Chapter 7

AVERAGE COST MODELS WITH LOST SALES

7.1. Introduction

This chapter is concerned with the long-run average cost minimization of a stochastic inventory problem with Markovian demand, fixed ordering cost, and convex surplus cost in the case of lost sales. The formulation of the problem is similar to that introduced in Chapter 4 except that we replace the discounted cost objective function by the long-run average cost objective function. To deal with this average cost problem, we apply the vanishing discount method to solve the dynamic programming equations defined for the problem, and establish the corresponding verification theorem.

The plan of this chapter is as follows. Required results for the discounted cost model derived in Chapter 4 are recapitulated in Section 7.3. In the next section, we provide a precise formulation of the problem. In Section 7.4, we obtain the asymptotic behavior of the differential discounted value function as the discount rate goes to zero. The vanishing discount approach to establish the average cost optimality equation is developed in Section 7.5. The associated verification theorem is proved in Section 7.6, and the theorem is used to show that a state-dependent (s, S) policy is optimal for the problem. Section 7.7 concludes the chapter with suggestions for future research.

7.2. Formulation of the Model

Consider an inventory problem over an infinite horizon. The demand in each period is assumed to be a random variable defined on a given probability space, and not necessarily identically distributed. To precisely define the demand process, we consider a finite collection

D. Beyer et al., *Markovian Demand Inventory Models*, International Series in Operations Research & Management Science 108,
DOI 10.1007/978-0-387-71604-6_7,

of demand states labeled $i \in \mathbb{I} = \{1, 2, \ldots, L\}$, and let i_k denote the demand state observed at the beginning of period k. We assume that i_k, $k = 0, 1, 2, \ldots$, is a Markov chain over $\mathbb{I}$, with the transition matrix $P = \{p_{ij}\}$. Thus,

$$0 \leq p_{ij} \leq 1, \; i \in \mathbb{I}, j \in \mathbb{I}, \;\; \text{and} \;\; \sum_{j=1}^{L} p_{ij} = 1, \; i \in \mathbb{I}.$$

Let the nonnegative random variable ξ_k denote the demand at the end of a given period $k \in \langle 0, N{-}1 \rangle$. Demand ξ_k depends only on the demand state in period k, by which we mean that it does does not depend on k and is independent of past demand states and past demands. We denote its probability density by $\varphi_i(x)$ and its probability distribution by $\Phi_i(x)$ when the demand state $i_k = i$.

We suppose that orders are placed at the beginning of a period, delivered instantaneously, and followed by the period's demand. Unsatisfied demands are lost.

In what follows, we list the assumptions that are needed to derive the main results of the chapter in Sections 7.5 and 7.6. Because not all the results proved in this chapter require all of these assumptions, we label them as follows so that we can specify the assumptions required in the statements of the specific results proved in this chapter.

(i) The production cost is given by $c(i, u) = K\mathbb{1}_{u>0} + c_i u$, where $K \geq 0$ is the fixed ordering cost and $c_i \geq 0$ is the variable cost.

(ii) For each i, the inventory cost function $f(i, \cdot)$ is convex, nondecreasing and of linear growth, i.e., $f(i, x) \leq C_f(1 + |x|)$ for some $C_f > 0$ and all x. Also, $f(i, x) = 0$ for all $x \leq 0$.

(iii) For each i, the shortage cost function $q(i, \cdot)$ is convex, nonincreasing and of linear growth, i.e., $q(i, x) \leq C_q(1 + |x|)$ for some $C_q > 0$ and all x. Also, $q(i, x) = 0$ for all $x \geq 0$.

(iv) There is a state $g \in \mathbb{I}$ such that $f(g, x)$ is not identically zero.

(v) The production and inventory costs satisfy for all i,

$$c_i x + \sum_{j=1}^{L} p_{ij} \int_0^{\infty} f(j, (x - z)^+) d\Phi_i(z) \to \infty \text{ as } x \to \infty. \tag{7.1}$$

(vi) The Markov chain $(i_k)_{k=0}^{\infty}$ is irreducible.

(vii) There is a state $h \in \mathbb{I}$ such that $1 - \Phi_h(\varepsilon) = \rho > 0$ for some $\varepsilon > 0$.

(viii) For each i, the inequality $q^{'-}(i,0) \leq \bar{f}^{'+}(i,0) - \bar{c}^i$ holds.

(ix) $\mathsf{E}\,\xi_k \leq D < \infty$ for all k.

REMARK 7.1 Assumptions (i)–(iii) reflect the usual structure of the production and inventory costs to prove the optimality of an (s_i, S_i) policy. Note that K is the same for all i. In the stationary case, this is equivalent to the condition (2.18) required in the nonstationary model for the existence of an optimal (s_i, S_i) policy; (see Chapter 2). Assumption (iv) rules out trivial cases where the optimal policy is never to order. Assumption (v) means that either the unit ordering cost $c_i > 0$ or the second term in (7.1), which is the expected holding cost, or both, go to infinity as the surplus level x goes to infinity. While related, Assumption (v) neither implies nor is implied by Assumption (iv). Assumption (v) is borne out of practical considerations and is not very restrictive. In addition, it rules out such unrealistic trivial cases as the one with $c_i = 0$ and $f(i,x) = 0, x \geq 0$, for each i, which implies ordering an infinite amount whenever an order is placed. Assumptions (iv) and (v) generalize the usual assumption made by Scarf (1960) and others that the unit inventory holding cost $h > 0$. Assumption (viii) replaces the more stringent assumption in the classical literature that the unit shortage cost is not smaller than the unit price.

REMARK 7.2 Assumptions (vi) and (vii) are needed to deplete any given initial inventory in a finite expected time. While Assumption (vii) says that in at least one state h, the expected demand is strictly larger than zero, Assumption (vi) implies that the state h would occur infinitely often with finite expected intervals between successive occurrences.

REMARK 7.3 Assumption (viii) means that the marginal shortage cost in one period is larger than or equal to the expected unit ordering cost less the expected marginal inventory holding cost in any state of the next period. If this condition does not hold, that is, if $-q_n^{'-}(i,0) < \bar{c}^i_{n+1} - \bar{f}^{'+}_{n+1}(i,0)$ for some i, a speculative retailer may find it attractive to meet a smaller part of the demand in period n than is possible from the available stock, carry the leftover inventories to period $(n+1)$, and order a little less as a result in period $(n+1)$, with the expectation that he will be better off. Thus, Assumption (viii) rules out this kind of speculation on the part of the retailer. But such a speculative behavior is not allowed in our formulation of the dynamics in any case, since the demand in any period must be satisfied to the extent of the availability of inventories. This suggests that it might be possible to prove our results without Assumption (viii).

The objective is to minimize the expected long-run average cost

$$J(i,x;U) = \limsup_{N\to\infty} \frac{1}{N}\mathsf{E}\Big\{ \sum_{k=0}^{N-1} [c(i_k,u_k) + f(i_k,x_k) + q(i_k, x_k + u_k - \xi_k)] \Big\}, \quad (7.2)$$

with $i_0 = i$ and $x_0 = x \geq 0$, where $U = (u_0, u_1, \ldots)$, $u_i \geq 0$, $i = 0, 1, \ldots,$ is a *history-dependent* or *nonanticipative* decision (order quantities) for the problem. Such a control U is termed admissible. Let $\mathcal{U}$ denote the class of all admissible controls. The surplus balance equations are given by

$$x_{k+1} = (x_k + u_k - \xi_k)^+, \ k = 0, 1, \ldots. \quad (7.3)$$

Our aim is to show that there exist a constant λ^*, termed the optimal average cost, which is independent of the initial i and x, and a control $U^* \in \mathcal{U}$ such that

$$\lambda^* = J(i,x;U^*) \leq J(i,x;U) \quad \text{for all} \quad U \in \mathcal{U}, \quad (7.4)$$

and

$$\lambda^* = \lim_{N\to\infty} \frac{1}{N}\mathsf{E}\Big\{ \sum_{k=0}^{N-1} [c(i_k,u_k^*) + f(i_k,x_k^*) + q(i_k, x_k + u_k - \xi_k)] \Big\}, \quad (7.5)$$

where x_k^*, $k = 0, 1, \ldots$, is the surplus process corresponding to U^* with $i_0 = i$ and $x_0 = x$.

To prove these results, we will use the vanishing discount approach. That is, by letting the discount factor α in the discounted cost problem approach one, we will show that we can derive a dynamic programming equation whose solution provides an average optimal control and the associated minimum average cost λ^*.

For this purpose, we recapitulate relevant results for the discounted cost problem obtained in Chapter 4.

Remark 7.4 Note that the objective function (7.2) is slightly, but not essentially, different from that used in the classical literature. Whereas we base the surplus cost on the initial surplus in each period, the usual practice in the literature is to charge the cost on the ending surplus levels, which means to have $f(i_k, x_{k+1})$ instead of $f(i_k, x_k)$ in (7.2). Note that

x_{k+1} is also the ending inventory in period k. It should be obvious that this difference in the objective functions does not change the long-run average cost for any admissible policy. By the same token, we can justify our choice to charge shortage costs at the end of a given period.

7.3. Discounted Cost Model Results from Chapter 4

Consider the model formulated above with the average cost objective (7.2) replaced by the extended real-valued objective function

$$J^{\alpha}(i,x;U) = \sum_{k=0}^{\infty} \alpha^k \mathsf{E}[c(i_k,u_k)+f(i_k,x_k)+q(i_k,x_k+u_k-\xi_k)], \quad 0 \le \alpha < 1. \tag{7.6}$$

Define the value function with $i_0 = i$ and $x_0 = x$ as

$$v^{\alpha}(i,x) = \inf_{U \in \mathcal{U}} J^{\alpha}(i,x;U). \tag{7.7}$$

Let $\mathbb{B}_0$ denote the class of all continuous functions from $I \times \mathbb{R}$ into $[0,\infty)$, and the pointwise limits of sequences of these functions; (see Feller (1971)). Note that it includes piecewise-continuous functions. Let $\mathbb{B}_1$ denote the space of functions in $\mathbb{B}_0$ that are of linear growth, i.e., for any $b \in \mathbb{B}_1$, $0 \le b(i,x) \le C_b(1+|x|)$ for some $C_b > 0$. Let $\mathbb{B}_2$ denote the subspace of functions in $\mathbb{B}_1$ that are uniformly continuous with respect to $x \in R$. For any $b \in \mathbb{B}_1$, we define

$$F(b)(i,y) = \sum_{j=1}^{L} p_{ij} \int_0^M b(j,(y-z)^+) d\Phi_i(z). \tag{7.8}$$

THEOREM 7.1 *Let Assumptions* (i)–(iii), (v), (viii), *and* (ix) *hold. Then, we have the following results.*

(a) *The value function $v^{\alpha}(\cdot,\cdot)$ is in $\mathbb{B}_2$, and it solves the dynamic programming equation*

$$\begin{aligned} v^{\alpha}(i,x) &= f(i,x) + \inf_{u \ge 0}\{c(i,u) + \mathsf{E}[q(i,x+u-\xi^i) \\ &\quad + \alpha \sum_{j=1}^{L} p_{ij} v^{\alpha}(j,(x+u-\xi^i)^+)]\} \\ &= f(i,x) + \inf_{u \ge 0}\{c(i,u) + \mathsf{E}q(i,x+u-\xi^i) \\ &\quad + \alpha F(v^{\alpha})(j,x+u)\}. \end{aligned} \tag{7.9}$$

(b) $v^\alpha(i,\cdot)$ *is* K*-convex and there are real numbers* (s_i^α, S_i^α), $s_i^\alpha \le S_i^\alpha$, *such that the feedback policy* $\hat{u}_k^\alpha(i,x) = (S_i^\alpha - x)\mathbb{1}_{x<s_i^\alpha}$ *is optimal.*

Proof. (Actually Theorem 7.1 has been stated but not proved in Chapter 4.) The proof of part (a) follows the lines of the proof of Theorem 3.3 in Chapter 3 by taking the limit of the n-period value function, for n tending to infinity. Part (b) immediately follows since the limit of a sequence of K-convex functions is K-convex. □

7.4. Limiting Behavior as the Discount Factor Approaches 1

Hereafter, we will omit the additional superscript α on the control policies for ease of notation. Thus, for example, $\hat{u}_k^\alpha(i,x)$ will be denoted simply as $\hat{u}_k(i,x)$. Since we do not consider the limits of the control variables as $\alpha \to 1$, the practice of omitting the superscript α will not cause any confusion. In any case, the dependence of controls on α will always be clear from the context.

To insure a "smooth" limiting behavior for $\alpha \to 1$, we prove in Lemma 7.2 that $v^\alpha(i,\cdot)$ is locally equi-Lipschitzian. For this we need some notation and a preliminary result. For any $y > 0$, let

$$\tau_y := \inf\{n : \sum_{k=0}^{n} \xi_k \ge y\}$$

be the first index for which the cumulative demand is not less than y. The following required result is proved in Chapter 5.

Lemma 7.1 *Let Assumptions* (vii) *and* (viii) *hold. Then, for any* $l \in \mathbb{I}$, *we have* $\mathsf{E}(\tau_y | i_0 = l) < \infty$.

Lemma 7.2 *Under Assumptions* (i)-(iii), (vi) , (vii), *and* (ix), $v^\alpha(i,\cdot)$ *is locally equi-Lipschitzian, i.e., for* $X > 0$ *there is a positive constant* $C^1 < \infty$, *independent of* α, *such that*

$$|v^\alpha(i,x) - v^\alpha(i,\tilde{x})| \le C^1|x - \tilde{x}| \quad \textit{for all} \quad x, \tilde{x} \in [0, X]. \tag{7.10}$$

Proof. Consider the case $\tilde{x} \ge x$. Let us fix an $\alpha \in [0,1)$. It follows from Theorem 7.1 that there is an optimal feedback strategy U. Use the strategy U with initial surplus x, and the strategy $\tilde{U}$ defined by

$$\tilde{u}_k = [u_k - (\tilde{x}_k - x_k)]^+ = \begin{cases} 0 & \text{if} \quad u_k \le \tilde{x}_k - x_k, \\ u_k + x_k - \tilde{x}_k & \text{if} \quad u_k > \tilde{x}_k - x_k, \end{cases}$$

with initial $\tilde{x}$, x_k, and $\tilde{x}_k$ denoting the inventory levels resulting from the respective strategies. It is easy to see that the inequalities

$$0 \le \tilde{x}_k - x_k \le \tilde{x} - x \quad \text{and} \quad \tilde{u}_k \le u_k$$

hold for all k. Let $\tilde{\tau} := \tau_{\tilde{x}}$ for ease of notation. If $\tilde{u}_k = 0$ for all $k \in [0, \tilde{\tau}]$, then $\tilde{x}_{\tilde{\tau}} = x_{\tilde{\tau}} = 0$, and the two trajectories are identical for all $k > \tilde{\tau}$. If $\tilde{u}_{k'} \neq 0$ for some $k' \in [0, \tilde{\tau}]$, then $\tilde{x}_{k'} = x_{k'}$, and the two trajectories are identical for all $k > k'$. In any case, the two trajectories are identical for all $k > \tilde{\tau}$. From Assumptions (i)–(iii), we have

$$c(i_k, \tilde{u}_k) \le c(i_k, u_k),$$

$$|f(i_k, \tilde{x}_k) - f(i_k, x_k)| \le C_f |\tilde{x} - x|,$$

$$|q(i_k, \tilde{x}_k + \tilde{u}_k - \xi_k) - q(i_k, x_k + u_k - \xi_k)| \le C_q |\tilde{x} - x|.$$

Therefore,

$$\begin{aligned}
& v^\alpha(i, \tilde{x}) - v^\alpha(i, x) \\
& \quad \le \quad J^\alpha(i, \tilde{x}; \tilde{U}) - J^\alpha(i, x; U) \\
& \quad = \quad \mathsf{E}\Big(\sum_{k=0}^{\tilde{\tau}} \alpha^k (f(i_k, \tilde{x}_k) - f(i_k, x_k) + q(i_k, \tilde{x}_k + \tilde{u}_k - \xi_k) \\
& \qquad\qquad - q(i_k, x_k + u_k - \xi_k) + c(i_k, \tilde{u}_k) - c(i_k, u_k)) \Big) \\
& \quad \le \quad \mathsf{E}\Big(\sum_{k=0}^{\tilde{\tau}} \alpha^k (C_f + C_q) |\tilde{x} - x| \Big) \\
& \quad \le \quad \mathsf{E}(\tilde{\tau} + 1)(C_f + C_q)|\tilde{x} - x|. \qquad (7.11)
\end{aligned}$$

It immediately follows from Lemma 7.1 that $\mathsf{E}(\tilde{\tau}+1) = \mathsf{E}(\tau_{\tilde{x}}+1|i_0=i) \le \mathsf{E}(\tau_x + 1|i_0 = i) < \infty$.

To complete the proof, it is sufficient to prove the above inequality for $\tilde{x} < x$. In this case, let us define the strategy $\tilde{U}$ by

$$\tilde{u}_k = \begin{cases} u_k + x - \tilde{x} & \text{if } u_k > 0, \\ 0 & \text{otherwise,} \end{cases}$$

It is easy to see that the inequalities

$$0 \ge \tilde{x}_k - x_k \ge \tilde{x} - x \quad \text{and} \quad \tilde{u}_k - u_k \le x - \tilde{x}$$

hold for all k. Let $\tau := \tau_x$ for ease of notation. If $u_k = 0$ for all $k \in [0, \tau]$, then $\tilde{x}_\tau = x_\tau = 0$, and the two trajectories are identical for all $k > \tau$. If $u_{k'} \neq 0$ for some $k' \in [0, \tau]$, then $\tilde{x}_{k'} = x_{k'}$, and the two trajectories are identical for all $k > k'$. In either case, the two trajectories are identical for all $k > \tau$.

From Assumptions (i)–(iii), we have

$$c(i_k, \tilde{u}_k) - c(i_k, u_k) \leq \max\{c_i\}|x - \tilde{x}|,$$

$$|f(i_k, \tilde{x}_k) - f(i_k, x_k)| \leq C_f|\tilde{x} - x|,$$

$$|q(i_k, \tilde{x}_k + \tilde{u}_k - \xi_k) - q(i_k, x_k + u_k - \xi_k)| \leq C_q|\tilde{x} - x|.$$

Therefore,

$$\begin{aligned}
&v^\alpha(i, \tilde{x}) - v^\alpha(i, x) \\
&\quad\leq \quad J^\alpha(i, \tilde{x}; \tilde{U}) - J^\alpha(i, x; U) \\
&\quad= \quad \mathsf{E}\Big(\sum_{k=0}^{\tau} \alpha^k (f(i_k, \tilde{x}_k) - f(i_k, x_k) + q(i_k, \tilde{x}_k + \tilde{u}_k - \xi_k) \\
&\qquad\qquad -q(i_k, x_k + u_k - \xi_k) + c(i_k, \tilde{u}_k) - c(i_k, u_k))\Big) \\
&\quad\leq \quad \mathsf{E}\Big(\sum_{k=0}^{\tau} \alpha^k (C_f + C_q + \max\{c_i\})|\tilde{x} - x|\Big) \\
&\quad\leq \quad \mathsf{E}(\tau + 1)(C_f + C_q + \max\{c_i\})|\tilde{x} - x|.
\end{aligned} \tag{7.12}$$

Lemma 7.1 implies again that $\mathsf{E}(\tau{+}1) = \mathsf{E}(\tau_x{+}1|i_0 = i) \leq \mathsf{E}(\tau_{\tilde{x}}{+}1|i_0 = i) < \infty$ and the proof is complete. □

Lemma 7.3 *Under Assumptions* (i)–(ix)*, there are constants* $\alpha_0 \in [0, 1)$ *and* $C_2 > 0$ *such that for all* $\alpha \geq \alpha_0$ *and for any* i *for which* $s_i^\alpha > 0$*, we have* $S_i^\alpha \leq C_2 < \infty$*.*

Proof. Let us fix the initial state $i_0 = i$ for which $s_i^\alpha > 0$. Fix $\alpha_0 > 0$ and a discount factor $\alpha \geq \alpha_0$. Let $U = (u(i_0, x_0), u(i_1, x_1), \ldots)$ be an optimal strategy with parameters (s_j^α, S_j^α), $j \in I$. Let us fix a positive real number Y and assume $S_i^\alpha > Y$. In what follows, we specify a value of Y, namely Y^*, in terms of which we will construct an alternative strategy $\tilde{U}$ that is better than U.

For the demand state g specified in Assumption (iv), let

$$\tau^g := \inf\{n > 0 : i_n = g\}$$

be the first period (not counting the period 0) with the demand state g. Furthermore, let d be the state with the lowest per unit ordering cost, i.e., $c_d \leq c_i$ for all $i \in I$. Then we define

$$\tau := \inf\{n \geq \tau^g : i_n = d\}.$$

Assume $x_0 = \tilde{x}_0 = \bar{x} := 0$, and consider the policy $\tilde{U}$ defined by

$$\begin{cases} \tilde{u}_k = 0, \; k = 0, 1, 2, \ldots, \tau - 1, \\ \tilde{u}_\tau = x_\tau + u(i_\tau, x_\tau), \\ \tilde{u}_k = u(i_k, x_k), \; k \geq \tau + 1. \end{cases}$$

The two policies and the resulting trajectories differ only in the periods 0 through τ. Therefore, we have

$$\begin{aligned} & v^\alpha(i, \bar{x}) - J^\alpha(i, \bar{x}; \tilde{U}) \\ = \;& J^\alpha(i, \bar{x}; U) - J^\alpha(i, \bar{x}; \tilde{U}) \\ = \;& \mathsf{E}\Big(\sum_{k=0}^{\tau} \alpha^k (f(i_k, x_k) - f(i_k, \tilde{x}_k) + q(i_k, x_k + u_k - \xi_k) \\ & -q(i_k, \tilde{x}_k + \tilde{u}_k - \xi_k) + c(i_k, u_k) - c(i_k, \tilde{u}_k))\Big) \\ = \;& \mathsf{E}\left(\sum_{k=1}^{\tau} \alpha^k (f(i_k, x_k) + q(i_k, x_k + u_k - \xi_k) - q(i_k, -\xi_k)) \right) \\ & + \mathsf{E}\left(\sum_{k=0}^{\tau} \alpha^k c(i_k, u_k) \right) - \mathsf{E}(\alpha^\tau c(i_\tau, \tilde{u}_\tau)). \end{aligned} \tag{7.13}$$

After ordering in period τ, the total accumulated ordered amount up to period τ is less for the policy $\tilde{U}$ than it is for U. Observe that the policy $\tilde{U}$ orders only in the period τ or later. The order of the policy $\tilde{U}$ is executed at the lowest possible per unit cost c_d in the period τ, which is not earlier than any of the ordering periods of policy U. Because $\tilde{U}$ orders only once and U orders at least once in periods $0, 1, \ldots, \tau$, the total fixed ordering cost of $\tilde{U}$ does not exceed the total fixed ordering cost of U. Thus,

$$\mathsf{E}\left(\sum_{k=0}^{\tau} \alpha^k c(i_k, u_k) \right) \geq \mathsf{E}(\alpha^\tau c(i_\tau, \tilde{u}_\tau)).$$

Furthermore, it follows from Assumptions (iii), (vi), and (ix) that

$$\mathsf{E}\left(\sum_{k=1}^{\tau} q(i_k, -\xi_t) \right) < \infty.$$

Because $\tau \geq \tau^g$, we obtain

$$\mathsf{E}\left(\sum_{k=1}^{\tau} \alpha^k (f(i_k, x_k) + q(i_k, x_k + u_k - \xi_k)) \right) \geq \mathsf{E}\left(\sum_{k=1}^{\tau^g} \alpha^k f(i_k, x_k) \right),$$

and because $S_i^\alpha \geq Y$, we obtain

$$x_k \geq Y - \sum_{t=1}^{k} \xi_t.$$

Irreducibility of the Markov chain $(i_n)_{n=0}^\infty$ implies existence of an integer m, $0 \leq m \leq L$, such that $\mathsf{P}(i_m = g) > 0$. Let m_0 be the smallest such m. It follows that $\tau^g \geq m_0$ and therefore, for all $\alpha \geq \alpha_0$

$$\begin{aligned} &\mathsf{E}\left(\sum_{k=1}^{\tau^g} \alpha^k f(i_k, x_k)\right) \\ &\geq \alpha^{m_0} \mathsf{E}(f(i_{m_0}, x_{m_0})) \\ &\geq \alpha_0^{m_0} \mathsf{E}(f(g, Y - \sum_{t=1}^{m_0} \xi_t) | i_{m_0} = g) \mathsf{P}(i_{m_0} = g). \end{aligned} \tag{7.14}$$

Using Assumptions (ii), (iv), and (ix), it is easy to show that the RHS of (7.14) tends to infinity as Y goes to infinity. Therefore, we can choose Y^*, $0 \leq Y^* < \infty$ such that for all $\alpha \geq \alpha_0$,

$$\begin{aligned} v^\alpha(i, \bar{x}) - J^\alpha(i, \bar{x}; \tilde{U}) \quad &\geq \quad \alpha_0^{m_0} \mathsf{E}(f(g, Y^* - \sum_{t=1}^{m_0} \xi_t) | i_{m_0} = g) \mathsf{P}(i_{m_0} = g) \\ &\quad - \mathsf{E}\left(\sum_{k=1}^{\tau} q(i_k, -\xi_t)\right) > 0. \end{aligned} \tag{7.15}$$

Note that the RHS of (7.15) is independent of α. Therefore, for $\alpha \geq \alpha_0$, a policy with $S_i^\alpha > C_2 := Y^*$ cannot be optimal. □

7.5. Vanishing Discount Approach

LEMMA 7.4 *Under Assumptions* (i)-(ix), *the differential discounted value function* $w^\alpha(i, x) := v^\alpha(i, x) - v^\alpha(1, 0)$ *is uniformly bounded with respect to* α *for all* x *and* i.

Proof. Since Lemma 7.2 implies

$$\begin{aligned} |w^\alpha(i, x)| \quad &= \quad |v^\alpha(i, x) - v^\alpha(1, 0)| \\ &\leq \quad |v^\alpha(i, x) - v^\alpha(i, 0)| + |v^\alpha(i, 0) - v^\alpha(1, 0)| \\ &\leq \quad C_3|x| + |w^\alpha(i, 0)|, \end{aligned}$$

it is sufficient to prove that $w^\alpha(i, 0)$ is uniformly bounded. Note that C_3 may depend on x, but it is independent of α.

First, we show that there is an $\underline{M} > -\infty$ with $w^\alpha(i,0) \geq \underline{M}$ for all α. Let α be fixed. From Theorem 7.1 we know that in this discounted case there is a stationary optimal feedback policy $U = (u(i,x), u(i,x), \ldots)$. With $k^* = \inf\{k : i_k = i\}$, we consider the cost for the initial state $(i_0, \tilde{x}_0) = (1,0)$, and the inventory policy $\tilde{U}$ that does not order in periods $0, 1, \ldots, k^* - 1$, and follows U starting from the period k^*, i.e., $\tilde{U}$ is defined by

$$\begin{cases} \tilde{u}_k = 0 \text{ for } k < k^*, \\ \tilde{u}_k = u(i_k, x_k) \text{ for } k \geq k^*. \end{cases}$$

The cost corresponding to this policy is

$$J^\alpha(1,0;\tilde{U}) = \mathsf{E}\left(\sum_{k=0}^{k^*-1} \alpha^k q(i_k, -\xi_k) + \alpha^{k^*} v^\alpha(i,0)\right). \tag{7.16}$$

Because of Assumptions (iii), (vi) and (ix), there exists a constant $\underline{M}$ such that

$$\mathsf{E}\left(\sum_{k=0}^{k^*-1} q(i_k, -\xi_k)\right) \leq -\underline{M} < \infty.$$

Therefore, we have

$$\begin{aligned} w^\alpha(i,0) &= v^\alpha(i,0) - v^\alpha(1,0) \geq v^\alpha(i,0) - J^\alpha(1,0;\tilde{U}) \\ &\geq v^\alpha(i,0) - \mathsf{E}\left(\sum_{k=0}^{k^*-1} \alpha^k q(i_k, -\xi_k) + \alpha^{k^*} v^\alpha(i,0)\right) \\ &\geq v^\alpha(i,0)(1 - \mathsf{E}(\alpha^{k^*})) + \underline{M} \geq \underline{M}. \end{aligned} \tag{7.17}$$

The validity of the inequality $w^\alpha(i,0) \leq \overline{M}$ is shown analogously by changing the role of the states 1 and i. Thus,

$$|w^\alpha(i,x)| \leq C_3|x| + \max\{\underline{M}, \overline{M}\},$$

and the proof is complete. □

LEMMA 7.5 *Under Assumptions* (iii) *and* (viii), $(1-\alpha)v^\alpha(1,0)$ *is uniformly bounded for* $0 < \alpha < 1$.

Proof. Consider the strategy $\mathbf{0} = (0, 0, \ldots)$. Then, because $\mathbf{0}$ is not necessarily optimal,

$$0 \leq v^\alpha(1,0) \leq J^\alpha(1,0;\mathbf{0}) = \mathsf{E}\left(\sum_{k=0}^{\infty} \alpha^k q(i_k, -\xi_k)\right).$$

Because of Assumptions (iii) and (ix), $\mathsf{E}q(i, -\xi_k)$ is bounded for all i and there is a $C_4 < \infty$ such that $\mathsf{E}(q(i_k, -\xi_k)) < C_4$. Therefore,

$$0 \leq (1-\alpha)v^{\alpha}(1,0) \leq (1-\alpha)\sum_{k=0}^{\infty} \alpha^k C_4 = C_4.$$

□

THEOREM 7.2 *Let Assumptions* (i)–(ix) *hold. There exist a sequence* $(\alpha_k)_{k=1}^{\infty}$ *converging to* 1*, a constant* λ^**, and a locally Lipschitz continuous function* $w^*(\cdot,\cdot)$*, such that*

$$(1-\alpha_k)v^{\alpha_k}(i,x) \to \lambda^* \quad \textit{and} \quad w^{\alpha_k}(i,x) \to w^*(i,x),$$

locally uniformly in x *and* i *as* k *goes to infinity. Moreover,* (λ^*, w^*) *satisfies the average cost optimality equation*

$$\begin{aligned} w(i,x) + \lambda \quad &= \quad f(i,x) + \inf_{u \geq 0}\{c(i,u) + \mathsf{E}q(i, x+u-\xi^i) \\ &\qquad + F(w)(i,x+u)\}. \end{aligned} \tag{7.18}$$

Proof. It is immediate from Lemma 7.2 and the definition of $w^{\alpha}(i,x)$ that $w^{\alpha}(i,\cdot)$ is locally equi-Lipschitzian for $\alpha \geq \alpha_0$, and therefore it is uniformly continuous on any finite interval. Additionally, according to Lemma 7.4, $w^{\alpha}(i,\cdot)$ is uniformly bounded, and by Lemma 7.5, $(1-\alpha)v^{\alpha}(1,0)$ is also uniformly bounded. Therefore, the Arzelà-Ascoli Theorem A.3.5 and Lemma 7.2, lead to the existence of a sequence $\alpha_k \to 1$, a locally Lipschitz continuous function $w^*(i,x)$, and a constant λ^* such that

$$(1-\alpha_k)v^{\alpha_k}(1,0) \to \lambda^* \quad \text{and} \quad w^{\alpha_k}(i,x) \to w^*(i,x),$$

with the convergence of $w^{\alpha_k}(i,x)$ to $w^*(i,x)$ being locally uniform.

It is easily seen that

$$\lim_{k\to\infty}(1-\alpha_k)v^{\alpha_k}(i,x) = \lim_{k\to\infty}(1-\alpha_k)(w^{\alpha_k}(i,x) + v^{\alpha_k}(1,0)) = \lambda^*.$$

Substituting $v^{\alpha_k}(i,x) = w^{\alpha_k}(i,x) + v^{\alpha_k}(1,0)$ into (7.9) yields

$$\begin{aligned} w^{\alpha_k}(i,x) \quad + \quad (1-\alpha_k)v^{\alpha_k}(1,0) &= f(i,x) + \inf_{u\geq 0}\{c(i,u) \\ &\quad + \mathsf{E}q(i,x+u-\xi^i) + \alpha_k F(w^{\alpha_k})(i,x+u)\}. \end{aligned} \tag{7.19}$$

Since $w^{\alpha_k}(i,x)$ converges locally uniformly with respect to x and i and since for any given x a minimizer u^* in (7.19) can be chosen such that,

$x + u^* - \xi \in [0, x + C_2]$, we can use Lemma 7.3 and pass to a limit on both sides of (7.19) to obtain (7.18). This completes the proof. □

LEMMA 7.6 *Let* λ^* *be defined as in Theorem* 7.2. *Let Assumptions* (i)–(ix) *hold. Then, for any admissible strategy* U, *we have* $\lambda^* \leq J(i, x; U)$.

Proof. Let $U = (u_0, u_1, \ldots)$ denote any admissible decision. Suppose

$$J(i, x; U) < \lambda^*. \tag{7.20}$$

Put

$$\tilde{f}(k) = \mathsf{E}[f(i_k, x_k) + q(i_k, x_k + u_k - \xi_k) + c(i_k, u_k)].$$

From (7.20), it immediately follows that $\sum_{k=0}^{n-1} \tilde{f}(k) < \infty$ for each positive integer n, since otherwise we would have $J(i, x; U) = \infty$. Note that

$$J(i, x; U) = \limsup_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \tilde{f}(k),$$

while

$$(1 - \alpha) J^\alpha(i, x; U) = (1 - \alpha) \sum_{k=0}^{\infty} \alpha^k \tilde{f}(k). \tag{7.21}$$

Since $\tilde{f}(k)$ is nonnegative for each k, the sum in (7.21) is well-defined for $0 \leq \alpha < 1$, and we can use the Tauberian Theorem A.5.2 to obtain

$$\limsup_{\alpha \uparrow 1} (1 - \alpha) J^\alpha(i, x; U) \leq J(i, x; U) < \lambda^*.$$

On the other hand, we know from Theorem 7.2 that $(1-\alpha_k) v^{\alpha_k}(i, x) \to \lambda^*$ for a subsequence $\{\alpha_k\}_{k=1}^\infty$ converging to one. Thus, there exists an $\alpha < 1$ such that

$$(1 - \alpha) J^\alpha(i, x; U) < (1 - \alpha) v^\alpha(i, x),$$

which contradicts the definition of the value function $v^\alpha(i, x)$. □

7.6. Verification Theorem

DEFINITION 7.1 Let (λ, w) be a solution of the average optimality equation (7.18). An admissible strategy $U = (u_0, u_1, \ldots)$ is called stable with respect to w if for each initial inventory level $x \geq 0$ and for each initial demand state $i \in I$,

$$\lim_{k \to \infty} \frac{1}{k} \mathsf{E}(w(i_k, x_k)) = 0,$$

where x_k is the inventory level in period k corresponding to the initial state (i, x) and the strategy U.

Lemma 7.7 *Let Assumptions* (i)-(ix) *hold. Then, there are constants* $S_i < \infty$ *and* $0 \leq s_i \leq S_i$, $i \in I$, *such that*

$$u^*(i, x) = \begin{cases} S_i - x, & x < s_i, \\ 0, & x \geq s_i, \end{cases}$$

attains the minimum on the RHS in (7.18) *for* $w = w^*$, *as defined in Theorem* 7.2. *Furthermore, the stationary feedback strategy* $U^* = (u^*, u^*, \ldots)$ *is stable with respect to any continuous function* w.

Proof. Let $\{\alpha_k\}_{k=0}^{\infty}$ be the sequence defined in Theorem 7.2. Let

$$G^{\alpha_k}(i, y) \quad = \quad c_i y + \mathsf{E} q(i, y - \xi^i) + \alpha_k F(w^{\alpha_k})(i, y) \tag{7.22}$$

and

$$G(i, y) \quad = \quad c_i y + \mathsf{E} q(i, y - \xi^i) + F(w^*)(i, y). \tag{7.23}$$

Because $w^*(i, \cdot)$ is K-convex, we know that a minimizer in (7.18) is given by

$$u^*(i, x) = \begin{cases} S_i - x, & x < s_i, \\ 0, & x \geq s_i, \end{cases}$$

where $0 \leq S_i \leq \infty$ minimizes $G(i, \cdot)$, and s_i solves

$$G(s_i) = K + G(S_i)$$

if a solution to this equation exists, or $s_i = 0$ otherwise. Note that if $s_i = 0$, it follows that $u^*(i, x) = 0$ for all nonnegative x. It remains to show that $S_i < \infty$.

We distinguish two cases.

- Case 1. If there is a subsequence, still denoted by $\{\alpha_k\}_{k=0}^{\infty}$, such that $s_i^{\alpha_k} > 0$ for all $k = 0, 1, ...$, then it follows from Lemma 7.3 that G^{α_k} attains its minimum in $[0, C_2]$ for all $\alpha_k > \alpha_0$. Thus, G^{α_k}, $k = 0, 1, ...$ are locally uniformly continuous and converge uniformly to G. Therefore, G attains its minimum also in $[0, C_2]$, which implies $S_i \leq C_2$.

- Case 2. If there is no such sequence, then there is a sequence, still denoted by $\{\alpha_k\}_{k=0}^{\infty}$, such that $s_i^{\alpha_k} = 0$ for all $k = 0, 1, ...$ It follows that for all $y > x$,

$$G^{\alpha_k}(x) < K + G^{\alpha_k}(y),$$

and therefore, in the limit,

$$G(x) < K + G(y).$$

This implies that the infimum in (7.18) is attained for $u^*(i,x) \equiv 0$, which is equivalent to $s_i = 0$. But if $s_i = 0$, we can choose S_i arbitrarily, say $S_i = C_2$.

It is obvious that the stationary policy U^* is stable with respect to any continuous function, since we have $x_k \in [0, \max\{C_2, x_0\}]$ for all $k = 0, 1, \ldots$ for such a policy. □

THEOREM 7.3 (Verification Theorem)
(a) *Let* $(\lambda, w(\cdot,\cdot))$ *be a solution of the average cost optimality equation* (7.18)*, with* w *continuous on* $[0,\infty)$*. Then,* $\lambda \leq J(i,x;U)$ *for any admissible* U*.*
(b) *Suppose there exists an* $\hat{u}(i,x)$ *for which the infimum in* (7.18) *is attained. Furthermore, let* $\hat{U} = (\hat{u}, \hat{u}, \ldots)$*, the stationary feedback policy given by* $\hat{u}$*, be stable with respect to* w*. Then,*

$$\begin{aligned} \lambda &= J(i,x;\hat{U}) = \lambda^* \\ &= \lim_{N\to\infty} \frac{1}{N}\mathsf{E}\left(\sum_{k=0}^{N-1} f(i_k,\hat{x}_k) + q(i_k, x_k + u_k - \xi_k) + c(i_k,\hat{u}_k)\right), \end{aligned}$$

and $\hat{U}$ *is an average optimal strategy.*
(c) *Moreover,* $\hat{U}$ *minimizes*

$$\liminf_{N\to\infty} \frac{1}{N}\mathsf{E}\left(\sum_{k=0}^{N-1} f(i_k,x_k) + q(i_k, x_k + u_k - \xi_k) + c(i_k,u_k)\right)$$

over the class of admissible decisions which are stable with respect to w*.*

Proof. We start by showing that

$$\lambda \leq J(i,x;U) \text{ for any } U \text{ stable with respect to } w. \tag{7.24}$$

We assume that U is stable with respect to w, and then follow the same approach used in deriving (2.15) to obtain

$$\mathsf{E}\{w(i_{k+1}, x_{k+1}) \,|\, i_0, \ldots, i_k, \xi_0, \ldots, \xi_{k-1}\} = F(w)(i_k, x_k + u_k) \quad \text{a.s.} \tag{7.25}$$

Because u_k does not necessarily attain the infimum in (7.18), we have

$$\begin{aligned} w(i_k, x_k) + \lambda &\leq f(i_k, x_k) + c(i_k, u_k) \\ &\quad + q(i_k, x_k + u_k - \xi_k) + F(w)(i_k, x_k + u_k) \quad \text{a.s.}, \end{aligned}$$

and from (7.25) we derive

$$\begin{aligned} w(i_k, x_k) + \lambda &\leq f(i_k, x_k) + q(i_k, x_k + u_k - \xi_k) + c(i_k, u_k) \\ &\quad + \mathsf{E}(w(i_{k+1}, x_k + u_k - \xi_k) \,|\, i_k) \quad \text{a.s.} \end{aligned}$$

By taking the expectation of both sides, we obtain

$$\begin{aligned} \mathsf{E}(w(i_k, x_k)) + \lambda &\leq \mathsf{E}(f(i_k, x_k) + q(i_k, x_k + u_k - \xi_k) + c(i_k, u_k)) \\ &\quad + \mathsf{E}(w(i_{k+1}, x_{k+1})). \end{aligned}$$

Summing from 0 to $n-1$ yields

$$\begin{aligned} n\lambda &\leq \mathsf{E}\left(\sum_{k=0}^{n-1} f(i_k, x_k) + q(i_k, x_k + u_k - \xi_k)\right. \\ &\quad \left. + c(i_k, u_k)\right) + \mathsf{E}(w(i_n, x_n)) - \mathsf{E}(w(i_0, x_0)). \end{aligned} \tag{7.26}$$

Divide by n, let n go to infinity, and use the fact that U is stable with respect to w, to obtain

$$\begin{aligned} \lambda &\leq \liminf_{n\to\infty} \frac{1}{n} \mathsf{E}\left(\sum_{k=0}^{n-1} f(i_k, x_k)\right. \\ &\quad \left. + q(i_k, x_k + u_k - \xi_k) + c(i_k, u_k)\right). \end{aligned} \tag{7.27}$$

Note that if the above inequality holds for 'liminf', it certainly also holds for 'limsup'. This proves (7.24).

On the other hand, if there exists a $\hat{u}$ for which the infimum in (7.18) is attained , we then have

$$\begin{aligned} w(i_k, \hat{x}_k) + \lambda &= f(i_k, \hat{x}_k) + q(i_k, \hat{x}_k + \hat{u}_k - \xi_k) + c(i_k, \hat{u}(i_k, \hat{x}_k)) \\ &\quad + F(w)(i_k, \hat{x}_k + \hat{u}(i_k, \hat{x}_k)), \quad \text{a.s.}, \end{aligned}$$

and we analogously obtain

$$\begin{aligned} n\lambda &= \mathsf{E}\left(\sum_{k=0}^{n-1} f(i_k, \hat{x}_k) + q(i_k, \hat{x}_k + \hat{u}_k - \xi_k) + c(i_k, \hat{u}(i_k, \hat{x}_k))\right) \\ &\quad + \mathsf{E}(w(i_n, \hat{x}_n)) - \mathsf{E}(w(i_0, \hat{x}_0)). \end{aligned} \tag{7.28}$$

Because $\hat{U}$ is assumed stable with respect to w, we get

$$\begin{aligned}\lambda &= \lim_{n\to\infty} \frac{1}{n} \mathsf{E} \left(\sum_{k=0}^{n-1} f(i_k, \hat{x}_k) + q(i_k, \hat{x}_k + \hat{u}_k - \xi_k) + c(i_k, \hat{u}(i_k, \hat{x}_k) \right) \\ &= J(i, x; \hat{U}). \end{aligned} \tag{7.29}$$

For the special solution (λ^*, w^*) defined in Theorem 7.2, and the strategy U^* defined in Lemma 7.7, we have

$$\lambda^* = J(i, x; U^*).$$

Since U^* is stable with respect to any continuous function by Lemma 7.7, it follows that

$$\lambda \leq J(i, x; U^*) = \lambda^*, \tag{7.30}$$

which, in view of Lemma 7.6, proves part (a) of the theorem.

Part (a) of the theorem, together with (7.29), proves the average optimality of $\hat{U}$ over all admissible strategies. Furthermore, since $\lambda = J(i, x; \hat{U}) \geq \lambda^*$ by (7.29) and Lemma 7.6, it follows from (7.30) that $\lambda = \lambda^*$, and the proof of Part (b) is completed.

Finally, Part (c) immediately follows from Part (a) and (7.27). □

Remark 7.5 It should be obvious that any solution (λ, w) of the average cost optimality equation and control u^* satisfying (a) and (b) of Theorem 7.3 will have a unique λ, since it represents the minimum average cost. On the other hand, if (λ, w) is a solution, then $(\lambda, w+c)$, where c is any constant, is also a solution. For the purpose of this chapter, we do not require w to be unique up to a constant. If w is not unique up to a constant, then u^* may not be unique. We also do not need w^* in Theorem 7.2 to be unique.

The final result of this section, namely, that there exists an average optimal policy of (s, S)-type, is an immediate consequence of Lemma 7.7 and Theorem 7.3.

Theorem 7.4 *Let Assumptions* (i)-(viii) *hold. Let s_i and S_i, $i \in I$ be defined as in Lemma 7.7. Then, the stationary feedback strategy $U^* = (u^*, u^*, \ldots)$ defined by*

$$u^*(i, x) = \begin{cases} S_i - x, & x < s_i, \\ 0, & x \geq s_i, \end{cases}$$

is average optimal.

7.7. Concluding Remarks and Notes

This chapter is based on Cheng and Sethi (1999b) and Beyer and Sethi (2005). We have proved a verification theorem for the average cost optimality equation, which we have used to establish the existence of an optimal state-dependent (s, S) policy.

As with the discounted cost models, the optimality of an (s, S) policy for a lost sales case is established only under the condition of zero ordering leadtime; (see references in Chapter 4). With nonzero leadtimes, the results for models with backlog do not generalize to the lost sales case. Specifically, an (s, S) policy is no longer optimal, and the form of the optimal policy is more complicated; (see, e.g., Zipkin (2008a) and Huh *et al.* (2008)). Nevertheless, an (s, S)-type policy is often used without the optimality proof; (see, e.g., Kapalka, *et al.* (1999)).