DENUMERABLE STATE CONTINUOUS TIME MARKOV DECISION PROCESSES WITH UNBOUNDED COST AND TRANSITION RATES UNDER AVERAGE CRITERION

XIANPING GUO[1] and WEIPING ZHU[2]

Abstract

In this paper, we consider denumerable state continuous time Markov decision processes with (possibly unbounded) transition and cost rates under average criterion. We present a set of conditions and prove the existence of both average cost optimal stationary policies and a solution of the average optimality equation under the conditions. The results in this paper are applied to an admission control queue model and controlled birth and death processes.

1. Introduction

Continuous time Markov decision processes (CTMDP) have received considerable attention because many optimization models are based on processes involving continuous time. Both average and discounted criteria in CTMDP are often used to determine optimal policies. In this paper, we consider denumerable state CTMDP with (possibly unbounded) transition rates and (possibly unbounded) cost rates under average criterion, which has only been discussed by Bather [2] as far as we know.

When the state space is finite, bounded solutions of the optimality equation (OE) for minimizing average cost and methods computing optimal policies have been investigated for CTMDP by Howard [10], Miller [13], Lembersky [12], and many others. Since then, most work has focused on CTMDP with denumerable state space [5, 11, 14, 18]. Under the conditions of bounded reward rates and bounded transition rates, Kakuman [11] showed that if there exists a constant and bounded function satisfying the average reward OE, then any stationary policy determined by the OE is average reward optimal. If the difference defined as $u_\alpha(i) = V_\alpha(i) - V_\alpha(0)$, where

[1]Department of Mathematics, Zhongshan University, P. R. China
[2]School of Computer Science, ADFA, The University of New South Wales, Canberra, ACT 2600, Australia; e-mail: weiping@cs.adfa.edu.au

541

$V_\alpha(i)$ is the α-discounted optimal reward value function, is uniformly bounded in state i with discounted rate factor $\alpha > 0$, then the existence of such a constant and bounded function was also proved in [11]. Dong [5] extended the results of Kakuman [11] to the case of all possibly randomized Markov policies and countable action space, and provided a value-iteration algorithm. This means that the average reward optimal policies have the optimality property of obtaining the maximal average reward relative to the class of all possibly randomized Markov policies. Song [18] further extended the results of [5] to the case of non-uniformly bounded transition rates, but his treatment is restricted to the case of bounded reward rates and uniformly bounded difference $u_\alpha(i)$. For the case of unbounded reward rates, Puterman [14] and Zheng [21] proved the existence of both an average reward optimal stationary policy having the optimality property and a solution to the OE, but their treatments are restricted to the case of bounded transition rates. For the case of unbounded cost rates and unbounded transition rates, to the best of our knowledge only Bather [2] addressed this issue, and gave a set of sufficient conditions to guarantee the existence of an average cost optimal stationary policy using unbounded solutions of the OE, but the proof of existence of optimal policies in [2] is based on the assumption of existence of unbounded solutions to the OE. In the same paper, Bather mentioned that it is not easy to prescribe a general method for constructing solutions of the OE. The aim of this paper is to prove the existence of solutions to the OE and optimal policies for the case when both transition and cost rates are possibly unbounded. In the spirit of [2, 14–16], we present a set of conditions under which the existence of a solution of the OE and an average cost optimal stationary policy are proved. Later, the results are illustrated by two examples. One is an admission control queue model with unbounded cost rates and bounded transition rates. The other is controlled birth and death processes with unbounded transition rates and unbounded cost rates.

The rest of this paper is organized as follows. In Section 2, we briefly present our model, notation and definitions. In Section 3 we deal with α-discounted cost optimality. In Section 4 we provide a theorem giving the existence of a solution of the OE and an average cost optimal stationary policy. The theory is illustrated by considering an admission control queuing model and controlled birth and death processes in Section 5.

2. Model, notation and definitions

The system under our consideration can be simply stated as follows: when the system is at state i of a denumerable state space S, the decision maker chooses an action a from a finite set $A(i)$ of available actions. There are two consequences: (1) the decision maker needs to pay the cost for implementing the action a at rate $r(i, a)$; (2) the system state moves to state j, $j \in S$, which is governed by the transition

rate $q(j \mid i, a)$. The goal of the decision maker is to choose a sequence of actions which causes the system to perform optimally with respect to some predetermined performance criterion V. So the model for this system can be denoted by a five-element tuple $\{S, (A(i), i \in S), r, q, V\}$ having the following characteristics:

(i) the state space S is denumerable;

(ii) every available action space $A(i)$ is a finite subset of an action space A;

(iii) the cost rate r is a bounded below function on $K := \{(i, a) \mid i \in S, a \in A(i)\}$;

(iv) the transition rate q satisfies $q(j \mid i, a) \geq 0, \forall i \neq j, a \in A(i), i, j \in S$ and $\sum_{j \in S} q(j \mid i, a) = 0, i, j \in S, a \in A(i)$;

(v) V is a discounted (or average) cost criterion which will be defined later.

For the average cost criterion defined with 'lim inf', it is not easy to extend the optimality property to the unbounded. So the treatment of Bather [2] is restricted to the class of all stationary policies F, that is, $F = \{f \mid f : i \mapsto f(i) \in A(i), \forall i \in S\}$. As in [2], we limit ourselves to the same class of policies F.

For any $f \in F$, let $q_{ij}(f) := q(j \mid i, f(i)), i, j \in S, r(i, f) := r(i, f(i)), i \in S$, and $Q(f) := (q_{ij}(f))$. The minimum transition matrix with respect to the Q-matrix $Q(f)$ is denoted by $P^{\min}(t, f) = (p_{ij}^{\min}(t, f)), t \geq 0$. In order to determine a unique standard transition matrix with respect to $Q(f)$, and discuss α-discounted cost optimality, we make the following assumptions which are also essential to the coming discussion on average cost optimality.

ASSUMPTION 1. *There exist k non-negative functions $w_n, n=1, \ldots, k$, such that*

(i) *for all $i \in S$ and $a \in A(i), n=1, \ldots, k-1, \sum_{j \in S} q(j \mid i, a)w_n(j) \leq w_{n+1}(i)$;*

(ii) *for all $i \in S$ and $a \in A(i), \sum_{j \in S} q(j \mid i, a)w_k(j) \leq 0$.*

ASSUMPTION 2. $W := (w_1 + \cdots + w_k) \geq 1$, *and for all $i \in S, t > 0, f \in F$,* $\int_0^t \sum_{j \in S} p_{ij}^{\min}(u, f)q_j(f)W(j)\,du < \infty$, *where $q_j(f) := -q_{jj}(f), j \in S$ and every w_n comes from Assumption 1.*

ASSUMPTION A. (i) *Assumptions 1 and 2 hold;*

(ii) $|r(i, a)| \leq M W(i), i \in S, a \in A(i)$, *for some $M > 0$.*

For the case of polynomial reward, Hou [8] presented the following condition which we denote (H-C).

H-C: There exist a function $w \geq 1$ on S, positive constants b and M, an integer $k \geq 0$, such that

(i) $|r(i, a)| \leq Mw(i)^k a, a \in A(i), i \in S$;

(ii) $\sum_{j \neq i} q(j \mid i, a)w(j)^n \leq -q(i \mid i, a)[w(i) + b]^n$, for $n = 1, \ldots, k$.

Obviously, H-C is different from Assumption A. However, comparing Assumption A with H-C, we can obtain the following conclusions.

PROPOSITION 1. (i) *If* $\sup_{a \in A(i) i \in S}(-q(i \mid i, a)) := \|q\| < \infty$, *then H-C implies Assumption 1*;

(ii) *If H-C holds and* $\sum_{j \in S} q(j \mid i, a) Q(j) \leq 0$, $a \in A(i)$, $i \in S$, *then Assumption 1 holds, here* $Q(j) := \sup_{a \in A(j)}(-q(j \mid j, a))$, $j \in S$;

(iii) *If H-C holds,* $\|q\| < \infty$ *and* $\int_0^t \sum_{j \in S} p_{ij}^{\min}(u, f) q_j(f) w(j)^k \, du < \infty$, *then Assumption A holds. Here, k and w come from H-C*;

(iv) *If Assumption A holds and there is an integer* $1 \leq m \leq k$ *such that* $w_{m+1}(i) \leq -q(i \mid i, a) b$, $|r(i, a)| \leq M w_m(i)$ *for positive constants b and M, then H-C holds.*

PROOF. (i) Under H-C, Assumption 1 is obviously valid for the case of $k = 0$. If $k \geq 1$, then we can derive that for any $1 \leq n \leq k$,

$$\sum_{j \in S} q(j \mid i, a) w(j)^n \leq \left(\|q\| \sum_{m=1}^n C_n^m b^m \right) w(i)^{n-1}.$$

Let $w_n = \left(\|q\| \sum_{m=1}^n C_n^m b^m \right)^{n-1} w^{k-n+1}$, $n = 1, \ldots, k + 1$. Then we can derive that Assumption 1 holds.

(ii) Under H-C, Assumption 1 is obviously valid if $k = 0$. Let $k \geq 1$, then we have $\sum_{j \in S} q(j \mid i, a) w(j) \leq q - q(i \mid i, a) b \leq Q(i) b$, $a \in A(i)$, $i \in S$. Let $w_1(i) = w(i)$, $w_2(i) = b Q(i)$, $i \in S$. Then we can obtain that Assumption 1 holds.

(iii) By part (i) and $w \geq 1$, we can derive that part (iii) is valid.

(iv) Part (iv) obviously holds.

In fact, the main results and method presented in this paper have nothing in common with those presented in [8].

Now we define the discounted and the average cost criteria and their optimal cost value functions, respectively, as follows: For any $f \in F$, $i \in S$ and $\alpha > 0$,

$$V_\alpha(f, i) := \int_0^\infty e^{-\alpha t} \sum_{j \in S} p_{ij}^{\min}(t, f) r(j, f) \, dt, \qquad V_\alpha^*(i) := \inf_{f \in F} V_\alpha(f, i),$$

$$\bar{V}(f, i) := \liminf_{T \to \infty} \frac{1}{T} \int_0^T \sum_{j \in S} p_{ij}^{\min}(t, f) r(j, f) \, dt, \quad \bar{V}^*(i) := \inf_{f \in F} \bar{V}(f, i).$$

A policy $f^* \in F$ is called α-discounted cost optimal, if $V_\alpha(f^*, i) = V_\alpha^*(i)$, $i \in S$. Similarly, we can define average cost optimal policies.

REMARKS. 1. Under Assumptions 1 and 2, we know that the transition matrix with respect to $Q(f)$ is unique and honest, denoted by $P(t, f)$. Hence we have $P^{\min}(t, f) = P(t, f)$, for any $t \geq 0$, $f \in F$. Since the cost function r is bounded below, by the above definitions, we may assume $r \geq 0$ without any loss.

2. Throughout this paper, we assume that every function on S is regarded as a vector and that any kind of operator on matrices and vectors corresponds to them on all components.

3. Discounted optimality

In this section we will prove the existence of discounted cost optimal policies and characterize the optimal discounted cost value V_α^*, which is essential in the following discussion on average cost optimality.

LEMMA 1. *If Assumption 1 holds, then for any $f \in F$ and $t > 0$,*

(i) $P^{\min}(t, f)W \le \left(1 + t + \cdots + \dfrac{t^{(k-1)}}{(k-1)!}\right) W;$

(ii) $\displaystyle\int_0^\infty e^{-\alpha t} P^{\min}(t, f)W \le \left(\dfrac{1}{\alpha} + \dfrac{1}{\alpha^2} + \cdots + \dfrac{1}{\alpha^k}\right) W.$

PROOF. Similar to the proof of [19, Theorem 1], we can obtain these conclusions.

LEMMA 2. *If one of the following conditions holds, then Assumption A holds*:

(i) $\|r\| \equiv \sup_{i \in S, a \in A(i)} r(i, a) < \infty$, $\|q\| < \infty$;

(ii) *Assumption 1 holds, $r \le W$ and $\|q\| < \infty$;*

(iii) *for all $i \in S \equiv \{0, 1, \dots\}$, $A(i) = \{0, 1\}$, $q(i - 1 \mid i, 0) = \mu i$, $q(i \mid i, 0) = -(\lambda + \mu)i$, $q(i + 1 \mid i, 0) = \lambda i$, $i \ge 1$, $q(0 \mid 0, 0) = 0$; $q(i - 1 \mid i, 1) = \mu i$, $q(i \mid i, 1) = -(\lambda + \mu)i + v$, $q(i + 1 \mid i, 1) = \lambda i + v$, $i \ge 1$, $q(0 \mid 0, 1) = -v$, $q(1 \mid 0, 1) = v$, $0 \le \lambda \le \mu$, $v \ge 0$. There are l positive numbers b_n, $n = 1, \dots, l$, such that $r(i, a) \le \sum_{n=1}^l b_n i^n$, for all $a \in A(i)$, $i \in S$.*

PROOF. Obviously, (i) is valid. By adding 1 to w_k in Assumption 1, from Lemma 1, we obtain (ii). Part (iii) can be proved by applying Lemma 1. The calculation is straightforward, but lengthy, and we shall omit the details.

With Assumption A we can define

$$B(S) := \{u \mid cW(i) \le u(i) \le cW(i), \forall i \in S \text{ for some } c > 0\}.$$

LEMMA 3. *If Assumption A holds, then we have*

(i) *for any $f \in F$, $t > 0$, $i \in S$, $\alpha > 0$*

(1) $\sum_{j \in S} p_{ij}^{\min}(t, f) = 1$,

(2) $V_\alpha^* \in B(S)$;

(ii) *for any $\alpha > 0$, the quantity $V_\alpha^*(i)$ $(i \in S)$ is a unique solution of the following OE within $B(S)$:*

$$\alpha V_\alpha^*(i) = \min_{a \in A(i)} \left\{ r(i, a) + \sum_{j \in S} q(j \mid i, a) V_\alpha^*(j) \right\}, \quad i \in S; \tag{3.1}$$

(iii) *for any $\alpha > 0$, there must exist an α-discounted cost optimal policy f_α^*;*

(iv) *a stationary policy $f \in F$ is optimal if and only if it realizes the minimum on the right-hand side of* (3.1).

PROOF. By [19, Theorems 2, 4 and 7], we can obtain these conclusions.

REMARK 3. Under Assumption A, Lemma 3 (i) means that the transition matrix with respect to $Q(f)$ is unique, standard and honest, for any $f \in F$.

We take an arbitrary, but fixed, function $m > 0$ on S such that $m(i) \geq Q(i)$, $i \in S$.

LEMMA 4. *Let Assumption A hold.*

(i) *If we let $u_0 := 0$ and*

$$u_{n+1}(i) := \min_{a \in A(i)} \left\{ \frac{r(i, a)}{m(i) + \alpha} + \frac{m(i)}{m(i) + \alpha} \sum_{j \in S} \left(\frac{q(j \mid i, a)}{m(i)} + \delta_{ij} \right) u_n(j) \right\},$$

$i \in S$ and $n \geq 0$, then $\lim_{n \to \infty} u_n(i) = V_\alpha^*(i)$, for $i \in S$.

(ii) *If $A(i) \equiv A$, $i \in S := \{0, 1, \ldots\}$, for any $a \in A$, $\alpha > 0$, $m(i)$ and $r(i, a)/(m(i) + \alpha)$ are increasing functions in i on S, and for any fixed $k \in S$, $a \in A$, $\tilde{q}(k \mid i, a) := \sum_{j \geq k} [q(j \mid i, a)/m(i) + \delta_{ij}]$ is increasing in i on S, then $V_\alpha^*(i)$ is increasing in i on S for any $\alpha > 0$.*

PROOF. (i) For $u \in B(S)$, let

$$Tu(i) = \min_{a \in A(i)} \left\{ \frac{r(i, a)}{m(i) + \alpha} + \frac{m(i)}{m(i) + \alpha} \sum_{j \in S} \left[\frac{q(j \mid i, a)}{m(i)} + \delta_{ij} \right] u(j) \right\}, \quad i \in S.$$

Then we have $u_{n+1} = Tu_n$, $u_n = T^n 0$, $u_n \leq u_{n+1}$, $n \geq 1$. Based on Assumption A, by induction, we can obtain that, for any $n \geq 1$,

$$u_n \leq \sum_{l=1}^{k} \alpha^{-l}(w_l + \cdots + w_k)M \leq \left(\frac{1}{\alpha} + \frac{1}{\alpha^2} + \cdots + \frac{1}{\alpha^k} \right) MW.$$

Hence by the control convergence theorem and Lemma 3, and noting that every $A(i)$ is finite, we have $u = Tu$. Hence (i) is valid.

To prove (ii), by (i), we need to prove that, for $i_1, i_2 \in S$, $i_1 \geq i_2$, $n \geq 0$,

$$u_n(i_1) \geq u_n(i_2). \tag{3.2}$$

By induction, when $n = 0$, (3.2) is obviously valid. Suppose (3.2) holds for $n = N$. With the denotation $u_N(-1) = 0$, for any $a \in A$, $i_1 \geq i_2$, $i_1, i_2 \in S$, we have

$$\sum_{j=0}^{\infty} \left(\frac{q(j \mid i_1, a)}{m(i_1)} + \delta_{i,j} \right) u_N(j)$$

$$= \sum_{j=0}^{\infty} \left[\frac{q(j \mid i_1, a)}{m(i_1)} + \delta_{i,j} \right] \left[\sum_{i=0}^{j} (u_N(i) - u_N(i-1)) \right]$$

$$= \sum_{j=0}^{\infty} (u_N(j) - u_N(j-1)) \left(\sum_{i=j}^{\infty} \left[\frac{q(i \mid i_1, a)}{m(i_1)} + \delta_{i,i} \right] \right)$$

$$= \sum_{j=0}^{\infty} (u_N(j) - u_N(j-1)) \tilde{q}(j \mid i_1, a)$$

$$\geq \sum_{j=0}^{\infty} (u_N(j) - u_N(j-1)) \tilde{q}(j \mid i_2, a) = \sum_{j=0}^{\infty} \left(\frac{q(j \mid i_2, a)}{m(i_2)} + \delta_{i,j} \right) u_N(j).$$

Hence

$$\frac{r(i_1, a)}{m(i_1) + \alpha} + \frac{m(i_1)}{m(i_1) + \alpha} \sum_{j \in S} \left(\frac{q(j \mid i_1, a)}{m(i_1)} + \delta_{i,j} \right) u_N(j)$$

$$\geq \frac{r(i_2, a)}{m(i_2) + \alpha} + \frac{m(i_2)}{m(i_2) + \alpha} \sum_{j \in S} \left[\frac{q(j \mid i_2, a)}{m(i_2)} + \delta_{i,j} \right] u_N(j).$$

Then $u_{N+1}(i_1) \geq u_{N+1}(i_2)$. That means (3.2) is valid for $n = N + 1$. Hence (ii) is also valid.

COROLLARY 1. *Let $\|q\| < \infty$ and Assumption A hold. If for any $i \in S \equiv \{0, 1, \dots\}$, $A(i) = A$, for any $a \in A$, $r(i, a)$ is increasing (or decreasing) in i on S and $\tilde{q}(k \mid i, a)$ is increasing in i on S for any fixed $k \in S$, then $V_\alpha^*(i)$ is increasing (or decreasing).*

PROOF. Let $m(i) = \|q\| + 1$, $i \in S$. Using a similar method to that used to prove Lemma 4, we can obtain these conclusions.

4. Average criterion

In this section, we always assume that Assumption A holds. By Lemma 3, we let f_α^* denote a discounted cost optimal stationary policy with respect to the discounted rate factor $\alpha > 0$.

By the Tychonoff theorem we have that F is a compact metric space. Hence for any sequence $\{f_{\alpha_n}^*, n \geq 1\} \subset F$, there must exist a convergent subsequence of

$\{f_{\alpha_n}^*, n \geq 1\}$. This means that there must exist a limit point f^* for any sequence $\{f_{\alpha_n}^*, n \geq 1\}$.

Let $k_0 \in S$ be a fixed state. For any $\alpha > 0$ and $i \in S$, let $u_\alpha(i) := V_\alpha^*(i) - V_\alpha^*(k_0)$. Now we present the conditions under which the existence of optimal policies and a solution of the OE can be proved.

ASSUMPTION B. *For some decreasing sequence $\{\alpha_n\}$ tending to zero and some $k_0 \in S$, there exist a non-negative function h and a constant N^*, such that for all $n \geq 1, a \in A(i), i \in S$*

(i) *$N^* \leq u_{\alpha_n}(i) \leq h(i)$;*

(ii) *$\sum_{j \in S} q(j \mid i, a) h(j) < \infty$;*

(iii) *there are a non-negative and decreasing sequence $\{\phi_j\}$ and a positive integer $\bar{u}$ such that, for any $a \in A(j)$, $r(j, a) \geq Q(j) h(j) \phi_j$, whenever $j \geq \bar{u}$, and $\sum_{j \in S} \phi_j = \infty$;*

(iv) *there exist positive integers $\bar{v}, \bar{w}$, constants $C > 0$, $\beta \in (0, 1)$ such that, for any $j \in S, a \in A(j)$, $q(k \mid j, a) h(k) \leq \beta^k C$, whenever $j > \bar{v}, k > \bar{w} + j$.*

ASSUMPTION C. *One of the following conditions holds:*

(i) *$\lim_{i \to \infty} \min_{a \in A(i)} r(i, a) = \infty$;*

(ii) *$\sum_{j \in S} \pi_{ij}(f) = 1, i \in S, f \in F$, here $(\pi_{ij}(f)) := \lim_{t \to \infty} P(t, f)$.*

To verify Assumption C, we have the following lemma.

LEMMA 5. *If one of the following conditions holds,*

(i) *there is a function G such that $\lim_{i \to \infty} G(i) = \infty$ and $r(i, a) \geq G(i), i \in S$, $a \in A(i)$;*

(ii) *for any $f \in F$, there exists a function $X(f) \geq 0$ such that $\lim_{i \to \infty} X(f)(i) = \infty$ and $Q(f) X(f) \leq 0$;*

(iii) *for any $f \in F$, the process $(P_{ij}(t, f))$ is irreducible and recurrent positive,*

then Assumption C holds.

PROOF. Obviously, the conclusion is valid under Condition (i). By [1, Propositions 5.1.7 and 5.4.8 and Theorem 5.1.6], we then obtain this conclusion under Conditions (ii) or (iii).

LEMMA 6. *For any $f \in F, i \in S$, we have*

(i) *if $\sum_{j \in S} \pi_{ij}(f) < 1$, $\lim_{k \to \infty} r(k, f(k)) = \infty$, then $\bar{V}(f, i) = \infty$;*

(ii) *$\sum_{j \neq k} \pi_{ij}(f) q(k \mid j, f(j)) \leq \pi_{ik}(f) q_k(f)$;*

(iii) *$\bar{V}(f, i) \geq \sum_{j \in S} \pi_{ij}(f) r(j, f(j))$;*

if there exist a function u bounded below and a constant g satisfying the equation

$$g = r(i, f(i)) + \sum_{j \in S} q(j \mid i, f(i)) u(j), \quad i \in S,$$

then $\bar{V}(f, i) \leq g$.

PROOF. Conditions (i) and (ii) are from [2, Lemmas 2.2.2 and 2.2.1], respectively. Condition (iv) is from [2, Theorem 2.1.3]. Since $r \geq 0$, by the Fatou lemma, we can prove (iii).

THEOREM 1. *If Assumptions A, B and C hold, then*

(i) *there exist a stationary policy f^*, a constant g^*, a function u^* on S and a decreasing sequence $\{\alpha_k\}$ tending to zero, such that for $i \in S$,*

(a) $g^* = \lim\limits_{k \to \infty} \alpha_k V_{\alpha_k}^*(i), \quad u^*(i) = \lim\limits_{k \to \infty} u_{\alpha_k}(i);$

(b)
$$g^* = r(i, f^*(i)) + \sum_{j \in S} q(j \mid i, f^*(i))u^*(j)$$

$$= \min_{a \in A(i)} \left\{ r(i, a) + \sum_{j \in S} q(j \mid i, a)u^*(j) \right\}; \qquad (4.1)$$

(ii) *f^* is an average cost optimal and satisfies $\bar{V}(f^*, i) = g^*, i \in S$;*

(iii) *any f realizing the minimum on the right-hand side of (4.1) is average cost optimal.*

PROOF. (i) For any $n \geq 1$, $i \in S$ and $a \in A(i)$, by Lemma 3, we have

$$\alpha_n V_{\alpha_n}^*(i) = r(i, f_{\alpha_n}^*(i)) + \sum_{j \in S} q(j \mid i, f_{\alpha_n}^*(i)) V_{\alpha_n}^*(j)$$

$$= \min_{a \in A(i)} \left\{ r(i, a) + \sum_{j \in S} q(j \mid i, a) V_{\alpha_n}^*(j) \right\}$$

$$= \min_{a \in A(i)} \left\{ r(i, a) + \sum_{j \in S} q(j \mid i, a)u_{\alpha_n}(j) \right\}. \qquad (4.2)$$

Hence

$$\alpha_n V_{\alpha_n}^*(k_0) + \alpha_n u_{\alpha_n}(i) = r(i, f_{\alpha_n}^*(i)) + \sum_{j \in S} q(j \mid i, f_{\alpha_n}^*(i))u_{\alpha_n}(j)$$

$$= \min_{a \in A(i)} \left\{ r(i, a) + \sum_{j \in S} q(j \mid i, a)u_{\alpha_n}(j) \right\}. \qquad (4.3)$$

Since f^* is a limit point of $\{f_{\alpha_n}^*\}$, there must exist a subsequence $\{\alpha_{n'}\}$ of $\{\alpha_n\}$ such that $\lim_{n' \to \infty} f_{\alpha_{n'}}^*(i) = f^*(i)$, for $i \in S$. By Assumption B (ii) and (4.2), to take any $a \in A(k_0)$, we have

$$0 \leq \alpha_{n'} V_{\alpha_{n'}}^*(k_0) \leq r(k_0, a) + \sum_{j \in S} q(j \mid k_0, a)h(j) - 2q(k_0 \mid k_0, a)(h(k_0) + |N^*|) < \infty.$$

So $|\alpha_{n'} V^*_{\alpha_{n'}}(k_0)|$ is bounded for any $n' \geq 1$. Hence there must exist a subsequence $\{\alpha_{m'}\}$ of $\{\alpha_{n'}\}$ such that $g^* = \lim_{m' \to \infty} \alpha_{m'} V_{\alpha_{m'}}(k_0)$. By Assumption B (i) and the Tychonoff theorem, we also have that $\{u_{\alpha_{m'}}\}$ is a sequence of the compact metric space $\prod_{i \in S}[N^*, h(i)]$. So there must exist a subsequence $\{\alpha_k\}$ of $\{\alpha_{m'}\}$ such that $\lim_{k \to \infty} u_{\alpha_k}(i) := u^*(i)$, $i \in S$. By Assumption B, we have $\lim_{k \to \infty} \alpha_k = 0$. Hence $\lim_{k \to \infty} \alpha_k u_{\alpha_k}(i) = 0$, for $i \in S$. Now we derive that $\lim_{k \to \infty} \alpha_k V^*_{\alpha_k}(i) = \lim_{k \to \infty} \alpha_k V^*_{\alpha_k}(k_0) = g^*$, for $i \in S$. From the above discussion, we get, for $i \in S$,

$$g^* = \lim_{k \to \infty} \alpha_k V^*_{\alpha_k}(i); \quad \lim_{k \to \infty} u_{\alpha_k}(i) := u^*(i); \quad \lim_{k \to \infty} f^*_{\alpha_k}(i) = f^*(i). \qquad (4.4)$$

From (4.3), for $k \geq 1$, we have

$$\frac{\alpha_k V^*_{\alpha_k}(k_0)}{m(i)} + \frac{\alpha_k u_{\alpha_k}(i)}{m(i)} + u_{\alpha_k}(i)$$

$$= \frac{r(i, f^*_{\alpha_k}(i))}{m(i)} + \sum_{j \in S}\left[\frac{q(j \mid i, f^*_{\alpha_k}(i))}{m(i)} + \delta_{ij}\right] u_{\alpha_k}(j)$$

$$= \min_{a \in A(i)}\left\{\frac{r(i, a)}{m(i)} + \sum_{j \in S}\left[\frac{q(j \mid i, a)}{m(i)} + \delta_{ij}\right] u_{\alpha_k}(j)\right\}. \qquad (4.5)$$

For any $i \in S$, since $A(i)$ is finite, there must exist an integer $N(i) > 0$ and an action $a^*(i) \in A(i)$ such that, for $k \geq N(i)$,

$$f^*_{\alpha_k}(i) = f^*(i) \qquad (4.6)$$

and $\max_{a \in A(i)}\left\{\sum_{j \in S} q(j \mid i, a)h(j)\right\} = \sum_{j \in S} q(j \mid i, a^*(i))h(j) < \infty$. Hence we have

$$\sum_{j \in S}\left(\frac{q(j \mid i, f^*(i))}{m(i)} + \delta_{ij}\right)|u_{\alpha_k}|(j) \leq \sum_{j \in S} q(j \mid i, a^*(i))\frac{h(j)}{m(i)} + h(i) + |N^*| < \infty.$$

By the control convergence theorem, (4.4)–(4.6) and noting that $A(i)$ is finite, for $i \in S$, we have

$$\frac{g^*}{m(i)} + u^*(i) = \frac{r(i, f^*(i))}{m(i)} + \sum_{j \in S}\left[\frac{q(j \mid i, f^*(i))}{m(i)} + \delta_{ij}\right] u^*(j)$$

$$= \min_{a \in A(i)}\left\{\frac{r(i, a)}{m(i)} + \sum_{j \in S}\left[\frac{q(j \mid i, a)}{m(i)} + \delta_{ij}\right] u^*(j)\right\}.$$

Thus

$$g^* = \min_{a \in A(i)}\left\{r(i, a) + \sum_{j \in S} q(j \mid i, a)u^*(j)\right\}. \qquad (4.7)$$

By (4.4) and (4.7), we have that (i) is valid.

(ii) First, we assume that $\sum_{j \in S} \pi_{ij}(f) = 1$, $i \in S$, $f \in F$. We prove

$$\sum_{j \in S} \pi_{ij}(f^*) r(j, f^*(j)) \geq g^*. \tag{4.8}$$

Obviously, we only need to consider $\sum_{j \in S} \pi_{ij}(f^*) r(j, f^*(j)) < \infty$. Since $\sum_{j \in S} q(j \mid i, f^*(i)) = 0$, and u^* is bounded below, the components of u^* may be increased by adding any constant without affecting the proof. So we may assume $u^* \geq 0$. From (4.1) and Lemma 6 (ii), we have that, for a large $N > 0$,

$$\sum_{j \leq N} \pi_{ij}(f^*)(g^* - r(j, f^*(j)))$$

$$= \sum_{j \leq N} \pi_{ij}(f^*) \left(\sum_{k \in S} q(k \mid j, f^*(j)) \right) u^*(k)$$

$$= \sum_{j \leq N} \pi_{ij}(f^*) \left[\sum_{k \neq j} q(k \mid j, f^*(j)) u^*(k) - q_j(f^*) u^*(j) \right]$$

$$= \sum_{j \leq N} \pi_{ij}(f^*) \left[\sum_{k > N} q(k \mid j, f^*(j)) u^*(k) + \sum_{\substack{k, j \leq N, \\ k \neq j}} \pi_{ij}(f^*) q(k \mid j, f^*(j)) u^*(k) \right]$$

$$\quad - \sum_{k \leq N} \pi_{ik}(f^*) q_k(f^*) u^*(k)$$

$$\leq \sum_{j \leq N} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k) + \sum_{k \leq N} \sum_{j \neq k} \pi_{ij}(f^*) q(k \mid j, f^*(j)) u^*(k)$$

$$\quad - \sum_{k \leq N} \pi_{ik}(f^*) q_k(f^*) u^*(k)$$

$$= \sum_{j \leq N} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k)$$

$$\quad - \sum_{k \leq N} u^*(k) \left[\pi_{ik}(f^*) q_k(f^*) - \sum_{j \neq k} \pi_{ij}(f^*) q(k \mid j, f^*(j)) \right]$$

$$\leq \sum_{j \leq N} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k).$$

Choosing $N > \bar{v}$ and using Assumption B (iv), we can obtain

$$\sum_{j \leq N} \pi_{ij}(f^*)(g^* - r(j, f^*(j)))$$

$$\leq \sum_{j \leq \bar{v}} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k) + \sum_{j = \bar{v}+1}^{N} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k)$$

$$= \sum_{j \le \bar{v}} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k)$$

$$+ \sum_{j = \bar{v}+1}^{N} \pi_{ij}(f^*) \left\{ \sum_{k=N+1}^{N+\bar{w}} q(k \mid j, f^*(j)) u^*(k) + \sum_{k=N+1+\bar{w}}^{\infty} q(k \mid j, f^*(j)) u^*(k) \right\}$$

$$\le \sum_{j \le \bar{v}} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k)$$

$$+ \sum_{k=N+1}^{N+\bar{w}} \sum_{j=\bar{v}+1}^{N} \pi_{ij}(f^*) q(k \mid j, f^*(j)) u^*(k) + \sum_{k=N+1+\bar{w}}^{\infty} \beta^k C$$

$$\le \sum_{j \le \bar{v}} \pi_{ij}(f^*) \sum_{k > N} q(k \mid j, f^*(j)) u^*(k) + \sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) q_k(f^*) u^*(k) + \frac{\beta^N C}{(1-\beta)}.$$

By Assumption B (ii), we can obtain $\sum_{k \in S} q(k \mid j, f^*(j)) u^*(k) < \infty$. So the first and last terms on the right-hand side of the above formula tend to zero as $N \to \infty$. Since $\sum_{j \in S} \pi_{ij}(f^*) = 1$, we have that

$$g^* \le \sum_{j \in S} \pi_{ij}(f^*) r(j, f^*(j)) + \liminf_{N \to \infty} \sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) u^*(k) q_k(f^*).$$

To prove that $\liminf_{N \to \infty} \sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) u^*(k) q_k(f^*) = 0$, it is enough to show that, given $\epsilon > 0$, there exist infinitely many values of N for which

$$\sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) u^*(k) q_k(f^*) < \epsilon.$$

Suppose, for contradiction, that we can find an integer $l \ge \bar{u}$ such that

$$\sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) u^*(k) q_k(f^*) \ge \epsilon$$

for $N = l\bar{w}, (l+1)\bar{w}, \ldots$. Then Assumption B (iii) shows that

$$\sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) r(k, f^*(k)) \ge \sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) Q(k) h(k) \phi_k$$

$$\ge \sum_{k=N+1}^{N+\bar{w}} \pi_{ik}(f^*) q_k(f^*) u^*(k) \phi_{N+\bar{w}} \ge \epsilon \phi_{N+\bar{w}}$$

for $N = l\bar{w}, (l+1)\bar{w}, \ldots$. It follows that $\sum_{k \in S} \pi_{ik}(f^*) r(k, f^*(k)) \ge \epsilon \sum_{j=l+1}^{\infty} \phi_{j\bar{w}}$. Since $\{\phi_j\}$ is a decreasing sequence and $\sum_{j \in S} \phi_j = \infty$, we have $\sum_{j=l+1}^{\infty} \phi_{j\bar{w}} = \infty$.

Hence $\sum_{k \in S} \pi_{ik}(f^*)r(k, f^*(k)) = \infty$. This contradiction shows that

$$\liminf_{N \to \infty} \sum_{k=N+1}^{N+\tilde{w}} \pi_{ik}(f^*)q_k(f^*)u^*(k) = 0.$$

So we have proved that $\sum_{j \in S} \pi_{ij}(f^*)r(j, f^*(j)) \geq g^*$. By Lemma 6, we have $\bar{V}(f^*, i) \geq g^*$.

For any $f \in F, i \in S$, by (4.1), we have $g^* \leq r(i, f(i)) + \sum_{j \in S} q(j \mid i, f(i))u^*(j)$. Similarly, we can prove $g^* \leq \bar{V}(f, i), f \in F$.

Second, if for some $i \in S, f \in F, \sum_{ij} \pi_{ij}(f) < 1$, by Assumption C and Lemma 6, we have $\bar{V}(f, i) = \infty \geq g^*$. Hence we have

$$\bar{V}(f, i) \geq g^*, \quad f \in F, \ i \in S. \tag{4.9}$$

On the other hand, since $g^* = r(i, f^*(i)) + \sum_{j \in S} q(j \mid i, f^*(i))u^*(j), \ i \in S$, by Lemma 6, we can have

$$\bar{V}(f^*, i) \leq g^*, \quad i \in S.$$

By (4.8) and (4.9), we have

$$\bar{V}(f^*, i) = g^* \leq \bar{V}(f, i), \quad i \in S, \ f \in F.$$

This means that (ii) is valid. Similarly, we can prove (iii).

5. Examples

In this section we will apply the results of previous sections to demonstrate the existence of optimal policies in two examples. One is an admission control queue model with unbounded cost and bounded transition rates. The other is controlled birth and death processes with unbounded cost and unbounded transition rates.

EXAMPLE 1. We observe continuously an admission control model for a queuing system. The system behaves as a single-server queuing system $M^X/M/1$. Let $p_k, k = 0, 1, \ldots, K < \infty$, denote the arrival probability of k tasks, $p_k \geq 0$ and $\sum_{k=0}^{K} p_k = 1$. The arrival rate of the system is λ. Let μ denote the exponential service rate of the system. At any arrival time, the controller decides whether to admit the arrival tasks. Rejected tasks are lost. Each accepted task generates a reward R. A non-decreasing function $r(i)$ denotes the cost rate for serving i tasks. Let $p > 0$ denote the cost rate of serving a single task. Hence we have $r(i) = pi$.

We formulate this model as continuous time Markov decision processes. The system state i denotes the number of tasks available for service in the system at any time (that is, the queue length). So $S = \{0, 1, 2, \dots\}$. For each $i \in S$, $i \geq 0$, $A(i) = \{0, 1\}$ with action 0 corresponding to rejecting and action 1 corresponding to accepting arrival tasks. The cost rate function r satisfies $r(i, 0) = r(i)$, $r(i, 1) = r(i) - R\lambda$. By the definition of average criterion, the cost function $r(i, a)$ may be increased by adding any constant without affecting the discussion of the average optimality. So we may take that $r(i, 0) = r(i) + R\lambda$, $r(i, 1) = r(i)$. The transition rates satisfy: $q(0 \mid 0, 0) = 0$, $q(i - 1 \mid i, 0) = \mu$, $q(i \mid i, 0) = -\mu$, $i > 0$; $q(k \mid 0, 1) = \lambda p_k$, $k = 1, \dots, K$, $q(0 \mid 0, 1) = -(\lambda p_1 + \cdots + \lambda p_K)$; $q(i - 1 \mid i, 1) = \mu$, $q(i + k \mid i, 1) = \lambda p_k$, $k = 1, \dots, K$, $q(i \mid i, 1) = -(\mu + \lambda p_1 + \cdots + \lambda p_K)$, $i > 0$. This model has the following properties:

(1)   Assumption A holds. In fact, we let $w_1(i) = pi + 1 + R\lambda$, $w_2(i) = p\lambda(\sum_{k=1}^{K} k p_k)$, $i \in S$. By Lemma 2, we can verify Assumption A.

(2)   Assumption B holds. In fact, by Lemma 3 (iii), we have, for any $\alpha > 0$,

$$\alpha V_\alpha^*(i) = \min_{a \in A(i)} \left\{ r(i, a) + \sum_{j \in S} q(j \mid i, a) V_\alpha^*(j) \right\}$$

$$\leq r(i, 0) + \mu V_\alpha^*(i - 1) - \mu V_\alpha^*(i), \quad i > 0.$$

So we have

$$V_\alpha^*(i) \leq \frac{r(i, 0)}{\mu} + V_\alpha^*(i - 1) \leq \frac{r(i, 0)}{\mu} + \frac{r(i - 1, 0)}{\mu} + V_\alpha^*(i - 2)$$

$$\leq \cdots \leq \sum_{s=1}^{i} \frac{r(s, 0)}{\mu} + V_\alpha^*(0) = \frac{p i(i + 1)}{2\mu} + \frac{i R\lambda}{\mu} + V_\alpha^*(0).$$

Hence we may take $k_0 = 0$ and thus $u_\alpha(i) \leq p i(i + 1)/(2\mu) + i R\lambda/\mu \doteq h(i)$, $i \in S$, $\alpha > 0$. To take $m(i) = \lambda + \mu$, $i \in S$, by Lemma 4, we have that V_α^* is an increasing function on S. Hence we have $u_\alpha(i) \geq 0$, for $i \in S$, $\alpha > 0$. So Assumptions B (i) and B (ii) hold. To verify Assumption B (iii), let $\bar{u} = 1$, $Q(i) \equiv \mu + \lambda$, $i \in S$; thus $\phi_i = p\mu/((\mu + \lambda)(p + R\lambda)(i + 1))$, $i \in S$ and thus we can obtain that Assumption B (iii) holds. Letting $C = 0$, $\tilde{v} = 0$, $\bar{w} = K$, we can verify Assumption B (iv). Obviously Assumption C holds. Hence by Theorem 1, we have the following conclusion:

For this admission control queue model, there must exist an average cost optimal stationary policy.

EXAMPLE 2. We consider controlled birth and death processes as follows: Let $S = \{0, 1, 2, \dots\}$, $A(i) \equiv \{0, 1\}$, $i \in S$, $q(0 \mid 0, 0) = 0$, $q(i - 1 \mid i, 0) = \mu i$,

$q(i \mid i, 0) = -\mu i, i > 0; q(0 \mid 0, 1) = -v, q(1 \mid 0, 1) = v, q(i - 1 \mid i, 1) = \mu i,$
$q(i \mid i, 1) = -(\lambda + \mu)i - v, q(i + 1 \mid i, 1) = \lambda i + v, \mu \geq \lambda > 0.$ Let $r(i, 0) =$
$a_1 i^2 + a_2 i + c, r(i, 1) = b_1 i^2 + b_2 i, i \in S, (c - 2a_1)(\lambda + \mu) \leq v(3a_1 + a_2), c, a_1, b_1 > 0,$
$a_2, b_2 \geq 0.$

Both cost and transition rates in this model are unbounded. Moreover, we can derive that:

(1) Assumption A holds. We shall now verify this conclusion. Let $w_1(i) = (a_1 + b_1)i^2$, $i \in S.$ Then we have $\sum_{j \in S} q(j \mid i, 0)w_1(j) \leq 0, i \in S;$ and

$$\sum_{j \in S} q(j \mid 0, 1)w_1(j) = v(a_1 + b_1) \leq (a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2), \quad i = 0;$$

$$\sum_{j \in S} q(j \mid i, 1)w_1(j)$$
$$= \mu i w_1(i - 1) - \mu i w_1(i) - \lambda i w_1(i) + \lambda i w_1(i + 1) + v w_1(i + 1) - v w_1(i)$$
$$= 2(a_1 + b_1)(-\mu + \lambda)i^2 + (a_1 + b_1)(\mu + \lambda)i + v(a_1 + b_1)(2i + 1)$$
$$\leq ((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2))(i + 1), \quad i \geq 1. \tag{5.1}$$

Let $w_2(i) = ((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2))(i + 1), i \in S$, then we have

$$\sum_{j \in S} q(j \mid i, a)w_1(j) \leq w_2(i) \quad \text{and} \quad \sum_{j \in S} q(j \mid i, 0)w_2(j) \leq 0, \quad i \in S, \ a \in \{0, 1\};$$

and

$$\sum_{j \in S} q(j \mid 0, 1)w_2(j) = v((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2)), \quad i = 0;$$

$$\sum_{j \in S} q(j \mid i, 1)w_2(j)$$
$$= \mu i w_2(i - 1) - \mu i w_2(i) - \lambda i w_2(i) + \lambda i w_2(i + 1) + v w_2(i + 1) - v w_2(i)$$
$$= ((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2))(-\mu i + \lambda i)$$
$$\quad + v((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2))$$
$$\leq v((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2)), \quad i \geq 1. \tag{5.2}$$

Let $w_3(i) = v((a_1 + b_1)(\mu + \lambda + 3v) + (a_2 + b_2)) + c + 1, i \in S$, then we have

$$\sum_{j \in S} q(j \mid i, a)w_2(j) \leq w_3(i), \quad i \in S, \ a \in \{0, 1\} \tag{5.3}$$

and

$$\sum_{j \in S} q(j \mid i, a)w_3(j) \leq 0, \quad i \in S, \ a \in \{0, 1\}.$$

Hence from (5.1), (5.2) and (5.3), Assumption 1 holds.

Similarly, we can verify that the function QR on S satisfies Assumption 1. Here $R = w_1 + w_2 + w_3$. By Lemma 2, we can obtain that Assumption 2 holds. Obviously, $R \geq 1$ and $r \leq R$.

Combining the above conclusions, we have that Assumption A holds. Here, we take $m(0) = 2(\lambda + \mu + v)$, $m(i) = 2((\lambda + \mu +)i + v)$, $i \geq 1$.

(2) The functions $m(i)$, $r(i, a)/(m(i) + \alpha)$ and $\tilde{q}(k \mid i, a)$ are all increasing in i on S, for any $\alpha > 0$, $a \in \{0, 1\}$. In fact, obviously, m is an increasing function on S. By the condition $(c - 2a_1)(\lambda + \mu) \leq v(3a_1 + a_2)$, we can verify that $r(i, a)/(m(i) + \alpha)$ is increasing in i on S, for any $\alpha > 0$, $a \in \{0, 1\}$. By the definition of $\tilde{q}(k \mid i, a)$, we can also verify that $\tilde{q}(k \mid i, a)$ is increasing in i on S, for any $\alpha > 0$, $a \in \{0, 1\}$, $k \in S$.

(3) The difference $u_\alpha(i) \geq 0$, here $k_0 = 0$. In fact, by Lemma 4 and (2), we have that $V_\alpha^*(i)$ is increasing in i on S. Hence $u_\alpha(i) \geq 0$.

(4) Let $h(i) = (a_2 + c)i/\mu + a_1 i(i + 1)/(2\mu)$, $i \in S$. Then we have $u_\alpha(i) \leq h(i)$, $i \in S$, $\alpha > 0$. By Lemma 3, we have $0 \leq \alpha V_\alpha^*(0) \leq r(0, 0) = c$, and, for $i \geq 1$,

$$\alpha V_\alpha^*(i) \leq r(i, 0) + i\mu V_\alpha^*(i - 1) - i\mu V_\alpha^*(i).$$

Hence

$$V_\alpha^*(i) \leq \frac{r(i, 0)}{i\mu} + V_\alpha^*(i - 1) \leq \frac{ia_1}{\mu} + \frac{a_2 + c}{\mu} + V_\alpha^*(i - 1)$$

$$\leq \cdots \leq \frac{i(a_2 + c)}{\mu} + \frac{a_1 i(i + 1)}{2\mu} + V_\alpha^*(0).$$

Also $u_\alpha(i) = V_\alpha^*(i) - V_\alpha^*(0) \leq i(a_2 + c)/\mu + a_1 i(i + 1)/(2\mu) \equiv h(i)$.

(5) Assumption B holds. By (3) and (4), we only need to verify Assumptions B (iii) and B (iv). Let $\bar{u} = 1$, $\bar{v} = 0$, $\bar{w} = 2$,

$$\phi_i = \frac{\mu \min\{a_1, b_1\}}{(a_1 + a_2 + c)(\lambda + \mu + v)(i + 1)}.$$

Then we have that Assumptions B (iii) and B (iv) are valid.

(6) Assumption C holds. In fact, this is obvious.

Hence by (1), (5), (6) and Theorem 1, we have the following conclusion:

For these controlled birth and death processes, there must exist an average cost optimal stationary policy.

Acknowledgement

References

[1] W. J. Andson, *Continuous time Markov chains* (Springer, New York, 1991).

[2] J. Bather, "Optimal stationary policies for denumerable Markov chains in continuous time", *Adv. Appl. Prob.* **8** (1976) 144–158.

[3] D. P. Bertsekas, *Dynamic programming: deterministic and stochastic models* (Prentice-Hall, Englewood Cliffs, NJ, 1987).

[4] K. L. Chung, *Markov chains with stationary transition probabilities* (Springer, Berlin, 1960).

[5] Z. Q. Dong, "Continuous time Markov decision programming with average reward criterion—countable state and action space", *Scientia Sinica* **SP ISS(II)** (1979) 11–148.

[6] J. A. Filar and K. Vrieze, *Competitive Markov decision processes* (Springer, New York, 1996).

[7] M. Haviv and M. L. Puterman, "Bias optimality in controlled queuing systems", *J. Appl. Prob.* **35** (1998) 16–150.

[8] B. Hou, "Continuous-time Markov decision processes programming with polynomial reward", Ph. D. Thesis, Institute of Appl. Math. Academic Sinica, Beijing, 1986.

[9] Z. T. Hou and X. P. Guo, *Markov decision processes* (Science and Technology Press of Hunan, Changsha, China, 1998).

[10] R. A. Howard, *Dynamic programming and Markov processes* (Wiley, New York, 1960).

[11] P. Kakuman, "Nondiscounted continuous time Markov decision processes with countable state space", *SIAM J. Control* **10** (1972) 210–220.

[12] M. Lembersky, "On maximal rewards and ε-optimal policies in continuous time Markov decision processes", *Ann. Stat.* **2** (1974) 159–169.

[13] R. L. Miller, "Finite state continuous time Markov decision processes with an infinite planning horizon", *J. Math. Anal. Appl.* **22** (1968) 552–569.

[14] M. L. Puterman, *Markov decision processes* (John Wiley & Sons, 1994).

[15] L. I. Sennott, "Average cost optimal stationary policies in infinite state Markov decision processes with unbounded cost", *Oper. Res.* **37** (1989) 623–633.

[16] L. I. Sennott, "Another set of conditions for average optimality in Markov decision processes", *Systems Control Lett.* **24** (1995) 147–151.

[17] R. Serfozo, "Optimal control of random walks, birth and death processes, and queues", *Adv. Appl. Prob.* **13** (1981) 61–83.

[18] J. S. Song, "Continuous time Markov decision programming with non-uniformly bounded transition rate", *Scientia Sinica* **12** (1987) 1258–1267.

[19] C. B. Wu, "Continuous time Markov decision processes with unbounded reward and non-uniformly bounded transition rate under discounted criterion", *Acta Math. Appl. Sinica* **20** (1997) 196–208.

[20] A. A. Yushkevich and E. A. Feinberg, "On homogeneous Markov model with continuous time and finite or countable state space", *Theory, Prob. Appl.* **24** (1979) 156–161.

[21] S. H. Zheng, "Continuous time Markov decision programming with average reward criterion and unbounded reward rate", *Acta Math. Appl. Sinica* **7** (1991) 6–16.