Critical Level Policies in Lost Sales Inventory Systems with Different Demand Classes

Aleksander Wieczorek[1,4] Ana Bušić[1] Emmanuel Hyon[2,3]

[1]INRIA/ENS, Paris, France

[2]Université Paris Ouest Nanterre, Nanterre, France

[3]LIP6, UPMC, Paris, France

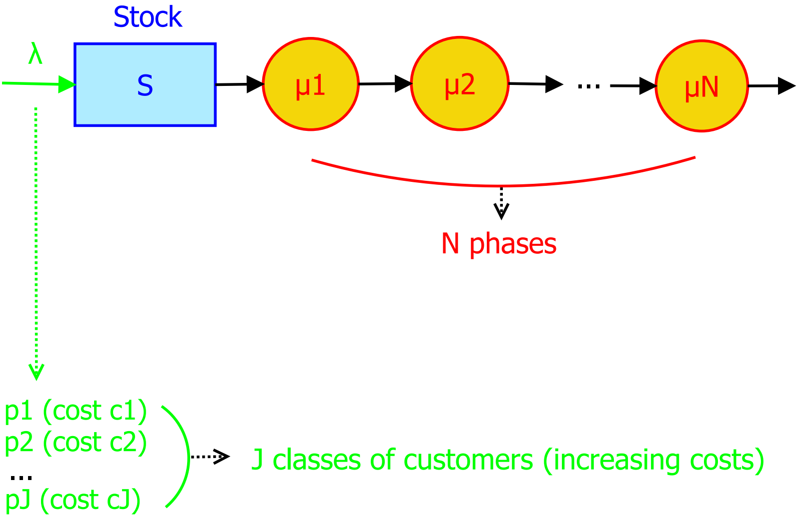
[4]Institute of Computing Science, Poznan University of Technology, Poznan, Poland

EPEW, Borrowdale, UK, October 13, 2011

Table of Contents

1. Markov Decision Processes
   - Definition
   - Optimal control
2. Model description
   - Admission control
   - Policies
   - Results
3. Extensions

Model presentation



Plan

1. Markov Decision Processes
   - Definition
   - Optimal control
2. Model description
   - Admission control
   - Policies
   - Results
3. Extensions

Markov Decision Process

Formalism and notation [3]

A collection of objects $(\mathcal{X}, \mathcal{A}, p(y|x,a), c(x,a))$ where:

$$
\begin{aligned}
&\mathcal{X} \text{ — state space,}\\
&\quad \mathcal{X} = \{1,\dots,S\} \times \{1,\dots,N\} \cup \{(0,1)\},\\
&\quad \forall (x,k) \in \mathcal{X}\ x \text{ — replenishment, } k \text{ — phase,}\\
&\mathcal{A} \text{ — set of actions,}\\
&\quad \mathcal{A} = \{0,1\},\\
&\quad 1 \text{ — acceptance, } 0 \text{ — rejection,}\\
&p(y|x,a) \text{ — probability of moving to state } y \text{ from state } x\\
&\qquad \text{when action } a \text{ is triggered,}\\
&c(x,a) \text{ — instantaneous cost in state } x \text{ when action } a \text{ is triggered.}
\end{aligned}
$$

Plan

1. Markov Decision Processes
   - Definition
   - Optimal control
2. Model description
   - Admission control
   - Policies
   - Results
3. Extensions

Optimal control problem

Policy

A policy π is a sequence of decision rules that maps the information history (past states and actions) to the action set $\mathcal{A}$.

Markov deterministic policy

A Markov deterministic policy is of the form $(a(\cdot), a(\cdot), \ldots)$ where $a(\cdot)$ is a single deterministic decision rule that maps the current state to a decision (hence, in our case $a(\cdot)$ is a function from $\mathcal{X}$ to $\mathcal{A}$).

Optimal control problem — optimality criteria

Minimal long-run average cost

$$\bar{v}^* = \min_{\pi} \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_y^{\pi}\left(\sum_{\ell=0}^{n-1} C(y_\ell, a_\ell)\right)$$

Policies π^* optimising some optimality criteria are called optimal policies (with respect to a given criterion).

Goal: characterise optimal policy π^* that reaches $\bar{v}^*$.

Optimal control problem — optimality criteria

Minimal (expected) *n*-stage total cost

$$V_n(y) = \min_{\pi(n)} \mathbb{E}_y^{\pi(n)} \left(\sum_{\ell=0}^{n-1} C(y_\ell, a_\ell) \right), \; y \in \mathcal{X}, \; y_0 = y$$

Convergence results [2], [3, Chapter 8]

The minimal *n*-stage total cost value function V_n does not converge when n tends to infinity.
The difference $V_{n+1}(y) - V_n(y)$ converges to the minimal long-run average cost ($\bar{v}^*$).

Relation between different optimality criteria [2], [3, Chapter 8]

The optimal *n*-stage policy (minimizing V_n) tends to the optimal average policy π^* (minimizing $\bar{v}^*$) when n tends to infinity.

Cost value function

Bellman equation

$V_{n+1} = TV_n$ where T is the dynamic programming operator:

$$(Tf)(y) = \min_a (\hat{T}f)(y,a) = \min_a \left(C(y,a) + \sum_{y' \in \mathcal{X}} \mathbb{P}\,(y'|(y,a))\, f(y') \right),$$

Decomposition of T

The dynamic programming equation is:

$$V_n(x,k) = T_{unif}\left(\sum_{i=1}^{J} p_i T_{CA(i)}(V_{n-1}), T_D(V_{n-1}) \right), \quad (1)$$

where $V_0(x,k) \equiv 0$ and T_{unif}, $T_{CA(i)}$ and T_D are the different event operators.

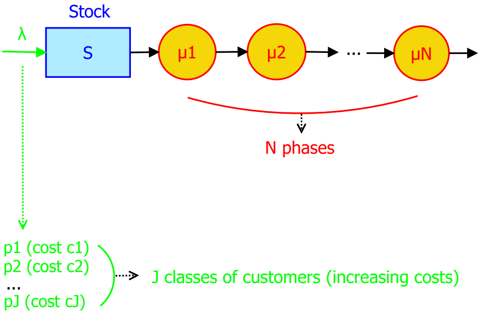
Plan

1 Markov Decision Processes
- Definition
- Optimal control

2 Model description
- Admission control
- Policies
- Results

3 Extensions

Description of operators



Controlled arrival operator of a customer of class i, $T_{CA(i)}$

$$T_{CA(i)}f(x,k) = \begin{cases} \min\{f(x+1,k),\ f(x,k)+c_i\} & \text{if } x < S, \\ f(x,k)+c_i & \text{if } x = S. \end{cases}$$

Description of operators

Let $\mu'_k = \mu_k / \alpha$.

Departure operator, T_D

$$T_D f(x,k) = \mu'_k \begin{cases} f(x, k+1) & \text{if } (k < N) \text{ and } (x > 0), \\ f((x-1)^+, 1) & \text{if } (k = N) \text{ or } (x = 0 \text{ and } k = 0) \end{cases} + (1 - \mu'_k) f(x,k).$$

Uniformization operator, T_{unif}

$$T_{unif}(f(x,k), g(x,k)) = \frac{\lambda}{\lambda + \alpha} f(x,k) + \frac{\alpha}{\lambda + \alpha} g(x,k).$$

Plan

1 Markov Decision Processes
- Definition
- Optimal control

2 Model description
- Admission control
- Policies
- Results

3 Extensions

Critical level policies

Definition (Critical level policy)

A policy is called a critical level policy if for any fixed k and any customer class j it exists a level $t_{k,j}$ in x, depending on phase k and customer class j, such that in state (x, k):
- for all $0 \leq x < t_{k,j}$ it is optimal to accept any customer of class j,
- for all $x \geq t_{k,j}$ it is optimal to reject any customer of class j.

Structural properties of policies

Assume a critical level policy and consider a decision for a fixed customer class j.

Definition (Switching curve)

For every k, we define a level $t(k) = t_{k,j}$ such that when we are in state (x, k) decision 1 is taken if and only if $x < t(k)$ and 0 otherwise. The mapping $k \mapsto t(k)$ is called a *switching curve*.

Definition (Monotone switching curve)

We say that a decision rule is of the monotone switching curve type if the mapping $k \mapsto t(k)$ is monotone.

Example — critical levels, switching curve

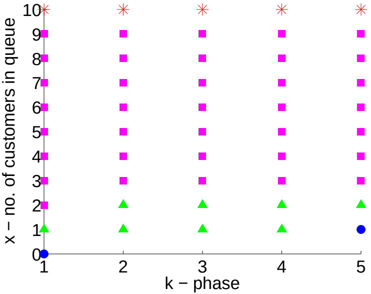


Figure: Acceptance points for different customer classes. Blue circle — all classes are accepted, green triangle — classes 2 and 3 are accepted, pink square — only class 3 is accepted, red asterisk — rejection of any class.

Properties of value functions

Definition (Convexity)

f is convex in x (denoted by $Convex(x)$) if for all $y = (x, k)$:

$$2f(x+1,k) \leq f(x,k) + f(x+2,k)\,.$$

Definition (Submodularity)

f is submodular in x and k (denoted by $\mathrm{Sub}(x,k)$) if for all $y = (x, k)$:

$$f(x+1,k+1) + f(x,k) \leq f(x+1,k) + f(x,k+1)\,.$$

Theorem (Th 8.1 [2])

Let $a(y)$ be the optimal decision rule:
i) If $f \in Convex(x)$, then $a(y)$ is decreasing in x.
ii) If $f \in Sub(x,k)$, then $a(y)$ is increasing in k.

Properties of value functions

Definition (Convexity)

f is convex in x (denoted by $Convex(x)$) if for all $y = (x, k)$:

$$2f(x+1,k) \leq f(x,k) + f(x+2,k)\,.$$

Definition (Submodularity)

f is submodular in x and k (denoted by $\mathrm{Sub}(x, k)$) if for all $y = (x, k)$:

$$f(x+1,k+1) + f(x,k) \leq f(x+1,k) + f(x,k+1)\,.$$

Theorem (Th 8.1 [2])

Let $a(y)$ be the optimal decision rule:
i) If $f \in Convex(x)$, then $a(y)$ is decreasing in x.
ii) If $f \in Sub(x, k)$, then $a(y)$ is increasing in k.

Plan

1. Markov Decision Processes
   - Definition
   - Optimal control
2. Model description
   - Admission control
   - Policies
   - Results
3. Extensions

Properties of the value function of the model

Let V_n be a n-steps total cost value function satisfying the definition of the model. $V_n(x,k) = T_{cost}\left(T_{unif}\left[\sum_{i=1}^{J} p_i T_{CA(i)}(V_{n-1}), T_D(V_{n-1})\right]\right)$,

Lemma

For all $n \geq 0$, V_n is in $Incr(\preceq) \cap AConvex(x) \cap Convex(x)$.

Lemma

For all $n \geq 0$ V_n is in $Sub(x,k) \cap BSub(x,k)$.

$f \in$ AConvex(x) means that $\forall k \in \{1, .., N\}$ $\quad f(0,1) + f(2,k) \geq 2f(1,k)$

$f \in$ BSub(x,k) means that $\forall 0 < x < S$ $\quad f(x,1) + f(x,N) \leq f(x-1,1) + f(x+1,N)$

Proofs are done by checking the preservation of all the properties by all the operators.

Main structural results

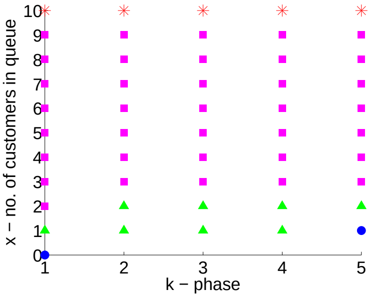
Theorem

The optimal policy is a critical level policy.

Theorem

For any critical level policy, if the rejection costs are nondecreasing ($c_1 \leq \cdots \leq c_J$), then the levels $t_{k,j}$ are nondecreasing with respect to customer class j, i.e. $t_{k,j} \leq t_{k,j+1}$.

Proofs: convexity (+ convergence).

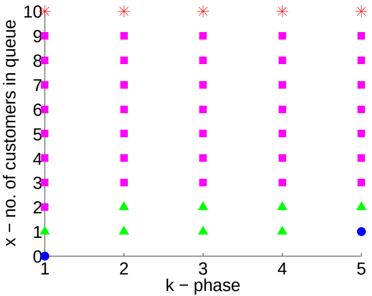


Main structural results

Theorem

The optimal policy defines an increasing switching curve.

Proof: submodularity (+ convergence).



Hyperexponential model

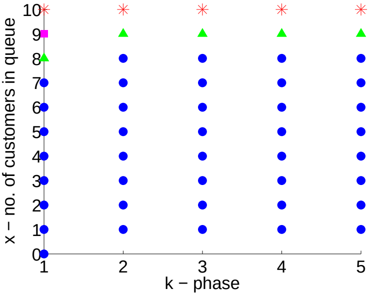


Figure: Acceptance points for different customer classes. Blue circle — all classes are accepted, green triangle — classes 2 and 3 are accepted, pink square — only class 3 is accepted, red asterisk — rejection of any class.

Extensions

Holding costs

The addition of holding costs breaks similarities between queueing models and inventory systems.

Holding cost operator, T_{cost}

$$T_{cost} f(x,k) = \frac{x}{\lambda + \alpha} + f(x,k)$$

Universality of the approach

The same reasoning can be applied to queueing models with holding costs resulting in the same properties of optimal policies.

References

A.Y. Ha.
Inventory rationing in a make-to-stock production system with several demand classes and lost sales.
Management Science, 43(8):1093–1103, 1997.

G. Koole.
Monotonicity in Markov reward and decision chains: Theory and applications.
Foundation and Trends in Stochastic Systems, 1(1), 2006.

M. Puterman.
Markov Decision Processes Discrete Stochastic Dynamic Programming.
Wiley, 2005.