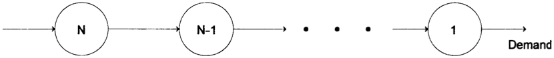
Modified Echelon (r, Q) Policies with Guaranteed Performance Bounds for Stochastic Serial Inventory Systems

Ming Hu
Rotman School of Management, University of Toronto
joint with
Yi Yang
School of Management, Zhejiang University

MSOM 2014 Supply Chain Management SIG

A Classic Problem

Continuous-review, N-stage serial inventory system



- Demand
  - Demand occurs at Stage 1
  - Poisson process with rate λ
  - Excess demand is fully backlogged with a penalty rate p
- Supply
  - Lead time L_i, $\forall i$
  - Setup cost K_i, $\forall i$, for each shipment from Stage $i+1$ to i
  - Echelon holding cost rate h_i
- Objective
  - To minimize the long-run average cost

Literature Review

- Zheng (1992): single-stage with $L > 0$, $K > 0$
  - Optimal policy: (r, Q) policy
- Clark and Scarf (1960): multi-stage with $L_i > 0$, $K_i = 0$
  - Optimal policy: echelon base-stock policy
- Clark and Scarf (1962): multi-stage with $L_i > 0$, $K_i > 0$
  - Optimal policy: extreme complex
- Graves (1985), Moinzadeh and Lee (1986), Jackson (1988), Svoronos and Zipkin (1988, 1991), Axsäter (1990, 1993), Chen and Zheng (1994a, 1998), Gallego and Özer (2005), Shang and Song (2003), Shang (2008), Shang and Zhou (2010), Shang et al. (2009) and Yang et al. (2011)
  - Simple and easy-to-implement heuristics

Literature Review: Performance Evaluation

- De Bodt and Graves (1985): multi-stage with $L_i > 0$, $K_i > 0$
  - A good approximation under mild assumptions
- Chen (1999): two-stage with $L_1 > 0$, $L_2 = 0$, $K_i > 0$
  - 94%-effective (r, Q) policy
- Cost-balancing techniques
  - Levi et al. (2007): single-stage finite-horizon periodic-review inventory system with nonstationary / correlated demand with $K = 0$ or $L = 0$: 2-optimal
  - Levi et al. (2008): single-stage capacitated system: 2-optimal
  - Levi et al. (2006): multiple-stage serial system: 3-optimal

Performance bound for multi-stage serial system with $L_i > 0$, $K_i > 0$?

Focus

- Old focus: Integer-ratioed nested (r, Q) policy
  - Integer-ratioed: Q_{i+1}/Q_i is a positive integer
  - Nested: whenever Stage $i+1$ orders, so does Stage i
  - Shipment in batches of Q_i

  In special settings, optimal policy is nested, e.g.,

  - Demand is deterministic (Atkins and Sun 1995)
  - Zero lead times at Stage 2 under demand uncertainty (Chen 1999)
  - Any order at Stage $i+1$ can be delayed until just before the next shipment to Stage i

  This is not true in general.

- New focus: Modified echelon (r, Q) policy
  - Need not to be integer-ratioed
  - Need not to be nested
  - Shipment may be of any size

Agenda

- Agenda
  - Provide an upper bound on the systemwide cost under any given modified echelon (r,Q) policy
  - Provide a performance bound for a specific modified echelon (r,Q) policy
  - Establish asymptotic optimality
- Assumption
  - Discrete units of inventories can be approximated by continuous quantities

Notation

$$
\begin{aligned}
IL_i(t) &= \text{echelon inventory level at stage } i \text{ at time } t \\
I_i(t) &= \text{echelon inventory at stage } i \text{ at time } t \\
IP_i(t) &= \text{echelon inventory position after ordering at stage } i \text{ at time } t \\
q_i(t) &= \text{order quantity at stage } i \text{ at time } t \\
B(t) &= \text{backorder level at stage 1 at time } t
\end{aligned}
$$

Sample Path Identity

$$
\begin{aligned}
I_i(t) &= IL_i(t) + B(t) \\
IL_i(t + L_i) &= IP_i(t) - D(t, t + L_i] \\
IP_i^-(t) &= IL_{i+1}(t) - OI_{i+1}(t) \\
IP_i(t) &= IP_i(t^-) + q_i(t)
\end{aligned}
$$

Formulation

Holding and shortage cost:

$$\sum_{i=1}^{N} h_i I_i(t) + pB(t) = \sum_{i=1}^{N} h_i IL_i(t) + (p + H_1)B(t),$$

where $H_1 \equiv \sum_{i=1}^{N} h_i$ is the installation holding cost rate at Stage 1

Problem ($\mathscr{B}$):

$$\begin{aligned} C^*_{\mathscr{B}} = \quad & \min_{\{IP_i(t), \forall i,t\}} \quad \lim_{T\to\infty} \frac{\mathbb{E}[\int_{t=0}^{T} \sum_{i=1}^{N} K_i \delta(IP_i(t) > IP_i^-(t)) + h_i IL_i(t) + (p + H_1)B(t)\, dt]}{T} \\ & \text{s.t.} \quad IP_i^-(t) \le IP_i(t) \le IL_{i+1}(t),\ i = 1, 2, \ldots, N-1 \\ & \qquad\ \ IP_N^-(t) \le IP_N(t) \end{aligned}$$

Policy

Definition (Modified Echelon $(\mathbf{r}, \mathbf{Q})$ Policy)

Stage $i+1$ ships to Stage i on the basis of its observation of the echelon inventory position at Stage i. In particular, if the echelon inventory position at Stage i is at or below r_i and Stage $i+1$ has positive on-hand inventory, then a shipment is sent to Stage i to raise its echelon inventory position as close as possible to $r_i + Q_i$.

Performance Bound

$$C_{\mathscr{B}}^{LB} \leq C_{\mathscr{B}}^{*} \leq C(\mathbf{r}, \mathbf{Q}) \leq C^{UB}(\mathbf{r}, \mathbf{Q})$$

Induced-Penalty Lower Bound (Chen and Zheng 1994)

- Stage 1
  - Cost rate: $G_1(y) = \mathbb{E}[h_1(y-D_1)^+ + (h_2+p)(y-D_1)^-]$
  - Objective function: $C_1(r,Q) = \frac{1}{Q}\left[\lambda K_1 + \int_r^{r+Q} G_1(y)dy\right]$
  - $(r_1^*, Q_1^*) = \arg\min_{(r,Q)} C_1(r,Q)$ and $C_1^* = C_1(r_1^*, Q_1^*)$
- Stage 2
  - Induced-penalty function:

$$G_1^2(y) = \begin{cases} G_1(y) - C_1^* & \text{if } y \le r_1^* \\ 0 & \text{o/w} \end{cases}$$

  - Cost rate: $G_2(y) = \mathbb{E}[h_2(y-D_2)] + \mathbb{E}[G_1^2(y-D_2)]$
  - Objective function: $C_2(r,Q) = \frac{1}{Q}\left[\lambda K_2 + \int_r^{r+Q} G_2(y)dy\right]$
  - $(r_2^*, Q_2^*) = \arg\min_{(r,Q)} C_2(r,Q)$ and $C_2^* = C_2(r_2^*, Q_2^*)$

Theorem (Lower Bound)

$$C_1^* + C_2^* \le C_{\mathscr{B}}^*$$

Upper Bound: Cycle and Shipment Period

We track the first shipment time of an order of Stage 2 being sent to Stage 1.

Definition (Cycle)

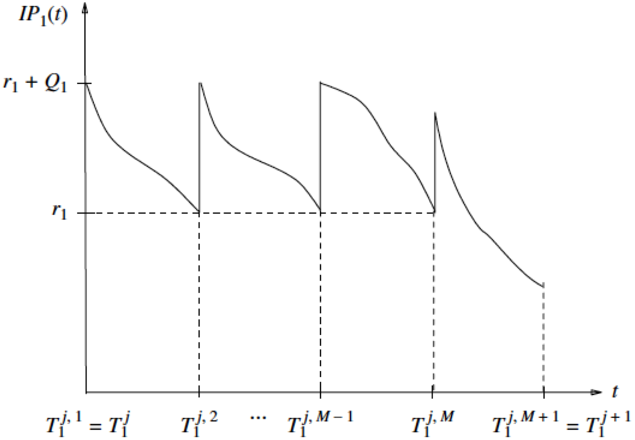
For any Stage $i = 1, 2$, we call $[T_i^j, T_i^{j+1})$, for any $j \in \mathbb{N}$, the j^{th} cycle, where T_i^j is the time epoch of the 1^{st} unit, contained in the j^{th} order of Stage 2, being sent to Stage i.

The j^{th} order of Stage 2 may be shipped to Stage 1 by one or multiple shipments.

Definition (Regular and Irregular Shipment Period)

For any shipment period $[T_1^{j,l}, T_1^{j,l+1})$, $l = 1, \ldots, M$, if $IP_1(T_1^{j,l}) = r_1 + Q_1$ and $IP_1^-(T_1^{j,l+1}) = r_1$, we call it a regular shipment period; otherwise, we call it an irregular shipment period.

Shipment Periods



Lemma (Irregular Shipment Frequency)

There exists at most one irregular shipment period in each cycle of Stage 1, whether the cycle is empty or not.

An Upper Bound of $C(\mathbf{r}, \mathbf{Q})$

Lemma (Cycle Length)

The long-run average expected cycle length $\lim_{j\to\infty} \mathbb{E}[(T_i^{j+1} - T_i^1)]/j = Q_2/\lambda.$

The long-run average setup cost for irregular shipment periods has an upper bound $\lambda K_1/Q_2$.

Theorem (Upper Bound)

$$C(\boldsymbol{r}, \boldsymbol{Q}) \leq \frac{1}{Q_2}\left[\lambda K_2 + \int_{r_2}^{r_2+Q_2} \Lambda(y)dy\right] + C_1(r_1, Q_1) + \frac{\lambda K_1}{Q_2}$$

where $\Lambda(y) = h_2\mathbb{E}[y - D_2] + \mathbb{E}[\hat{G}_1(y - D_2)]$

$$\hat{G}_1(y) = \begin{cases} G_1(y) - C_1(r_1, Q_1) & \text{if } y \leq r_1 \\ \max\{0, G_1(\omega_1) - C_1(r_1, Q_1)\} & \text{o/w} \end{cases}$$

and $\omega_1 \equiv \arg\max_{r_1 < z \leq r_1 + Q_1} \{G_1(z)\}$.

A Specific $(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ Policy

To compare with the lower bound, we construct a heuristic as follows:

$$(\hat{\mathbf{r}}, \hat{\mathbf{Q}}) = \left(r_1^*, Q_1^*, \tilde{r}_2^*, \tilde{Q}_2^*\right)$$

where

$$(r_1^*, Q_1^*) = \arg\min_{r_1, Q_1} C_1(r_1, Q_1)$$

$$(\tilde{r}_2^*, \tilde{Q}_2^*) = \arg\min_{r_2, Q_2} \tilde{C}_2(r_2, Q_2) \equiv \arg\min_{r_2, Q_2} C_2(r_2, Q_2) + \frac{\lambda K_1}{Q_2}$$

Theorem ((K_1, K_2)-Dependent Performance Bound)

The modified echelon $(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ policy is $(1 + K_1/K_2)$-optimal.

Performance Bounds

Theorem (Alternative Performance Bound)

The modified echelon $(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ *policy is* $\left(1 + \frac{1}{2(\beta_1^* + \sqrt{\beta_1^*})}\right)$*-optimal, where* $\beta_1^* \equiv Q_2^*/Q_1^*$.

Corollary

If $Q_2^*/Q_1^* \geq \frac{1+\frac{1}{\rho}-\sqrt{1+\frac{2}{\rho}}}{2}$*, then the modified echelon* $(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ *policy is at least* $(1+\rho)$*-optimal. In particular, we have the performance bounds listed as follows:*

$Q_2^*/Q_1^* \in$	$[0.0429, 0.134)$	$[0.134, 0.382)$	$[0.382, 1)$	$[1, 3.21)$	$[3.21, \infty)$
Bound	*3-opt.*	*2-opt.*	*1.5-opt.*	*1.25-opt.*	*1.1-opt.*

Asymptotic Optimality Results

Deterministic counterparts:

$$Q_1^* \approx Q_1^d = \sqrt{2\lambda K_1 \left(\frac{1}{h_1} + \frac{1}{p+h_2}\right)}$$

and

$$Q_2^* \approx Q_2^d = \sqrt{2\lambda K_2 \left(\frac{1}{h_2} + \frac{1}{p}\right)}$$

If $K_2 \geq K_1$ and $h_2 \leq h_1$,

$$Q_1^d \leq Q_2^d$$

K_2: international, K_1: domestic
h_2: suburban, h_1: downtown

Theorem (Asymptotic Optimality)

The modified echelon $(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ *policy is asymptotically optimal if one of the following conditions holds:*
(i) $K_2/K_1 \to \infty$.
(ii) $h_2/h_1 \to 0$.

Numerical Results: Comparison with Optimal (r, nQ)

Table 1 of Chen and Zheng (1994):

$L_1 = 1$, $L_2 = 2$, $K_1 = 10$, $h_1 = 0.5$, $h_2 = 1$, $p = 5$

λ	K_2	$C(\mathbf{r}^n, \mathbf{Q}^n)$	$C(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ or UB	LB	$\xi(\mathbf{r}^n, \mathbf{Q}^n)(\%)$	$\xi(\hat{\mathbf{r}}, \hat{\mathbf{Q}})(\%)$
	5	8.3828	8.3948	8.0216	4.50	4.65
	100	17.2446	17.1451	17.1317	0.66	0.08
1	200	22.5396	22.3690	22.3692	0.77	0.009
	400	30.0028	29.8456	29.8456	0.53	0
	5	21.4394	21.5758	20.6433	3.86	4.51
	100	41.3823	41.1538	41.1448	0.58	0.02
5	200	53.1838	52.8492	52.8355	0.66	0.026
	400	69.9265	69.5690	69.5613	0.52	0.01
	5	33.2192	33.3534	32.1126	3.45	3.86
	100	61.4478	61.1632	61.1264	0.53	0.06
10	200	78.1643	77.6698	77.6614	0.65	0.01
	400	101.8232	101.3087	101.3039	0.51	0.005
	5	43.4355	43.6150	42.0567	3.28	3.70
	100	78.0018	77.6335	77.6118	0.50	0.03
15	200	98.4661	97.8904	97.8676	0.61	0.02
	400	127.4852	126.8234	126.8215	0.52	0.005

Multiple Stages: Upper Bound

We define $\Lambda_1(y) = G_1(y)$, $\hat{C}_1(\hat{r}_1, \hat{Q}_1) = C_1(\hat{r}_1, \hat{Q}_1)$ and for $i = 2, 3, \ldots, N$,

$$\hat{G}_{i-1}(y) = \begin{cases} \Lambda_{i-1}(y) - \hat{C}_{i-1}(\hat{r}_{i-1}, \hat{Q}_{i-1}) & \text{if } y \le \hat{r}_{i-1} \\ \max\{0, \Lambda_{i-1}(\omega_{i-1}(y)) - \hat{C}_{i-1}(\hat{r}_{i-1}, \hat{Q}_{i-1})\} & \text{o/w} \end{cases}$$

$$\Lambda_i(y) = h_i \mathbb{E}[y - D_i] + \mathbb{E}[\hat{G}_{i-1}(y - D_i)]$$

$$\hat{C}_i(\hat{r}_i, \hat{Q}_i) = \frac{1}{\hat{Q}_i} \left[\lambda K_i + \int_{\hat{r}_i}^{\hat{r}_i + \hat{Q}_i} \Lambda_i(y) dy \right]$$

where D_i is a leadtime demand over $[0, L_i)$.

Theorem (Upper Bound)

For any given modified echelon $(\hat{\boldsymbol{r}}, \hat{\boldsymbol{Q}})$ policy, the expected long-run average systemwide cost $C(\hat{\boldsymbol{r}}, \hat{\boldsymbol{Q}})$ is bounded by

$$C(\hat{\boldsymbol{r}}, \hat{\boldsymbol{Q}}) \le \sum_{\nu=1}^{N} \hat{C}_\nu(\hat{r}_\nu, \hat{Q}_\nu) + \sum_{\nu=1}^{N-1} \frac{\lambda K_\nu}{\hat{Q}_{\nu+1}}.$$

Multiple Stages: Performance Bound

We construct a modified echelon $(\hat{\mathbf{r}}, \hat{\mathbf{Q}})$ heuristic as follows:

$$(\hat{r}_i, \hat{Q}_i) = (r_i^*, Q_i^*), \quad i = 1, 2, \ldots, N.$$

For $i = 1, 2, \ldots, N-1$, define $\beta_i^* = \frac{Q_{i+1}^*}{Q_i^*}$ and $\beta^* = \min_i \{\beta_i^*\}$.

Theorem (Performance Bound)

(i) *The modified echelon $(\hat{\boldsymbol{r}}, \hat{\boldsymbol{Q}})$ policy is $(1 + \frac{1}{2\beta^*})$-optimal, e.g., if $Q_{i+1}^* \geq Q_i^*$ for $i = 1, 2, \ldots, N-1$, then it is 1.5-optimal.*

(ii) *The modified echelon $(\hat{\boldsymbol{r}}, \hat{\boldsymbol{Q}})$ policy is asymptotically optimal, as Q_{i+1}^*/Q_i^* goes to infinity for any $i = 1, 2, \ldots, N-1$.*

Summary

We consider the classical serial inventory system with N stages, Poisson demand, fixed shipment costs and positive leadtimes.

- Introduce modified echelon (r, Q) policy
- Establish an upper bound on the systemwide cost under an arbitrarily given modified echelon (r, Q) policy
- Demonstrate that the effectiveness of a modified echelon $(\hat{r}, \hat{Q})$ heuristic depends on the ratios between Q_i^* and Q_{i+1}^* that are obtained from decomposed single-stage problems
  - If $Q_i^* \leq Q_{i+1}^*$, the heuristic is 1.25-optimal for two stages and 1.5-optimal for multiple stages
- Obtain some asymptotical optimality results

Features of Modified Echelon (r, Q) Policy

- Provably primitive-dependent performance bounds
- Simple and efficiently computable (e.g., free from the clustering step required in other heuristics)
- Numerically perform well
- Asymptotical optimal

Q & A

Thank you!