Chapter 6. Long Run Average Cost Problem

Let $\{X_n\}$ be an MCP with state space S and control set $U(x)$ for $x \in S$. To fix notation, let $J_\beta(x; \{u_n\})$ be the cost associated with the infinite horizon β-discount problem; i.e.,

$$J_\beta(x; \{u_n\}) \doteq E_x \left[\sum_{j=0}^{\infty} \beta^j c(X_j; u_j) \right].$$

Its value function, which is the infimum of J_β over all control $\{u_n\}$, is denoted by $V_\beta(x)$. Our main interest in this chapter is the *long run average cost.* More precisely, we will define

$$I(x; \{u_n\}) \doteq \limsup_{n \to \infty} \frac{1}{n} E_x \left[\sum_{j=0}^{n-1} c(X_j; u_j) \right].$$

The goal is to minimize this quantity, and the value function is denoted by $\Lambda(x)$; that is,

$$\Lambda(x) \doteq \inf_{\{u_n\}} I(x; \{u_n\}).$$

Remark **1** Sometimes, we will also define another long run average cost

$$I_*(x; \{u_n\}) \doteq \liminf_{n \to \infty} \frac{1}{n} E_x \left[\sum_{j=0}^{n-1} c(X_j; u_j) \right].$$

But mostly, we will focus on cost function I and value function Λ. ■

We will assume throughout that

Condition **1** *The running cost c is non-negative.*

Interested students may want to construct a control problem and a control policy such that $I \neq I_*$.

1 Finite state MCP with finite controls

We have the following useful lemma connecting the infinite horizon problem with the long-run average cost problem.

Lemma 1 *Under condition 1, for any control process $\{u_n\}$ and initial state $x \in S$ we have*

$$I_*(x; \{u_n\}) \le \liminf_{\beta \uparrow 1} (1-\beta) J_\beta(x; \{u_n\}) \le \limsup_{\beta \uparrow 1} (1-\beta) J_\beta(x; \{u_n\}) \le I(x; \{u_n\}).$$

Furthermore, if

$$\lim_{\beta \uparrow 1} (1-\beta) J_\beta(x; \{u_n\})$$

exists and is finite, then all the inequalities are equalities.

Proof. The proof is the direct consequence of the Tauberian theorem. ■.

From this lemma, we have the following lemma concerning the stationary policies of a finite state MCP with finite controls.

Lemma 2 *Suppose $\{X_n\}$ is an MCP with finite state space S and for every x, the control set $U(x)$ is finite. Let $\phi : x \mapsto U(x)$ be a stationary control policy; that is $u_n = \phi(X_n)$. Then*

$$I(x; \phi) = I_*(x; \phi) = \lim_{\beta \uparrow 1} (1-\beta) J_\beta(x; \phi).$$

Proof. The quantity $J_\beta(x; \phi)$, viewed as a function of β, is rational; see the proof of Theorem 1 in Chapter 5. Furthermore, thanks to finiteness, the cost c is bounded from above, say by constant C. It follows that

$$(1-\beta) J_\beta(x; \phi) \le (1-\beta) \sum_{j=0}^{\infty} C\beta^j = C.$$

It follows that $(1-\beta) J_\beta(x; \phi)$ converges as $\beta \uparrow 1$ and the limit is finite. We complete the proof by Lemma 1. ■

The following result connects the Blackwell optimality and the long-run average cost optimality.

Theorem 1 *Suppose $\{X_n\}$ is an MCP with finite state space S and for every x, the control set $U(x)$ is finite. Then the Blackwell optimal stationary policy, say ϕ^*, is optimal for the long run average cost; that is $\Lambda(x) = I(x; \phi^*)$.*

Moreover, if we further assume that for any $x, y \in S$, there exists a stationary policy such that $x \to y$ under this stationary policy. Then $\Lambda(x) \equiv \lambda$ for some constant λ.

Proof. Let $\{u_n\}$ be an arbitrary control. We have

$$I(x;\{u_n\}) \geq \limsup_{\beta\uparrow 1}(1-\beta)J_\beta(x;\{u_n\}) \geq \limsup_{\beta\uparrow 1}(1-\beta)J_\beta(x;\phi^*) = I(x;\phi^*).$$

This yields that $I(x;\phi^*) = \Lambda(x)$.

Let $x, y \in S$ be such that x and y achieve the maximum and the minimum respectively over $\Lambda(z)$ over $z \in S$. Assume that $x \to y$ under the stationary policy ϕ, and without incurring confusion, let P denote the transition probability matrix under policy ϕ. It follows that there exists $k \in \mathbb{N}$ such that $P^{(k)}_{xy} > 0$.

Let M denote the upper bound for the running cost c. The DPE for the infinite horizon discount problem implies that

$$V_\beta(x) \leq c(x;\phi(x)) + \beta \sum_{z\in S} P_{xz} V_\beta(z) \leq M + \sum_{z\in S} P_{xz} V_\beta(z).$$

Repeating this, we arrive at

$$V_\beta(x) \leq kM + \sum_{z\in S} P^{(k)}_{xz} V_\beta(z).$$

Multiplying both sides by $(1-\beta)$ and letting $\beta \uparrow 1$, it follows from the definition of x and y that

$$\Lambda(x) \leq \sum_{z\in S} P^{(k)}_{xz}\Lambda(z) \leq P^{(k)}_{xy}\Lambda(y) + [1 - P^{(k)}_{xy}]\Lambda(x),$$

or equivalently $\Lambda(x) \leq \Lambda(y)$. This completes the proof. ■

It could happen that the value function $\Lambda(x)$ is not a constant across $x \in S$, as the following pathological example shows.

Example 1 Consider an MCP with state space $S = \{0, 1, 2, 3\}$. The states $\{1, 2, 3\}$ is uncontrolled, and transition probability matrix satisfies $P_{11} = P_{23} = P_{32} = 1$. At state 0, $U(0) = \{A, B, C\}$ with policy A gives

$$P^A_{01} = P^A_{03} = 1/2, \quad P^B_{01} = P^B_{03} = 1/2, \quad P^C_{01} = 1/8 = 1 - P^C_{03}.$$

The associated cost is $c(1) = 0$, $c(2) = 2$, $c(3) = 1$, and

$$c(0;A) = 0, \quad c(0;B) = 1, \quad c(0;C) = 1/8.$$

The corresponding minimization problem is easy to solve. Let V_β denote the value function for the β-discount problem. Clearly, $V_\beta(1) = 0$, $V_\beta(2) =$

$(2+\beta)/(1-\beta^2)$, and $V_\beta(3) = (2\beta+1)/(1-\beta^2)$. It can be shown that A is β-optimal and $V_\beta(0) = \beta(2\beta+1)/[2(1-\beta^2)]$.

It follows that

$$\Lambda : \{0,1,2,3\} \mapsto \left\{\frac{3}{4}, 0, \frac{3}{2}, \frac{3}{2}\right\},$$

which is not a constant. In this example clearly there is no stationary policy such that $1 \to 2$ or $1 \to 3$. ■

1.1 The DPE associated with the long-run average problem

The DPE for the long-run average problem is not as straightforward as the total expected cost problem. New quantities other than the value function is to be introduced into the equation.

Theorem 2 *Assume that the state space S is finite and that $U(x)$ is finite for every $x \in S$. Suppose $\Lambda(x) \equiv \lambda$ for some constant λ. Then there exists a vector $h : S \to \mathbb{R}$ such that*

$$\lambda + h(x) = \min_{u \in U(x)} \left[c(x;u) + \sum_{y \in S} P_{xy}(u) h(y) \right],$$

where $P(u)$ denotes the transition probability matrix under control u, and the minimum is attained at $u = \phi^*(x)$.

Conversely, suppose there exists a constant λ and a vector h satisfies the above equation, then $\Lambda(x) \equiv \lambda$ and the minimizer $u^* : x \to U(x)$ defines an optimal (stationary) policy.

Proof. Let ϕ^* be the Blackwell optimal policy and $P \doteq P(\phi^*)$ the probability transition probability matrix corresponding to ϕ^*. It follows that for β close to one,

$$V_\beta(x) = c(x;\phi^*(x)) + \beta \sum_{y \in S} P_{xy}(\phi^*) V_\beta(y)$$

Or equivalently, define vectors $V_\beta \doteq [V_\beta(x)]$ and $C = [c(x;\phi^*(x))]$, then

$$V_\beta = (I - \beta P)^{-1} C.$$

A result from linear algebra (see Proposition 3) asserts that there exists matrices Q and P such that

$$(1-\beta)(I-\beta P)^{-1} = Q + (1-\beta)H + o(1-\beta).$$

It follows that

$$(1-\beta)V_\beta = QC + (1-\beta)HC + o(1-\beta).$$

Letting $\beta \to 1$ we have $QC = \Lambda \equiv \lambda$. Define $h = HC$, we have

$$V_\beta = (1-\beta)^{-1}\lambda + h + o(1),$$

which yields

$$(1-\beta)^{-1}\lambda + h(x) = c(x;\phi^*(x)) + \beta \sum_{y \in S} P_{xy} \left[(1-\beta)^{-1}\lambda + h(y)\right] + o(1),$$

or equivalently,

$$\lambda + h(x) = c(x;\phi^*(x)) + \beta \sum_{y \in S} P_{xy} \left[h(y) + o(1)\right].$$

Letting $\beta \uparrow 1$, we have

$$\lambda + h(x) = c(x;\phi^*(x)) + \sum_{y \in S} P_{xy}(\phi^*) h(y).$$

As for the inequality

$$\lambda + h(x) \le c(x;u) + \sum_{y \in S} P_{xy}(u) h(y),$$

given an arbitrary $u \in U(x)$. The proof is exactly the same and thus omitted.

The second half of the theorem is just a verification argument. For each control $\{u_n\}$, construct the process

$$Z_n \doteq \sum_{j=0}^{n-1} c(X_j;u_j) + h(X_n) - n\lambda,$$

which is in general a submartingale. We omit the details. ■

Corollary 1 *Suppose the MCP $\{X_n\}$ has finite state space S and the control set $U(x)$ is finite for every $x \in S$. Assume that $\Lambda(x) \equiv x$. Then*

$$\lim_n \frac{v_n(x)}{n} \equiv \lambda.$$

Here v_n is the value function for the corresponding finite horizon problem with horizon n and without discounting.

The proof again is the same verification argument and thus omitted.

1.2 Probabilistic interprtation of h

The function $h : S \to \mathbb{R}$ has some nice probabilistic interpretations. For simplicity, we should assume the following condition throughout this subsection.

Condition **2** *The MCP $\{X_n\}$ is irreducible under arbitrary stationary control policy.*

Under this condition, clearly $\Lambda(x) \equiv \lambda$ for some constant λ; see Theorem 1. This further implies that there exist a vector h such that

$$\lambda + h(x) = \min_{u \in U(x)} \left[c(x; u) + \sum_{y \in S} P_{xy}(u) h(y) \right]. \tag{1}$$

We have the following result regarding the uniqueness of the vector h.

Proposition 1 *Suppose $\{X_n\}$ has a finite state space and for each $x \in S$ the control set $U(x)$ is finite. If Condition 2 holds, then the vector h, as a solution to the DPE (1), is unique up to an additive constant.*

Proof. Suppose (λ, h) and $(\lambda, \bar{h})$ are both solutions to the DPE (1). Assume that $u^* : x \mapsto U(x)$ is the optimal policy (i.e., minimizer) corresponding to (λ, h). We have

$$\begin{aligned} \lambda + h(x) &= c(x; u^*(x)) + \sum_{y \in S} P_{xy}(u^*) h(y) \\ \lambda + \bar{h}(x) &\leq c(x; u^*(x)) + \sum_{y \in S} P_{xy}(u^*) \bar{h}(y). \end{aligned}$$

Letting $s \doteq \bar{h} - h$, it follows that

$$s(x) \leq \sum_{y \in S} P_{xy}(u^*) s(y).$$

Repeating this to obtain that for any n,

$$s(x) \leq \sum_{y \in S} \left[\frac{1}{n} \sum_{j=0}^{n-1} P^j_{xy}(u^*) \right] s(y) \ \rightarrow \ \sum_{y \in S} \pi(y) s(y),$$

where π is the stationary distribution corresponding to the probability transition matrix $P(u^*)$. Due to irreducibility, each component of the vector π

is non-zero. This clearly implies that s is a constant. This completes the proof. ■

Assume that the conditions of Proposition 1 hold. In this case, h admits the following representation. Let $\phi^* : x \mapsto U(x)$ be the Blackwell optimal policy. Consider the MCP $\{X_n\}$ under the transition probability matrix $P(\phi^*)$. The proof of Theorem 2 implies that ϕ^* is a minimizer of DPE corresponding to h.

Fix a generic state $z \in S$. For an arbitrary state $x \in S$, define

$$\tau_x \doteq \inf\{n \geq 0 : \; X_n = z \mid X_0 = x\}.$$

Then

$$h(x) = h(z) + E_x \sum_{j=0}^{\tau_x - 1} [c(X_j; \phi^*(X_j)) - \lambda]. \tag{2}$$

The remainder of this subsection is to prove this representation (2). We should define h as in (2) with $h(z)$ an arbitrary constant. We need to show that (λ, h) solves the DPE. Assume from now on $x \neq z$, and $X_0 = x$. We have, for β close to one enough,

$$V_\beta(x) = E_x \left[\sum_{j=0}^{\tau_x - 1} \beta^j c(X_j; \phi^*(X_j)) \right] + E_x \left[\beta^{\tau_x} V_\beta(z) \right].$$

It follows that

$$V_\beta(x) - V_\beta(z) = E_x \left[\sum_{j=0}^{\tau_x - 1} \beta^j c(X_j; \phi^*(X_j)) \right] - (1-\beta) V_\beta(z) E_x \left[\frac{1 - \beta^{\tau_x}}{1 - \beta} \right].$$

Letting $\beta \uparrow 1$, we have

$$\lim_{\beta \uparrow 1} [V_\beta(x) - V_\beta(z)] = E_x \sum_{j=0}^{\tau_x - 1} [c(X_j; \phi^*(X_j)) - \lambda] = h(x) - h(z).$$

The DPE for V_β implies that

$$(1-\beta) V_\beta(x) = \inf_{u \in U(x)} \left[c(x; u) + \beta \sum_{y \in S} P_{xy}(\phi^*) \left[V_\beta(y) - V_\beta(x) \right] \right].$$

Letting $\beta \to 1$, we have

$$\lambda = \inf_{u \in U(x)} \left[c(x; u) + \sum_{y \in S} P_{xy}(\phi^*) \left[h(y) - h(x) \right] \right].$$

This says that h satisfies the DPE. ■

Sometimes the difference $x \mapsto V_\beta(x) - V_\beta(z)$ is called the *relative value function*. Thus h also characterizes the limit of this relative value function as $\beta \uparrow 1$.

1.3 Computational issues

Corollary 1 gives an algorithm for computing the value function $\Lambda(x)$. However, such an algorithm suffers from slow convergence, lack of information on $h : S \to \mathbb{R}$ and the optimal policy. The convergence of infinite horizon value function to the long-run cost value function and the Blackwell optimal policy being long-run cost optimal give rise to another possible algorithm. But it is usually very inconvenient working with the infinite horizon problems.

We consider an alternative computation method based on the finite horizon problem v_n. Fix a state $z \in S$, and define $h_n(x) \doteq v_n(x) - v_n(z)$. The DPE for $\{v_n\}$ gives that, for every $x \in S$,

$$[v_{n+1}(z) - v_n(z)] + h_{n+1}(x) = \inf_{u \in U(x)} \left[c(x;u) + \sum_{y \in S} P_{xy}(u) h_n(y) \right].$$

Suppose we can show that $\{v_{n+1}(z) - v_n(z)\}$ and $\{h_n\}$ converge, then the limit pair is a solution to the DPE associated with the long-run average cost problem, whence the limit of the former is the value for the long-run average problem and the limit of $\{h_n\}$ is just h with $h(z) = 0$.

Condition **3** *Every stationary control policy leads to a Markov process $\{X_n\}$ that is irredcucible and aperiodic.*

Proposition 2 *Suppose Condition 3 holds. Thus $\Lambda(x) \equiv \lambda$ and let (λ, h) is a solution pair to the DPE associated with the long-run average cost problem such that $h(z) = 0$. Then for every $z \in S$,*

$$\lim_n [v_{n+1}(z) - v_n(z)] = \lambda,$$

and for every $x \in S$,

$$\lim_n h_n(x) = \lim_n [v_n(x) - v_n(z)] = h(x).$$

Proof. Define $F_n(x) \doteq n\lambda + h(x) - v_n(x)$. We first show that $\{F_n\}$ is uniformly bounded. The lower bound is just a verification theorem (very similar to Corollary 1. As for the upper bound, let $\{u_n^*\}$ be the optimal policy for the finite-horizon problem with horizon n. It is clear that

$$V_\beta(x) \le v_n(x) + \beta^n E[V_\beta(X_n)].$$

Thus

$$V_\beta(x) - V_\beta(z) \le v_n(x) + \beta^n E[V_\beta(X_n) - V_\beta(z)] - \frac{1-\beta^n}{1-\beta}(1-\beta)V_\beta(z).$$

Letting $\beta \uparrow 1$, we have

$$h(x) \le v_n(x) + Eh(X_n) - n\lambda.$$

The upper bound of $\{F_n\}$ follows readily.

Let (λ, h) solve the DPE with u^* the minimizer of the DPE (whence optimal). It follows that

$$\lambda + h(x) = c(x; u^*(x)) + \sum_{y \in S} P_{xy}(u^*)h(y).$$

However, we also have

$$v_{n+1}(x) \le c(x; u^*(x)) + \sum_{y \in S} P_{xy}(u^*)v_n(y).$$

It follows then

$$F_{n+1}(x) \ge \sum_{y \in S} P_{xy}(u^*)F_n(y),$$

which further implies that

$$F_{n+m}(x) \ge \sum_{y \in S} P_{xy}^m(u^*)F_n(y).$$

Consider now an arbitrary subsequence of $\{F_n\}$, say $\{F_{n_k}\}$, that converges, say to a vector $F : S \to \mathbb{R}$. We first show that F is constant valued. Fix $j \in \mathbb{N}$ and let $k \ge j$. It follows that

$$F_{n_k}(x) \ge \sum_{y \in S} P_{xy}^{n_k - n_j}(u^*)F_{n_j}(y).$$

Letting $k \to \infty$ we have

$$F(x) \ge \sum_{y \in S} \pi(y)F_{n_j}(y),$$

where π is the stationary distribution corresponds to transition probability matrix $P(u^*)$. Now letting $j \to \infty$, we have

$$F(x) \geq \sum_{y \in S} \pi(y) F(y).$$

Thus it follows easily that F is constant valued.

Suppose now that $F_{n_k} \to F$ and $F_{m_k} \to F'$, where $\{n_k\}$ and $\{m_k\}$ are two subsequences. A similar argument yields that

$$F \geq \sum_{y \in S} \pi(y) F'(y) = F'.$$

Exchange the role of F and F' we have $F' \geq F$, or $F = F'$. This implies that $\{F_n\}$ converges to vector F. The claims of proposition follow trivially. ■

The algorithm goes as follows: Fix a state $z \in S$ and set $h_0(x) \equiv 0$, and then compute

$$\inf_{u \in U(x)} \left[c(x; u) + \sum_{y \in S} P_{xy}(u) h_n(y) \right] \doteq W_{n+1}(x).$$

It follows that $h_{n+1}(x) = W_{n+1}(x) - W_{n+1}(z)$ and $v_{n+1}(z) - v_n(z) = W_{n+1}(z)$.

Remark **2** We have not discussed the optimal policy yet. Indeed, one can easily show that any limit point of the minimizing policy $\{u_n\}$ in the above infimum is long-run average optimal. We say ϕ is a limit point of $\{u_n\}$ if there exists a subsequence $\{n_k\}$ such that $u_{n_k}(x) \equiv \phi(x)$ for sufficiently large k. Due to the finiteness, such a limit point always exists.

A Appendix. The proof of a Tauberian theorem

Consider a sequence of non-negative numbers $c_n \in [0, \infty]$. For every $\beta \in [0, 1)$, define

$$V(\beta) \doteq \sum_{j=0}^{\infty} c_j \beta^j$$

and

$$S_n \doteq \sum_{j=0}^{n-1} c_j.$$

The Tauberian theorem states that

Theorem 3 *The following inequality holds.*

$$\liminf_n \frac{S_n}{n} \leq \liminf_{\beta \uparrow 1} (1-\beta) V(\beta) \leq \limsup_{\beta \uparrow 1} (1-\beta) V(\beta) \leq \limsup_n \frac{S_n}{n}.$$

Furthermore, if

$$\lim_{\beta \uparrow 1} (1-\beta) V(\beta)$$

exists and is finite, then all the inequalities are indeed equalities.

Proof. Without loss of generality, we can assume that every c_n is finite, otherwise all the quantities are infinity.

It follows that

$$V(\beta) = \sum_{j=0}^{\infty} c_j \beta^j = \sum_{j=0}^{\infty} (S_{j+1} - S_j) \beta^j = \sum_{j=1}^{\infty} (\beta^{j-1} - \beta^j) S_j.$$

For every $N \in \mathbb{N}$, we have

$$\begin{aligned} V(\beta) &\leq \sum_{j=1}^{N-1} (\beta^{j-1} - \beta^j) S_j + \left(\sup_{j \geq N} \frac{S_j}{j} \right) \sum_{j=N}^{\infty} (\beta^{j-1} - \beta^j) j \\ &\leq \sum_{j=1}^{N-1} (\beta^{j-1} - \beta^j) S_j + \left(\sup_{j \geq N} \frac{S_j}{j} \right) \sum_{j=0}^{\infty} (\beta^{j-1} - \beta^j) j \\ &= \sum_{j=1}^{N-1} (\beta^{j-1} - \beta^j) S_j + \left(\sup_{j \geq N} \frac{S_j}{j} \right) \frac{1}{1-\beta}. \end{aligned}$$

This yields

$$\limsup_{\beta \uparrow 1} (1-\beta) V(\beta) \leq \sup_{j \geq N} \frac{S_j}{j}$$

for every N. Letting $N \to \infty$, we complete the rightmost inequality. Similarly, we have for every $N \in \mathbb{N}$,

$$V(\beta) \geq \left(\inf_{j \geq N} \frac{S_j}{j}\right) \sum_{j=N}^{\infty} (\beta^{j-1} - \beta^j) j \geq \left(\inf_{j \geq N} \frac{S_j}{j}\right) \cdot \frac{\beta^N}{1-\beta}.$$

Letting $\beta \to 1$ and then $N \to \infty$ we complete the leftmost inequality.

Now assume that

$$\lim_{\beta \uparrow 1} (1-\beta) V(\beta) \doteq L$$

is finite. We claim that for every continuous function F on interval $[0, 1]$,

$$\lim_{\beta \uparrow 1} (1-\beta) \sum_{j=0}^{\infty} c_j F(\beta^j) \beta^j = L \int_0^1 F(x)\, dx. \tag{3}$$

This is true for every polynomial. Indeed, for $F(x) = x^n$ with $n \in \mathbb{N}$, we have

$$\begin{aligned} \lim_{\beta \uparrow 1} (1-\beta) \sum_{j=0}^{\infty} c_j F(\beta^j) \beta^j &= \lim_{\beta \uparrow 1} \frac{1-\beta}{1-\beta^{n+1}} (1-\beta^{n+1}) \sum_{j=0}^{\infty} c_j (\beta^{n+1})^j \\ &= L \lim_{\beta \uparrow 1} \frac{1-\beta}{1-\beta^{n+1}} = \frac{L}{n+1} = L \int_0^1 x^n dx. \end{aligned}$$

Since every continuous function can be uniformly approximated by polynomials, the claim follows easily from the standard $\varepsilon - \delta$ language.

Now consider a function

$$F(x) \doteq \begin{cases} 0 & ; \text{ if } 0 \leq x < e^{-1} \\ 1/x & ; \text{ if } e^{-1} \leq x \leq 1 \end{cases}.$$

F is not continuous, but we still have the equality (3). The reason is that one can always find sequences of continuous functions $\{g_n\}$ and $\{G_n\}$ such that $g_n \leq F \leq G_n$ and

$$\lim_n \int_0^1 g_n(x) dx = \lim_n \int_0^1 G_n(x) dx = \int_0^1 F(x) dx = 1.$$

For this function F, we have

$$(1-\beta) \sum_{j=0}^{\infty} c_j F(\beta^j) \beta^j = (1-\beta) \sum_{j=0}^{\lfloor -(\log \beta)^{-1} \rfloor} c_j = (1-\beta) S_{\lfloor -(\log \beta)^{-1} \rfloor}.$$

Let $\beta = e^{-1/n}$, we have

$$(1-\beta)\sum_{j=0}^{\infty} c_j F(\beta^j)\beta^j = (1-e^{-1/n})S_n.$$

Letting $n \uparrow \infty$, we have $\beta \uparrow 1$, and

$$L = L\int_0^1 F(x)dx = \lim_n (1-e^{-1/n})S_n = \lim_n \frac{S_n}{n}.$$

This completes the proof. ■

B Appendix. A result from linear algebra

Proposition 3 *Let P be an arbitrary probability transition matrix, and $\beta \in (0,1)$. Then*

$$\lim_{\beta\uparrow 1}\left[(I-\beta P)^{-1} - \frac{Q}{1-\beta} - H\right] = 0,$$

with

$$Q = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} P^k, \qquad H = (I-P+Q)^{-1} - Q.$$

Furthermore,

$$Q + H = I + PH = I + HP.$$

Proof. Consider the matrix

$$M(\beta) \doteq (1-\beta)(I-\beta P)^{-1}.$$

Clearly each component of M is a rational function of β. Since

$$(I-\beta P)^{-1} = I + \beta P + \beta^2 P^2 + \cdots,$$

each component of $M(\beta)$ is bounded by

$$(1-\beta)(1+\beta+\beta^2+\cdots) = 1.$$

Thus, as $\beta \to 1$, $M(\beta)$ has a limit. Define

$$Q \doteq \lim_{\beta\uparrow 1} M(\beta).$$

The Taylor expansion then implies, for some matrix H,

$$M(\beta) = Q + (1-\beta)H + o(1-\beta).$$

Here $o(1)$ denotes a matrix converges to 0 as $\beta \uparrow 1$. It follows that

$$H = \lim_{\beta \uparrow 1} \left[(I - \beta P)^{-1} - \frac{Q}{1-\beta} \right].$$

Multiplying both sides by $(I - \beta P)$ to the right, we have

$$H - \beta HP = I - \lim_{\beta \uparrow 1} \frac{Q - \beta QP}{1 - \beta}.$$

We have $Q = QP$, and

$$H + Q = I + HP.$$

Analogously, multiplying both sides by $(I-\beta P)$ to the left, we have $Q = PQ$ and

$$H + Q = I + PH.$$

This gives $P^n H + P^n Q = P^n + P^{n+1} H$, or

$$Q = P^n + (P^{n+1} - P^n) H.$$

Thus

$$nQ = \sum_{j=0}^{n-1} P^j + (P^n - I) H.$$

Dividing both sides by n and letting $n \to \infty$, we have

$$Q = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} P^j,$$

since $(P^n - I)H/n \to 0$.

Finally, since

$$(I - \beta P) Q = Q - \beta PQ = (1 - \beta) Q,$$

we have $Q = M(\beta) Q$. Letting $\beta \uparrow 1$ we have $Q = Q^2$. It is easy to see now that

$$(P - Q)^n = P^n - Q.$$

This in turn implies that

$$\begin{aligned} H &= \lim_{\beta \uparrow 1} \left[(I - \beta P)^{-1} - \frac{Q}{1-\beta} \right] = \lim_{\beta \uparrow 1} \sum_{j=0}^{\infty} \beta^j (P^j - Q) \\ &= \lim_{\beta \uparrow 1} \left[I - Q + \sum_{j=1}^{\infty} \beta^j (P - Q)^j \right] = \lim_{\beta \uparrow 1} [I - \beta(P - Q)]^{-1} - Q \\ &= (I - P + Q)^{-1} - Q. \end{aligned}$$

This completes the proof. ■