theorem 3.2. For each $t\geqslant 0$ and $f\in D(Q)$, define

$$Y_n(t)=f(X(\tau_n\wedge t))-\int_0^{\tau_n\wedge t}(Qf)(X(s))ds,\quad n\geqslant 0. \tag{17}$$

Then, for each $i\in E$, $[Y_n(t)\}_{n=0}^\infty$ is a $(\Omega,\{\mathscr{F}_{\tau_n\wedge t}\}_{n\geq 0}, \boldsymbol{P}_i)$-martingale. In particular,

$$\boldsymbol{E}_i[Y_n(t)]=f_i,\quad i\in E,\ n\geqslant 0,\ t\geqslant 0. \tag{18}$$

Proof. Use induction. First,

$$\boldsymbol{E}_i[Y_0(t)]=f_i,\quad i\in E,\ t\geqslant 0.$$

Suppose that

$$\boldsymbol{E}_i[Y_n(t)]=f_i,\quad i\in E,\ t\geqslant 0$$

for some n, then

$$\begin{aligned}&\boldsymbol{E}_i[Y_{n+1}(t)]\\ &=\boldsymbol{E}_i\Big[f(X(\tau_{n+1}\wedge t))-\int_0^{\tau_{n+1}\wedge t}(Qf(X(s))ds\\ &=\boldsymbol{E}_i\Big[I_{[\tau_1>t]}\Big[f(X(\tau_{n+1}\wedge t))-\int_0^{\tau_{n+1}\wedge t}(Qf)(X(s))ds\Big]\Big]\\ &\quad+\boldsymbol{E}_i\Big[I_{[\tau_1<t]}\Big[f(X(\tau_{n+1}\wedge t))-\int_0^{\tau_{n+1}\wedge t}(Qf(X(s))ds\Big]\Big]\end{aligned}$$

$$=(f_i-(Qf)_it)e^{-q_it}-\int_0^t q_ie^{-q_is}(Qf)_is\,ds$$
$$+\boldsymbol{E}_i\Big[I_{[\tau_1<t]}\boldsymbol{E}_{X(\tau_1)}\Big[f(X(\tau_n\wedge(t-\tau_1(\cdot))))-\int_0^{\tau_n\wedge(t-\tau_1(\cdot))}(Qf)(X(s))ds\Big]\Big]$$
$$=(f_i-(Qf)_it)e^{-q_it}+(Qf)_i[e^{-q_is}s|_0^t-e^{-q_is}ds]$$
$$+\sum_{j\neq i}\frac{q_{ij}}{q_i}\int_0^t q_ie^{-q_is}\boldsymbol{E}_j\Big[f(X(\tau_n\wedge(t-s)))-\int_0^{\tau_n(t\wedge s)}(Qf)(X(\sigma))d\sigma\Big]ds$$
$$=[f_i-(Qf)_it]e^{-q_it}+(Qf)_ite^{-q_it}+\frac{(Qf)_i}{q_i}e^{-q_it}-\frac{(Qf)_i}{q_i}+\sum_{j\neq i}\frac{q_{ij}}{q_i}\int_0^t q_ie^{-q_is}f_j\,ds$$
$$=f_i,\quad i\in E,\ t\geqslant 0.$$

Hence (18) holds. It is now easy to check the martingale property.

Similarly, one can prove the following two results:

Theorem 3.3. For each $f\in C_b^+(E)$, set

$$Z_n=f(X(\tau_n))\exp\Big[-\int_0^{\tau_n}\Big(\frac{Qf}{f}\Big)(X(t))\,dt\Big],\quad n\geqslant 0.$$

Then, for each $i\in E$, $\{Z_n\}_0^\infty$ is a $(\Omega,\{\mathscr{F}_{\tau_n}\}_{n\geqslant 0},\boldsymbol{P}_i)$-martingale. In particular,

$$\boldsymbol{E}_i[Z_n]=f_i,\quad i\in E,\ n\geqslant 0.$$

Theorem 3.4. For each $t\geqslant 0$ and $f\in C_b^+(E)$, put

$$Z_n(t)=f(X(\tau_n\wedge t))\exp\Big[-\int_0^{\tau_n\wedge t}\Big(\frac{Qf}{f}\Big)(X(s))\,ds\Big],\quad n\geqslant 0.$$

Then, for each $i\in E$, $\{Z_n(t)\}_{n=0}^\infty$ is a $(\Omega,\{\mathscr{F}_{\tau_n\wedge t}\}_{n\geqslant 0},\boldsymbol{P}_i)$-martingale. In particular,

$$\boldsymbol{E}_i[Z_n(t)]=f_i,\quad i\in E,\ n\geqslant 0,\ t\geqslant 0.$$

Up to now, we have presented four results on martingale approach for Markov chains. The results may have their own interesting, even though we need only the last one which can be proved directly.

To state our main result in this section, we use the notations given in § 2.

Theorem 3.5. Let $Q=(q_{ij})$ be a regular Q-matrix with no absorbing states. Suppose that there exists a sequence $\{f_k\}_1^\infty\subset C_b^+(E)$ satisfying

$$f_k\geqslant c>0,\quad k\geqslant 1;\quad \sup_k f_k(i)<\infty,\quad i\in E \tag{19}$$

such that, for each $i\in E$,

$$-\Big(\frac{Qf_k}{f_k}\Big)_i\to\Phi(i),\quad \text{as } k\to\infty, \tag{20}$$

where Φ is a function on E, $\lim_{i\to\infty}\Phi(i)=+\infty$. Then, for each closed set C in $\mathfrak{M}_1(E)$, we have

$$\overline{\lim_{t\to\infty}}\frac{1}{t}\log Q_{t,i}(C)\leqslant-\inf_{\mu\in C}I(\mu).$$

Proof. By Theorem 3.4, we have

$$\boldsymbol{E}_i\Big[f_k(X(\tau_n\wedge t))\exp\Big[-\int_0^{\tau_n\wedge t}\frac{Qf_k}{f_k}(X(s))\,ds\Big]\Big]=f_k(i),\quad i\in E,\ k\geqslant 1,$$

and so

$$E_i\left[\exp\left[-\int_0^{\tau_n\wedge t}\frac{Qf_k}{f_k}(X(s))\,ds\right]\right]\leqslant\frac{1}{c}\sup_l f_l(i),\quad i\in E,\ k\geqslant 1.$$

Letting $k\to\infty$, then $n\to\infty$, and using Fatou lemma twice, one sees that

$$E_i\left[\exp\left[\int_0^{\tau_n\wedge t}\Phi(X(s))ds\right]\right]\leqslant\frac{1}{c}\sup_k f_k(i)$$

and that

$$E_i\left[\exp\left[\int_0^t\Phi(X(s))ds\right]\right]\leqslant\frac{1}{c}\sup_k f_k(i).$$

Next, we modify the function Φ so that Φ is non-decreasing. To this end, let

$$-M=\inf_{i\in E}\Phi(i)$$

$$\Psi_i=\inf_{k\geqslant i}\Phi(k),\quad i\in E$$

$$\tilde{\Phi}(i)=\Psi_i+M,\quad i\in E$$

then

$$-M\leqslant\Psi_i\uparrow\infty,\quad \text{as } i\uparrow\infty$$

and so

$$0\leqslant\tilde{\Phi}(i)\uparrow\infty,\quad \text{as } i\uparrow\infty.$$

Moreover,

$$\begin{aligned}
&E_i\left(\exp\left[\int_0^t\tilde{\Phi}(X(s))ds\right]\right)\\
&=e^{tM}E_i\left(\exp\left[\int_0^t\Psi(X(s))ds\right]\right)\\
&\leqslant e^{tM}E_i\left(\exp\left[\int_0^t\Phi(X(s))ds\right]\right)<\infty.
\end{aligned}$$

Now the assertion follows from [9; Theorem (8.12)]

Corollary 3.6. Let $Q=(q_{ij})$ be a regular Q-matrix having properties:

(i) $q_i>0$, $i\in E$;

(ii) there is a positive integer n such that

$$q_{i,i+n+k}=0,\quad i\in E,\ k\geqslant 1;$$

(iii) there exist an $\alpha>n$ and an $N\in E$ such that

$$\sum_{j<i}q_{ij}\geqslant\alpha\sum_{j>i}q_{ij},\quad i\geqslant N;$$

(iv) $\sum_{j<i}q_{ij}\to\infty$, as $i\to\infty$.

Then the upper estimate of large deviation holds.

Proof. Take an $\bar{\alpha}\in\left(1,\left(\frac{\alpha}{n}\right)^{\frac{1}{n}}\right)$ and define

$$f_k(i)=\begin{cases}\bar{\alpha}^i, & \text{if } i<N+k,\\ \bar{\alpha}^{N+k}, & \text{if } i\geqslant N+k,\end{cases}\quad i\in E,\ k\geqslant 1,$$

$$F_k(i)\equiv-\left(\frac{Qf_k}{f_k}\right)_i,\quad i\in E,\ k\geqslant 1.$$

Then $C_b(E)\ni f_k\geqslant 1$, $k\geqslant 1$; $\sup\limits_k f_k(i)<\infty$, $i\in E$ and

$$\begin{aligned}\Phi(i)&\equiv\lim_{k\to\infty}F_k(i)\\&=\sum_{j<i}q_{ij}(1-\bar{\alpha}^{j-i})+\sum_{i<j<i+n}q_{ij}(1-\bar{\alpha}^{j-i})\\&\geqslant\sum_{j<i}q_{ij}(1-(\bar{\alpha})^{-1})+\sum_{i<j<i+n}q_{ij}(1-\bar{\alpha}^{n})\\&\geqslant\sum_{j<i}q_{ij}\left[1-\frac{1}{\bar{\alpha}}+\frac{1}{\alpha}(1-\bar{\alpha}^{n})\right]\\&=(\sum_{j<i}q_{ij})(\bar{\alpha}-1)\left[\frac{1}{\bar{\alpha}}-\frac{\bar{\alpha}^{n-1}+\cdots+\bar{\alpha}+1}{\alpha}\right]\\&\geqslant(\sum_{j<i}q_{ij})\frac{\bar{\alpha}-1}{\alpha\bar{\alpha}}[\alpha-n\bar{\alpha}^{n}]\to\infty,\quad\text{as } i\to\infty.\end{aligned}$$

Therefore, Theorem 3.5 is now available.

Example 3.7. Schlögl model in one-dimension. Take

$$q_{i,i+1}=\lambda_1 a\binom{i}{2}+\lambda_4 b,\quad i\geqslant 0$$

$$q_{i,i-1}=\lambda_2\binom{i}{3}+\lambda_3 i,\quad i\geqslant 1$$

$$q_{ij}=0,\quad\text{other } i\neq j,$$

where a, b, λ_1, $\cdots$, $\lambda_4>0$. Then the assumptions of Corollary 3.6 hold.

To show that the condition (ii) in Corollary 3.6 is not necessary, we now present two results:

Corollary 3.8. Let $Q=(q_{ij})$ be a regular Q-matrix with no absorbing states. Suppose that there is a positive sequence (c_i) and an nonnegative sequence (d_i), such that

(i) $q_{ij}\leqslant c_i d_j$ for all large enough i and all $j>i$;

(ii) $\sum\limits_j \bar{c}_j d_j<\infty$;

(iii) $\sum\limits_{j<i}\left(1-\dfrac{\bar{c}_j}{\bar{c}_i}\right)q_{ij}\to\infty$, as $i\to\infty$,

where $\bar{c}_i=\max\{c_k:\ k\leqslant i\}$. Then the upper estimate of large deviation holds.

Proof. Since $\bar{c}_i\geqslant c_0>0$ and the condition (ii), we have

$$\sum_j d_j\leqslant\frac{1}{c_0}\sum_j\bar{c}_j d_j<\infty$$

and

$$\bar{c}_i\sum_{j>i}d_j\leqslant\sum_{j>i}\bar{c}_j d_j\to 0,$$

as $i\to\infty$. Set $f(i)=\bar{c}_i\left(1+\sum\limits_1^i d_k\right)$ and $f_k(i)=f(i\wedge k)$, then for every $k>i$,

$$-\left(\frac{Qf_k}{f_k}\right)_i\to\sum_{j<i}q_{ij}\left(1-\frac{f(j)}{f(i)}\right)+\sum_{j>i}q_{ij}\left(1-\frac{f(j)}{f(i)}\right)\equiv\Phi(i),$$

as $k\to\infty$. Because of f is increasing, the conclusion follows from Theorem 3.5 and

$$\Phi(i)\geqslant\sum_{j<i}q_{ij}\left(1-\frac{\bar{c}_j}{\bar{c}_i}\right)+\sum_{j>i}d_j(f(i)-f(j))\to\infty,\quad \text{as } i\to\infty.$$

Corollary 3.9. Let $Q=(q_{ij})$ be a regular Q-matrix with no absorbing states. Suppose that there is an $\alpha\in(0,\infty)$ and an $\bar{\alpha}\in(0,\alpha)$ such that

$$\left(\frac{\bar{\alpha}}{1+\bar{\alpha}}\right)^i\sum_{j<i}q_{ij}\to\infty,\quad \text{as } i\to\infty,$$

and $\sum_{j<i}q_{ij}\geqslant\alpha\sum_{j>i}q_{ij}$ for all large enough i. Then the upper estimate holds.

Proof. Choose $0<d<D<\infty$ and define

$$f(i)=D-\left(\frac{\bar{\alpha}}{1+\bar{\alpha}}\right)^i(D-d),\quad i\in E.$$

Then $f(0)=d$; $f(i)\uparrow D$, as $i\uparrow\infty$. Finally,

$$\begin{aligned}\Phi(i)&\equiv-\left(\frac{Qf}{f}\right)_i\\&=\sum_{j<i}q_{ij}\left(1-\frac{f(j)}{f(i)}\right)+\sum_{j>i}q_{ij}\left(1-\frac{f(j)}{f(i)}\right)\\&\geqslant f(i)^{-1}\Big[\sum_{j<i}q_{ij}(f(i)-f(j))+\sum_{j>i}q_{ij}(f(i)-D)\Big]\\&\geqslant f(i)^{-1}\left[\sum_{j<i}q_{ij}\left(f(i)-f(i-1)+\frac{f(i)-D}{\alpha}\right)\right]\\&=\frac{D-d}{f(i)}\left(\frac{\bar{\alpha}}{1+\bar{\alpha}}\right)^{i-1}\Big(\sum_{j<i}q_{ij}\Big)\left[1-\frac{\bar{\alpha}}{1+\bar{\alpha}}-\frac{1}{\alpha}\cdot\frac{\bar{\alpha}}{1+\bar{\alpha}}\right]\to\infty,\quad \text{as } i\to\infty.\end{aligned}$$

The next and final result enables us to return to Case (I):

Theorem 3.10. Let $Q=(q_{ij})$ be a regular Q-matrix with at most finite absorbing states, and satisfy the order-preserving condition (see [2] and [3]). Define

$$\tilde{q}_{i,i+1}=1,\quad \text{if } q_i=0;$$
$$\tilde{q}_{i,j}=0,\quad j\neq i,\ i+1,\ i\in E,$$
$$\tilde{q}_i=\sum_{j\neq i}\tilde{q}_{ij},\quad i\in E,$$

and put $\bar{Q}=Q+\tilde{Q}$. Then $\bar{Q}$ is regular. If $\bar{Q}$ satisfies the assumptions of Theorem 3.5, then the upper estimate for Q holds. In particular, Corollary 3.6 implies Theorem 2.10.

Proof. Once we have proved the first assertion, the second assertion follows easily from the basic couplings for Markov chains. Using the coupling in the proof of Theorem 2.10 and the second assertion, one can easily gets the last assertion. We now prove the first assertion. To this end, we have to check the equation:

$$\begin{cases}\bar{u}_i=\sum_{k\neq i}\dfrac{\bar{q}_{ik}}{\lambda+\bar{q}_i}\bar{u}_k,\\ 0\leqslant\bar{u}_i\leqslant1,\quad i\in E\end{cases}$$

has only zero solution. But this is a directly consequence of the following result:

Lemma 3.11. (Comparison Lemma). Let $\pi_i(x, dy)$ $(i=1, 2)$ be nonnegative measurable kernels on a measurable space $(E, \mathscr{E})$, and $g_i(i=1, 2)$ be nonnegative measurable functions so that

$$\pi_i(x, E) + g_i(x) \leqslant 1, \quad x \in E, \ i = 1, 2;$$
$$\pi_1(x, B) \geqslant \pi_2(x, B), \quad g_1(x) \geqslant g_2(x), \quad x \in E, \ B \in \mathscr{E}.$$

Denote by $\bar{f}_1$ the maximal solution to

$$\begin{cases} f_1 = \int \pi_1(\cdot, dy) f_1(y) + g_1, \\ 0 \leqslant f_1 \leqslant 1. \end{cases}$$

Then for every solution f_2:

$$\begin{cases} f_2 = \int \pi_2(\cdot, dy) + g_2, \\ 0 \leqslant f_2 \leqslant 1, \end{cases}$$

we have

$$f_2(x) \leqslant \bar{f}_1(x), \quad x \in E.$$

In particular, the above two equations have, or have no non-zero solutions simultaneously.

To prove this lemma, first note that if we set

$$f_1^{(1)} = g_1, \ f_1^{(n+1)} = \int \pi_1(\cdot, dy) f_1^{(n)}(y) + g_1, \quad n \geqslant 1,$$

then $f_1^{(n)} \downarrow \bar{f}_1$. Next use induction on n to show that

$$f_2(x) \leqslant f_1^{(n)}(x)$$

for each $x \in E$ and $n \geqslant 1$.

Acknowledgment. We would like to thank Professors L. Accardi, S. Olla, S. Orey, D. W. Stroock (from him the first author learnt the subject) and S. R. S. Varadhan for sending us their books and preprints, and thank Professor S. J. Yan for his helpful comments.

References

[1] Accardi, L., S. Olla, Donsker and Varadhan's theory for stationary process, preprint.
[2] Chen, M. F., Basic couplings for Markov chains, *Beijing Norm. Univ. J.*, **4** (1984), 3—10, in Chinese.
[3] Chen, M. F. Couplings for jump processes, *Acta Math. Sinica, New Ser.*, **2** (1986), 123—136.
[4] Chen, M. F., Zheng, X. G., Uniqueness criterion for q-processes, *Sci. Sinica*, **26** (1983), 11—24.
[5] Donsker, M., S. R. S., Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, *Comm. Pure Appl. Math.*, Pt. 1 **28** (1975), 1—47; Pt. 2, **29** (1976), 279—301; Pt. 3, **29** (1976), 389—461; Pt. 4, **36** (1983), 183—212.
[6] Hou, Z. T., Uniqueness criterion for Q-processes, *Sci. Sinica*, **15** (1974), 141—159.
[7] Hou, Z. T., Guo, Q. F., *Homogeneous Denumerable Markov Processes*, Science Press, Beijing, 1978, in Chinese.
[8] Orey, S., Large deviations and Shannon-McMillian theorems, preprint.
[9] Stroock, D. W., *An Introduction to the Theory of Large Deviations*, Springer-Verlag, Berlin, Heidelberg, New York, 1984.
[10] Stroock, D. W., S. R. S. Varadhan, *Multi-dimensional Diffusion Processes*, Grund. Ser. **223**, Springer-Verlag, Berlin, 1979.
[11] Varadhan, S. R. S., Large deviation and application, *SIAM CBMS-NSF Reg. Conf. in Appl. Math. Ser.* **46**, 1984.
[12] Wang, Z. K., *Stochastical Processes*, Science Press, Beijing, 1965, in Chinese.
[13] Yan, S. J., Chen, M. F., Multi-dimensional Q-processes, *Chin. Ann. of Math.*, **7B** (1986), 90—110.