Singular perturbation analysis for countable Markov chains

Liu Yuanyuan (刘源远)

School of Mathematics and Statistics,
Central South University, Changsha

The 12th Workshop on Markov Processes and Related Topics, Xuzhou, July 13-17, 2016, BNU and JSNU

中南大學 CENTRAL SOUTH UNIVERSITY

Outline

Perturbed DTMC

Let $\widetilde{X}_n$ be a time-homogeneous discrete-time Markov chain (DTMC) on a countable state space $\mathbb{E}$ with an irreducible stochastic transition matrix $\widetilde{P}$. Suppose that $\widetilde{P}$ is nearly decomposable, i.e.

$$\widetilde{P} = P + \Delta,$$

- the (small) perturbation matrix Δ satisfies that $\Delta \boldsymbol{e} = 0$.
- the unperturbed transition matrix P is decomposable (reducible), the state space $\mathbb{E}$ is decomposed into denumerable irreducible and ergodic classes $\mathbb{E}_n$ for $n \in \hat{\mathbb{E}}$, where the set $\hat{\mathbb{E}} := \{0, 1, \cdots, \ell\}$, $0 \leq \ell \leq \infty$. Thus $\mathbb{E} = \bigcup_{n \in \hat{\mathbb{E}}} \hat{\mathbb{E}} \times \mathbb{E}_n$.

The unperturbed transition matrix P can be written as

$$P = \begin{bmatrix} P_0 & 0 & 0 & \cdots \\ 0 & P_1 & 0 & \cdots \\ 0 & 0 & P_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

When $\ell = 0$, the perturbation is called regular perturbation; when $\ell \geq 1$, the perturbation is called singular perturbation.

Perturbed CTMC

Let $\widetilde{X}_t$ be a time-homogeneous CTMC on a countable space $\mathbb{E}$ with an irreducible, conservative and possibly unbounded generator $\widetilde{Q}$. Suppose that $\widetilde{Q}$ is nearly decomposable, i.e.

$$\widetilde{Q} = Q + \Delta,$$

where Δ is the small and Q is decomposable:

$$Q = \begin{pmatrix} Q_0 & 0 & 0 & \cdots \\ 0 & Q_1 & 0 & \cdots \\ 0 & 0 & Q_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Similarly, regular perturbation when $\ell = 0$, and singular perturbation when $\ell \geq 1$

For singular perturbation, the elements of P (Q) and Δ reflect the strong and weak interactions, respectively.

A wide range of large-scale systems are distinguished by this feature and hence are nicely modeled with the help of decomposable Markov chains, such as

▷ Markov decision processes, e.g., Bielecki and Stetlner (1998), Yin and Zhang (2000);

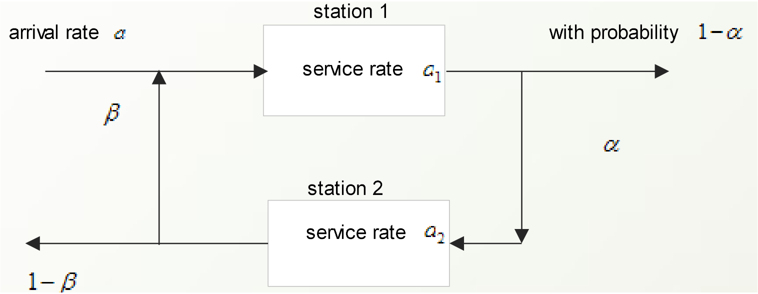
▷ control problems, e.g., Delebecque and Quadrat (1981), Yin and Zhang (2003);

▷ queueing networks, e.g., Latouche and Schweitzer (1995), Yin and Zhang (2008)

......

Example 1: a 2d discrete-time queue

Consider a DT queueing system with two stations:



which is modified from Yin and Zhang (2008) by letting that Station 1 has an unlimited room and the arrival customers follow a geometric distribution with a constant parameter.

Let $\widetilde{X}_j(n)$ be the number of customers at Station j at time n. Then $(\widetilde{X}_2(n), \widetilde{X}_1(n))$ is a two-dimensional DTMC on the state space $\mathbb{E} = \bigcup_{n\in\hat{\mathbb{E}}} \hat{\mathbb{E}} \times \mathbb{E}_n$, with $\hat{\mathbb{E}} = \mathbb{E}_n = \{0, 1, 2, \cdots\}$.

Its transition matrix $\widetilde{P}$ is given by $\widetilde{P} = P + \Delta$, where P is decomposable and $P_n, n \geq 1$ and Δ are determined by

$$P_0 = \begin{bmatrix} a & 1-a & 0 & \cdots \\ \mu_1 & \theta_0 & \lambda_1 & \cdots \\ 0 & \mu_1 & \theta_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad P_n = \begin{bmatrix} c & 1-c & 0 & \cdots \\ \mu_2 & \theta_1 & \lambda_2 & \cdots \\ 0 & \mu_2 & \theta_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

$$\Delta = \begin{bmatrix} \overline{\Delta}_0 & \overline{\Delta}_1 & 0 & \cdots \\ \Delta_{-1} & \Delta_0 & \Delta_1 & \cdots \\ 0 & \Delta_{-1} & \Delta_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \overline{\Delta}_0 = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 0 & -\widehat{\gamma}_{12} & 0 & \cdots \\ 0 & 0 & -\widehat{\gamma}_{12} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\overline{\Delta}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ a\widehat{\gamma}_{12} & (1-a)\widehat{\gamma}_{12} & 0 & 0 & \cdots \\ 0 & a\widehat{\gamma}_{12} & (1-a)\widehat{\gamma}_{12} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\Delta_{-1} = \begin{bmatrix} a\gamma_2 & \widehat{\delta}_0 & \overline{\delta}_0 & 0 & 0 & \cdots \\ a\gamma_1\gamma_2 & \delta_1 & \widehat{\delta}_1 & \overline{\delta}_1 & 0 & \cdots \\ 0 & a\gamma_1\gamma_2 & \delta_1 & \widehat{\delta}_1 & \overline{\delta}_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\Delta_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ aa_2\widehat{\gamma}_{12} & (1-a)a_2\widehat{\gamma}_{12} & 0 & 0 & \cdots \\ 0 & aa_2\widehat{\gamma}_{12} & (1-a)a_2\widehat{\gamma}_{12} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

and

$$\Delta_0 = \begin{bmatrix} -\overline{\gamma}_2 & 0 & 0 & 0 & \cdots \\ a\widehat{\gamma}_{12}\gamma_2 & \widehat{\theta}_1 & (1-a)\widehat{\gamma}_{12}\widehat{\gamma}_{21} & 0 & \cdots \\ 0 & a\widehat{\gamma}_{12}\gamma_2 & \widehat{\theta}_1 & (1-a)\widehat{\gamma}_{12}\widehat{\gamma}_{21} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Assume that parameters α and $(1 - a_2)$ are small$\Rightarrow$ Δ is small.

$\Rightarrow$ the changes of the queue length process and departure process corresponding to Station 2 are relatively slow compared with those corresponding to Station 1, which results in that the transition probability between any two levels is very small.

Example 2: a M M Birth-death process

Consider a Markov modulated state dependent birth-death process.

The environment process is a slowly varying M/M/1 queue, in which customers arrive at this queue according to a Poisson process with rate εq and are served at rate εp. Let $X(t)$ be the number of customers in the queue at time t.

Define a birth-death process $Y(t)$ as follows: if at time t, $X(t) = n$ and $Y(t) = i$, then $Y(t)$ jumps up to $i + 1$ at the birth rate $r_i b_n$ and jumps down to $i - 1$ at the death rate $s_i a_n$ for non-negative functions r_i and and s_i on $\{0, 1, 2, \cdots\}$.

The process $(X(t), Y(t))$ is a two-dimensional level dependent QBD process with infinitely many levels and phases.

Its generator is $\tilde{Q} = Q + \Delta$, where Q is decomposable with $Q_n, n \geq 0$ given by

$$Q_n = \begin{pmatrix} -r_0 b_n & r_0 b_n & 0 & \cdots \\ s_1 a_n & -r_1 b_n - s_1 a_n & r_1 b_n & \cdots \\ 0 & s_2 a_n & -r_2 b_n - s_2 a_n & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

and Δ has the same form as that of Example 1 with

$$\Delta_1 = \begin{pmatrix} \varepsilon q & 0 & 0 & \cdots \\ 0 & \varepsilon q & 0 & \cdots \\ 0 & 0 & \varepsilon q & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \ \Delta_{-1} = \begin{pmatrix} \varepsilon p & 0 & 0 & \cdots \\ 0 & \varepsilon p & 0 & \cdots \\ 0 & 0 & \varepsilon p & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $\overline{\Delta}_1 = \Delta_1$, $\overline{\Delta}_0 = -\overline{\Delta}_1$, $\Delta_0 = -(\Delta_{-1} + \Delta_1)$.

Our goal

Understand how perturbation affect the long-run behavior:

(i) Find conditions on Δ and P_i to ensure that $\widetilde{P}$ is positive recurrent with stationary distribution $\widetilde{\boldsymbol{\pi}}$, and

$$\widetilde{\boldsymbol{\pi}} = \pi^{(0)} \sum_{n=0}^{\infty} U^n,$$

for some matrix U and probability vector $\pi^{(0)}$.

(ii) When (i) is addressed well, we consider the bound on the difference $\widetilde{\boldsymbol{\pi}} - \pi^{(0)} = \pi^{(0)} \sum_{n=1}^{\infty} U^n$.

Literature

The singular perturbation problems are investigated by posing the Dobelin condition, see, e.g.

$\diamond$: Korolyuk and Turbin (1993), Bielecki and Stetlner (1998) for singular perturbation of Markov chains on general measurable state,

$\diamond$: Yin and Zhang (2002, 2008) for two-time scales singular perturbation.

The Dobelin condition is quite restrictive for DTMCs on an infinitely countable state space, see e.g. Hou and L (2004).

Altman et al. (2004): first adopt geometric ergodicity condition.

$\diamond$: the bounds are not well investigated,

$\diamond$: $\widetilde{P}$ and P_n are aperiodic,

$\diamond$: the perturbation Δ is linear (i.e. $\Delta = \varepsilon G$),

$\diamond$: the generator $\widetilde{Q}$ and Q are bounded.

We will investigate these issues by extending the ideas for regularly perturbed MCs, see e.g.

$\diamond$: Kartashov (1986), Mouhoubi and Aissani (2010), L (2012) for DTMCs,
$\diamond$: Mitrophanov (2006), Heidergott etal. (2010), L (2015) for CTMCs.

Outline

1. Background
2. Singular perturbation for DTMCs
3. Singular perturbation for CTMCs

Notations

Recall that each class P_n is positive recurrent, whose stationary probability vector is π_n and stationary probability matrix is Π_n.

Define the fundamental matrix by $R_n = (I - P_n + \Pi_n)^{-1}$.

Let

$$\mathbb{\Pi} = \begin{bmatrix} \Pi_0 & 0 & 0 & \cdots \\ 0 & \Pi_1 & 0 & \cdots \\ 0 & 0 & \Pi_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \mathbb{R} = \begin{bmatrix} R_0 & 0 & 0 & \cdots \\ 0 & R_1 & 0 & \cdots \\ 0 & 0 & R_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

To introduce the aggregated CTMC (e.g. Delebecque 1983), define the matrices M and W by

$$M = \begin{bmatrix} \pi_0 & 0 & 0 & \cdots \\ 0 & \pi_1 & 0 & \cdots \\ 0 & 0 & \pi_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{\hat{\mathbb{E}}\times\mathbb{E}}, \quad W = \begin{bmatrix} \mathbf{e}_0 & 0 & 0 & \cdots \\ 0 & \mathbf{e}_1 & 0 & \cdots \\ 0 & 0 & \mathbf{e}_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{\mathbb{E}\times\hat{\mathbb{E}}}.$$

The generator $\hat{Q}$ of the aggregated chain is defined by

$$\hat{Q} = M\Delta W, \quad (\hat{q}_{ij} = \pi_i \Delta_{ij} \mathbf{e}_j, \; i,j \in \hat{\mathbb{E}}),$$

whose stationary distribution and deviation matrix are denoted respectively by $\hat{\pi}$ and $D_{\hat{Q}}$ if they exist.

a key matrix

Introduce a key matrix U as follows

$$U = \Delta(\mathbb{R} - \mathbb{\Pi})(I + \Delta W D_{\hat{Q}} M).$$

Remark:

(i) When $\ell = 0$, $U = \Delta(R - \Pi)$, which plays a key role for regular perturbation analysis.

(ii) Compared with Altman et al. (2004), this definition permits that $\widetilde{P}$ and each class P_n are periodic and the perturbation Δ is general.

v-norm and Assumption 1

Let $\boldsymbol{v} \geq 1$. For a measure $\boldsymbol{\mu}$, its $\boldsymbol{v}$-norm is defined by $\|\boldsymbol{\mu}\|_{\boldsymbol{v}} = \sum_{i\in\mathbb{E}} |\mu_i| v_i$. For a matrix A, its $\boldsymbol{v}$-norm is defined by

$$\|A\|_{\boldsymbol{v}} = \sup_{i\in\mathbb{E}} v_i^{-1} \sum_{j\in\mathbb{E}} |A_{ij}| v_j.$$

Assumption 1: Let $\boldsymbol{v} \geq 1$. Assume that

(i) all the matrices P, Δ, Π, $\hat{\boldsymbol{\pi}} M$, $\mathbb{R}$ and $WD_{\hat{Q}}M$ have finite $\boldsymbol{v}$-norm; and

(ii) there exists a finite positive integer N and a constant $\delta_N \in (0,1)$ such that $\delta_N = \|U^N\|_{\boldsymbol{v}} < 1$.

Main results

Theorem 1: Suppose that Assumption 1 holds. Then
(i) $\widetilde{P}$ has a unique stationary distribution, given by

$$\tilde{\pi} = \hat{\pi} M \sum_{k=0}^{\infty} U^k.$$

(ii) For $N = 1$ (the case $N > 1$ can be obtained)

$$\left\| \widetilde{\pi} - \hat{\pi} M \sum_{k=0}^{m} U^k \right\|_{\boldsymbol{v}} \leq \|\hat{\pi} M\|_{\boldsymbol{v}} \frac{\|U\|_{\boldsymbol{v}}^{m+1}}{1 - \|U\|_{\boldsymbol{v}}}.$$

Remark:

▷ Note that for regular perturbation $\hat{\boldsymbol{\pi}} M = \pi$ $(\pi P = \pi)$, while for singular perturbation $\hat{\boldsymbol{\pi}} M = (\hat{\pi}_0 \boldsymbol{\pi}_1, \hat{\pi}_1 \boldsymbol{\pi}_1, \cdots)$.

▷ We call $\sum_{k=0}^{m} \hat{\boldsymbol{\pi}} M U^k$ the $(m{+}1)$-th order approximation of the stationary distribution $\widetilde{\boldsymbol{\pi}}$.

▷ For regular perturbation, the first approximation $\hat{\boldsymbol{\pi}} M$ (i.e., $m = 0$) becomes the stationary distribution of P, which has caused much concern in the literature.

▷ To prove this theorem, we need to make use of associativity for the multiplications of (possibly negative) infinite matrices. (i) of Assumption 1 can guarantee the associativity.

We now state Lyapunov drift conditions (see Chen (2003)), which are equivalent to geometric and exponential ergodicity, respectively.

D1($\boldsymbol{V}, \lambda, b, \{i\}$): For a transition matrix P, suppose that there exist a finite vector $\boldsymbol{V}$, $\boldsymbol{V} \geq \boldsymbol{e}$, some state i and positive numbers $b < \infty, \lambda < 1$ such that

$$P\boldsymbol{V} \leq \lambda \boldsymbol{V} + b\mathbb{1}_{\{i\}}.$$

D2($\boldsymbol{V}, \lambda, b, \{i\}$): For a q-matrix Q, suppose that there exist a finite function $\boldsymbol{V}$, $\boldsymbol{V} \geq \boldsymbol{e}$, some state i and positive constants $\lambda > 0, b < \infty$ such that

$$Q\boldsymbol{V} \leq -\lambda \boldsymbol{V} + b\mathbb{1}_{\{i\}}.$$

Assumption 2: Assume that
(i) **D1(**$V, \lambda, b, \{i_0(n)\}$**)** holds uniformly for each P_n; and
(ii) **D2(**$\hat{V}, \hat{\lambda}, \hat{b}, \{\hat{i}_0\}$**)** holds for the generator $\hat{Q}$.

Theorem 2: Suppose that Assumption 2 holds and $\|\Delta\|_{\boldsymbol{v}} < \infty$, where $v_{ni} := \hat{V}_n V_i$, $n \in \hat{\mathbb{E}}, i \in \mathbb{E}_n$. Then Assumption 1 holds and

$$\|U\|_{\boldsymbol{v}} \leq x(1 + y\|\Delta\|_{\boldsymbol{v}} + z),$$

where x, y, z are given by

$$x = \frac{c_1}{1-\lambda}, \quad y = \frac{c_2(1+c_3)^2}{\hat{\lambda}}, \quad z = c_2 c_3$$

with $c_1 := \|\Delta(I - \Pi)\|_{\boldsymbol{v}}$, $c_2 := \sup_{n \in \hat{\mathbb{E}}} \pi_n(V)$ and $c_3 := \hat{\pi}(\hat{V})$.
Note that $c_1 \leq (1 + c_2)\|\Delta\|_{\boldsymbol{v}}$ $c_2 \leq b/(1-\lambda)$, and $c_3 \leq \hat{b}/\hat{\lambda}$.

Back to Example 1

Assumption-1: Suppose that (i) $a_1 + a\alpha < a + aa_1\alpha$; and (ii) $a_2 + \alpha < 1 + aa_2\alpha$.

▷ Note that (i) of Assumption-1 is equivalent to $\rho := \frac{\lambda_1}{\mu_1} < 1$.

▷ Under (i) of Assumption-1, $\mathbf{D1}(\boldsymbol{V}, \lambda, b, \{i_0(n)\})$ holds uniformly in n for $i_0(n) = 0$, $\boldsymbol{V}_j = \left(\sqrt{\frac{1}{\rho}}\right)^j, j \geq 0$, $\lambda = 1 - \left(\sqrt{\lambda_1} - \sqrt{\mu_1}\right)^2$ and $b = a + (1-a)\sqrt{\frac{1}{\rho}} - \lambda_0$.

$$\pi_{n,0} = \frac{\frac{a_1}{\rho}}{\frac{a_1}{\rho} + \frac{1}{1-\rho}}, \quad \pi_{n,i} = \frac{\rho^{i-1}}{\frac{a_1}{\rho} + \frac{1}{1-\rho}}, \quad i \geq 1.$$

The generator of the aggregated CTMC $\hat{Q}$ is

$$\hat{Q} = \begin{bmatrix} d_0 & d_1 & 0 & \cdots \\ c_0 & c_1 & c_2 & \cdots \\ 0 & c_0 & c_1 & \cdots \\ \vdots & \vdots & \ddots & \cdots \end{bmatrix},$$

where $d_i = \boldsymbol{\pi}_0 \Delta_{0i} \mathbf{e}$, and $c_i = \boldsymbol{\pi}_0 \Delta_i \mathbf{e}$.

▷ Note that (ii) of Assumption-1 is equivalent to $\hat{\rho} := \frac{c_2}{c_0} < 1$.

▷ When $\hat{\rho} < 1$, $\mathbf{D2}(\hat{\boldsymbol{V}}, \hat{\lambda}, \hat{b}, \{\hat{i}_0\})$ holds for $\hat{i}_0 = 0$, $\hat{V}_j = \left(\sqrt{\frac{1}{\hat{\rho}}}\right)^j, j \geq 0, \hat{\lambda} = \left(\sqrt{c_2} - \sqrt{c_0}\right)^2, \hat{b} = d_0 + d_1\sqrt{\frac{1}{\hat{\rho}}} + \overline{\lambda}_d$.

Define the drift vector $\boldsymbol{v}$ for the whole space by

$$\boldsymbol{v}_{ni} = \hat{V}_n V_i = \left(\sqrt{\frac{1}{\hat{\rho}}}\right)^n \left(\sqrt{\frac{1}{\rho}}\right)^i.$$

Assumption-2: Both parameters α and $(1 - a_2)$ are small enough such that $x\,(1 + y\|\Delta\|_v + z) < 1$.

Under the above two assumptions, by Theorem 2, we have

$$\left\| \widetilde{\pi} - \hat{\pi} M \sum_{k=0}^{m} U^k \right\|_{\boldsymbol{v}} \leq \frac{\left[x\,(1 + y\|\Delta\|_{\boldsymbol{v}} + z)\right]^{m+1}}{1 - x\,(1 + y\|\Delta\|_{\boldsymbol{v}} + z)} \|\hat{\pi} M\|_{\boldsymbol{v}},$$

Note that $\|\Delta\|_{\boldsymbol{v}}$, $\|\hat{\pi} M\|_{\boldsymbol{v}}$, x, y, and z can be determined by the model parameters.

Outline

1. Background
2. Singular perturbation for DTMCs
3. Singular perturbation for CTMCs

Singular perturbation for CTMCs

Let π_n and D_n respectively be the stationary distribution and the deviation matrix corresponding to Q_n whenever they exist.

Let

$$D_Q = \begin{bmatrix} D_0 & 0 & 0 & \cdots \\ 0 & D_1 & 0 & \cdots \\ 0 & 0 & D_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The generator $\hat{Q}$ of the aggregated CTMC is defined by

$$\hat{Q} = M\Delta W.$$

Let $\hat{\pi}$ and $D_{\hat{Q}}$ respectively be the stationary distribution and the deviation matrix of $\hat{Q}$ whenever they exist.

To establish the counterpart of the discrete-time perturbation results, define

$$U = \Delta D_Q (I + \Delta W D_{\hat{Q}} M).$$

Assumption 1': Let $\boldsymbol{v} \geq 1$ be a vector on $\mathbb{E}$.
(i) All the matrices Q, Δ, Π, $\hat{\pi} M$, D_Q and $W D_{\hat{Q}} M$ have finite $\boldsymbol{v}$-norm;
(ii) there exists a finite number N and a constant $\delta_N \in (0, 1)$ such that $\delta_N = \| U^N \|_{\boldsymbol{v}} < 1$.

Theorem 1' Suppose that Assumption 1' holds. Then the same results as that in Theorem 1 hold.

Assumption 2': Assume that
(i) **D2**($\boldsymbol{V}, \lambda, b, \{i_0(n)\}$) holds uniformly in n for the same $\boldsymbol{V}, \lambda$ and b; and
(ii) **D2**($\hat{\boldsymbol{V}}, \hat{\lambda}, \hat{b}, \{\hat{i}_0\}$) holds for $\hat{Q}$.

Theorem 2' Suppose that Assumption 2' and $\|\Delta\|_{\boldsymbol{v}} < \infty$, where $v_{ni} := \hat{V}_n V_i$ for $n \in \hat{\mathbb{E}}, i \in \mathbb{E}_n$. Then (i) of Assumption 1' holds and

$$\|U\|_{\boldsymbol{v}} \le x \left[1 + \frac{(1 + \hat{\pi}(\hat{\boldsymbol{V}}))^2}{\hat{\lambda}} \sup_{n \in \hat{\mathbb{E}}} \boldsymbol{\pi}_n(\boldsymbol{V}) \|\Delta\|_{\boldsymbol{v}}\right],$$

where the value of x is given by

$$x = \|\Delta D_Q\|_{\boldsymbol{v}} \le \frac{(1 + \sup_{n \in \hat{\mathbb{E}}} \boldsymbol{\pi}_n(\boldsymbol{V}))^2}{\lambda} \|\Delta\|_{\boldsymbol{v}}.$$

To prove the above result, one key point is to use the following result

Theorem [L 2015] For a CTMC with a (possibly unbound) Q-matrix. Suppose that $\mathbf{D2}(\boldsymbol{V}, \lambda, b, \{i\})$ holds. Then we have

$$\|D\|_{\boldsymbol{v}} \le \frac{(1+\boldsymbol{\pi}(\boldsymbol{V}))^2}{\lambda} \le \frac{(\lambda+b)^2}{\lambda^3}.$$

Back to Example 2

Assumption: Assume that (i) $\inf_{n\geq 0}(a_n - b_n) > 0$, $\sup_{n\geq 0} b_n < \infty$; and (ii) $p < q$.

(i) of this assumption implies that $\mathbf{D2}(V, \lambda, b, \{i_0(n)\})$ holds uniformly in n for $i_0(n) = 0$, $V_j = j+1, j \geq 0$, $\lambda = \frac{\inf_{n\geq 0}(a_n - b_n)}{2}$, and $b = \sup_{n\geq 0} b_n$.

The stationary distribution $\{\boldsymbol{\pi}_n, n \in \hat{\mathbb{E}}\}$, dependent on n, is given by

$$\pi_{n,0} = \frac{1}{1 - \log(1 - b_n/a_n)}, \quad \pi_{n,k} = \frac{1}{k}\left(\frac{b_n}{a_n}\right)^k \pi_{n,0}, \quad k \geq 1.$$

The generator $\hat{Q}$ of the aggregated CTMC is given by

$$\hat{Q} = \begin{pmatrix} -\varepsilon q & \varepsilon q & 0 & \cdots \\ \varepsilon p & -\varepsilon(p+q) & \varepsilon q & \cdots \\ 0 & \varepsilon p & -\varepsilon(p+q) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

(ii) of the assumption implies that $\hat{\rho} := \frac{q}{p} < 1$, and **D2**$(\hat{\boldsymbol{V}}, \hat{\lambda}, \hat{b}, \{\hat{i}_0\})$ holds for $\hat{i}_0 = 0$, $\hat{V}_j = \left(\sqrt{\frac{1}{\hat{\rho}}}\right)^j, j \geq 0$, $\hat{\lambda} = \varepsilon\left(\sqrt{p} - \sqrt{q}\right)^2$, $\hat{b} = \varepsilon(p - \sqrt{pq})$. Furthermore, we have

$$\hat{\pi}_0 = 1 - \hat{\rho}, \quad \hat{\pi}_n = (1-\hat{\rho})\hat{\rho}^{n-1}, \quad n \geq 1, \quad \hat{\boldsymbol{\pi}}(\hat{\boldsymbol{V}}) = 1 + \sqrt{\hat{\rho}}.$$

Define the drift vector $\boldsymbol{v}$ for the whole space by

$$\boldsymbol{v}_{ni} = \hat{V}_n V_i = \left(\sqrt{\frac{1}{\hat{\rho}}}\right)^n (i+1), \quad n \geq 0, i \geq 0.$$

Let $c = \sup_{n\geq 0} \boldsymbol{\pi}_n(\boldsymbol{V})$ and

$$u = \frac{(1+c)^2(\sqrt{p}+\sqrt{q})^2}{\lambda}\left[1 + c(2+\sqrt{q/p})^2\frac{(\sqrt{p}+\sqrt{q})^4}{(p-q)^2}\right].$$

Under the above assumption, if $0 < \varepsilon < 1/u$, then we have

$$\left\|\widetilde{\boldsymbol{\pi}} - \hat{\boldsymbol{\pi}} M \sum_{k=0}^{m} U^k\right\|_{\boldsymbol{v}} \leq c(1+\sqrt{\hat{\rho}})\frac{(\varepsilon u)^{m+1}}{1-\varepsilon u}.$$

Thank you for your attention!